











**MALLA REDDY ENGINEERING COLLEGE**  
(Autonomous)

**Faculty Publications in the academic year 2015-2016**

S.No	Department	Conferences		International Journals	Total
		International	National		
1	Civil Engineering	03	Nil	36	<b>39</b>
2	Electrical and Electronics Engineering	11	Nil	20	<b>31</b>
3	Mechanical Engineering	Nil	Nil	22	<b>22</b>
4	Electronics and Communication Engineering	01	Nil	19	<b>20</b>
5	Computer Science and Engineering	20	02	43	<b>65</b>
6	Humanities & Sciences	01	Nil	09	<b>10</b>
7	Master of Business Administration	07	01	01	<b>09</b>
<b>Total</b>		<b>43</b>	<b>03</b>	<b>150</b>	<b>196</b>



# JOURNALS



































**Department of  
Civil Engineering**















































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































































## REFERENCES

- 1] L. Labecki, L.C.Ganippa. Effect of injection parameters and EGR on Combustion and emission characteristics of rapeseed oil and its blends in diesel engines, april2012, fuel 98(2012) 15-28.
- 2] Orkun Ozener, Levent Yuksek, Alp Tekin Ergenic, Muhammer Ozkan. Effect of soya bean biodiesel on a DI engine performance, emissions and combustion characteristic, december2012.
- 3] Sanjay patil, Dr. M.M.akarte. Effect of injection pressure on CI engine Performance fuelled with biodiesel and its blends, International journal of Scientific & Engineering Research, volume3, issue3, march2012, ISSN 2229-5518.
- 4] G.R. Kanan, R.Anand. Effect of injection pressure and injection timing on DI diesel engine fuelled with biodiesel from waste cooking oil, august2012, Biomass and Bioenergy 46 (2012) 343-352.
- 5] Cenk sayin, metin gumus, Mustafa canakci. Effect of fuel injection pressure on the injection, combustion and performance characteristics of a DI diesel engine fueled with canola oil methyl esters-diesel fuel blends. Biomass and Bioenergy 46 (2012) 435-446.
- 6] Venkanna BK, Reddy CV, W adawadagi SB. Performance. Emission and combustion characteristics of direct injection diesel engine running on rice bran oil/diesel fuel blend. Int J chem. Biomol Engg 2009; 2-131-7.



**International Journal of Innovative Research in Science,  
Engineering and Technology**

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2015

# **Experimental Analysis on Friction Stir Welding of Aluminium Alloy 6082**

E.V.Renuka Rao<sup>1</sup>, Heruthunnisa Shaik<sup>2</sup>

P.G. Student, Department of Mechanical Engineering, Mallareddy Engineering College, Hyderabad, India<sup>1</sup>

Assistant Professor, Department of Mechanical Engineering, Mallareddy Engineering College, Hyderabad, India<sup>2</sup>

**ABSTRACT:** Friction stir welding is a solid state welding process used for welding similar and dissimilar materials. In the current work, friction stir welding is carried out on aluminium alloy AA6082. The friction stir welding process parameters such as tool rotation speed, tool travel speed play an important role in deciding the joining quality. These parameters are adjusted so that the interface is heated into the plastic temperature range where welding can take place. The work has been carried out to study the effect of process parameters like tool travel speed and tool rotation speed on the tensile strength of the weldments.

**KEYWORDS:** Friction Stir Welding, Process Parameters, Tensile Strength

## **I. INTRODUCTION**

Friction stir welding is also a solid state welding processes; this remarkable upgradation of friction welding was invented in 1991 in The Welding Institute (TWI). In friction stir welding (FSW) a cylindrical, shouldered tool with a profiled probe is rotated and slowly plunged into the joint line between two pieces butted together. The parts have to be clamped onto a backing bar in a manner that prevents the abutting joint faces from being forced apart. Frictional heat is generated between the wear resistant welding tool and the material of the work pieces. This heat causes the latter to soften without reaching the melting point and allows traversing of the tool along the weld line. The maximum temperature reached is of the order of 80% of the melting temperature of the material. The plasticized material is transferred from the leading edge of the tool to the trailing edge of the tool probe and is forged by the intimate contact of the tool shoulder and the pin profile. It leaves a solid phase bond between the two pieces. The process can be regarded as a solid phase keyhole welding technique since a hole to accommodate the probe is generated, then filled during the welding sequence.

The advantages of FSW technique is that it is environment friendly, energy efficient, there is no necessity for gas shielding for welding Aluminium. Mechanical properties as proven by fatigue, tensile tests are excellent. There is no fume, no porosity, no spatter and low shrinkage of the metal. Joining dissimilar and previously unweldable metals can be attempted by this unique process.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2015

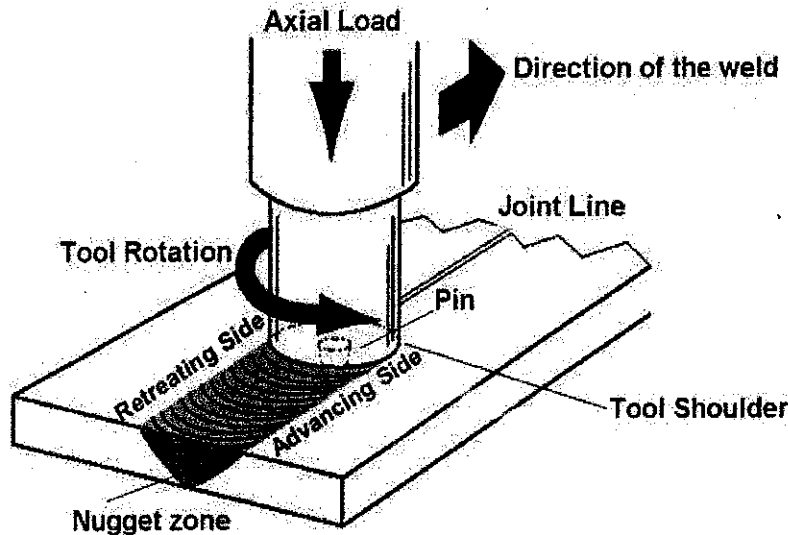


Figure 1 Schematic diagram of FSW process

Figure 1 show briefly about friction stir welding process and process parameters involved during friction stir welding process like axial load, tool rotation and direction of weld etc.

## II. LITERATURE REVIEW

Amancio-Filhoa et al. (2007) described that aircraft Aluminium alloys generally present low weld ability by traditional fusion welding process. The development of the friction stir welding has provided an alternative improved way of satisfactory producing Aluminium joints, in a faster and reliable manner.

Riahi (2010) in this research residual stress is lower in friction stir welding (FSW) compared with other melting weldment processes. This is due to being solid-state process in its nature. There are several advantages in utilizing stir welding process. Lower fluctuation and shrinkage in weldment metal enhanced mechanical characteristics, less defects, and ability to weld certain metals otherwise impractical by other welding processes are to name just a few of these advantages. In the prediction of results of residual stress, only heat impact was studied. This was recognized as the main element causing minor difference in results obtained for simulation in comparison with that of actual experiment.

Liu a (2013) In their research, the 4 mm thick 6061-T6 Aluminium alloy was self-reacting friction stir welded at a constant tool rotation speed of 600 r/min. The specially designed self-reacting tool was characterized by the two different shoulder diameters. The results of transverse tensile test indicated that the elongation and tensile strength of joints increased with increasing welding speed. The defect-free joints were obtained at lower welding speeds and the tensile fracture was located at the heat affected zone adjacent to the thermal mechanically affected zone on the advancing side.

## III. EXPERIMENTAL WORK

The 6xxx alloys are heat treatable, and have moderately high strength coupled with excellent corrosion resistance. They are readily welded. A unique feature is their extrudability, making them the first choice for architectural and structural members where unusual or particularly strength- or stiffness-criticality is important.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2015

The chemical composition of AA6082 is shown in Table1. All the elements are shown with percentage wise. The major elements involved in AA6082 are silicon, ferrous, copper, manganese, magnesium, zinc, titanium and chromium.

Element	% present
Si	0.7 – 1.3
Fe	0.0 – 0.5
Cu	0.0 – 0.1
Mn	0.4 – 1.0
Mg	0.6 – 1.2
Zn	0.0 – 0.2
Ti	0.0 – 0.1
Cr	0.0 – 0.25
Al	Balance

Table 1: Chemical composition of AA6082

In present experiment we select 4mm thickness of Aluminium sheet. Aluminium sheet was cut with the dimensions of 4mmx70mmx100mm using shearing machine and the edges of the plate are filed to get smooth surface finish. After the plates are cut into required sizes, fix the plates side by side on work table rigidly by clamps to form butt joint. The tool material selected for this experiment is H13 Tool Steel due to its excellent combination of high toughness and fatigue resistance. Tool profile is threaded pin profile.

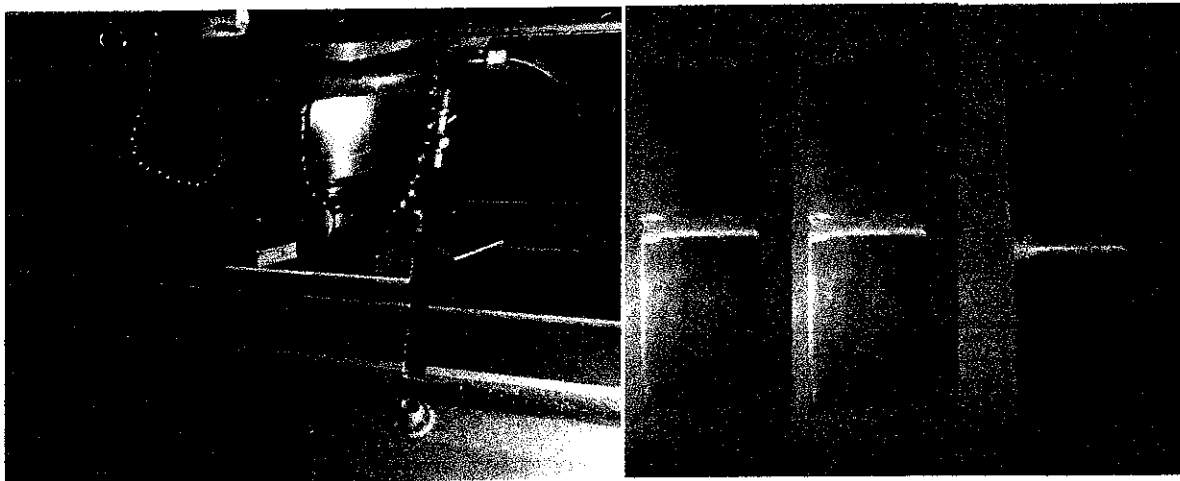


Figure 2 Experimental set up and weld pieces

Figure 2 shows the experimental set up on CNC milling machine. The aluminium alloy work pieces are rigidly clamped on work table. By using Tool with a threaded pin profile experiment is done with the combination of welding process parameters. After welding the weld pieces are shown in figure.

Table 2 shows the combinations of process parameters, friction stir welding experiment were done on CNC Vertical Milling Machine. By following the sequence in below table we do experiment one by one. The speed and feed (Tool Travel Speed) are increased with specified incremental rate.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2015

Specimen Number	Speed (RPM)	Feed (mm/minute)
01.	750	50
02.	750	100
03.	1000	50
04.	1000	100
05.	1250	50
06.	1250	100

Table 2: Process parameters for FSW

## IV. EXPERIMENTAL RESULTS

The weld pieces are subjected to tensile tests on a universal testing machine to evaluate the strength of welded joints. The weld pieces are prepared as per ASTM E8/E8M standard. The results of tensile test which were observed during the experiment were tabulated which are shown in below table 3. Increase in rotational speed has resulted in increase of tensile strength. Primary reason is, the higher the speed, the higher will be the deformation and heat generation in the weld. This will result in finer grain structures, because of which tensile strength is increases.

Specimen Number	Speed (RPM)	Feed (mm/minute)	UTS (Mpa)
01.	750	50	112.1
02.	750	100	130.0
03.	1000	50	139.0
04.	1000	100	145.3
05.	1250	50	145.7
06.	1250	100	152.4

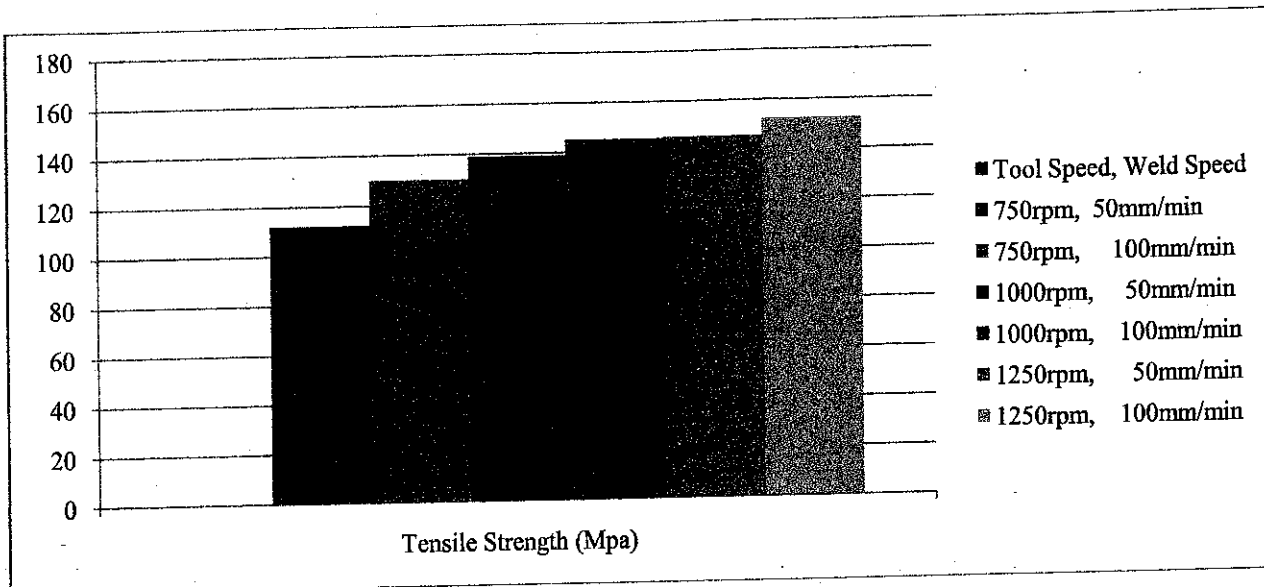
Table 3: Ultimate Tensile Strength Results

Table 3 shows the ultimate tensile test results coming from tensile test done on universal testing machine. With increasing tool speed and tool travel speed the ultimate tensile strength is increased. Also with increasing tool speed and tool travel speed the surface finish is increased.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2015



Graph 1: Graphical Representation of Tensile Test Results

## V. CONCLUSION

The following conclusions can be drawn from present study

1. Tensile strength is found to increase with increase in rotational speed and weld speed. Maximum Tensile strength of 152 Mpa was observed at 1250 rpm (for 100 mm/min feed). This indicates that for AA6082 Al Alloys, higher range of rotational speed is best suited to achieve maximum tensile strength.
2. Surface finish is increased by increasing tool rotation speed and tool travel speed.

## REFERENCES

- [1] R.S Mishra, Z.Y.Ma, Friction stir welding and processing, Materials Science and Engineering, R 50 (2005) 1-78.
- [2] W.M Thomas, E.D. Nicholas, S.D. Smith, in: S.K Das, J.G. Kaufman, T.J. Lienett (Eds.), Aluminium 2001 – Proceedings of the TMS 2001 Aluminium Automotive and Joining Sessions, TMS, 2001, p.213.
- [3] K. Kumar, Satish V. Kailas, The role of friction stir welding tool on material flow and weld formation Materials Science and Engineering: A, Volume 485, Issues 1-2, Pages 367-374, 25 June 2008.
- [4] Karan Singla, Dr. Deepak Bhardwaj, Vikas Burman, International journal of research in Aeronautical and Mechanical engineering.
- [5] Yousuf ul haq, Influence of Friction Stri Welding on parameter and strength of weldments, pages 47-56, April 2009.
- [6] For example, edited by Japan Welding Society, Welding Handbook, Maruzen, 1990, p. 496.
- [7] Ákos Meilinger, The Importance Of Friction Stir Welding Tool, Inre Török University of Miskolc, Department of Mechanical Technology.
- [8] V. Balasubramanian et al, (2010). "An assessment of microstructure, hardness, tensile and impact strength of friction stir welded ferritic stainless steel joints" Materials and Design. 3: 4592–4600.





## OPTIMIZATION OF CAR AIR CONDITIONING SYSTEM USING ENERGY FROM EXHAUST GASES

RATHLAVATH RAJENDER<sup>1</sup>, BHARADWAJA .K<sup>2</sup>

<sup>1</sup>M.Tech Scholar, <sup>2</sup>Asst.Professor

Department of Mechanical Engineering  
Malla Reddy Engineering College (Autonomous)  
Hvderabad. India



RATHLAVATH  
RAJENDER

### ABSTRACT

Due to the international attempt to find alternative energies, absorption refrigeration has become a prime system for many cooling applications. Where thermal energy is available the absorption refrigerator can very well substitute the vapour compression system. It is a well-known fact that a large amount of heat energy associated with the exhaust gases from an engine is wasted. A rough energy balance of the available energy in the combustion of fuel in a motor car engine shows that one third is converted into shaft work, one third is lost at the radiator and one third is wasted as heat at the exhaust system. Even for a relative small car-engine, 15 kW of heat energy can be utilized from the exhaust gas. This heat is enough to power an absorption refrigeration system to produce a refrigeration capacity of 5 kW. In this thesis, energy from the exhaust gas of an internal combustion engine is used to power an absorption refrigeration system to air-condition an ordinary passenger car. In this thesis an absorption refrigeration system is designed and parts of the refrigeration system are analyzed. Modeling is done in Pro/Engineer and analysis is done in Ansys. Thermal analysis is done on condenser and evaporator using different materials Copper, Aluminum alloy 204, Brass and Stainless Steel.

**Key words:** Cooling Fluid, Cylinder fins, Aluminum Alloy 204, Aluminum Alloy 6061, Magnesium

©KY PUBLICATIONS

### INTRODUCTION

#### INTRODUCTION TO AIR CONDITIONING

Air conditioning is the process of altering the properties of air (primarily temperature and humidity) to more favorable conditions. More generally, air conditioning can refer to any form of technological cooling, heating, ventil. An air conditioner (often referred to as air con, AC or A/C, and not to be confused with the abbreviation for alternating current) is a major or home appliance, system, or mechanism designed to change the air

temperature and humidity within an area (used for cooling and sometimes heating depending on the air properties at a given time). The cooling is typically done using a simple refrigeration cycle, but sometimes evaporation is used, commonly for comfort cooling in buildings and motor vehicles. In construction, a complete system of heating, ventilation and air conditioning is referred to as "HVAC".

## EXPERIMENTAL IC ENGINE

In an IC engine, fuel (usually petrol or diesel) is combusted inside the cylinder due to which the piston moves outward and rotates the crank, and hence the engine produces work. In IC engines the combustion of the fuel produces heat, which is converted to mechanical work using the piston and crank arrangement. From the heat produced from combustion of fuel only 30% (approx) of heat is converted into useful mechanical work. The remaining heat energy is wasted into the atmosphere in the form of:

- (i) Heat carried away by the cooling water,
- (ii) Heat taken away by the exhaust gases,
- (iii) Heat carried away by the lubricating oil,
- (iv) Heat lost by radiation.

The cooling water and exhaust gases carry away the maximum amount of heat from the engine, i.e. around 60% (approx). This heat is called the low grade energy of the engine.

## THE ENGINE

Let us consider an engine of an automobile on which the vapour absorption refrigeration system is to be implemented.

## Engine parameters

The IC engine based on which the calculations are done is

- Make - Hindustan Motors
- Model - Ambassador
- No of cylinders,  $n = 4$ .
- Power,  $P = 60$  bhp at 2000 rpm.
- Capacity,  $V = 1717$ cc.
- No of strokes = 4.
- Fuel used = diesel.
- Air-fuel ratio,  $A/F = 15:1$

## WASTE HEAT OF THE ENGINE

The main two areas through which the heat is exhausted into the atmosphere from the engine are the cooling water and the exhaust gases. It is necessary to calculate the amount of heat energy carried away by the exhaust gases and the cooling water.

## Exhaust gas heat

Mass flow rate air into the cylinder,  
$$m_a = V N E_{vo} i / 2$$

flow rate of fuel,

$$m_f = m_a / (A/F \text{ ratio})$$

Flow rate of exhaust gas

$$m_e = m_a + m_f$$

Heat available at exhaust pipe

$$Q_e = m_e C_{p_e} (t_e - t_a)$$

## Cooling water heat

Heat carried away by cooling water

$$Q_w = m_w C_{p_w} (t_{co} - t_{ci})$$

## FINAL VALUES

To produce 0.5 TR inside the automobile cabin it is required to have

- (a) mass flow rate of refrigerant,  $m_r = 0.71$  gm/s
- (b) mass flow rate of strong solution from absorber generator,  $m_g = 8.804$  gm/sec
- (c) mass flow rate of weak solution from generator to absorber,  $m_a = 8.094$  gm/sec
- (d) Co-efficient of Performance,  $COP = 0.2$

## DESIGN

Designing involves developing each components of the system that has to be installed on to the automobile to produce the required cooling effect which involves generator, condenser and evaporator

## Design of pre-heater

Outside Diameter of the tube,  $D_o = 0.012$  m

Inside Diameter of the tube,  $D_j = 0.01$  m

Length of the tube,  $L = 2$  m

## Design of generator

Outside Diameter of the exhaust gas tube,

$$D_o = 0.04 \text{ m}$$

Taking inside diameter of the exhaust gas tube,  $D_i = 0.038$  m

Length of the tube required for the required heat transfer,  $L = 1$  m

## Design of condenser

### Dimensions of the designed condenser

Diameter of the tube,  $d = 0.018$  m Thickness of the tube,  $a = 0.005$  m Length of the tube,  $L = 7.45$  m

## Design of Evaporator

Outside Diameter of the tube,  $D_o = 0.01$  m

Inside Diameter of the tube,  $D_j = 0.008$  m Length of the tube,  $L = 6.26$  m

## Design of absorber

Outside diameter of the absorber,  $D_o = 76$  mm Total length of the absorber,  $L = 205$  mm Outer diameter of the fins,  $D_f = 109$  mm No. of fins,  $n = 7$

## ASSEMBLY OF VAPOUR ABSORPTION SYSTEM

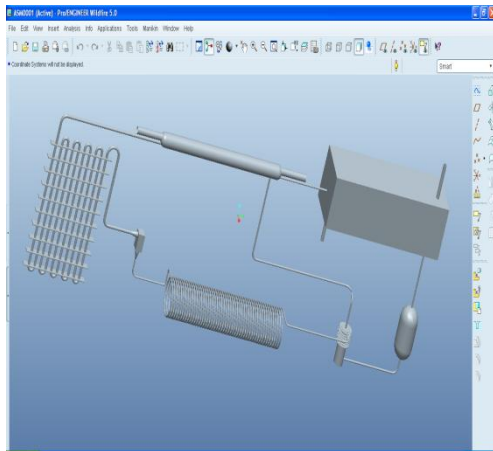


Fig 2.1 Vapour absorption system

## 2D DRAWINGS

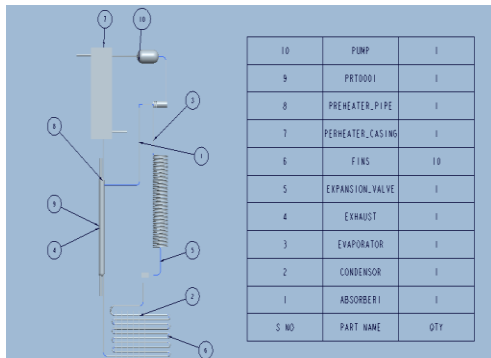


Fig 2.2 2D- Drawing of Vapour absorption system

## THERMAL ANALYSIS OF CONDENSER

### BRASS

Material Properties:

Thermal Conductivity – 200 W/mK

Specific Heat – 350 J/Kg K

Density - 0.0000850 Kg/mm<sup>3</sup>

Nodal temperature

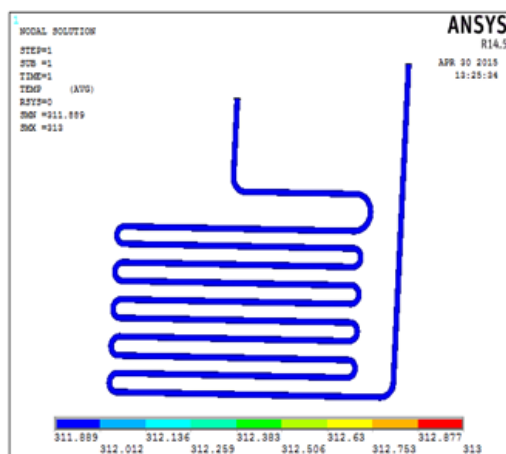


Fig 2.3 Nodal temperature of Condenser

## Thermal Gradient

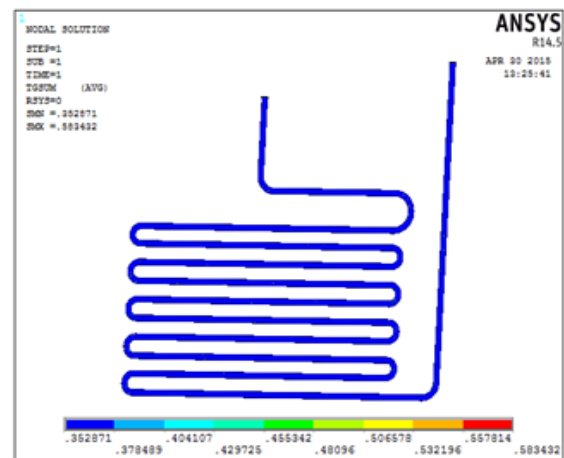


Fig 2.4 Thermal Gradient of Condenser

## Thermal Flux

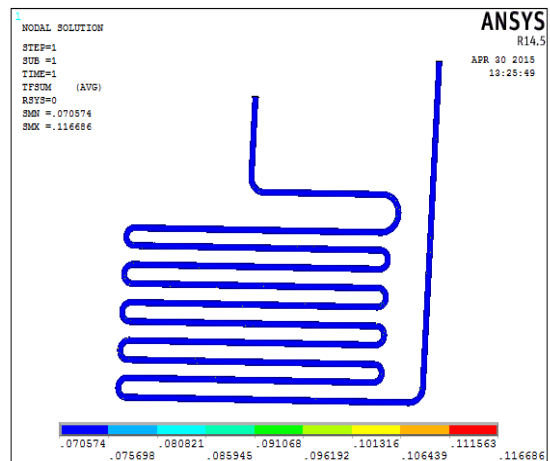


Fig 2.2 Thermal flux Condenser

## EVAPORATOR

### BRASS

Nodal Temperature

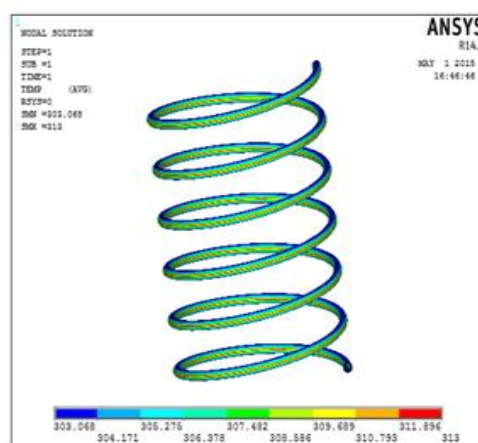


Fig 3.1 Nodal temperature of evaporator

	Temp eratur e (K)	Thermal gradient (K/mm)	Heat Flux (W/mm <sup>2</sup> )
Coper	313	0.308508	0.118776
Aluminiu m-204	313	0.768531	0.11528
Brass	313	0.583432	0.116686
Stainless Steel	313	0.577865	0.116729

Thermal Flux  
Thermal Gradient  
CONDENSER

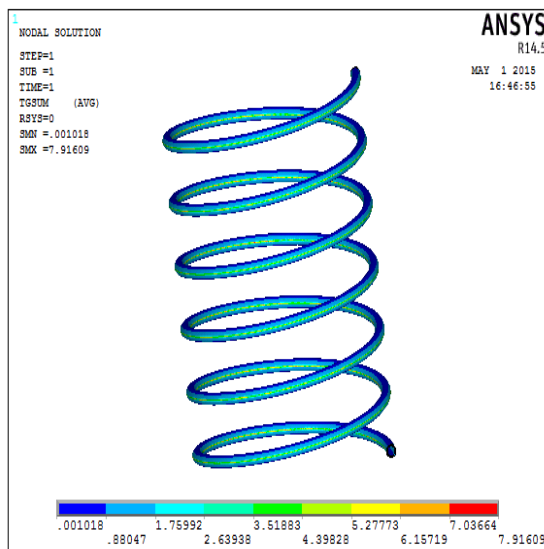


Fig 3.2 Thermal Gradient of evaporator  
EVAPORATOR

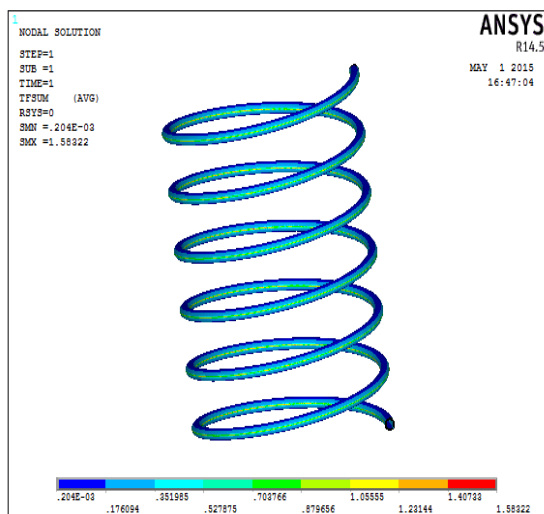


Fig 3.3 Thermal flux of evaporator

#### RESULT TABLE

	Temperature (k )	Thermal gradient( Kg/mm)	Heat Flux (W/mm <sup>2</sup> )
Copper	313	4.6619	1.79483
Aluminium- 204	313	9.77183	1.46577
Brass	313	7.91609	1.58322
Stainless Steel	313	7.91609	0.557693

#### CONCLUSION

In the exhaust gases of motor vehicles, there is enough heat energy that can be utilised to power an air -conditioning system. Therefore, if air -conditioning is achieved without using the engine's mechanical output, there will be a net reduction in fuel consumption and emissions.

Once a secondary fluid such as water or glycol is used, the aqua-ammonia combination appears to be a good candidate as a working fluid for an absorption car air-conditioning system. This minimises any potential hazard to the passengers.

The low COP value is an indication that improvements to the cycle are necessary. A high purity refrigerant would give a higher refrigeration effect, while the incorporation of a solution heat exchanger would reduce the input heat to the generator. The present system has both a reflux condenser and a heat exchanger.

However, the reflux condenser is proved inadequate to provide high purity of the refrigerant and needs to be re-addressed. The evaluation of the COP, with and without the heat exchanger also proves that unless there is a high purity refrigerant, the effect of the heat exchanger to the generator's heat is small.

In this project a pre heater, condenser, generator, evaporator, absorber for required capacity are designed by using theoretical calculations. Thermal analysis is conducted on condenser and evaporator by using materials copper, aluminum alloy 204, Brass and Stainless steel y using refrigerant water ammonia mixture.

By observing the analysis results, thermal flux is more for Copper or both so heat transfer rate is more for both condenser and evaporator.

#### REFERENCES

- [1]. Ananthanarayanan P N (2005) 'Refrigeration & Air Conditioning', Tata McGraw-Hill. 3<sup>rd</sup> edition pp. 398 - 424.
- [2]. Arora ; Domkundwar. (2004) 'A course in Refrigeration & Air Conditioning', Dhanpat Rai & Co., 7<sup>th</sup> edition, pp. 6.1 - 6.23.
- [3]. Arora C P (2002) 'Refrigeration & Air Conditioning', Tata McGraw Hill., 2<sup>nd</sup> edition, pp. 301 - 314, 339 - 356, 427 - 456.
- [4]. Ballaney P L (2003), Refrigeration & Air Conditioning' Khanna Publishers., 13<sup>th</sup> edition, pp. 483 - 542.
- [5]. Domkundwar & Domkundwar. (2005) 'Heat & Mass Transfer- Data Book', Dhanpat Rai & Co., 2<sup>nd</sup> edition,
- [6]. Horuz I (August 1999), 'Vapor Absorption Refrigeration in Road Transport Vehicles', Journal of Energy Engrg., Volume 125, Issue 2, pp. 48-58.
- [7]. Manohar Prasad. (2000) 'Refrigeration & Air Conditioning', New Age International Publishers., 3<sup>rd</sup> edition, pp. 188 - 225.
- [8]. Manohar Prasad. (2000) 'Refrigeration & Air Conditioning - Data Book'. New Age International Publishers., 2<sup>nd</sup> edition.
- [9]. Kothandraman C P. (2002) 'Steam Tables', New Age International Publishers., 2<sup>nd</sup> edition.
- [10]. Rajput R K (2006) 'A textbook of Internal combustion engines', Laxmi Publications (P) Ltd., 1<sup>st</sup> edition.
- [11]. Rajput R K (2007) 'Heat And Mass Transfer', S. Chand & Company Ltd., 3<sup>rd</sup> edition.



**Department of  
Electronics and  
Communication Engineering**



## Obstacle Detection in Textured Environment by using Color Coherence

V. DATTA KIRAN<sup>1</sup>, DR. M. J. C. PRASAD<sup>2</sup>

<sup>1</sup>PG Scholar, Dept of ECE, Malla Reddy Engineering College, JNTUH, Hyderabad, TS, India,  
E-mail: dattakiran2010@gmail.com.

<sup>2</sup>Associate Professor, Dept of ECE, Malla Reddy Engineering College, JNTUH, Hyderabad, TS, India,  
E-mail: jagdissh@mrec.ac.in.

**Abstract:** Obstacle detection algorithm that makes use of color information and color coherence vectors for robust obstacle detection. The algorithm makes use of color cue to classify a pixel in an image as an obstacle or a path. Color is one of the prominent image features. Color information is readily available as input from a color camera. Our algorithm makes use of coherence vectors for representation and matching instead of histograms. A color histogram provides no spatial information. It merely describes the color information present in an image. Color coherence vectors represent pixels as either coherent or incoherent. Color coherence vectors prevent coherent pixels from getting matched with incoherent pixels. The color histogram cannot make such fine distinction. This paper a novel algorithm for obstacle detection is proposed that makes use of color cue and color coherence vectors for robust obstacle detection. The algorithm detects obstacles based on the appearance of individual pixels. Whether an individual pixel belongs to an obstacle or a navigable path is determined by a color classifier. The coherence vectors are better than histograms as they incorporate spatial information. Moreover, our approach does not have any learning requirement prior to the use of classifier.

**Keywords:** Obstacle Detection Algorithm, Color Histogram, Color Coherence Vectors.

### I. INTRODUCTION

To get the perception of the environment around them, humans depends upon five senses – vision, hearing, smell, touch and taste. Out of these, vision is undoubtedly the one that people depend upon the most for the performance of their everyday activities. Most people cannot even imagine what life would be if they lose it- unable to read, requiring assistance for daily activities and difficulty in recognizing near and dear ones. This is however, a hardcore reality for nearly 45 million people worldwide, who are blind. A further 135 million people have severely impaired vision in both the eyes and additional 100 million people suffer from monocular vision loss. These visually impaired people experience serious difficulties in leading an independent life due to the reduced perception of the environment. Most of them confront serious difficulties in mobility and navigation

when they find themselves in new environment. Obstacle detection is one of the major problems to be solved to ensure safe navigation for visually impaired users. The visually impaired people experience serious difficulties in leading an independent life due to the reduced perception of the environment.

Most of them confront serious difficulties in mobility and navigation when they find themselves in new environment. Obstacle detection is one of the major problems to be solved to ensure safe navigation for visually impaired users. To avoid this problem we proposed this novel algorithm. In this paper, a novel algorithm for obstacle detection is proposed that makes use of color cue and color coherence vectors for robust obstacle detection. The algorithm detects obstacles based on the appearance of individual pixels. Whether an individual pixel belongs to an obstacle or a navigable path is determined by a color classifier. The color classifier makes use of color information which is readily available as input in a color image. Color is one of the prominent image features and easy to process. Our approach makes use of coherence vectors instead of histogram for color representation and matching. The coherence vectors are better than histograms as they incorporate spatial information. Moreover, our approach does not have any learning requirement prior to the use of classifier. Whatever is left of this paper is organized as follows. Segment II presents the literature survey. Segment III locations the proposed system taking into account and HVS color coherence analysis, results and discussion are given in Section IV. At last, conclusions are made in Section VI.

### II. LITERATURE SURVEY

In this chapter, we will discuss about the information found by study and research that is critical and have an important value in the contribution of the whole paper. It also gives some basic knowledge or theoretical base and is used as a foundation to successfully achieve the main objectives. Most of the literatures are from the related articles, journals, books and previous works of the same fields. These literatures are then compiled and use as a guidance to the work of this paper. In the implementation of certain human vision properties in the real time image

processing for vision substitution by acoustic transform of visual information. An experimental prototype has been developed with vision sensor fixed in a headgear and stereo earphone and a laptop computer interconnected. Captured image from vision sensor is processed and transformed into a stereo sound through the earphone. The image processing is critically designed as a model of human vision. It involves edge detection, background suppression, normalization and stepped weightage. The scheme is tested with a blind volunteer and his suggestions in formatting the sound pattern are considered. The software module suits most of the commercially available vision sensors.

In [2], Color histograms are used to compare images in many applications. Their advantages are efficiency, and insensitivity to small changes in camera viewpoint. However, color histograms lack spatial information, so images with very different appearances can have similar histograms. For example, a picture of fall foliage might contain a large number of scattered red pixels; this could have a similar color histogram to a picture with a single large red object. We describe a histogram-based method for comparing images that incorporates spatial information. We classify each pixel in a given color bucket as either coherent or incoherent, based on whether or not it is part of a large similarly-colored region. A color coherence vector (CCV) stores the number of coherent versus incoherent pixels with each color. By separating coherent pixels from incoherent pixels, CCV's provide distinctions than color histograms. CCV's can be computed at over 5 images per second on a standard workstation. A database with 15,000 images can be queried for the images with the most similar CCV's in less than 2 seconds. We show that CCV's can give superior results to color histograms for image retrieval.

### A. Color Coherence Vector

Color Coherence Vector (CCV) is a more complex method than Color Histogram. It classifies each pixel as either coherent or incoherent. Coherent pixel means that it's part of a big of connected component (CC) while incoherent pixel is part of a small connected component. Of course first we define the criteria which we use to measure whether a connected component is big or not.

#### Feature Extraction Algorithm

- Blur the image (by replacing each pixel's value with the average value of the 8 adjacent pixels surrounding that pixel).
- Discretize the color-space (images' colors) into n distinct color.
- Classify each pixel either as coherent or incoherent. This is computed by
  - Find connected components for each discretized color.
  - Determine tau's value (Tau is a user-specified value (Normally it's about 1% of image's size)).
  - Any Connected Component has number of pixels more than or equal to tau then its pixels are considered coherent and the others are incoherent.

- For each color compute two values (C and N).
  - C is the number of coherent pixels.
  - N is the number of incoherent pixels.

It's clear that the summation of all color's C and N = number of pixels.

### B. Matching Function

To compare 2 images a, b.

$C_i$ : number of coherent pixels in color i.

$N_i$ : number of incoherent pixels in color i.

$$D(a, b) = \sum_{i=0}^n (|a_{C_i} - b_{C_i}|) + (|a_{N_i} - b_{N_i}|) \quad (1)$$

Let's take this example to make algorithm's steps clear. Assuming that the image has 30 colors instead of 16777216 colors (256\*256\*256).

13	15	11	17	15
27	10	12	25	22
23	4	25	11	9
29	3	21	23	8
1	9	24	2	7

Now we'll discredited the colors to only three colors (0:9, 10:19, 20, 29).

1	1	1	1	1
2	1	1	2	2
2	0	2	1	0
2	0	2	2	0
0	0	2	0	0

Assuming that our tau is 4

For color 0 we have 2 CC (8 coherent pixels)

For color 1 we have 1 CC (8 coherent pixels)

For color 2 we have 2 CC (6 coherent pixels and 3 incoherent pixels)

So finally our feature vector is

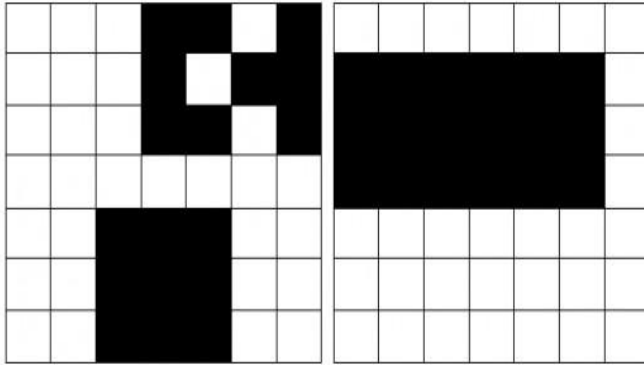
Color	0	1	2
C	8	8	6
N	0	0	3

### C. Drawbacks of Color Coherence Vector

Now we see that Color Coherence Vector method considers information about color spatial distribution between pixels in its coherence component. But this method has some drawbacks. The remaining part of this post will discuss two main drawbacks of it. Coherent pixels in CCV represent the

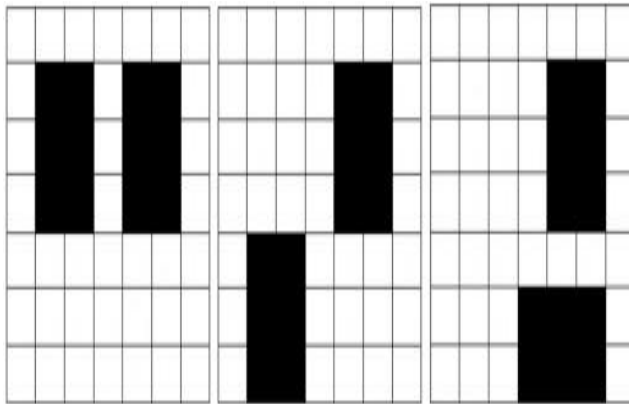
### Obstacle Detection in Textured Environment by using Color Coherence

pixels which are inside remarkable components in image. But what we combined these entire components into one component as shown in Fig.1. We will have only one component the number of its pixels will be obviously equal to the number of pixels in the remarkable components. To make it clear look at these pictures assuming tau equals to 8.



**Fig.1. Combining entire components into one component.**

Although they are different pictures but they have the same CCV. Another problem we may encounter is the positions of these remarkable connected components relative to each other. These pictures have the same CCV with different appearance as shown in Fig.2.



**Fig.2. The above pixels having same CCV with different appearance.**

There are many solutions to these problems. Most of them add another dimension in feature vector which is components' position relative to each other. So this dimension is used in the comparison in order to differentiate between pictures that have the same CCV. The color coherency vector (Color Coherence Vector, CCV) is histogram - method was based for the determination of the similarity between two pictures. The disadvantage of the pure color histograms is the absence of spatial information. Differently appearing pictures can possess very similar color histograms, then for example a picture with a chessboard sample possesses the same histogram as one in that half of the picture is black and the other one knows. Color coherency vectors do not know the exact positions of the objects in the display space, but they can differentiate

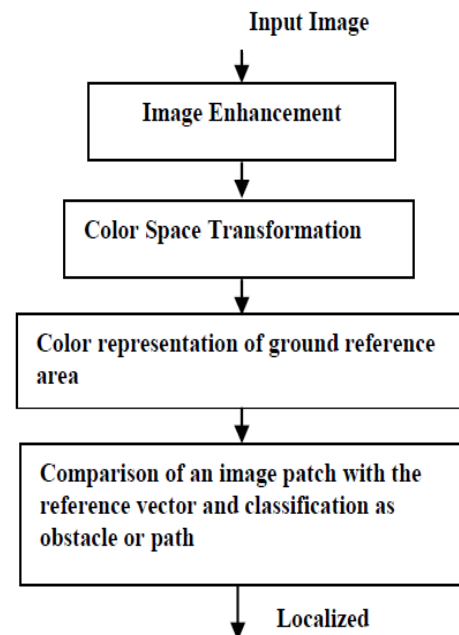
whether a certain color with respect to few large or with respect to many small ranges occur. Thus this method makes a finer distinction possible between pictures contrary to pure color histograms.

### III. PROPOSED METHODOLOGY

The system consists of two parts- storing ground representation in form of color coherence vectors and a classifier. The classifier compares a patch of pixels in the acquired images with the stored ground representation to decide if a patch is a path or an obstacle. The system makes the following assumptions that are reasonable for a variety of indoor environments:

- Ground plane is flat.
- Obstacles differ in appearance from the ground, protrude out from the ground plane and there are no overhanging obstacles.
- Initially, a small area ahead of the user is free from obstacles.

The work assumes that the ground plane is flat and is free from steps, stairs and ramps. Steps, stairs and ramps can be better detected by making use of range based stereo vision instead of appearance based monocular vision. The second assumption allows us to distinguish obstacles from the ground and estimate the distance of obstacles from the user. If the first two assumptions are satisfied then, the distance is monotonically increasing function of the pixel height in the image.



**Fig.3. CCV Algorithm Steps.**

The algorithm consists of steps listed below and shown in Fig.3.

- Image enhancement.
- Transformation to color space.
- Color representation of reference area in form of coherence vectors.

- Comparison of an image patch with reference vectors for obstacle detection.

### A. RGB Color Model

The RGB (Red, Green, Blue) color model is an especially important one in digital image processing because it is used by most digital imaging devices (e.g., monitors and color cameras). In the RGB model, a color is expressed in terms that define the amounts of Red, Green and Blue light it contains. In a 24-bit, color image, pure red would be represented as 255/000/000, where 255 represents the highest level of red light possible, untainted by any green (000) or blue (000) light. Various combinations of the Red, Green and Blue values allow us to define 224 (over 16 million) colors. (See also RGB image class.)

### B. HIS Color Model

The HSI (Hue, Saturation, Intensity) color model describes a color in terms of how it is perceived by the human eye. "Hue" is what an artist refers to as "pigment"; it is what we think of as "color" -- yellow, orange, cyan and magenta are examples of different hues. An artist usually starts with a highly saturated (i.e., pure), and intense (i.e., bright) pigment, and then adds some white to reduce its saturation and some black to reduce its intensity. Red and Pink are two different "saturation" of the same hue, Red. The HSI model is useful when processing images to compare two colors, or for changing a color from one to another. For example, changing a value from Cyan to Magenta is more easily accomplished in an HSI model; only the H value needs to be changed (from 180 to 300). Making the same change in an RGB view is less intuitive; since you must know the correct amounts of Red, Green and Blue needed to create Magenta. The HSI model is also a more useful model for evaluating or measuring an object's color characteristics, such as the "redness" of a berry or the "yellowness" of an autumn leaf.

### C. HSV Color Model

Hue, Saturation, Value or HSV is a color model that describes colors (hue or tint) in terms of their shade (saturation or amount of gray) and their brightness (value or luminance). The HSV color wheel may be depicted as a cone or cylinder. Instead of Value, the color model may use Brightness, making it HSB (Photoshop uses HSB).

- Hue is expressed as a number from 0 to 360 degrees representing hues of red (starts at 0), yellow (starts at 60), green (starts at 120), cyan (starts at 180), blue (starts at 240), and magenta (starts at 300).
- Saturation is the amount of gray (0% to 100%) in the color.
- Value (or Brightness) works in conjunction with saturation and describes the brightness or intensity of the color from 0% to 100%.
- The hue (H) of a color refers to which pure color it resembles. All tints, tones and shades of red have the same hue.

Hues are described by a number that specifies the position of the corresponding pure color on the color wheel, as a fraction between 0 and 1. Value 0 refers to red; 1/6 is yellow; 1/3 is green; and so forth around the color wheel.

- The saturation (S) of a color describes how white the color is. A pure red is fully saturated, with a saturation of 1; tints of red have saturations less than 1; and white has a saturation of 0.
- The value (V) of a color, also called its lightness, describes how dark the color is. A value of 0 is black, with increasing lightness moving away from black.

This diagram, called the single-hexcone model of color space, can help you visualize the meaning of the H, S, and V parameters as shown in Fig.4.

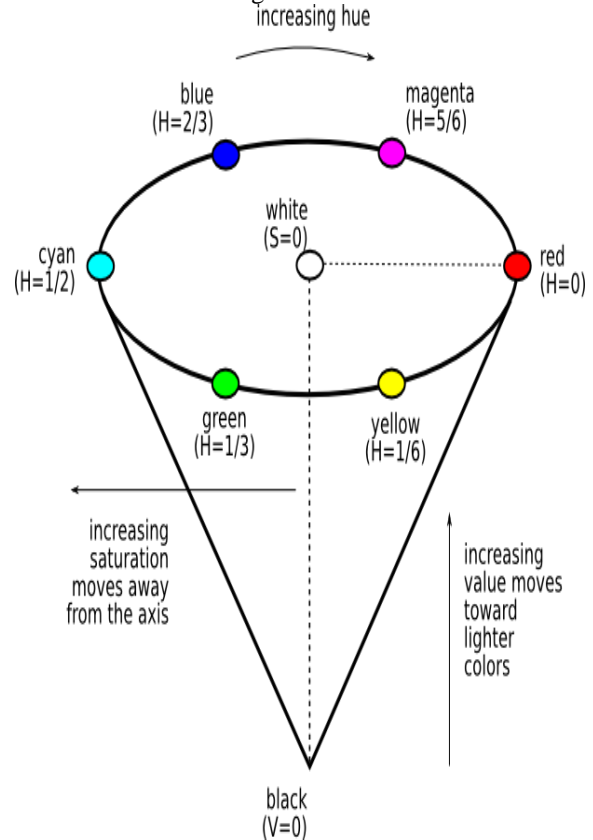


Fig.4. Single hexcone model of color space.

- The outer edge of the top of the cone is the color wheel, with all the pure colors. The H parameter describes the angle around the wheel.
- The S (saturation) is zero for any color on the axis of the cone; the center of the top circle is white. An increase in the value of S corresponds to a movement away from the axis.
- The V (value or lightness) is zero for black. An increase in the value of V corresponds to a movement away from black and toward the top of the cone.

## IV. SIMULATION RESULTS

In this chapter all the simulation results which are done using Xilinx ise 9.1 are shown in below Figs.5 to 10.



## V. CONCLUSION

An efficient algorithm for obstacle detection that is using the reference of color cue and color coherence vector is presented. The algorithm is used in guidance system for visually impaired and helps them in navigating both indoor and outdoor environments. The algorithm makes use of color coherence vectors which can be computed in linear time. The color coherence vector matching is computationally inexpensive and hence the algorithm has good efficiency. Moreover, the algorithm does not have any learning requirements. We are using color cue and coherence as our main area of concern rather than the color histogram, which gives no spatial information. HIS color model is used in this project for convert the original image into HSI image and we also proposed HSV color model for same obstacle detection.

## VI. REFERENCES

- [1] E. T. Cunningham and J. Blindness, "A global priority for the twenty-first century," presented at Special theme report on blindness, Bulletin of the World Health Organization, 2001.
- [2] J. P. Whitcher and C. Blindness: "A global perspective," presented at Special theme report on blindness, Bulletin of the World Health Organization, 2001.
- [3] G. Pass, R. Zabih, and J. Miller, "Comparing Color images using Color Coherence Vectors," in Proc. 4th ACM International conference on Multimedia, 1997.
- [4] G. Sainarayanan, "Incorporating certain human vision properties in vision substitution by stereo acoustic transform," Int. symp. on Signal processing and its applications, pp. 13-16, Aug. 2001.
- [5] W. Miled, "Robust Obstacle detection based on Dense disparity maps," Int. conf. on computer aided systems theory- EUROCAST, 2007.
- [6] F. Ferrari et al., "A stereo-vision system for real time obstacle avoidance in unknown environment," Int. workshop on Intelligent Robots and Systems, 1990.
- [7] David Aldavert, "Obstacle Detection and alignment using an stereo camera pair," Technical report, Artificial Intelligence research institute, Spain, 2008.
- [8] T. Hong, "Obstacle detection using a TOF Range Camera for Indoor AGV Navigation," Int. Conf. on advanced robotics, 2005.
- [9] N. Chumerin, "Ground plane estimation based on dense stereo disparity," Int. Conf. on neural networks and artificial intelligence, 2008.
- [10] I. Ulrich, "Appearance based obstacle detection with Monocular color vision," AAAI National Conference on Artificial Intelligence, July-August 2000.
- [11] C. N. Viet, "An efficient obstacle detection algorithm using color and texture," World Academy of Science and Technology, 2009.

# Space-Time Block Coding (STBC) for Multiple Transmit Antennas over Time Selective Fading Channels

Ramaswamy T.  
Assistant Professor  
MREC, Hyderabad

K. Chennakesava Reddy,  
PhD  
Principal  
BITSW, Hyderabad

Manda Divyasri  
PG Scholar  
MREC, Hyderabad

## ABSTRACT

A study of space-time block coding in analytical form-multiple input and multiple output (STBC-MIMO) system was done in various cases. Initially, a basic Alamouti scheme for encoding was illustrated. Then, the performance of proposed STBC-MIMO was studied under varying a number of transmitting antennas, varying number of receiving antennas, varying channels, various modulations and also at varying rates. The simulation results were obtained for all cases. At each and every stage, a new encoding matrix was derived.

## Keywords

Space-Time Block Coding, OSTBC, Rayleigh fading channel, AWGN, MIMO, BER.

## 1. INTRODUCTION

In digital wireless communication required for reliability, quality and speed has tempted the notice of huge communication networks and technologist community, to contain a look at the MIMO (multiple input multiple output) technology, plays the main role to fulfill their requirements [1]. Now a day's wireless networks are used generally: wireless mesh networks (WMNs), personal area networks (PANs), cellular networks, wireless Local Area Networks (WLANs) and other type of networks like wireless sensor networks (WSNs), and now a day's these networks demand growing to have twisted spectrum into a valuable resource. Because of this cause, these approaches require anytime to set additional bits per Hz. The main solution the researcher's has fixed a notice to utilize multiple antennas at receiver (RX) and as well as a transmitter (TX). To get back, winters in [2] the MIMO used for raising capacity dates. This type of approach is called MIMO (Multiple-Input Multiple-Output) method. Here it is very important to note is that every antenna element on an MIMO method manages in the similar frequency, and it doesn't need additional bandwidth. Also for fair evolution, the single antenna system is greater than or equal to the total power via all antenna elements, i.e.

$$\sum_{k=1}^N p_k \leq P \quad (1)$$

If system was having a single antenna at that time P is power, N is total antenna elements number, and  $p_k$  is allocated power via kth antenna element [4]. Efficiently, (1) MIMO doesn't require more power because it has multiple antenna elements. A number of advantages are there in MIMO model, but the only disadvantage is, when MIMO rank increases, the design of model becomes more complex. It can be realized by dividing a single into multiple models, called diversity coding. Diversity coding techniques used high range communication of MIMO model brings extra complexity. This paper gives a complete explanation on MIMO techniques involves the system model, channel models, and capacity analysis. After

that importance is given to special diversity, particularly to STBC (space-time block codes). In this STBC, the concentration on orthogonal STBC and Alamouti STBC for three as well as four transmits antennas and at last shown analysis and simulation results. The remaining of paper is as follows: MIMO system model complete details shown in Section 2, STBC transmitting antennas models for two, three, four and eight are shown in Section 3, simulation results give Section 4 and finally, Section 5 gives the conclusions.

## 2. MIMO SYSTEM MODEL

Conventional wireless techniques are precious by multipath broadcast. In MIMO, but, this multipath result is demoralized to advantage the user. In fact, an occurrence of the rich multipath divide of parallel stream depended on. The cause for this result will be noticeable in system model explain in Section II-A.

### 2.1 System Model

MIMO model was having three elements those are the transmitter (TX), channel (H), receiver (RX). In this,  $N_t$  represents the transmitter of a number of antennas, and  $N_r$  represents the receivers of a number of antennas. Fig.1 illustrates the MIMO system model block diagram. It is significant that system model were explained in the channel model. For e.g. the multiple inputs are placed at the input of the channel (H), and multiple output were placed at the output of the channel (H).

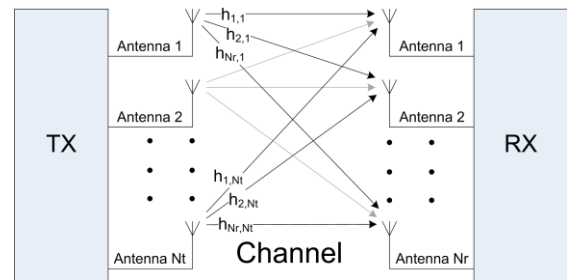


Fig. 1. MIMO (Multiple-Input Multiple-Output) system block diagram.

$N_t$  is channel inputs, and  $N_r$  is channel outputs is

The channel with  $N_r$  outputs and  $N_t$  inputs is denoted as a  $N_r \times N_t$  matrix:

$$H = \begin{pmatrix} h_{1,1} & h_{1,2} & \cdots & h_{1,N_t} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_r,1} & h_{N_r,2} & \cdots & h_{N_r,N_t} \end{pmatrix} \quad (2)$$

The place each and every entry  $h_{i,j}$  denotes the attenuation and segment shift (transfer operate) between the jth transmitter and the ith receiver. In this paper, the MIMO

channel behaves in a “quasi-static” nature, i.e. the channel changes randomly between burst to burst, however, constant inside a transmission. This is a rule which is reasonable and assumes, as it represents an indoor channel the place the time of exchange is constant and negligible in comparison with the time of a burst of data [10]. The MIMO signal can be described as

$$\vec{r} = H\vec{s} + \vec{n} \quad (3)$$

Where  $\vec{r}$  the bought vector of size  $N_r \times 1$ ,  $H$  is the channel matrix of measurement  $N_r \times N_t$ ,  $\vec{s}$  is the transmitted vector of size  $N_t \times 1$ , and  $\vec{n}$  is the noise vector of size  $N_r \times 1$ . Each noise detail is on the whole modeled as independent identically disbursed (i.e. I.D.) white Gaussian noise [1], [4] with variance  $N_t = (2 \cdot \text{SNR})$  [11]. A cause of this signal is as follows. The transmitted alerts are blended in the channel because they use the identical service frequency. On the receiver part, the got sign is composed of a linear mixture of every transmitted signal plus noise. The receiver can clear up for the transmitted indicators by using treating (3) as a procedure of linear equations [3]. If the channel  $H$  is correlated, the process of linear equations could have more unknowns than equations. One purpose correlation between indicators can occur is because of the spacing between antennas. To hinder correlation as a result of the spacing, which is spaced with a minimum value of  $\lambda_c/2$ , where  $\lambda_c$  is the wavelength of the service frequency [1]. The 2d rationale correlation can occur is because of lack of multipath components. It's for this reason that wealthy multipath is desirable in MIMO systems. The multipath output can be understood by means of receive antenna being with an extra channel. Accordingly, the rank of an MIMO channel is defined as the number of impartial equations provided. It is most important to notice that

$$\text{Rank}(H) \leq \min(N_r, N_t) \quad (4)$$

And for that reason the maximum number of streams that a MIMO system can help is higher-bounded by  $\min(N_r, N_t)$ . When the efficiency of MIMO systems is dependent enormously on the channel matrix, it is a main signal to the channel matrix realistically. The following section supplies an overview of usual channel units used for computer simulations.

## 2.2 Channel Models

Channel units for MIMO methods will also be either simple or very elaborate, relying on the environment modeled and the desired accuracy. There are two unique tactics for modeling MIMO channels. One approach is to calculate the MIMO channel matrix consistent with a physical representation of the environment. The channel matrix in physical model would rely upon important parameters like the angle of arrival (AOA), angle of departure (AOD), and time of arrival (TOA) [12]. In [13], a physical MIMO model and provides normal physical parameters for each macro and other micro environments. As expected, these styles of deterministic models are highly complex. An additional system to MIMO model has no. of channels logically. Such a system considers all channels b/w every transmit antenna to every receiving antenna when single input single output channels. Like this system considers that channels should be i.e. I.D (independent and identically dispersed). But, it depends on environment system, this consideration is true not all the times. The reason is that MIMO channels can expertise spatial correlation between links [12]. It is viable to generate an MIMO channel with a specified correlation matrix. The channel correlation

matrix is in most cases measured within the discipline, and it's tied to the environment setup equivalent to spacing between antennas, antenna element patterns, and surrounding reflectors [12].

As one of the main themes of this paper is that differentiate the presentation of dissimilar STBC methods, such that the selected channel model link doesn't involve with the presentation. Another, channel models are the case of LOS (line-of-sight) and the case of NLOS (non-line-of-sight) correspondingly.

i) Non-line-of-sight Environment: A typical model used in research to model NLOS scenarios is the Rayleigh model. The Rayleigh system considers non-line-of-sight, and with the huge amount of scatters is used for environments. The Rayleigh model has independent identically distributed (i.i.d.) complex, zero mean, unit variance channel elements and is given by [10]:

$$h_{i,j} = \frac{1}{\sqrt{2}} (\text{Normal}(0,1) + \sqrt{-1} \cdot \text{Normal}(0,1)) \quad (5)$$

This model results in an approximation that improves as the spacing between antennas become large compared to the wavelength  $\lambda$ .

ii) Line-of-sight Model: The MIMO channel matrix for the Line-of-sight setting is given by [12]:

$$H = \sqrt{\frac{K}{1+K}} H_{LOS} + \sqrt{\frac{1}{1+K}} H_{NLOS} \quad (6)$$

Where:

$$K(\text{dB}) = 10 \log_{10} \frac{P_{LOS}}{P_{NLOS}} \quad (7)$$

In (6), HLOS is a rank-one matrix akin to the LOS element, and the HNLOS corresponds to the NLOS components. In (7), PLOS is the vigour because of the LOS element, and PNLOS is due to the power of the NLOS element [12]. The HNLOS aspect is in most cases modeled as (5) [3]. In SISO techniques, the better the ok element, the smaller the fade fringe needed. For MIMO model, the K facto is high, will be the more leading the rank-one HLOS element, and thus, will be the less leading the HNLOS element. But, for the setting of rich multipath, measurements and simulations showed that the Line-of-sight element infrequently controls [3].

## 3. STBC PROPOSED APPROACH

One of the methodologies for exploiting the ability in MIMO method consists of using the further variety of MIMO programs, specifically spatial range, to combat channel fading. This will be attained with a sending quantity of units of the similar information via every antenna. Even as doing these matters dropping probability understanding reduces exponentially [3]. The antennas in an MIMO approach are used for aiding a transmission of an SISO approach because the distinctive expense of is that of an SISO approach. The variety order or range obtains of an MIMO system is defined as the number of impartial receptions of the same sign. A MIMO process with  $N_t$  transmit antennas and  $N_r$  acquire antennas has possibly full variety (i.e. Maximum diversity) attain equal to  $N_t N_r$ . The dissimilar replicas transmit for developing diversity is producing with a space-time encoder it encodes a single via time with transmitting every symbol at dissimilar times. This type of coding called STC (space-time coding). Their simple decoding, most leading type of space-time coding is STBC (space-time block coding). In next, a

space-time block coding system is dissimilar and has contrast for presentation in Section 4.

### 3.1 Alamouti's STBC

In [14], Alamouti published his manner on transmit diversity. Historically, Alamouti's scheme was once the primary STBC [4]. The simplicity and basic structure of the Alamouti STBC have placed the scheme in each the W-CDMA and CDMA-2000 standards [3]. The Alamouti STBC scheme makes use of two transmit antennas and  $N_r$  obtain antennas and can accomplish a maximum range order of  $2N_r$  [14]. In addition to the Alamouti system has rated 1 (i.e. full rate) as each 2-time intervals it transmits two symbols. After, that Alamouti system is obtained for one and two received antennas, followed with a normal term used for the decoding method for the case of receive antennas  $N_r$ .

(i) Explanation:

As acknowledged prior, Alamouti STBC uses two transmit antennas whatever the number of acquiring antennas. The Alamouti scheme encoding operation is obtained through (eight). In this paper, the rows of each and every coding scheme represent an extra time instant, even as the columns characterize the transmitted symbol by way of each special antenna. On this case, the first and 2nd rows signify the transmission at the first and second time instant respectively. At a time  $t$ , the logo  $s_1$  and image  $s_2$  are transmitted from antenna 1 and antenna two respectively. Assuming that each and every symbol has a period  $T$ , and then at time  $t + T$ , the symbols are  $-s_2^*$  and  $s_1^*$ , where  $(\cdot)^*$  denotes the problematic conjugate, are transmitted from antenna 1 and antenna two respectively.

$$\mathcal{G}_2 = \text{time} \begin{pmatrix} s_1 & s_2 \\ -s_2^* & -s_1^* \end{pmatrix} \quad (8)$$

ii) Case of 1 Receive Antenna:

The amount of available receiving antennas of the decoding and response of the signal are depended on. For case of one received antenna signals are [14]:

$$\begin{aligned} r_1^{(1)} &= r_1(t) = h_{1,1}s_1 + h_{1,2}s_2 + n_1^{(1)} \\ r_1^{(2)} &= r_1(t+T) = -h_{1,1}s_2^* + h_{1,2}s_1^* + n_1^{(2)} \end{aligned} \quad (9)$$

The place at one antenna acquired sign is  $r_1$ , channel switch perform is  $h_{i,j}$  is transmitted antenna  $j_{th}$  and acquire antenna  $i_{th}$  mentioned in chapter 3, at antennal tricky random variable  $n_1$ , and represents  $x$  at instantaneous time  $k$ . Got signal are transmitted before handiest to the decoder, they blended as follows [14]:

$$\begin{aligned} \tilde{s}_1 &= h_{1,1}^* r_1^{(1)} + h_{1,2} r_1^{*(2)} \\ \tilde{s}_2 &= h_{1,2}^* r_1^{(1)} + h_{1,1} r_1^{*(2)} \end{aligned} \quad (10)$$

And substituting (9) in (10) yields:

$$\begin{aligned} \tilde{s}_1 &= (\alpha_{1,1}^2 + \alpha_{1,2}^2) s_1 + h_{1,1}^* n_1^{(1)} + h_{1,2} n_1^{*(2)} \\ \tilde{s}_2 &= (\alpha_{1,1}^2 + \alpha_{1,2}^2) s_2 - h_{1,1} n_1^{*(2)} + h_{1,2}^* n_1^{(1)} \end{aligned} \quad (11)$$

Where  $\alpha_{i,j}^2$  is the squared magnitude of the channel transfer function  $h_{i,j}$ . The calculated  $\tilde{s}_1$  and  $\tilde{s}_2$  are then transmit to an

ML (Maximum Likelihood) decoder to approximation the sender symbols  $s_1$  and  $s_2$  respectively [14].

iii). Case of two Receive Antennas:

For the case considered with two receive antennas, the received symbols are represented as [14]:

$$\begin{aligned} r_1^{(1)} &= h_{1,1}s_1 + h_{1,2}s_2 + n_1^{(1)} \\ r_1^{(2)} &= -h_{1,1}s_2^* + h_{1,2}s_1^* + n_1^{(2)} \\ r_2^{(1)} &= h_{2,1}s_1 + h_{2,2}s_2 + n_2^{(1)} \\ r_2^{(2)} &= -h_{2,1}s_2^* + h_{2,2}s_1^* + n_2^{(2)} \end{aligned} \quad (12)$$

and the combined signals are [14]:

$$\tilde{s}_1 = h_{1,1}^* r_1^{(1)} + h_{1,2} r_1^{*(2)} + h_{2,1}^* r_2^{(1)} + h_{2,2} r_2^{*(2)} \quad (13)$$

$$\tilde{s}_2 = h_{1,2}^* r_1^{(1)} + h_{1,1} r_1^{*(2)} + h_{2,2}^* r_2^{(1)} + h_{2,1} r_2^{*(2)}$$

Which, after substituting (11) becomes:

$$\begin{aligned} \tilde{s}_1 &= (\alpha_{1,1}^2 + \alpha_{1,2}^2 + \alpha_{2,1}^2 + \alpha_{2,2}^2) s_1 + h_{1,1}^* n_1^{(1)} + h_{1,2} n_1^{*(2)} + h_{2,1}^* n_2^{(1)} + h_{2,2} n_2^{*(2)} \\ \tilde{s}_2 &= (\alpha_{1,1}^2 + \alpha_{1,2}^2 + \alpha_{2,1}^2 + \alpha_{2,2}^2) s_2 - h_{1,1} n_1^{*(2)} + h_{1,2}^* n_1^{(1)} - h_{2,1} n_2^{*(2)} + h_{2,2}^* n_2^{(1)} \end{aligned} \quad (14)$$

iv). Decoding decision statistic for  $N_r$  receives antennas:

The ML decoder decision statistic decodes in desire of  $s_1$  and  $s_2$  over all possible values of  $s_1$  and  $s_2$  such that (15) and (16) are minimized where is given with the aid of (17) for  $N_t = 2$  [15], [11].

$$|\sum_{i=1}^{N_r} (r_i^{(1)} h_{i,1}^* + r_i^{*(2)} h_{i,2}) - s_1|^2 + \psi |s_1|^2 \quad (15)$$

$$|\sum_{i=1}^{N_r} (r_i^{(1)} h_{i,2}^* - r_i^{*(2)} h_{i,1}) - s_2|^2 + \psi |s_2|^2 \quad (16)$$

$$\psi = (-1 + \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} |h_{i,j}|^2) \quad (17)$$

Alamouti STBC does not require CSI at the transmitter. Also, the Alamouti STBC can be utilized with two transmitted antennas and 1 obtain antenna even as engaging in the entire variety of 2. That is a foremost characteristic of Alamouti STBC as it reduces the outcomes of fading at cellular stations with best requiring additional antenna factors on the base station, the place it's less expensive than having multiple antennas on the receivers [14]. However, if having more antennas at the receivers is just not a difficulty, this scheme can be utilized with two transmitted antennas and  $N_r$  acquires antennas whilst undertaking a  $2N_r$  full diversity. The case space time block codes for  $N_r > 2$  are discussed within the following part.

### 3.2 Orthogonal Space-Time Block Codes

The Alamouti scheme mentioned in chapter III. It is a part of a common classification of STBCs often called orthogonal space-Time Block Codes (OSTBCs) [4]. The authors of [15] observe the mathematical framework of orthogonal designs to construct each real and problematic orthogonal code that acquire full range. For the case of actual orthogonal codes, it has been shown that a full fee code will also be developed

[15]. Nevertheless, for the case of complicated orthogonal codes, it's unknown if a full cost and entire variety codes exist for  $N_t > 2$  [3]. Elaborate modulations methods are of interest in this paper and hence real orthogonal codes aren't discussed. In subsequent sections, the full diversity complex orthogonal codes presented in [15] for special premiums are briefly introduced.

#### 1) Orthogonal STBC for $N_t = 3$ :

For the case of three transmit antennas. Assemble block codes for the with half of and 3/4 coding expense and whole diversity  $3N_r$ .

##### i). $N_t = 3$ with Rate 1/2:

The full diversity, rate 1/2 code for  $N_t = 3$  are given by [15], [11]:

$$\mathcal{G}_3 = \begin{pmatrix} s_1 & s_2 & s_3 \\ -s_2 & s_1 & -s_4 \\ -s_3 & s_4 & s_1 \\ -s_4 & -s_3 & s_2 \\ s_1^* & s_2^* & s_3^* \\ -s_2^* & s_1^* & -s_4^* \\ -s_3^* & s_4^* & s_1^* \\ -s_4^* & -s_3^* & s_2^* \end{pmatrix} \quad (18)$$

This code transmits four symbols each eight time intervals, and for that reason has expense half. The choice metric to shrink via the decoder for detecting  $s_1, s_2, s_3, s_4$  are given with the aid of (20), (21), (22), (23) respectively the place

$$\xi = (-1 + 2 \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} |h_{i,j}|^2) \quad (19)$$

For  $N_t=3$  [11].

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,1}^* + r_i^{(2)} h_{i,2}^* + r_i^{(3)} h_{i,3}^* + r_i^{*(5)} h_{i,1} + r_i^{*(6)} h_{i,2} + r_i^{*(7)} h_{i,3}) \right] - s_1 \right|^2 + \xi |s_1|^2 \quad (20)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,2}^* - r_i^{(2)} h_{i,1}^* + r_i^{(4)} h_{i,3}^* + r_i^{*(5)} h_{i,2} - r_i^{*(6)} h_{i,1} + r_i^{*(8)} h_{i,3}) \right] - s_1 \right|^2 + \xi |s_2|^2 \quad (21)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,3}^* - r_i^{(3)} h_{i,1}^* + r_i^{(4)} h_{i,2}^* + r_i^{*(5)} h_{i,3} - r_i^{*(7)} h_{i,1} - r_i^{*(8)} h_{i,2}) \right] - s_3 \right|^2 + \xi |s_3|^2 \quad (22)$$

$$\left| \left[ \sum_{i=1}^{N_r} (-r_i^{(2)} h_{i,3}^* + r_i^{(3)} h_{i,2}^* - r_i^{(4)} h_{i,1}^* - r_i^{*(6)} h_{i,3} + r_i^{*(7)} h_{i,2} - r_i^{*(8)} h_{i,1}) \right] - s_4 \right|^2 + \xi |s_4|^2 \quad (23)$$

##### ii) $N_t = 3$ with Rate 3/4:

A higher rate code with  $N_t=3$  is given by [15], [11]

$$\mathcal{H}_3 = \begin{pmatrix} s_1 & s_2 & \frac{s_3}{\sqrt{2}} \\ -s_2^* & s_1^* & \frac{s_3}{\sqrt{2}} \\ \frac{s_3^*}{\sqrt{2}} & \frac{s_3}{\sqrt{2}} & \frac{-s_1 - s_1^* + s_2 - s_2^*}{2} \\ \frac{s_3^*}{\sqrt{2}} & -\frac{s_3}{\sqrt{2}} & \frac{s_2 + s_2^* + s_1 - s_1^*}{2} \end{pmatrix} \quad (24)$$

As can be determined, (23) transmits 3 symbols each 4 time intervals, and therefore has rate three/4. The choice statistic to diminish for detecting  $s_1, s_2$ , and  $s_3$  are given by (25), (26),

$$\text{and (27) respectively: } \left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,1}^* + r_i^{*(2)} h_{i,2} + \frac{(r_i^{(4)} - r_i^{(3)}) h_{i,3}^*}{2} - \frac{(r_i^{(3)} + r_i^{(4)})^* h_{i,3}}{2}) \right] - s_1 \right|^2 + \psi |s_1|^2 \quad (25)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,2}^* - r_i^{*(2)} h_{i,1} + \frac{(r_i^{(4)} + r_i^{(3)}) h_{i,3}^*}{2} + \frac{(-r_i^{(3)} + r_i^{(4)})^* h_{i,3}}{2}) \right] - s_2 \right|^2 + \psi |s_2|^2 \quad (26)$$

$$\left| \left[ \sum_{i=1}^{N_r} \left( \frac{(r_i^{(1)} + r_i^{(2)}) h_{i,3}^*}{\sqrt{2}} + \frac{r_i^{*(3)} (h_{i,1} + h_{i,2})}{\sqrt{2}} + \frac{r_i^{*(4)} (h_{i,1} - h_{i,2})}{\sqrt{2}} \right) \right] - s_3 \right|^2 + \psi |s_3|^2 \quad (27)$$

#### 2) Orthogonal STBC for $N_t = 4$ :

For the case of 4 transmit antennas, [15] provide block codes of price 1=2 and 3=4, each of which have full diversity  $4N_r$ .

i)  $N_t = 4$  with Rate 1=2: In the case of 4 transmit antennas, the rate 1/2 code block is given by [15], [11]:

$$\mathcal{G}_4 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 \\ -s_2 & s_1 & -s_4 & s_3 \\ -s_3 & s_4 & s_1 & -s_2 \\ -s_4 & -s_3 & s_2 & s_1 \\ s_1^* & s_2^* & s_3^* & s_4^* \\ -s_2^* & s_1^* & -s_4^* & s_3^* \\ -s_3^* & s_4^* & s_1^* & -s_2^* \\ -s_4^* & -s_3^* & s_2^* & s_1^* \end{pmatrix} \quad (28)$$

Where, just like (18), has price 1/2 as 4 symbols are transmitted in 8 time intervals. To decode, the ML decoder minimizes the selection metric (22), (23), (24), and (25) for decoding  $s_1, s_2, s_3$ , and  $s_4$  respectively the place is given by means of (18) for  $N_t = \text{four}$  [11]. The decoding selection metric (32) for decoding  $s_4$  differs from that of [11] considering that the writer found out a mistake within the metric offered.

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,1}^* + r_i^{(2)} h_{i,2}^* + r_i^{(3)} h_{i,3}^* + r_i^{*(4)} h_{i,4} + r_i^{*(5)} h_{i,1} + r_i^{*(6)} h_{i,2} + r_i^{*(7)} h_{i,3} + r_i^{*(8)} h_{i,4}) \right] - s_1 \right|^2 + \xi |s_1|^2 \quad (29)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,2}^* - r_i^{(2)} h_{i,1}^* - r_i^{(3)} h_{i,4}^* + r_i^{(4)} h_{i,3}^* + r_i^{*(5)} h_{i,2} - r_i^{*(6)} h_{i,1} - r_i^{*(7)} h_{i,4} + r_i^{*(8)} h_{i,3}) \right] - s_2 \right|^2 + \xi |s_2|^2 \quad (30)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,3}^* + r_i^{(2)} h_{i,4}^* - r_i^{(3)} h_{i,1}^* - r_i^{*(4)} h_{i,2}^* + r_i^{*(5)} h_{i,3} + r_i^{*(6)} h_{i,4} - r_i^{*(7)} h_{i,1} - r_i^{*(8)} h_{i,2}) \right] - s_3 \right|^2 + \xi |s_3|^2 \quad (31)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,4}^* - r_i^{(2)} h_{i,3}^* + r_i^{(3)} h_{i,2}^* - r_i^{(4)} h_{i,1}^* + r_i^{*(5)} h_{i,4} - r_i^{*(6)} h_{i,3} + r_i^{*(7)} h_{i,2} - r_i^{*(8)} h_{i,1}) \right] - s_4 \right|^2 + \xi |s_4|^2 \quad (32)$$

##### ii). $N_t = 4$ with Rate 3/4:

A four transmit antenna block code with rate 3/4 is given by [15], [11]:

$$\mathcal{H}_4 = \begin{pmatrix} s_1 & s_2 & \frac{s_3}{\sqrt{2}} & \frac{s_3}{\sqrt{2}} \\ -s_2^* & s_1^* & \frac{s_3}{\sqrt{2}} & -\frac{s_3}{\sqrt{2}} \\ \frac{s_3}{\sqrt{2}} & \frac{s_3}{\sqrt{2}} & \frac{-s_1-s_1^*+s_2-s_2^*}{2} & \frac{-s_2-s_2^*+s_1-s_1^*}{2} \\ \frac{s_3}{\sqrt{2}} & -\frac{s_3}{\sqrt{2}} & \frac{s_2+s_2^*+s_1-s_1^*}{2} & \frac{-s_1+s_1^*+s_2-s_2^*}{2} \end{pmatrix} \quad (33)$$

The decision metrics to minimize for decoding  $s_1$ ,  $s_2$ , and  $s_3$  are given by (34), (35), and (36) respectively

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,1}^* + r_i^{*(2)} h_{i,2} + \frac{(r_i^{(4)} - r_i^{(3)})(h_{i,3}^* - h_{i,4}^*)}{2} - \frac{(r_i^{(3)} + r_i^{(4)})^* (h_{i,3} + h_{i,4})}{2}) \right] - s_1 \right|^2 + \psi |s_1|^2 \quad (34)$$

$$\left| \left[ \sum_{i=1}^{N_r} (r_i^{(1)} h_{i,2}^* - r_i^{*(2)} h_{i,1} + \frac{(r_i^{(4)} + r_i^{(3)})(h_{i,3}^* - h_{i,4}^*)}{2} + \frac{(-r_i^{(3)} + r_i^{(4)})^* (h_{i,3} + h_{i,4})}{2}) \right] - s_2 \right|^2 + \psi |s_2|^2 \quad (35)$$

$$\left| \left[ \sum_{i=1}^{N_r} \left( \frac{(r_i^{(1)} + r_i^{(2)}) h_{i,3}^*}{\sqrt{2}} + \frac{(r_i^{(1)} - r_i^{(2)}) h_{i,4}^*}{\sqrt{2}} + \frac{r_i^{*(3)} (h_{i,1} + h_{i,2})}{\sqrt{2}} + \frac{r_i^{*(4)} (h_{i,1} - h_{i,2})}{\sqrt{2}} \right) \right] - s_3 \right|^2 + \psi |s_3|^2 \quad (36)$$

iii)  $N_t = 8$  with Rate 1/2:

Within the case of 8 transmit antennas, the fee half code block is given with the aid of eq.(37). Similarly the real transmission matrix is given through  $N_t=8$

$$G_8 = \begin{pmatrix} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 & s_7 & s_8 \\ -s_2 & s_1 & s_4 & -s_3 & s_6 & -s_5 & -s_8 & s_7 \\ -s_3 & -s_4 & s_1 & s_2 & s_7 & s_8 & -s_5 & -s_6 \\ -s_4 & s_3 & -s_2 & s_1 & s_8 & -s_7 & s_6 & -s_5 \\ -s_5^* & -s_6^* & -s_7^* & -s_8^* & s_1^* & s_2^* & s_3^* & s_4^* \\ -s_6^* & s_5^* & -s_8^* & s_7^* & -s_2^* & s_1^* & -s_4^* & s_3^* \\ -s_7^* & s_8^* & s_5^* & -s_6^* & -s_3^* & s_4^* & s_1^* & -s_2^* \\ -s_8^* & -s_7^* & s_6^* & s_5^* & -s_4^* & s_3^* & s_2^* & s_1^* \end{pmatrix} \quad (37)$$

On the receiver finish, the obtained expressions are headquartered on Alamouti's model with the simplicity of getting only actual symbols and consequently no conjugate symbol in the equations. Accordingly the acquired expressions for any number of got antennas end up  $r_1, N_R = r_{1,N_R} t + h_{1,N_R} s_1 + n_1, N_R$ ,

$$r_2, N_R = r_{2,N_R} t + T = -h_{2,N_R} s_2 + h_{1,N_R} s_1 + n_2, N_R, \quad (38)$$

Where  $n_1, N_R, n_2, N_R$  are independent noise samples,  $h_{i,N_R}$  is the channel transfer operate from the  $i_{th}$  transmit antenna and  $N_R$  denotes the MT acquire antenna. Acquired alerts are then mixed for two transmit antennas as

$$\tilde{s}_1 = \sum_{N_R=1}^{M_R} r_1, N_R h_{1,N_R} + r_2, N_R h_{2,N_R}$$

$$\tilde{s}_2 = \sum_{N_R=1}^{M_R} r_1, N_R h_{2,N_R} - r_2, N_R h_{1,N_R} \quad (39)$$

Similarly, the acquired sign for four and eight transmit antennas will also be derived. Alamouti STBC don't require CSI on the transmitter and can be utilized with two transmit antennas and 1 obtain antenna with accomplishment of full

variety of two. It reduces the outcomes of fading at receiver station at the rate of some additional antenna factors at the transmitter end. If having extra antennas will not be a trouble, then this scheme is correct for getting full variety of 2NR with two transmit antennas. Now that the block codes and their decoding metrics were awarded, the simulation setup is mentioned, adopted by using results and evaluation, and subsequently the conclusion.

## 4. EXPERIMENTAL RESULTS

An experimental result shows the different constant  $N_r$ ,  $N_t$  diversity order and rate for different cases of QPSK, 16-QAM and 64-QAM.

### Case 1: Rayleigh channel

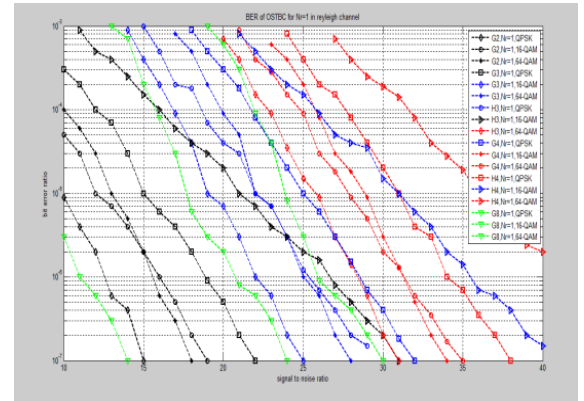


Fig.2 Bit error rate versus SNR of OSTBC for  $N_r = 1$ .

As we shown in the above figure represents the BER V/s SNR with constant  $N_r=1$ . In this case G8 shows the best performance compared to the G2, G3, G4, H3 and H4. Compare to the G8 of QPSK G2, G3, H3, G4, H4 of QPSK is high SNR, the G8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are high SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the low SNR value so that it is the best performer when compare to the other.

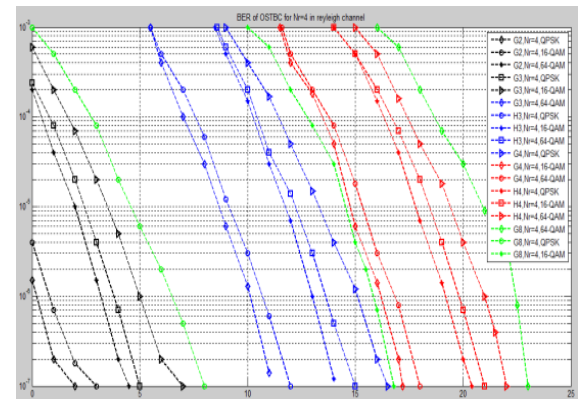
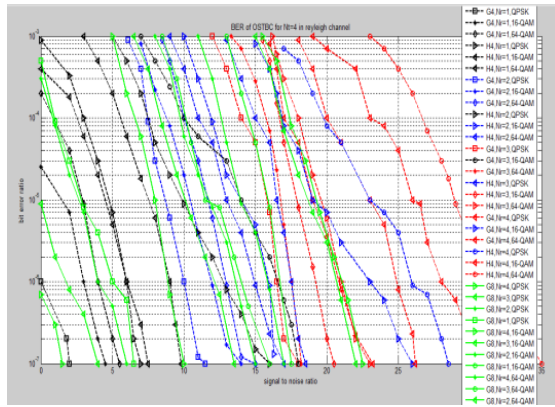


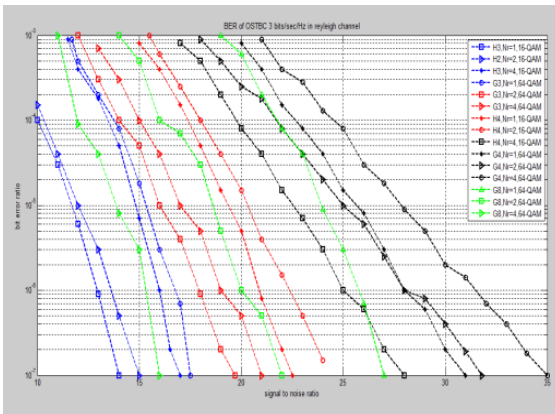
Fig.3 Bit error rate versus SNR of OSTBC for  $N_r = 4$ .

As we shown in the above figure represents the BER V/s SNR with constant  $N_r=4$ . In this case G8 shows the best performance compared to the G2,G3,G4, H3 and H4. Compare to the G8 of QPSK G2, G3, H3, G4, H4 of QPSK is low SNR, the G8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are low SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the high SNR value so that it is the best performer when compare to the other.



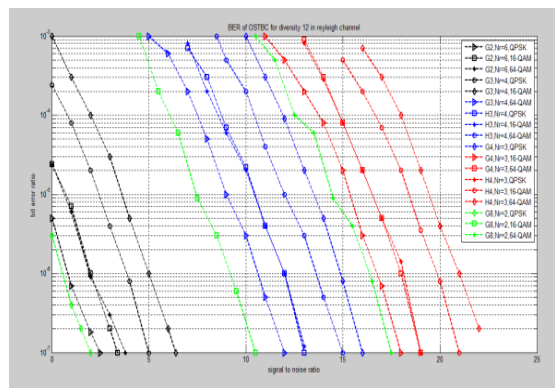
**Fig.4 Bit error rate versus SNR of OSTBC for  $N_t = 4$**

As we shown in the above figure represents the BER V/s SNR with different  $N_r=1,2,3,4$ . In this case G8 shows the best performance compared to the G2, G3, G4, H3 and H4. Compare to the G8 of QPSK G2, G3, H3, G4, H4 of QPSK is high SNR, the G8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are high SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the low SNR value so that it is the best performer when compare to the other.



**Fig.5 Bit error rate versus SNR of OSTBC at 3 bits/sec/Hz.**

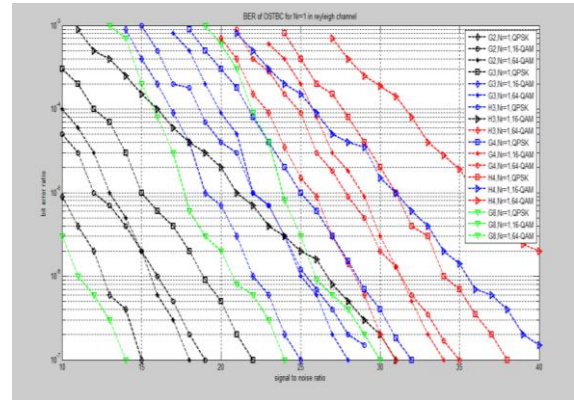
As we shown in the above figure represents the BER V/s SNR with different  $N_r=1, 2, 4$ . In this case G8 shows the best performance compared to the G3, G4. Compare to the G8, 64-QAM G3, G4 of 64-QAM are low SNR, so that the results shown that G8 is best performance compare to the other. G8 having the high SNR value so that it is the best performer when compare to the other.



**Fig.6 Bit error rate versus SNR of OSTBC with spatial diversity 12.**

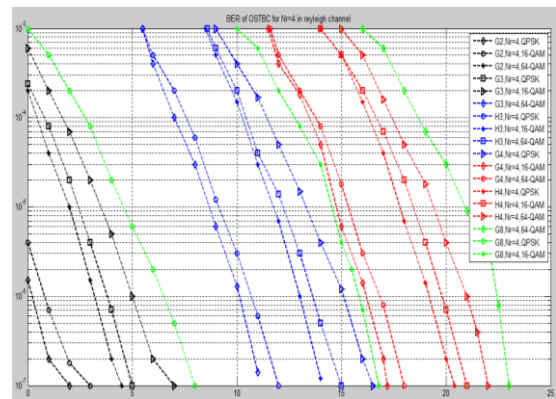
As we shown in the above figure represents the BER V/s SNR with different  $N_r=6, 4, 3, 2$ . In this case G8 shows the best performance compared to the G2, G3, G4, H3 and H4. Compare to the G8 of QPSK G2, G3, H3, G4, H4 of QPSK is high SNR, the G8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are high SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the low SNR value so that it is the best performer when compare to the other. Increase the number of  $N_t$  transmit antennas at the base station than increasing the number of  $N_r$  receive antennas at all mobile stations.

#### Case 2: Rician channel



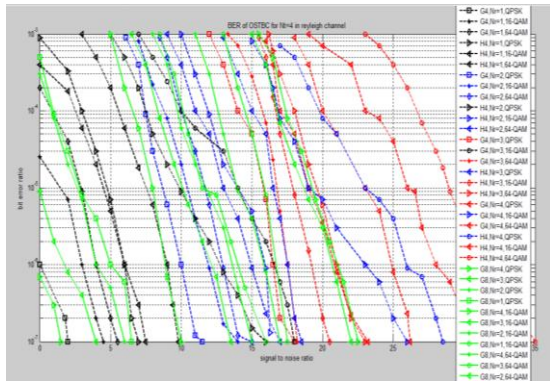
**Fig.7 Bit error rate versus SNR of OSTBC for  $N_r = 1$ .**

As we shown in the above figure represents the BER V/s SNR with constant  $N_r=1$ . In this case G8 shows the best performance compared to the G2, G3, G4, H3 and H4. Compare to the G8 of QPSK G2, G3, H3, G4, H4 of QPSK is high SNR, the G8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are high SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the low SNR value so that it is the best performer when compare to the other.



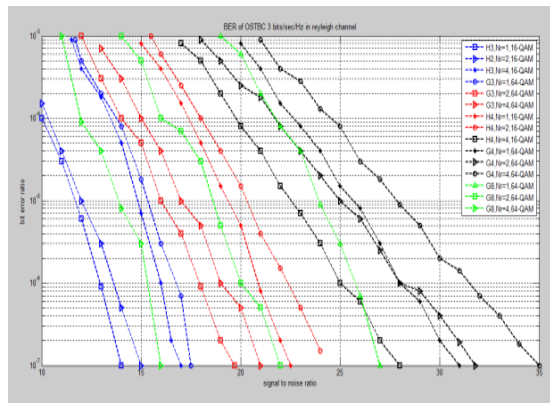
**Fig.8 Bit error rate versus SNR of OSTBC for  $N_r = 4$ .**

As we shown in the above figure represents the BER V/s SNR with constant  $N_r=4$ . In this case G8 shows the best performance compared to the G2, G3, G4, H3 and H4. Compare to the G8 of QPSK G2, G3, H3, G4, H4 of QPSK is low SNR, the G8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are low SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the high SNR value so that it is the best performer when compare to the other.



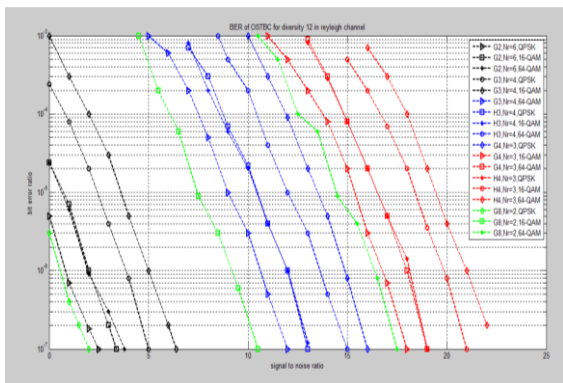
**Fig.9 Bit error rate versus SNR of OSTBC for  $N_t = 4$**

As we shown in the above figure represents the BER V/s SNR with different  $N_r=1,2,3,4$ . In this case G8 shows the best performance compared to the G2,G3,G4, H3 and H4. Compare to theG8 of QPSK G2, G3, H3, G4, H4 of QPSK is high SNR, theG8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of 16-QAM and 64-QAM are high SNR also, so that the results shown that G8 is best performance compare to the other. G8 having the low SNR value so that it is the best performer when compare to the other.



**Fig.10 Bit error rate versus SNR of OSTBC at 3 bits/sec/Hz.**

As we shown in the above figure represents the BER V/s SNR with different  $N_r=1,2,4$ . In this case G8 shows the best performance compared to theG3,G4. Compare to theG8, 64-QAM G3, G4 of 64-QAM are low SNR, so that the results shown that G8 is best performance compare to the other. G8 having the high SNR value so that it is the best performer when compare to the other.



**Fig.11 Bit error rate versus SNR of OSTBC with spatial diversity 12.**

As we shown in the above figure represents the BER V/s SNR with different  $N_r=6, 4, 3, 2$ . In this case G8 shows the best performance compared to the G2, G3, G4, H3 and H4. Evaluate to theG8 of QPSK G2, G3, H3, G4, H4 of QPSK is high SNR, theG8 of 16-QAM and 64-QAM G2, G3, H3, G4, H4 of sixteen-QAM and sixty four-QAM are excessive SNR also, so that the results shown that G8 is first-rate efficiency compare to the other. G8 having the low SNR value so that it is the exceptional performer when compare to the opposite. Expand the number of  $N_t$  transmit antennas on the base station than increasing the quantity of  $N_r$  obtain antennas in any respect mobile stations.

## 5. CONCLUSION

This was especially proven when G3 outperformed all others for equal range obtain. The penalty of having more transmit antennas, which consequently reduces the power per transmit antenna was once observed. Also, we determined diminishing returns for every scheme because the number of bought antennas elevated. It was once specifically interesting to search out that even though H3 and H4 have better cost than G3 and G4, the efficiency of G3 and G4 is bigger and might consequently be favored in some situations. In the end, we conclude that it's most desirable to make use of a low constellation order with high code fee than excessive constellation order with low code expense.

## 6. REFERENCES

- [1] A. Molisch, Wireless Communications. Wiley-IEEE Press, 2005.
- [2] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," IEEE Journal on Selected Areas in Communications, vol. 5, no. 5, pp. 871–878, 1987.
- [3] D. Gesbert, M. Shafi, D. Shiu, P. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," IEEE Journal on selected areas in Communications, vol. 21, no. 3, pp. 281–302, 2003.
- [4] G. Tsoulos, MIMO system technology for wireless communications. CRC Press, 2006.
- [5] L. Dai, S. Zhou, H. Zhuang, and Y. Yao, "Closed-loop MIMO architecture based on water-filling," Electronics Letters, vol. 38, no. 25, pp. 1718–1720, 2002.
- [6] K. Zheng, L. Huang, W. Wang, and G. Yang, "TD-CDM-OFDM: Evolution of TD-SCDMA toward 4G," IEEE Communications Magazine, vol. 43, no. 1, pp. 45–52, 2005.
- [7] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: design, performance, and field trial results," IEEE Communications Magazine, vol. 40, no. 9, pp. 143–149, 2002.
- [8] "IEEE P802.11n/D5.0," May 2008.
- [9] H. Niu and Ngo, "Diversity and Multiplexing Switching in 802.11 n MIMO Systems," in Signals, Systems and Computers, 2006.ACSSC'06. Fortieth Asilomar Conference on, 2006, pp. 1644–1648.
- [10] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," Wireless personal communications, vol. 6, no. 3, pp. 311–335, 1998.

- [11] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE Journal on selected areas in communications*, vol. 17, no. 3, pp. 451–460, 1999.
- [12] T. Kaiser, *Smart Antennas—State of the Art*. Hindawi Publishing Corporation, 2005.
- [13] A. Molisch, "A generic model for MIMO wireless propagation channels in macro-and microcells," *IEEE Transactions on Signal Processing*, vol. 52, no. 1, pp. 61–71, 2004.
- [14] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on selected areas in communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [15] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1456–1467, 1999.
- [16] L. Cortes-Pena, "Orthogonal Space-Time Coding Matlab Source Files," Software online: [http://users.ece.gatech.edu/\\_Cortes/STBC Matlab.html](http://users.ece.gatech.edu/_Cortes/STBC Matlab.html), 2009.

# A Maximum Likelihood (ML) based OSTBC-OFDM System over Multipath Fading Channels

Ramaswamy T.  
Assistant Professor  
MREC, Hyderabad

K. Chennakesava Reddy, PhD  
Principal  
BITSW, Hyderabad

Jaggari Manasa  
PG Scholar  
MREC, Hyderabad

## ABSTRACT

This paper proposes a space-time block-coded orthogonal frequency-division multiplexing (STBC-OFDM) scheme for frequency-selective fading channels which does not require channel knowledge either at the transmitter or at the receiver. This paper proposed a generalized maximum likelihood estimation decoding algorithm. Due to the orthogonality nature of STBC, the rule for decoding was reduced to a single step. The performance investigation of the proposed system was done over Rayleigh fading channel and also over the Rician fading channel. Simulation results reveal the performance of proposed STBC communication system is near optimum.

## Keywords

MIMO, STBC-OFDM, Alamouti Scheme, Maximum likelihood estimator, BER.

## 1. INTRODUCTION

The target of next generation wireless communication is to achieve high data rates with low bandwidth. It should be power efficient. At present time, Orthogonal Frequency Division Multiplexing is widely used for its bandwidth efficiency property because of its orthogonal characteristic and more data can be transmitted at a certain amount of bandwidth compare to the other systems. Orthogonal frequency division multiplexing (OFDM) is being considered to be the most promising multiplexing techniques to support the future wireless multimedia communication system. It is because of its bandwidth efficiency performance. The affects of Inter Symbol interference (ISI) is also very less than the system compare to the other multiplexing techniques. OFDM has been adopted and implemented in wire and wireless communication system. Severe attenuation in a multipath wireless environment makes it extremely difficult for the receiver to determine the transmitted signal unless the receiver is provided with some form of diversity i.e. some less-attenuated replica of the transmitted signal is provided to the receiver. In some applications, the only practical means of achieving diversity is deployment of antenna array at the transmitter and/or at the receiver end. In recent days, the communication systems demands for highly bandwidth-efficient and also highly power-efficient techniques are very valued for wireless communication in next generation systems. Compared to single antenna systems the multiple antenna systems offer an increased capacity of data rates. STBC was a bandwidth-efficient and also a power-efficient technique for communication through fading channel by using multiple transmitting antennas systems. Diversity gain enhancement is another technique for good performance achievement. Through MIMO-Space Time Block Code (STBC), the antenna diversity gain can be increased. This paper studies the performance of STBC-OFDM in the MIMO communication system under various channel constraints,

Rayleigh fading channel, Gaussian channel, etc. This paper also gives the details about the effect of modulation, power conditions, antenna selection techniques, fast and slow fading channel conditions. The remaining part of paper is ordered as follows: Section 2 gives the details about the system mode of STBC-OFDM. The decoding algorithm used to decode the information at the receiver is given in Section 3. Section 4 demonstrates the results of the proposed STBC system and finally conclusions are provided in section 5.

## 2. SYSTEM MODEL

The block diagram of STBC-OFDM is illustrated in Fig.1. The binary data stream is first modulated and mapped to a sequence of complex modulation symbols. The modulated sequence is then passed through a serial-to-parallel converter producing  $L$  data streams of length  $M$  in each transmit branch. An  $M$ -point Inverse Discrete Fourier Transform (IDFT) is performed on each serial data stream. IDFT gives the output as a block of  $M$ -time domain samples of a respective OFDM frame. To sidestep the effects of inter-symbol-interference, a cyclic prefix (CP) is added to output samples. Then, every OFDM frame is nurtured into a space-time block encoder. This paper has used the simple Alamouti's scheme [2] for simplicity, even the proposed approach can be applicable to the space-time block code from the orthogonal family design [3]. The  $l$ th frames of first and second transmitting antennas are represented by,

$$(X_l^1(0), \dots, X_l^1(M-1)) \quad \dots (1)$$

$$(X_l^2(0), \dots, X_l^2(M-1)) \quad \dots (2)$$

For the subsequent OFDM frame interval, the OFDM frame can be represented as

$$(X_l^{1*}(0), \dots, X_l^{1*}(M-1)) \quad \dots (3)$$

This is transmitted from the second antenna and

$$(-X_l^{2*}(0), \dots, -X_l^{2*}(M-1)) \quad \dots (4)$$

Transmits from the first antenna according to the signaling structure of Alamouti's scheme, where  $(.)^*$  denotes complex conjugation.

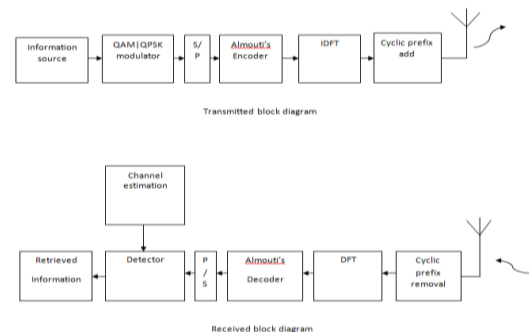


Fig.1 Overall system Model

In the proposed system model, a FIR filter with the memory was used to represent channels belonging to the 2 transmit antennas. At the receiver, first the cyclic prefix is removed, and then the received signal is passed through a DFT operation. It can be shown that the OFDM demodulator output can be represented as,

$$Y_l(k) = \sum_{m=1}^2 X_l^m(k) H^m(k) + N_l(k), l = 1, 2, \dots, 2L, k = 0 \dots M - 1 \quad \dots (5)$$

Where  $N_l(k)$  are independent samples of a Gaussian random variable with variance  $NO$  and  $H^m(k)$  is the frequency response of the channel from the  $m$ th transmit antenna to the single receive antenna at the  $k$ th tone. After performing DFT operation, by considering the  $k$ th frequency component of each and every frame, the receiver tries to construct a tone stream of  $2L$  frequency-domain samples. Such that,  $2L$  streams each of length  $M$  are realigned onto  $M$  streams, each of length  $2L$ . The obtained stream was fed to the detecting algorithm and explained in the next section in detail. By considering the coding structure, let us define the following sub-matrices

$i=1, 2 \dots L$

$$Y_i(k) = \begin{pmatrix} Y_{2i-1}(k) \\ Y_{2i}(k) \end{pmatrix}, \quad X_i(k) = \begin{bmatrix} X_i^1(k) & X_i^2(k) \\ -X_i^{2*}(k) & X_i^{1*}(k) \end{bmatrix} \quad \dots (6)$$

Then based on the above matrices and the system model the received signal receiver section can be written as

$$Y(k) = X(k)H(k) + N(k) \quad \dots (7)$$

### 3. DECODING ALGORITHM

In this approach, the decoding algorithm is based on the generalized maximum likelihood estimation [9], just a two-step process. For every toned frequency  $k$ ,  $k = 0, 1, \dots, M-1$ , the rules of decoding are to maximize and the respective likelihood logarithmic function is given as

$$\ln p(Y(k)|X(k), H(k)) = -\|Y(k) - X(k)H(k)\|^2 \quad (8)$$

Maximizing (8) with respect to  $H(k)$  yields the channel estimate (i.e. the  $k$ th DFT coefficient of the channel impulse response),

$$\hat{H}(k) = (X'(k)X(k))^{-1}X'(k)Y(k) \quad \dots (9)$$

In general for Alamouti's STBC coding, space-time block codes are based on orthogonality conditions, the primary term in eq.(6) represents a scale identity matrix, thus the eq.(9) reduces to

$$\hat{H}(k) = X'(k)Y(k) \quad \dots (10)$$

Introducing the channel estimate in eq.(5) produces the simple decoding rule

$$\hat{X}(k) = \arg \max_{X(k)} \|X'(k)Y(k)\|^2, k = 0, 1, \dots, M - 1 \quad \dots (11)$$

Where  $\hat{X}(k)$  consists of the estimates for the  $k$ th component in each OFDM frame. By the collection of all the outputs of all  $M$  processors, the original signal can be reconstructed from both antennas. The decoding rule creates phase ambiguity at decoder. This ambiguity can be solved by setting the samples of decoder input to a fixed value. This procedure requires the

starting frame of both antennas to be reconstructed with the help of pilot symbols.

### 4. RESULTS

This section gives the details about the performance evaluation of proposed approach. Simulations are done in MATLAB using the Rayleigh fading channel model. For each simulation, blocks of 1000 symbols are simulated. The proposed system was simulated with two transmitter antennas, two receiver antennas and also at maximum Doppler frequency of 50Hz ( $f=50$ Hz). We have used channel conditions, which have two independent paths with path delays in seconds and average path gains =  $[0 -18]$  dB. The simulation was done under various constraints like various types of modulation schemes, power allocation techniques and also over various fading channel assumptions. The respective figures are shown figure.2, figure.3 and figure.4 respectively. Table.1 gives the details about the summarization of simulation parameters.

#### Case 1: STBC 2X2 MIMO with Rayleigh channel

Table.1: Simulation Parameters

Parameter	Value
Number of Symbols	1000
Number of Transmitters	2
Number of Receivers	2
Fading Effect	Rayleigh
Noise	AWGN
Doppler Frequency	50Hz
SNR	0-30
System Model	MIMO

The order of the modulation stipulates the number of the diverse symbols that can be transmitted in a digital communication system. There are various modulation schemes like PSK, FSK and AM. In this paper, the proposed approach was tested over various PSK modulations like, BPSK, QPSK, 16-PSK and 32-PSK over Rayleigh fading channel and the respective results are illustrates in figure.2.

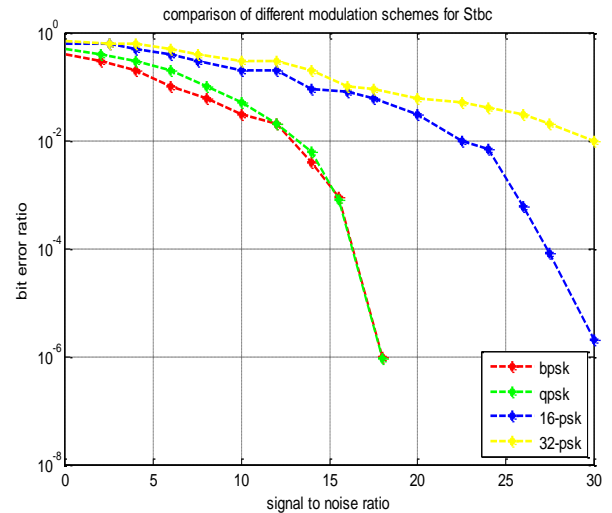
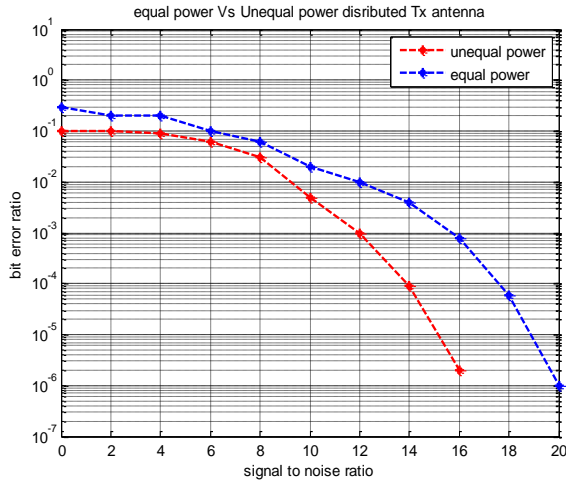


Fig.2. Performance comparison of 2\*2 STBC-OFDM with M-PSK Modulation

The above figure illustrates the details about the BER performance of proposed system with varying modulation technique for the same data stream. The above figure shows

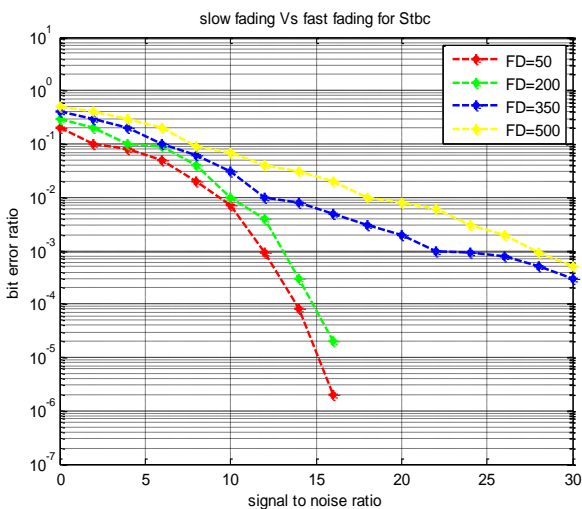
that there is a significant loss in the SNR with increasing in the modulation order and the proposed system has high BER. Here, one more comparison was illustrated though power allocation with respect to equal and non-equal powers allocated for TX antennas. In the case of equal power, both antennas having same power and in the case of unequal second antenna will receive less power compare to first antenna.



**Fig.3 Performance comparison of 2\*2 STBC-OFDM with equal and unequal power distributions**

From the fig. 3, it can be noticed that equal power system gives a 3dB penalty when compared with unequal power system. Fading is unconventionality of the attenuation that a carrier modulated telecommunication signal experiences over a particular propagation media. The terms, fast and slow fading demonstrates the rate at which the phase and magnitude of the signal was changing by the channel. The coherence time was the time above which the value of correlation function is above 0.5. When the delay constraint of the channel was small compared to coherence time of channel, it is called slow fading channel. Fast fading channel can be defined as the channel at which the channel delay constraint is higher compared with coherence time of channel.

The Simulation results for different Doppler frequency (fd) 50, 200, 300, 500 are presented.



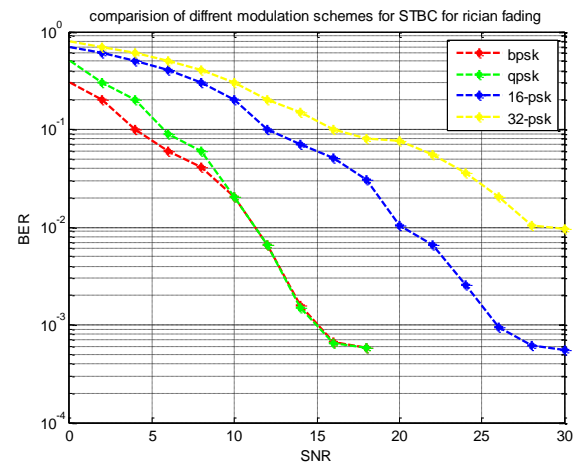
**Fig.4. Performance comparison of 2\*2 STBC-OFDM at various Doppler frequencies**

From the performance result in Figure 4, it can be seen that the performances of STBC-OFDM degrade in fast fading conditions.

#### Case 2: STBC 2X2 MIMO with Rician fading channel

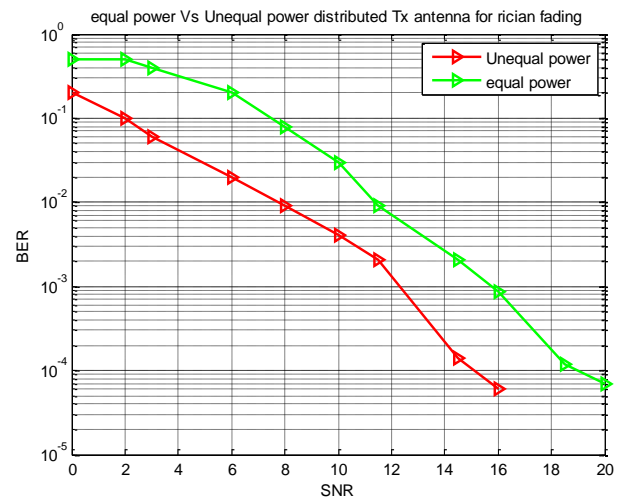
**Table.2: Simulation Parameters**

Parameter	Value
Number of Symbols	1000
Number of Transmitters	2
Number of Receivers	2
Fading Effect	Rician
Noise	AWGN
Doppler Frequency	50Hz
SNR	0-30
System Model	MIMO



**Fig.5. Performance comparison of 2\*2 STBC-OFDM with M-PSK Modulation**

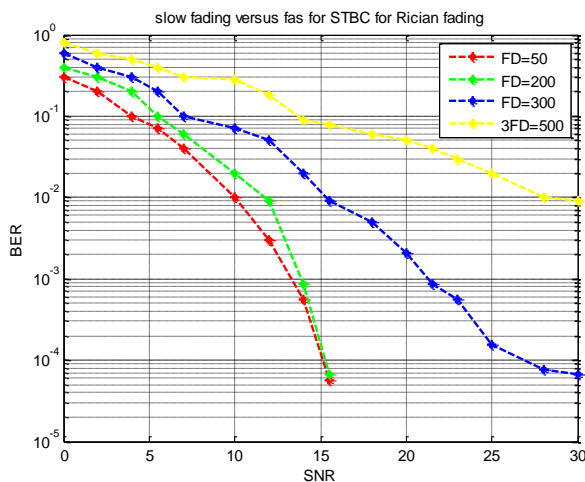
The modulation order specifies the number of the different symbols that can be transmitted by a digital communication system. The simulation results in figure.5 show that increasing the order of modulation will result in a significant SNR loss over rician fading channel. Therefore, it can be said that for the higher order of modulation order, the system will have higher BER.



**Fig.6 Performance comparison of 2\*2 STBC-OFDM with equal and unequal power distributions**

Fig.6 illustrates the results of the proposed approach over rician fading channel with equal an un-equal power distributions. From the fig. 6, it can be noticed that equal power system gives a 3dB penalty when compared with unequal power system.

The coherence time was the time above which the value of correlation function is above 0.5. When the delay constraint of the channel was small compared to coherence time of channel, it is called slow fading channel. Fast fading channel can be defined as the channel at which the channel delay constraint is higher compared with coherence time of channel. The Simulation results for different Doppler frequency ( $f_d$ ) 50, 200, 300, 500 are presented in figure.7. Low values of  $f_d$  denotes slow fading channel and high values denotes fast fading channel.



**Fig.7. Performance comparison of 2\*2 STBC-OFDM with various Doppler frequencies**

From the performance result in Figure.7, it can be observed that the performance of STBC-OFDM degrades in fast fading conditions.

## 5. CONCLUSIONS

This paper proposed an Alamouti based STBC-OFDM approach for various frequency selective fading channels. This approach didn't require channel state information at receiver or at the transmitter. This approach also used MLE for decoding for recursive decoding. The simulation results were done for two types of channels, Rayleigh and rician fading channels. Simulation results show the performance of the proposed recursive-type receiver to be nearly optimum.

## 6. REFERENCES

- [1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction", *IEEE Trans. Inform. Theory*, vol.44, no.2, pp.744-765, Mar. 1998.
- [2] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications", *IEEE J. Select. Areas Comm.*, vol.16, no.8, pp.1451-1458, Oct. 1998.
- [3] V. Tarokh, H. Jafarkhani and A. R. Calderbank, "Space-time block codes from orthogonal designs", *IEEE Trans. Inform. Theory*, vol.45, no.5, pp. 1456-1467, July 1999.
- [4] D. Agrawal, V. Tarokh, A. Naguib and N. Seshadri, "Space-time coded OFDM for high data rate wireless communication over wideband channels", *Proc. IEEE VTC'98*, pp.2232-2236, 1998.
- [5] S. Mudulodu and A. Paulraj, "A transmit diversity scheme for frequency selective fading channels", *Proc. IEEE GLOBECOM' 00*, pp.1089-1093, 2000.
- [6] Z. Liu, G. B. Giannakis, A. Scaglione and S. Barbarossa, "Decoding and equalization of unknown multipath channels based on block precoding and transmit antenna diversity", *Proc. 33rd Asilomar Conf.*, pp.1557-1561, 1999.
- [7] Y. Li, N. Seshadri and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels", *IEEE J. Select. Areas Comm.*, vol.17, no.3, pp.461-471, Mar. 1999.
- [8] M. Uysal and C. N. Georgiades, "Efficient implementation of a maximum likelihood detector for space-time coded systems", *Proc. Allerton Conf*, Oct 2000.
- [9] R. Raheli, A. Polydoros and C. Tzou, "Per-survivor processing: A general approach to MLSE in uncertain environments", *IEEE Trans. Comm.*, vol.43, no.2/3/4, Feb./Mar./Apr. 1995, p.354-364.
- [10] V. Tarokh, N. Seshadri and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction", *IEEE Trans. Inform Theory*, vol. 44, n° 2, pp. 744-765, Mar. 1998.
- [11] S. Baro, G. Bauch and A. Hansmann, "Improved codes for spacetime trellis-coded modulation", *IEEE Comm. Letters*, vol.4 pp. 20-22, Jan. 2000.
- [12] H. Jafarkhani, N. Seshadri, "Super-Orthogonal Space-Time Trellis Codes", *IEEE Trans. Inform. Theory*, vol. 49, n° 4, April 2003, pp.937-951.
- [13] S. Siwamogsatham and M. P. Fitz, "Improved high rate space-time codes via expanded STBC-MTCM constructions" *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Lausanne, Switzerland, June/July 2002, p.106.
- [14] G. Ungerboeck, "Channel coding for multilevel/phase signals", *IEEE Trans. Inform. Theory*, vol. IT 28, pp. 55-67, Jan. 1982.
- [15] W. Huiming, X. Xiang-Gen, Y. Qinye and L. Bin, "A family of space-time block codes achieving full diversity with linear receivers," *IEEE Transactions on Communications*, vol. 57, pp. 3607-3617, 2009.
- [16] A. Slaney and Y. Sun, "Space-time coding for wireless communications: an overview," *IEE Proceedings in Communications*, vol. 153, pp. 509-518, 2006.
- [17] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis and A. R. Calderbank, "Great expectations: the value of spatial diversity in wireless networks," *Proceedings of the IEEE*, vol. 92, pp. 219-270, 2004.
- [18] D. Agrawal, V. Tarokh, A. Naguib and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *IEEE Conference on Vehicular Technology*, vol.3, pp. 2232-2236, 1998.
- [19] D. Torrieri and M. C. Valenti, "Efficiently decoded full-rate space-time block codes," *IEEE Transactions on Communications*, vol. 58, pp. 480-488, 2010.



# A Comprehensive Study on Power Reduction Techniques in Deep Submicron Technologies

V.Kalpna Bharathi<sup>1</sup>, T.Srinivas Reddy<sup>2</sup>

<sup>1</sup>Assistant Professor, Dept. of ECE, MallaReddy Engineering College for Women, Telangana, India

<sup>2</sup>Associate Professor, Dept. of ECE, Malla Reddy Engineering College (Autonomous), Telangana, India

**ABSTRACT:** The use of Very Large Scale Integration (VLSI) technologies in high performance computing, wireless communication and consumer electronics has been growing at a very fast rate. For the most recent CMOS technologies static power dissipation i.e. leakage power dissipation has become a challenging area for VLSI chip designers. A comprehensive study and analysis of various leakage power minimization techniques have been presented in this paper. The present research study and its corresponding analysis are mainly focusing on circuit performance parameters.

**KEYWORDS:** Sub threshold leakage current, Leakage power consumption, Sub-threshold Leakage, Threshold voltage.

## 1. INTRODUCTION

Power consumption is one of the top issues of VLSI circuit design, for which CMOS is the primary technology. Today's focus on low power is not only because of the recent growing demands of mobile applications. Even before the mobile era, power consumption has been a fundamental problem. To solve the power dissipation problem, many researchers have proposed different ideas from the device level to the architectural level and above. However, there is no universal way to avoid tradeoffs between power, delay and area and thus, designers are required to choose appropriate techniques that satisfy application and product needs. With recent advancements in semiconductor technology the density of transistors in Integrated Circuits is still growing, which in turn demands expensive cooling and packaging technologies. Keeping this in view, the supply voltages are scaled down for reducing the switching power dissipation. Moreover, the threshold voltage is also scaled down for the performance tradeoffs. However, the scaling of threshold voltage has resulted in exponential increase of subthreshold leakage current causing leakage (static) power dissipation. Static power dissipation is now growing considerably proportional to the switching dynamic power dissipation in deep submicron technologies and battery operated devices. The longer the battery lasts, the better the leakage power savings [2]-[3]. The four main sources of leakage current in a CMOS transistor are i) Reverse-biased junction leakage current ii) Gate induced drain leakage iii) Gate direct-tunnelling leakage and iv) Subthreshold (weak inversion) leakage current. The subthreshold leakage current being the most predominant amongst all the leakage current sources becomes extremely challenging for research in current and future silicon technologies. In a short channel device, however, the source and drain depletion width in the vertical direction and the source drain potential have a strong effect on the band bending over a significant portion of the device. Therefore, the threshold voltage and consequently the sub threshold current of short channel devices vary with the drain bias. This effect is referred to as Drain-Induced Barrier Lowering (DIBL)

$$I_{ds} = \mu_0 * C_{ox} * \frac{W}{L} * (m - 1) * (V_{th}^2) * e^{\frac{V_{gs} - V_{th}}{(m - 1)V_t}} * (1 - e^{-V_{ds}/V_t})$$

(Where  $m = 1 + C_{dm}/C_{ox}$ ,  $m$  is the sub-threshold swing co-efficient,  $C_{dm}$  is capacitance of the depletion layer,  $C_{ox}$  is the capacitance of the oxide layer,  $\mu_0$  is the mobility,  $V_{th}$  is the threshold voltage,  $V_{gs}$  is the gate to source voltage,  $V_{ds}$  is the drain to source voltage and  $V_t$  is the thermal voltage. According to technology trend for the high density the transistor size should be scaled down and for high performance the threshold voltage ( $V_{th}$ ) scaling down. In the case of sub-threshold leakage if scaling down of  $V_{th}$ , the leakage power increases exponentially as  $V_{th}$  decreases and the short channel effect of the channel controlled by drain. But in case of gate oxide leakage gate tunneling due to thin oxide and high  $k$ -dielectric could be a possible solution and another motivation for leakage power reduction technique can potentially increase the battery life. With application of dual threshold voltage ( $V_{th}$ ) techniques the sleep, zigzag and sleepy stack approaches result in orders of magnitude sub-threshold leakage power reduction. There are several current components which are responsible for the leakage power dissipation in VLSI circuits. The modernised cooling and packaging strategies are of little help to the rapid increase of the power consumption in today's chips. The diagram of leakage current mechanism of deep-submicrometre transistors is depicted below:

# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

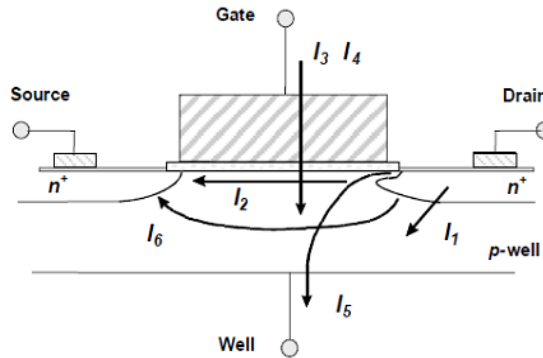


Figure1. Leakage Current Mechanism of deep-submicron transistor

- I1 = Reverse-bias p-n junction diode leakage current
- I2 = Subthreshold leakage current
- I3 = Gate Oxide tunnelling current
- I4 = Hot-carrier injection current
- I5 = Gate induced drain-leakage current (GIDL)
- I6 = Channel punch-through

The dependence of subthreshold current [6] on the transistor parameters are listed in the Table 1.

Table 1: Dependence of sub threshold leakage on device parameters

Parameter	Dependence
Temperature(T)	Exponential increase
Transistor length(L)	Inversely proportional
Transistor width (L)	Directly proportional
Input voltage(V <sub>gs</sub> )	Exponential increase
Transistor Threshold voltage(V <sub>th</sub> )	Increases by an order of magnitude with 100mV decrease

There are various leakage power reduction techniques based on modes of operation of systems. The two operational modes are a) active mode and b) standby (or) idle mode. Most of the techniques aim at power reduction by shutting down the power supply to the system or circuit during standby mode.

## II. LITERATURE SURVEY

In today's VLSI circuits, low power is an important consideration factor along with high performance and high density. There are several techniques to reduce the leakage power but disadvantage of each technique limits its implementation.

**A. Dual VT and MTCMOS:** This is a basic approach to reduce the leakage power. MTCMOS reduces the leakage by introducing the high threshold NMOS gating between pull down network and ground terminal, in series to low threshold voltage circuitry. As stated in [6] Dual VT technique is a variation in MTCMOS, in which high threshold voltage can be assigned to transistors of non-critical path to reduce leakage current and low threshold voltage transistors are used in critical paths. An additional mask layer is required due to VT (Threshold voltage) variation, thereby making fabrication process complicated. This technique suffers from latency period i.e. it needs some time to get into normal operating mode after reactivation. The structure for dual VT and MTCMOS technique is shown as:

# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

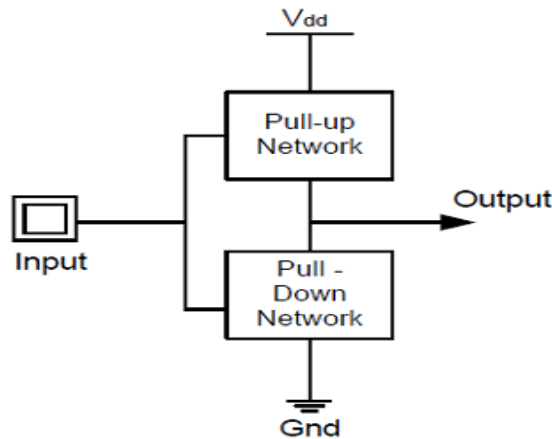


Figure 2: Dual VT and MTCMOS Structure

**B. Sleep Transistor Technique:** Addition to the MTCMOS technique, high VT sleep transistor is introduced between VDD (supply voltage) and pull up network, and between pull down network and ground for high switching speed, where low VT transistors are used in circuit [10]. Efficient power management is done by sleep control mechanism. This modified MTCMOS technique can only reduce the standby leakage power and the introduced MOSFETs results increase in area and delay.

During stand-by mode both sleep transistors get turned off, introducing large resistance in conduction path and thus, leakage current is low. Isolation between VDD and ground path is necessary for leakage reduction. This technique faces a problem for data retention purpose during sleep mode. The Wakeup time and energy of the sleep technique have a significant impact on the efficiency of the circuit.

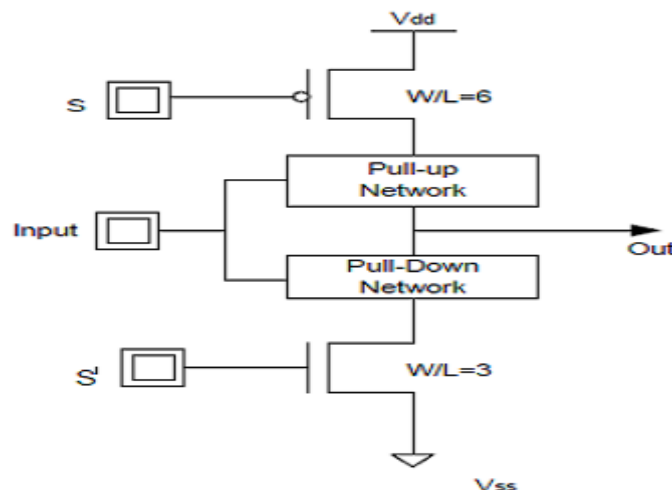


Figure 3: Sleep Transistor Approach

**C. Input vector control (IVC):** The strong dependence of leakage power values on the input combination is given by Abdollahi et al [11] by citing an example of 2-input NAND gate to illustrate the concept of transistor stacking. The minimum leakage causing input vector is identified by an automation process and is applied to the circuit under sleep mode. An algorithm to obtain the minimum leakage vector (MLV) is given by [11].

# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

**D. Variable threshold CMOS (VTMOS):** This technique involves dynamically modifying the threshold voltage during active mode, which is classically known as standby power reduction (SPR). In this method the threshold voltage  $V_{th}$  is raised during the standby mode by connecting the substrate voltage either lower than (for N transistors) or higher than ground (for P transistors). The major drawback of this technique is that it requires an additional power supply, which may not be appropriate in some commercial designs.

## III. PROPOSED ALGORITHM

In LECTOR [8], the concept of effective stacking transistors has been introduced between the VDD and GND for the leakage power reduction. In this technique two leakage control transistors i.e. P-type and N-type are inserted between the pull up and pull down network of a circuit, in which each LCT gate is controlled by the source of other, hence termed as self-controlled stacked transistors. Since it is a self-controlled technique so no external circuit is required for controlling purpose. These LCT produces high resistance path between the VDD and GND by turning more than one transistor OFF, thereby reducing the leakage current. This technique has a very low leakage but there is no provision of sleep mode of operation for state retention.

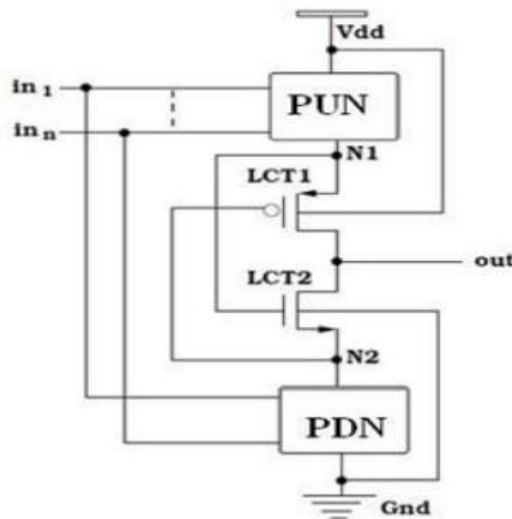


Figure 4: LECTOR Approach

### A. LECTOR Stack State Retention (LSSR)

**LECTOR Stack State Retention (LSSR)** This technique combines the feature of both, LECTOR approach and the Forced Stack Technique with the additional feature of state retention in circuit. The circuit configuration includes [1], two leakage control transistors are added between the pull up and pull down network, and the stack effect is introduced to pull up and pull down network by replacing each existing transistor with two half-sized transistors. It provides the limitation of area because of usage of extra transistors to preserve the circuit state during sleep mode. But this technique provides good leakage current reduction without any delay penalty.



# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 12, December 2015

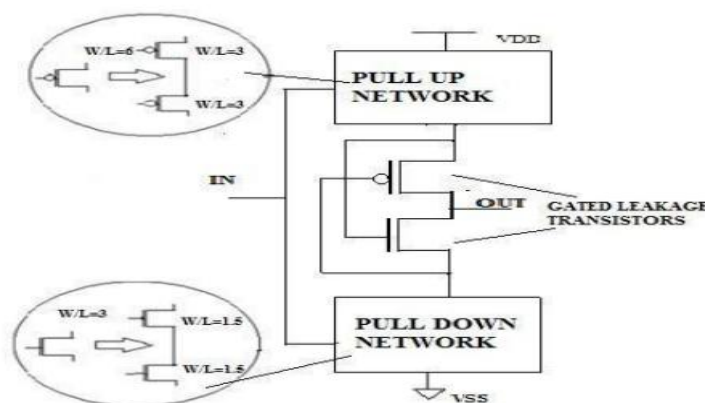


Figure 5: LECTOR Stack State Retention (LSSR) Approach

## IV.CONCLUSION

The present study provides an appropriate choice for leakage power minimization technique for a specific application by a VLSI circuit designer based on sequential analytical approach. To solve the problem, several approaches have been implanted and still work is in progress on many more. The designers, therefore, have to select particular technique depending on application and product requirements. In this paper, we have presented these several leakage power reduction techniques along with their respective advantages and disadvantages. We conclude that LECTOR and the new approach LSSR circuit may lead to much large reduction of leakage power than the previously introduced techniques.

## REFERENCES

- [1] Praveen Kumar, Pradeep SR, Pratibha SR, "LSSR: LECTOR Stacked State Retention Technique a novel leakage reduction and state retention technique in low power VLSI design," IJERT, vol. 2, pp. 1-4, October 2013.
- [2] V.De and S.Borkar, "Technology and design challenges for low power and high performance," in Proc. Int. Symp. Low Power Electronics and Design, 1999, pp.163-168.
- [3] K.Roy and S.C.Prasad, "Low-power CMOS VLSI circuit design". New York: Wiley, 2000, ch.5, pp.214-219.
- [4] Y.Taur, T.H. Ning, "Fundamentals of Modern VLSI Devices", Cambridge University Press, New York, 1998.
- [5] Piguet, C. "Low-Power Electronics Design". CRC Press 2005.
- [6] K. Roy, S. Mukhopadhyay, H. Mahmoodi-Meimand, "Leakage Current Mechanisms and Leakage Reduction Techniques in Deep-Submicrometer CMOS Circuits", In Proc. IEEE, vol. 91, pp. 305-327, Feb., 2003.
- [7] Mohab Anis, Member, IEEE, Shawki Areibi, Member, IEEE, and Mohamed Elmasry, "Design and Optimization of Multithreshold CMOS (MTCMOS) Circuits", IEEE Transactions On Computer-Aided Design Of Integrated Circuits And Systems, Vol. 22, No. 10, October 2003.
- [8] N. Hanchate and N. Ranganathan, "Lector: A technique for leakage reduction in CMOS circuits", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 12, no. 2, pp. 196-205, February 2004
- [9] Kaushal Kumar Nigam, Ashok Tiwari, "Zigzag Keepers: A New Approach For Low Power CMOS Circuit" International Journal of Advanced Research in Computer and Communication Engineering Vol. 1, Issue 9, pp. 694- 699, November 2012
- [10] Khushboo Kumari, Arun Agarwal, Jayvrat, Kabita Agarwal, "Review of Leakage Power Reduction in CMOS Circuits" American Journal of Electrical and Electronic Engineering, 2014, Vol. 2, No. 4, 133-136.
- [11] Abdollahi, A., Fallah, F., and Pedram, M. "Leakage Current Reduction in CMOS VLSI Circuits by Input Vector Control," IEEE Transactions on Very Large Scale Integration (VLSI) Systems 12, February 2004, pp.140-154.

## BIOGRAPHY



**V.Kalpana Bharathi** received B.Tech (ECE) in Alfa college of Engineering and technology, JNTUH, in the year 2005 and M.Tech in Embedded Systems in Rajeev Gandhi Memorial Engineering College, JNTUA, 2011. She is currently working as Assistant Professor, Dept. of ECE, MRCW (Mall Reddy College of Engg for Women), India.



ISSN (Print) : 2320 – 3765  
ISSN (Online): 2278 – 8875

# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 12, December 2015**



**T.SrinivasReddy** received B.Tech (ECE) from (Nagarjuna Institute of Technology and Sciences, JNTUH, 2004) and M.E (Master of Engineering) in Systems Signal Processing (University College of Engg., Osmania University, 2011). He is currently working as Associate Professor. ECE Dept. Malla Reddy Engineering College (Autonomous), India.

# Implementation of a Self-Driven Robot for Remote Surveillance

**Boini Jyothi**

PG Scholar, Dept of ECE  
Malla Reddy Engineering College (Autonomous), Hyderabad, India  
*jyothi.boini@gmail.com*

**T.Srinivas Reddy**

Associate Professor, Dept of ECE  
Malla Reddy Engineering College (Autonomous), Hyderabad, India  
*Srinivasreddy.14@mrec.ac.in*

**Abstract:** *A self-driven mobile robot that can be controlled by itself automatically with the help sensors as well controllable with a smart remote control. The proposed robotic system is intended for remote surveillance in case of natural disasters and explosions, etc. The implemented robotic system is based on ARM 7 architecture along with IEEE 802.15 Bluetooth standard. The image acquisition device attached with the robotic vehicle is capable of streaming the video through IP address based communication channel. The smart remote control feature can be accomplished by a mobile or a laptop with inbuilt Bluetooth. The performance evaluation of the robot has done in terms of range of video streaming and remote accessing.*

**Keywords:** *Bluetooth, Surveillance, Remote access, Self-driven*

## 1. INTRODUCTION

The evolution of mobile robotics from the past few years has been creating many modern robotic applications in the aspects of surveillance, disaster rescue and security, etc. The robotic technology is always intended for mimicking the various human characteristics which we can't say replacing the human being in a specific application. In several applications, robotics will perform the activities where a human being can't even enter into the places. For example, when the natural disasters like earth quakes occurred, there might be a huge loss of property and human lives, and some victims may fight for the life under the pieces of distracted constructions and need someone's help. In such cases, human being can't enter into the narrow places at least to recognize the victim existence, and a robot can enter into such small areas easily because of it is minimized and robust architecture.

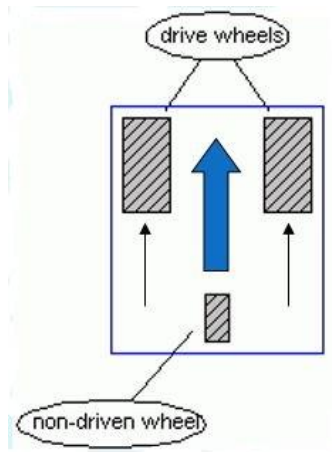
The artificial intelligence in robotics makes it accurate and faster compared to human brains. Robotics are the artificial machines to perform specialized tasks where a human being is ill-suited to perform it. An important concern in the design of robotic machines is about locomotion. The locomotion will depends on the mechanical design of a robot. The locomotion unit of a robot involves a rotational devices drives with a power source and it needs an electrical to mechanical conversion of energy source and that can be achieved by using the motors. Motors will helps the robot to move forward, backward and take turns. The issue is to choose the motor with a sufficient torque and speed.

Surveillance is a security aspect for monitoring the behaviour of objects. The surveillance applications require an image acquisition device to record the video stream or to capture the images of the particular place where an accident or disaster occurred. The places where disasters occurred are usually consists a lot of obstacles for robot movement. In some case, the robots have to move ascending and descending paths and they might be plane surfaces or shapeless surfaces. In such cases, design of the robot should capable of climbing up and down by escaping from obstacles. The design concerns also includes torque and gravity while choosing the motors for ascending movement and for descending movement, payload and mass have to consider. In military applications, the robots have to carry weights, and for that, torque of the motor should be high. Generally the motors with high torque will give a less speed.

## 2. DESIGN STRATEGY

The Design strategy of the proposed system firstly describes about the locomotion unit. For this robotic system, among various types of driving methods, we preferred differential drive. In differential

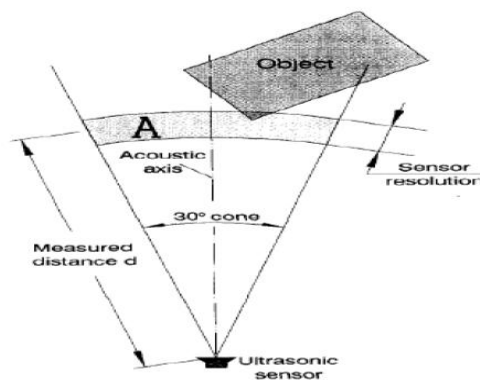
drive method, robot consists of two wheels in the back side and one free wheel in its front. The two back wheels contain one motor on each wheel. The motors connected with the wheels are 60rpm DC motors which will give high torque. The free wheel moves accordingly in any direction based on the programmatically instructions given by the microcontroller. The mechanical body of the robotic system should be good enough for placing the items to be carried on it based on the application. The wheel placement of the differential driving method is shown in the below figure.



**Fig1.** Differential Drive

The power supply is also an important concern for a mobile robot. The robot requires a continuous power source to run the robot for a longer times in remote areas where a human can't give the power supply manually and the power source must be uninterrupted. The motors connected with the robot wheels requires more current compared to other peripherals connected to it. For amplifying the current given from the logic values of microcontroller IO pins, we use a driver circuit between the microcontroller and motors.

The self-driving feature for the proposed robotic system can be accomplished by monitoring the obstacles while moving the robot in any direction, the change in moving direction should be taken based on the direction of and distance from the detected obstacles. The obstacles can be detected by any of the proximity sensing methods. By studying about all the proximity methods, we considered the ultrasonic sensors. The ultrasonic sensors will works based RADAR working principle. The sensor emits microwave signals to the free space and if any obstacles detected by the sensors, then the distance and direction can be analysed by the microcontroller and the instruction can be given to the locomotion unit to take turn by skipping the obstacle. The distance measurement from the obstacle is shown in the below figure.



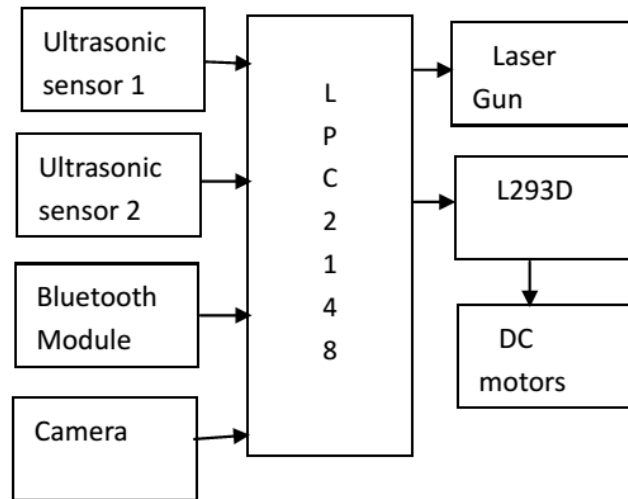
**Fig2.** Ultrasonic Sensor working

In the proposed system, we used 2 ultrasonic sensors to get precise obstacle avoidance and self-driving. The surveillance application can be achieved by using a digital web cam, which is placed on the robot in forward direction and that can be communicated through its IP address. The user can watch and record the video stream from the camera through its IP address in the web browser of

laptop or any mobile device. The range of video streaming depends on the communication range of the established local area network.

### 3. SYSTEM ARCHITECTURE

The system architecture involves a detailed description of hardware modules used in the proposed system.



**Fig3.** Block Diagram of system design process

It is an ARM7 based 32 bit microcontroller with a simplified RISC architecture. It is having many on-board features like SPI, I2C and UART etc.

The communication between the robot and the smart remote gadget can be established with IEEE 802.15 Bluetooth standard. The Bluetooth module connected with the microcontroller will be paired up with the internal Bluetooth in the mobile phone or laptop. The ultrasonic sensors used is HCSR04, which is a 4 pin sensor device. It is having two sections in the sensor, one is trigger and another is echo receiver. The triggering section emits the microwave signals into space, and the echo receiver will wait for receiving the echo signal. The received echo signal will be continuously checked by the microcontroller to obtain the distance and depending upon a specific distance we process the communication with robot in different directions.

The motors used here are gearless DC motors and they require a power source with more power source. To amplify the current drawn from the microcontroller IO lines, we used a driver IC L293D, which can be worked with a source voltage range of 5 to 24v. The speed of the motor will vary depends on the supply voltage given to it.

Additionally, the robotic system consists of a laser gun as a self-rescue feature. In the war field applications, if a human soldier is not able to enter into some narrow places where he/she found an enemy hidden there, he can be able to fire with a laser gun fixed on robot by controlling with his/her mobile phone.

### 4. SOFTWARE IMPLEMENTATION

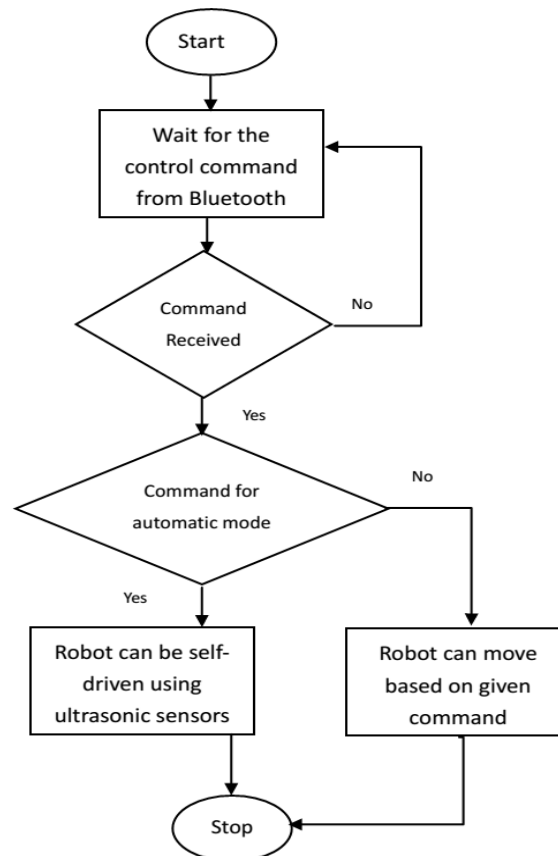
We used 2 software tools for programming the LPC2148. They are

- Keil u Vision
- Flash Magic

Keil uVision is an IDE used for programming the many microcontrollers. It is having a text editor, C/C++ code compiler, assembler and simulator. All the required header files and libraries of various microcontrollers will be included in this IDE. A powerful programming environment will be provided by Keil uVision along with the hex file generation of the compiled C/C++ code which is required to flash into the microcontroller.

Flash magic is a code writing tool used for writing the code in the form of hex file into microcontroller's flash memory. It is also having a terminal application which is used to send/receive serial data to the devices communicated with serial port.

The algorithm followed in this robot is shown in the below flow diagram.

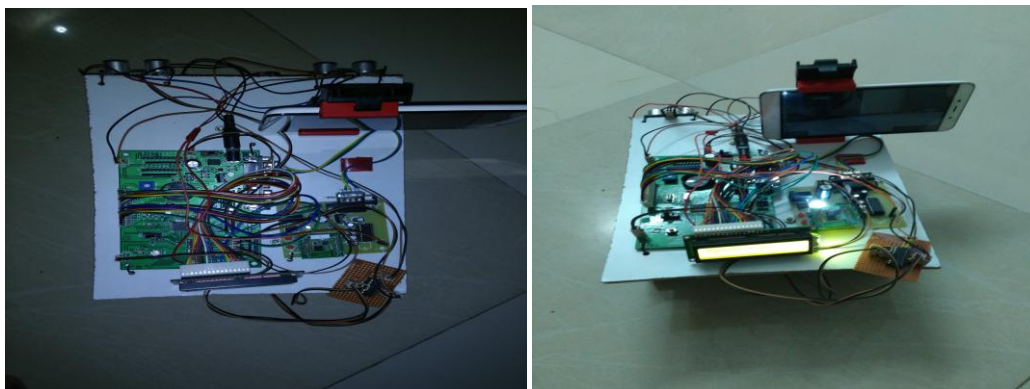


**Fig4.** Flow Chart

## 5. RESULTS

The robotic system was experimentally tested in the aspects of communication range and video streaming distance.

The experimental results of the implemented robotic system is as shown in the below picture.



**Fig5.** Designed Robotic System

## 6. CONCLUSION

The proposed system of self-driven robot is an economic solution for surveillance applications. The robot will accomplish the tasks which can't be done by a human being in some special cases. Due to

usage of robotics in surveillance applications in war fields, we can reduce the man power losses for an unnecessary cause. The system is experimentally tested for video streaming and it is getting a communication range of nearly 50m, which is good enough for many surveillance applications. We can use this system in few more areas like industrial & disaster management by adding fire sensors, temperature sensors.

### **REFERENCES**

- [1] S. Pratheepa and P. Srinivasan, Surveillance Robot for Tracking multiple targets, International Journal for Scientific and Engineering Research, vol.2, issue 3, pp. 1-4,2011.
- [2] T. Hellström, "On the moral responsibility of military robots". Ethics and Information Technology, **15** (2): 99–107, 2013.
- [3] K. Brannen, "Army kills off MULE unmanned vehicle", Military Times, 15 August 2011.
- [4] Robotics.org, accessed on May 10th 2014.
- [5] Parallax.com
- [6] C. Gerald, "Mobile Robot Navigation, Control and remote Sensing", Wiley Publisher, 2011
- [7] W. Budiharto, et al, "A New Obstacle Avoidance Method for Service Robots in Indoor Environments", Journal of Engineering and Technological Science, Vol. 44, No. 2, pp.148-167, 2012. DOI Number :10.5614/itbj.eng.sci.2012.44.2.4
- [8] O. Khatib, , Real-time Obstacle Avoidance for Manipulator and Mobile Robots, International Journal of Robotics Research 5(1), pp.90-98, 1986.
- [9] E. Masehian and Y. Katebi, Robot Motion Planning in Dynamic Environments with Moving Obstacles and Target, Int. Journal of Mechanical Systems Science and Engineering, 1(1), pp. 20-25, 2007.

### **AUTHOR'S BIOGRAPHY**



**Boini Jyothi**, presently pursuing final semester of M. Tech in Embedded Systems in Malla Reddy Engineering College, Hyderabad, Telangana, India.



**Mr. T.Srinivas Reddy**, obtained Masters Degree from University College of Engineering, Osmania University, Hyderabad. Presently working as an Associate Professor, Department of ECE, Malla Reddy Engineering College, Hyderabad, Telangana, India.

## Novel Error-Detection and Correction Technique for Memory Application

Dandem Sudhakar<sup>1</sup>, K.S.Indrani<sup>2</sup>

<sup>1</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (PG Scholar)

<sup>2</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (Associate Professor)

### ABSTRACT

In Advanced digital communication, the testing of Memory System is more complicated. As the technology, dimensions and operating voltages of the computer electronics are reduced to satisfy the consumer's which leads to soft errors. The detection and correction of soft errors in the memory system is more susceptible. To protect memory cells from soft errors, we need more Advanced Errors correction codes. One-step Majority of Logic Decodable codes is suitable for Memory Application to detect and correct large number of errors during communication. One type of Euclidean Geometry Low – Density Parity Check (EG-LPDC) Codes are used for Error correction, because it has fault- secure detector capability. In this paper, an Enhanced Majority Logic Decoder/Detector (MLDD) is proposed to detect silent data errors (SDE) using additional logic and in order to reduce the area of Majority gate, the Sorting Network is designed. Thus, the proposed Method reduces the decoding time, area and power consumption. Hence the proposed Method Simulation Results are Shown as Power saving & Area Utilization compared to existing Method (One-step Majority of Logic Decodable codes).

**Keywords:** Majority Logic Decoder, EG-LPDC, Registers, Counters, Soft Errors, Silent Data Errors (SDE).

### INTRODUCTION

Now a day, digital communication has becoming essential part of life and a lot of data is being transferred. Many communication channels are leads to channel noise. Networks must be able to transfer data from one device to another with acceptable accuracy. For most applications, a system must guarantee that the data received are identical to the data transmitted. The corresponding data will effected by error, which is either soft error or hard error. The Hard error is represented by Hardware Mechanism. Hence the soft error is represented by corrupted data. This Project is developed for Memory Application. When the data is stored in the memory it may be affected from error, this type of error is called soft error. To detect these types of soft errors, many error correction and detection codes are needed. At any time data are transmitted from one node to the next node, they can become corrupted in passage. Many factors can alter one or more bits of message. Some application required a mechanism for detecting and correcting large number of errors. Data can be corrupted during transmission. Some application are required that errors can be detected and corrected. There are different methods to detect and correct errors to keep secure and accuracy data communication. Some applications can tolerate a small level of errors. For example, random errors in audio or video transmission may be tolerable, but when transfer text; we expect a very high level of accuracy.

### Type of Errors

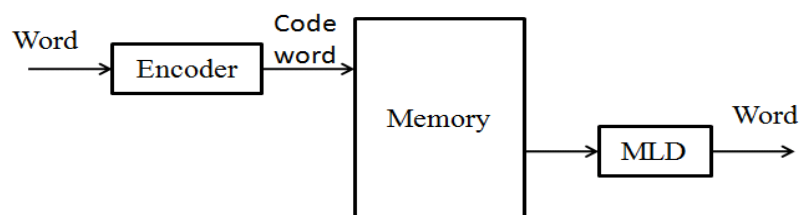
The error is soft because it will change the logic value of memory cells without damaging the circuit or device. The soft errors referred as a Single Event Error/Upset (SEU). If the radiation event is of high energy, more than a single bit may be affected by error, i.e. Multi Bit Error/Upset (MBU) [2]. For reliable communication, errors must be detected and corrected. Some multi-bit

*\*Address for correspondence:*

sudhakar.dandem@gmail.com

error correction codes are BCH codes, Reed Solomon codes but in the following algorithm, it is very difficult and complexity. These codes can correct a Multi bit or large number of errors, but need complex decoders [10], [11]. Among the error correction codes, cyclic block codes have higher capability error detection, decoding complexity is very low and which is called as a Majority Logic (MLD) decoder. The low-density parity-check (LDPC) codes are the linear error correcting code. Which is used to avoid a high decoding complexity [6]-[9]?

In this paper, one specific type of low density parity check codes, namely Euclidean Geometry-LDPC codes [1] are used due to the Faults secure detector capability, higher reliability and lower area overhead. Various error detection techniques are used to avoid the soft error[10].The method is majority logic decoder which used to detect and correct the error in simple way but it requires large decoding time. This method uses the first iteration of majority logic decoding to detect the error present in the codeword is EG-LPDC that are one step majority logic decoder. The main disadvantage is it will detect only single error. If there are no errors, then the decoding process can be stopped without completing the remaining iterations [1].The main reason for using Majority Logic Decoding (MLD) is that it is very easy to implement and it has low complexity [11]. The major drawback of this method is increase the average latency of the decoding process because it depends on the size of the code, thus increases the memory access time. Another method is syndrome fault detector [11] which is an XOR matrix that calculates the syndrome based on the parity check matrix. This method results in a quite complex module and large decoding process with a large amount of hardware and power consumption in the system. The parallel encoders and decoders have been implemented to overcome the drawback of majority logic decoder in which it takes number of cycles to detect the error [11].In this paper, the Majority Logic Decoder/ Detector (MLDD) method [11] used to detect the error in memory device itself, so the data corruption during processing has been eliminated easily to improve the system performance. The MLDD is used the control unit for detecting the error. This method did not detect the silent data error [12].



**Fig1.** General memory system with MLD

The general schematic of memory system implemented with majority logic decoder is shown in figure1. Initially the data words are encoded and then it fed in to the memory. When the memory is read, the corresponding code word is stored in the cyclic shift register with in a Majority Logic Decoder circuit (MLD) before sent to the output. In this decoding process, the code word is corrected if there is 1 bit or more than 1bit change in the data word stored in the memory.

The proposed enhanced MLDD method uses additional error detection technique to detect the silent data error (SDE) in MLDD. To produce the accurate result of MLDD, this addition logic is used to detect the error, which is not detected by the first three iteration of the MLDD. To reduce the number of gates in the majority gate, a sorting network is used. Thus reduces the area of the majority gate and also reduces the power consumption. The main description of this paper is organized as follows. Section II gives an a overview of existing system of Serial one-step Majority Logic decoding; Section

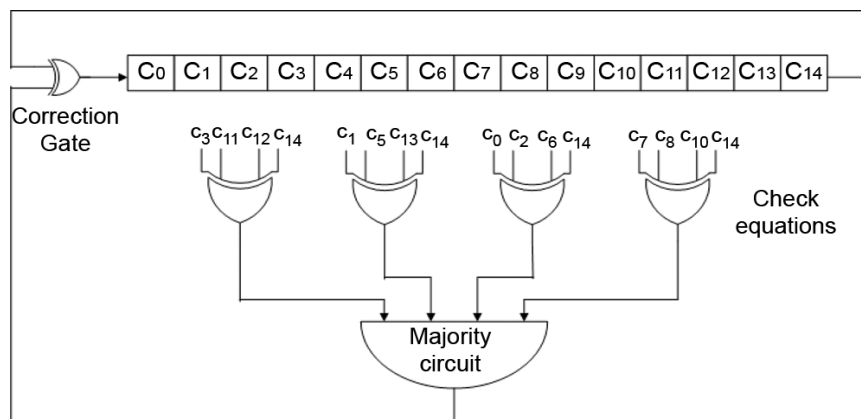
III presents the proposed novel improved ML decoder/detector (MLDD) using Euclidean Geometry Low Density Parity Check Codes (EG-LDPC); Section IV deal with results and discussion for proposed method with RTL Schematics, Simulation Results and comparison table between Existing and proposed Method and power consumption; Finally, Section VI deals with Conclusions.

## EXISTING SYSTEM

These section deals with the Existing method that is Serial One Step Majority Logic Decoder of EG-LPDC used for error detection and correction. This Method detects if the code word is error. The disadvantage of this method is, it will take 15 cycles to detect and correct the entire code. when there are no errors in the data, decoding process will end without completing the rest of cycles and directly sent to the output of MLD. Hence most of the code word is free from error and the decoding time is greatly reduced.

### Serial One-Step Majority Logic Decoder

The Serial one Step Majority Logic Decoder of EG-LPDC, Majority-logic is a simple and effective decoder capable of correcting multiple bit flips depending on the number of parity check sum equations. It consists of four parts: 1) a cyclic shift register; 2) an XOR matrix; 3) a majority gate; 4) an EXOR gate for error correction, as shown in figure2.

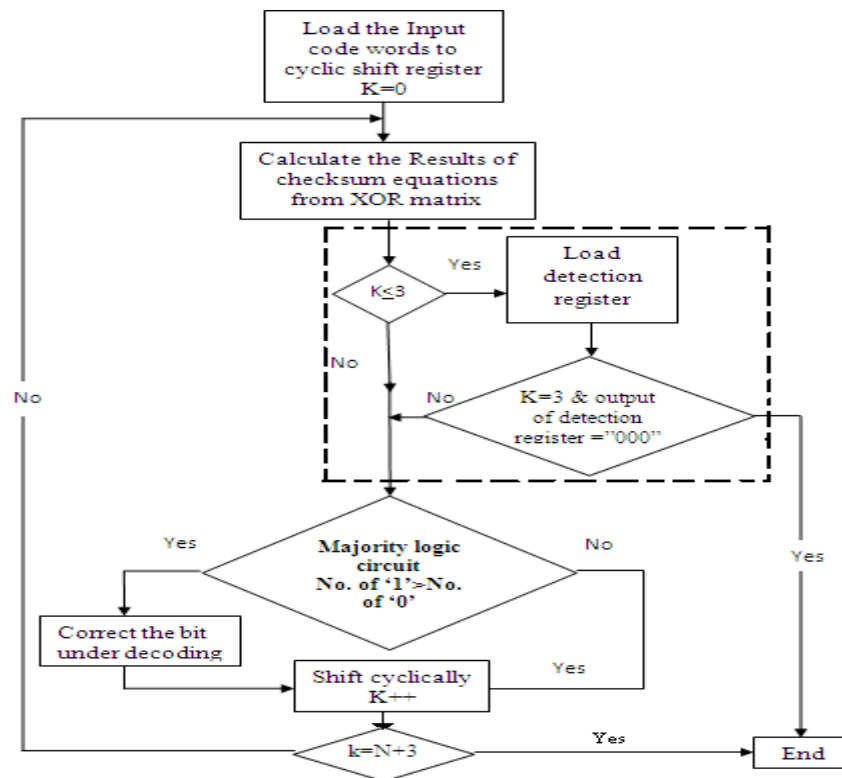


**Fig2.** Serial One- step Majority Decoder for (15, 7) EG-LDPC codes

In one step majority logic decoding [1], initially the code word is loaded in to the cyclic shift register. Then the check equations are computed. The resulting sums are then forwarded to the majority gate for evaluating its correctness. If the number of 1's received in is greater than the number of 0's which means that the current bit under decoding is wrong, and a signal to correct it would be triggered. Otherwise the bit under decoding is correct and no extra operations would be needed on it. In next, the content of the registers are rotated and the above procedure is repeated until codeword bits have been processed. Finally, the parity checksums should be zero if the code word has been correctly decoded. In this process, each bit may be corrected only once, As a result, the decoding circuit is simple, but it require as long decoding time if the code word is large. Thus, by one-step majority-logic decoding, the code is capable of correcting any error pattern with two or few errors. For example, for a code word of 15-bits, the decoding would take 15 cycles, which would be excessive for most applications. The Main disadvantage of serial one-step Majority Logic decoder of EG-LPDC it will take 15 cycles to detect and correct the entire code word. If the code word is large, the decoding process of serial one-step Majority Logic Decoder of EG-LPDC is low. The Speed of the operation is also low and time delay is more.

## PROPOSED SYSTEM (ENHANCED MLDD)

The proposed method is developed in order to overcome the drawback of serial one step Majority Logic Decoder of EG-LPDC, The Proposed method Schematic for 15 bit code word is shown in fig 4. Initially the following code word from Memory is load in to the cyclic shift register. Then the following code is fed into the XOR Matrix for check sum. If the following code word contains error then it performs decoding process to detect and correct the error in the code word. The main advantage of proposed method is it will correct upto five error within three iteration. The proposed version uses the same decoding algorithm as the one in plain ML decoder version. The advantage is that, proposed method stops intermediately in the third cycle when there is no error in data read, [2] as illustrated in Figure. 4, instead of decoding it for the whole codeword size of N. The xor matrix is evaluated for the first three cycles of the decoding process, and when all the outputs  $\{B_j\}$  is “0,” the codeword is determined to be error-free and forwarded directly to the output. On other hand, the proposed method would continue the whole decoding process to eliminate the errors [2] if the  $\{B_j\}$  contain at least a “1” in any of the three cycles. we can clearly understand by seeing the below flow chart of figure 3 figure 3. The proposed Method will detect five errors within the three iterations.



**Fig3.** A detailed schematic of the proposed design for 15 bit code word

A detailed schematic of the proposed design for 15 bit code word is shown in Figure 4. The figure shows the basic ML decoder with a 15-tap shift register, an XOR array to calculate the orthogonal parity check sums and a majority logic circuit which will decide whether the current bit under decoding is erroneous and the need for its inversion. The plain ML decoder [2] shown in Figure 1 is also having the same schematic structure up to this stage. The additional hardware [2] intended for fault detection illustrated in Figure 4 are: a) the control logic unit and b) the output tristate buffers. The control unit triggers a finish flag when there is no errors are detected in data read. The output tristate buffers are always in high impedance state until the control unit sends the finish signal so that the current values are directly forwarded to the output from the shift register.

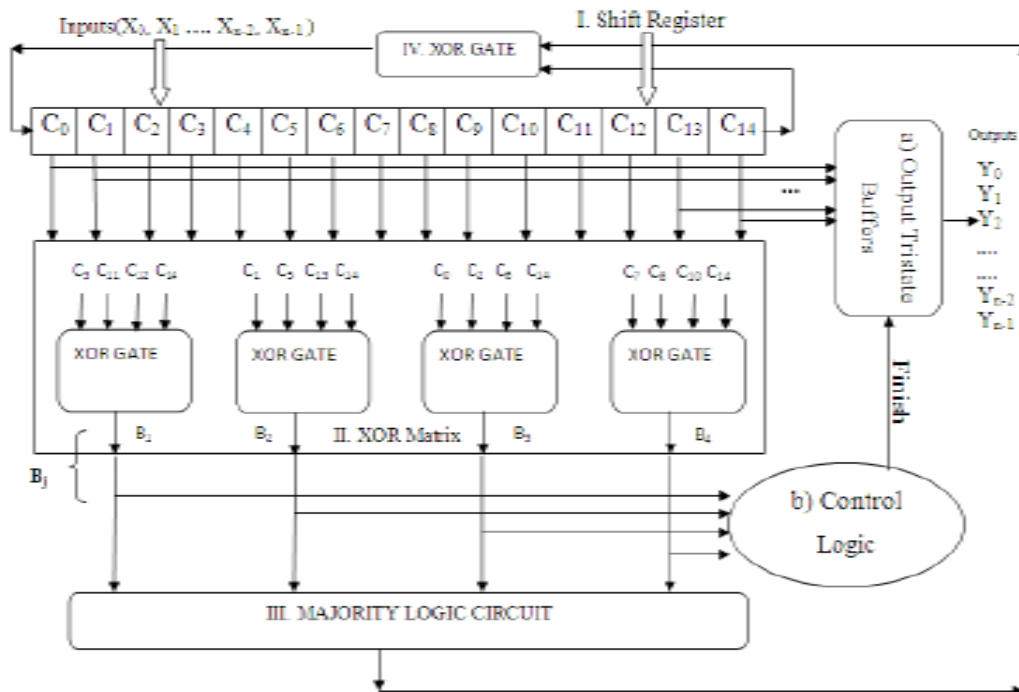


Fig4. Schematic for Proposed Enhanced MLDD

The control logic shown along with proposed Enhanced Majority Logic Decoder/Detector schematic [2] is illustrated in Figure 4. Hence the proposed control unit is shown in below Figure 5. The detection process is managed by the control unit. For distinguishing the first three iterations of the Majority Logic decoding, a counter is used in control unit as shown in figure 5, which counts up to three cycles. The control unit evaluates the output from xor matrix  $\{B_j\}$  by giving it input to the OR 1 gate. This output value is fed to two shift registers which the results of the previous stages stored in it. The values are shifted accordingly. The third coming input is directly forwarded to the OR 2 gate and finally all are evaluated in the third cycle in the OR 2 gate. If the result is “0,” a finish signal is send by the FSM (Finite State Machine) which indicates that the processed code word is error-free, then it send the finish signal to tristate buffer. Hence the following codeword in the register is directly fed to the output. The ML decoding process runs until the end, if the result is “1”.

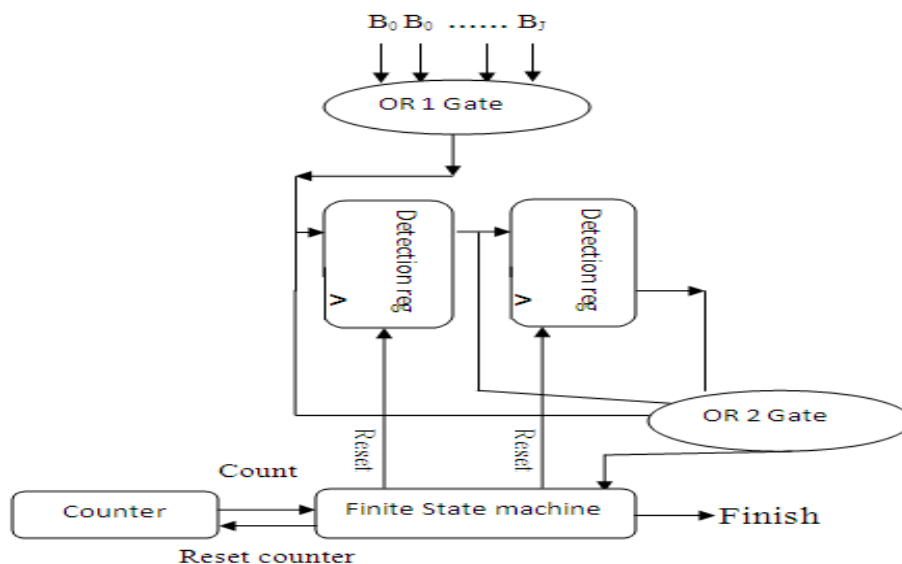


Fig5. Schematic of the Proposed Control Unit

## RESULTS AND DISSCUSION

Hence the simulation results and RTL Schematics of proposed system i.e. Enhanced Majority Logic Decoder/Detector is shown in below, hence the proposed method will correct up to five errors in three decoding cycles. Thus, the proposed Method reduces the decoding time, area and power consumption by comparing with existing system. Hence the proposed Method Simulation Results are Shown as Power saving & Area Utilization compared to codes).existing Method (One step Majority of Logic Decodable. Hence The UART BIST architecture simulation was done through the Xilinx ISE using VERILOG HDL. The data address-bit verification can also to be done through this simulation and the waveform could be verified by using the XILINX. The corresponding simulation and RTL Schematic is shown in the below figure.

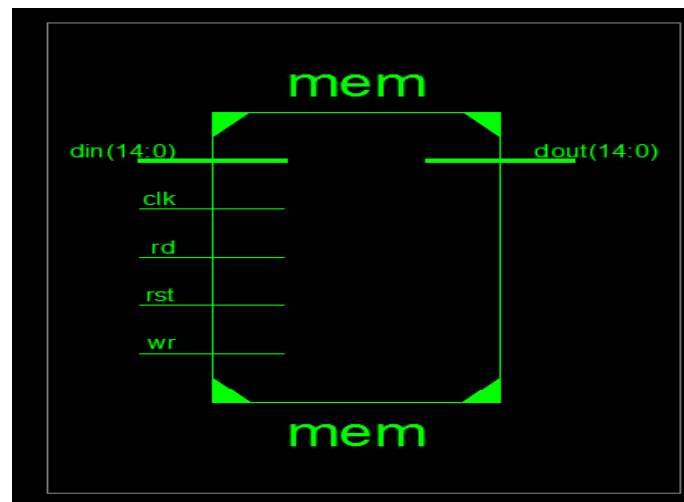


Fig6. Memory Block

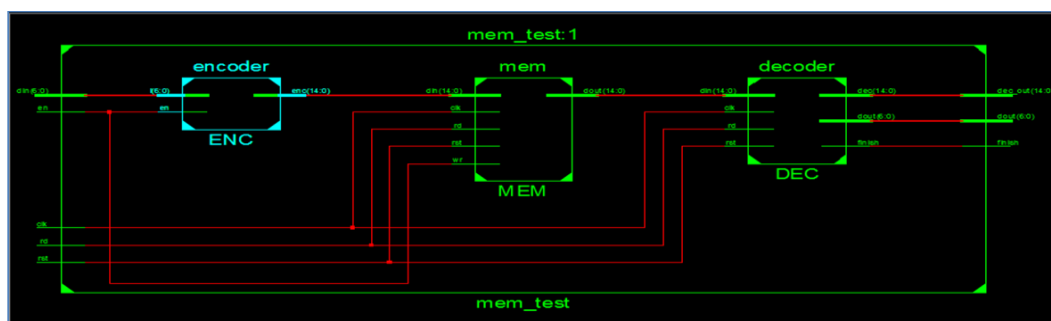


Figure7. RTL for Proposed Method

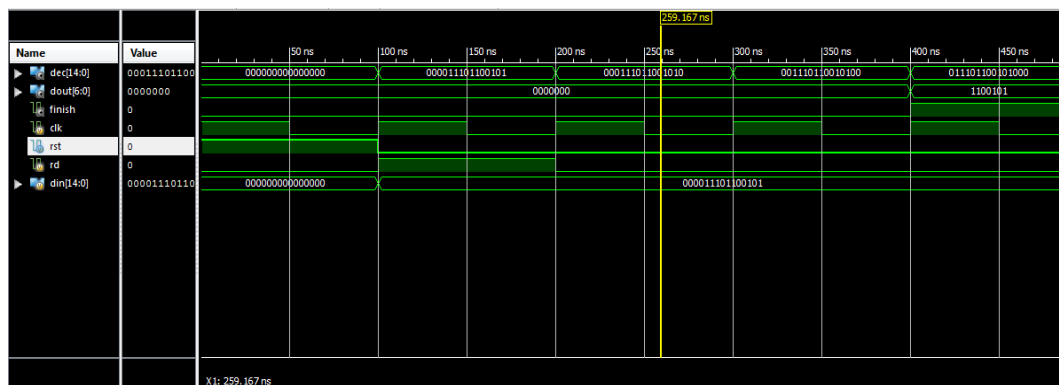


Figure8. Waveform for proposed Method

Comparing between One Step MLD EG-LPDC and Enhanced MLDD

Design	Block Size	Information Bits	No of .MLD Chech Equation	No of Errors that the code can Correct
One Step MLD EG-LPDC	15	7	4	2
Enhanced MLDD	15	7	4	5

**Table1.** Comparisons between Existing Method and Proposed Method

The Results for Power Consumption, Decoding cycles, Area Utilization and Delay

Design	N	Decoding Cycle	Power (in watts)	Area (in no. of slices)	Delay
One Step MLD EG-LPDC	15	15	0.081	120	9.718 ns
Enhanced MLDD	15	3	0.036	25	7.337 ns

## CONCLUSION

In this paper, the detection and correction of errors during the first iteration of serial one step Majority Logic Decoding of EG-LPDC codes has been studied. Hence the proposed method is An Enhanced Majority Logic Decoder/Detector, which will detect and correct up to five errors within three decoding cycles, if no errors are found then the decoding process will be stopped; hence the following code is directly fed to the output. Hence in the proposed system, the decoding time, power, area, delay is reduced by stopping the decoding process when no errors are detected. This papers presents the UART based BIST Architecture using VERILOG HDL. Most of the researchers have been used to implement this testing algorithm in VERILOG for stable, compact and reliable transmission. The structural details have been recognized and it can be integrated into the chip could be easier. The UART transmission could be relatively used in all the devices for the reliable transmission of data's from the structure where it could be converter and tested as a bit files generation. This design function can be adopted as a technical preserving data's for communication. The BIST controller as a device uses as an efficient bit generation for the chip implementation.

## ACKNOWLEDGEMENT

I am very much thankful to my project guide Mrs. K.S.INDRANI, Associate Professor, for his extensive patience and guidance throughout my project work.

I am grateful to our Project Coordinator, Mr.Kesavan Gopal, Associate Professor, for extending his support and cooperation throughout my project work.

I would like to thank Dr. M. JAGADISSH CHANDRA PRASAD, Head of the ECE Department, for providing the freedom to use all the facilities available in the department, for successful completion of project.

I whole-heartedly express my deep sense of gratitude to our beloved Principal, Dr. S.SUDHAKAR REDDY, Principal, MALLA REDDY ENGINEERING COLLEGE (Autonomous) for providing me adequate facilities for my successful completion of both course work and project work.

I sincerely thank all the staff of the Department, for their timely suggestions, healthy criticism and motivation during the course of my project work.

I would also like to thank all my friends for providing help and moral support at the right timing. With great affection and love, I thank my parents who were the backbone behind my deeds.

## REFERENCES

- [1] Pedro Reviriego, Juan A. Maestro, and Mark F. Flanagan, "Error Detection in Majority Logic Decoding of Euclidean Geometry Low Density Parity Check (EG-LDPC) Codes "IEEE Trans. Very Large Scale Integration (VLSI) Systems, Vol.21, No.1, January2013.
- [2] R.C.Baumann, "Radiation-induced soft errors in advanced semiconductor technologies, "IEEE Trans. Device Mater. Reliab., vol.5, no.3, pp.301–316, Sep.2005.
- [3] M.A.Bajura, Y.Boulghas soul, R.Naseer, S.Das Gupta, A.F.Witulski, J.Sondeen, S.D.Stans berry, J.Draper, L.W.Massengill, and J.N.Damoulakis, "Models and algorithmic limits for an ECC-based approach to hard ening sub-100-nm.
- [4] SRAMs," IEEE Trans. Nucl. Sci.,vol. 54, no.4, pp. 935–945, Aug.2007.
- [5] R.Naseer and J.Draper, "DEC ECC design to improve memory reliability insub-100nm Technologies," Proc. IEEE IC ECS, pp.586–589, 2008.
- [6] S.Ghosh and P.D.Lincoln, "Dynamic low-density parity check codes for fault-tolerant nano-scale memory," presented at the Foundations Nano sci. (FNANO), Snowbird, Utah,2007.
- [7] S.Ghosh and P.D.Lincoln, "Low-density parity check codes for error correction in nano scale memory,"SRI Computer Science Lab., Menlo Park, CA, Tech.Rep.CSL-0703,2007
- [8] H.Naeimi and A.DeHon, "Fault secures encoder and decoder for memory applications, "in Proc.IEEE Int.Symp. Defect Fault Toler.VLSI Syst., 2007, pp.409–417.
- [9] B.Vasic and S.K.Chilappagari, information theoretical frame work for analysis and design of nano scale fault-tolerant memories based on low-density parity-checkcodes," IEEETrans. Circuits Syst.I, Reg. Papers, vol.54, no.11, Nov.2007.
- [10] H.Naeimi and A.DeHon, "Fault secure encoder and decoder for nano memory applications," IEEE Trans. Very Large Scale Integer .(VLSI) Syst., vol.17, no.4, pp.473– 486, Apr.2009.
- [11] S.Lin and D.J.Costello, Error Control Coding, 2<sup>nd</sup> ed. Engle wood Cliffs, NJ: Prentice-Hall,2004.
- [12] S.Liu, P.Reviriego, and J.Maestro, "Efficient majority logic fault detection with difference-set codes for memory applications,"IEEE Trans. Very Large Scale Integer.(VLSI) Syst., vol.20, no.1, pp.148–156, Jan.2012
- [13] R.Gallager, "Low-density parity-check codes," IRE Trans. Information Theory, pp.21-28, Jan.1962.

## AUTHOR'S BIOGRAPHY



**DANDEM SUDHAKAR**, B.Tech Passed in 2012 in ECE branch from Kommuri Pratap Reddy institute of Technology, JNTUH, Ranga Reddy. M.Tech from Malla Reddy Engineering College (AUTONOMOUS)

Branch: Digital Systems and Computer Electronics, Dept of ECE.

## 64 Bit Pipelined Hybrid Sparse Kogge-Stone Adder Using Different Valance

Gangula Thirupathi Reddy<sup>1</sup>, Dr. M. K Charan<sup>2</sup>

<sup>1</sup>M. Tech, ECE Department, Malla Reddy Engineering College (Autonomous), Hyderabad, India

<sup>2</sup>Associate Professor, Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India

<sup>1</sup>thiru.steve@gmail.com, <sup>2</sup>kavicharan@mrec.ac.in

**Abstract:** *Present architectures concerns to the advancement in the area, power and speed of different adders, different designs starting from the ripple carry adder to the carry select adder and carry tree adders. There are many carry tree structures like kogge-stone, sparse kogge-stone, spanning tree, brunt Kung and linder-fishcher but these consider only two input black cell or white cell there by the stages will be increased when input length increases. In order to overcome the disadvantages that occurred in previous adder methods a novel method called different valency block cells is adopted. With this we design basic adders along with Valance structure and characteristic evaluation can be carried out. Finally 64 bit hybrid sparse kogge-stone adder with different valency block cells can be designed.*

### 1. INTRODUCTION

Adder is an important operational block that is associated with many applications which demands advance accelerated adder.

Basically ripple carry adder is the primary parallel adder structure. This gives good result when processing bit length is small. As the bit length increases the design produces long carry propagation chain introducing latency [7].

The next level carry skip and carry chain adders that were proposed also could not meet the requirements. There by for better performance we make use of carry select adder and carry tree adder. Even though the carry select adder performs well it requires two pairs of the adder with default carry '0' and '1' as inputs thereby increasing the area considerably [6].

Hence carry tree adders are promising alternatives. In carry tree adder also many structures are proposed which are used for achieving better performance comparatively. On the whole these much one foot back towards the number of levels (black cell and white cell).

For the advancement referring to the reduction of the number of black cells we propose the advanced concept called as the Valance node for the longer bit length. This has the additional feature of pipelining. In our proposed method we designed pipelined hybrid carry tree structure based on the Valance nodes.

This paper proceeds as follows: Section I deals with the basic introduction followed by Section II dealing with the literature survey of the carry tree adder. Section III pertains to the proposed method, Section IV deals with the results generated and discussion followed by Section V of conclusion and references.

## 2. CARRY TREE ADDERS

Carry tree adder is also referred as the parallel prefix adder. It basically consists of the three sub sections namely [2].

- Pre-processing
- Carry generation
- Final addition

In preprocessing, propagate and generate of the corresponding inputs can be calculated. The equations are

$$G=a\&b$$

$$P=a^{\wedge}b$$

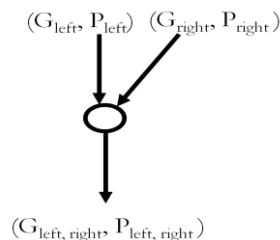
If A, B are the inputs, generate G is produced by performing logical AND operation to the inputs while propagate P is produced by performing the logical XOR operation to the inputs.

The carry generation stage consists of the black cells and white cells.

Black cell takes GP value-pairs as  $(G_{\text{left}}, P_{\text{left}})$  and  $(G_{\text{right}}, P_{\text{right}})$ , then the combined block has the GP values as follows:

$$G_{\text{left, right}} = G_{\text{left}} \mid (P_{\text{left}} \& G_{\text{right}})$$

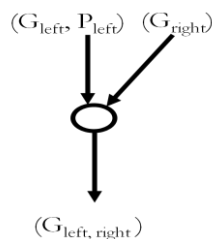
$$P_{\text{left, right}} = P_{\text{left}} \& P_{\text{right}}$$



**Fig1. Black Cell**

White cell is similar to that of the black except the propagate output i.e. it also takes two pair of GP but produces only the G.

Black cell is connected to another black cell or white cell bit. Black cell is not the final state but white cell is the final state.



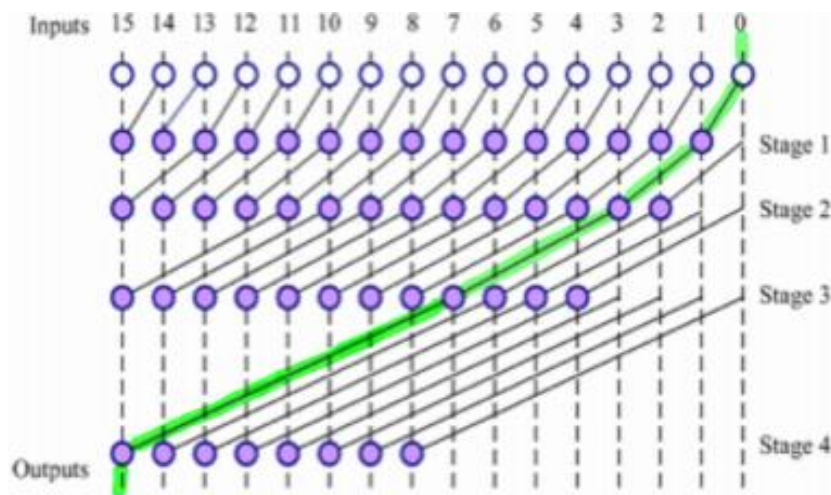
**Fig2. White Cell**

### 2.1. Kogge-Stone Adder

The kogge-stone adder can be constructed using the combination of black and white cells based on the fundamental carry operator (fco)

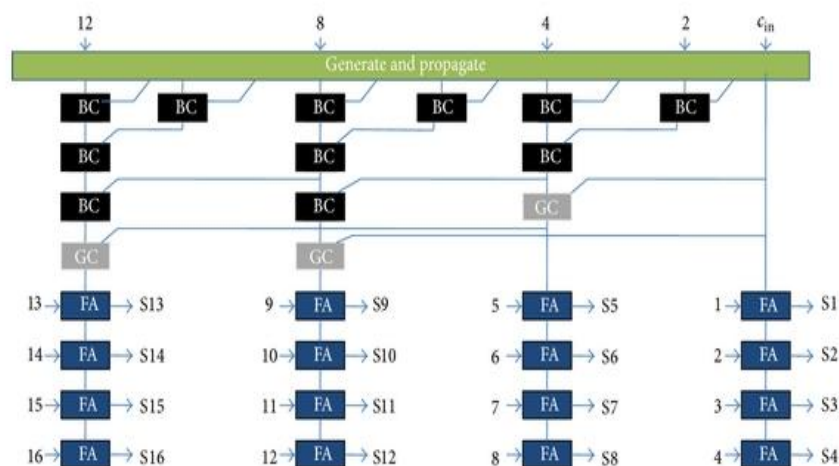
$$(gL, pL) \circ (gR, pR) = (gL + pL \cdot gR, pL \cdot pR)$$

From the design shown below we can easily observe that the count of the black cell and white cell are high. It is designed in such a way that the carry will be generated as the old traditional method. Thereby the carry is generated for the all inputs. This increases the number of stages thereby increasing the area and power [3].



**Fig3.** 16 bit Kogge-Stone adder

By segmenting all the inputs as sub blocks the carry will be generated for the specified blocks only. Those sub blocks will propagate carry forming the full structure. Thus, sparse kogge-stone adder was introduced. [10],[6].



**Fig4.** 16-bit SparseKogge-Stone adder

### 3. PROPOSED METHOD

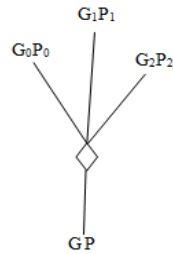
The black cell will be basic building block for the carry tree adder design. Presently it takes only two pairs of propagates and generates and increases as the higher bits are increased.

In order to overcome the demerits we espoused new method as the advance of the black node where we increase the input pairs of PG to reduce the number of levels.

The corresponding equations for the node can be shown below

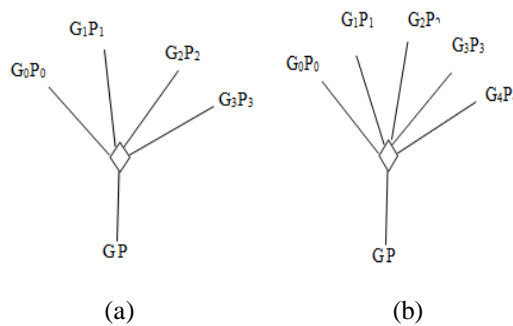
$$G = (g_0 | (g_1 \& p_0) | (g_2 \& p_0 \& p_1))$$

$$P = p_1 \& p_2 \& p_0$$



**Fig5.** Black cell with valency-3

Different valencies like valency-3, valency-4, valency-5 etc. can be constructed to form a huge architecture made pipeline with less stages.



**Fig6.** a) Black cell with valency-4, b) Black cell with valency-5

The equations for the valency-4 node can be shown below

$$G = (g1 \mid (g2 \& p1) \mid (g3 \& p1 \& p2) \mid (g4 \& p1 \& p2 \& p3))$$

$$P = p1 \& p2 \& p3 \& p4$$

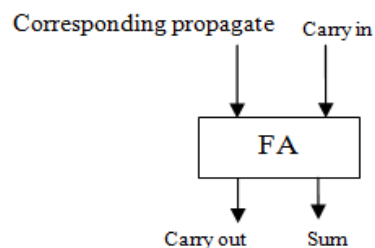
The equations for the valency-5 node can be shown below

$$G = (g1 \mid (g2 \& p1) \mid (g3 \& p2 \& p1) \mid (g4 \& p1 \& p2 \& p3) \mid (g5 \& p1 \& p2 \& p3 \& p4));$$

$$P = p1 \& p2 \& p3 \& p4 \& p5;$$

In our design we are not using the traditional full adder rather we are using a modified full adder.

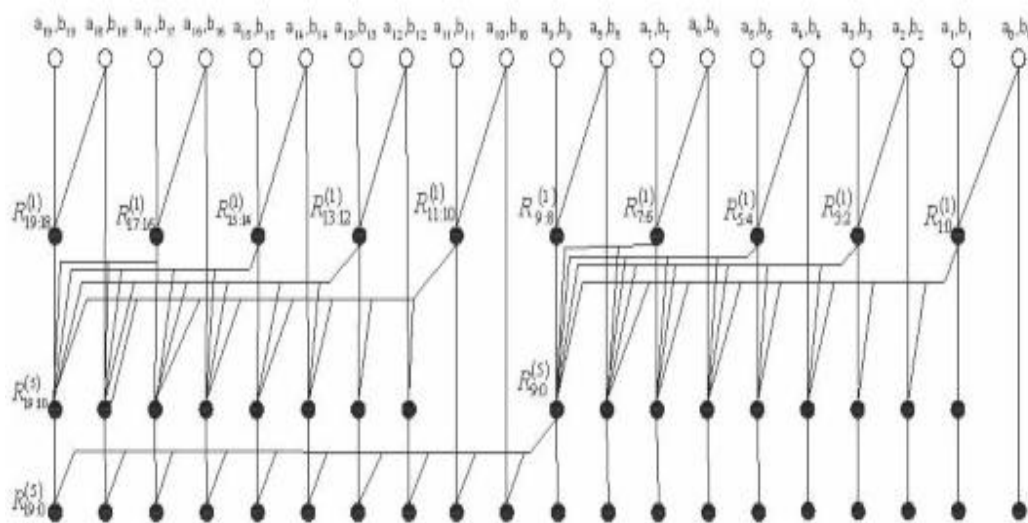
Our full adder takes only two inputs carry and propagate input which was calculated at the first stage of the operation.



**Fig7.** Modified full adder

The proposed 20 bit adder is designed using the different valency. We use the design methods based on the two, five, two valency. In second stage it includes valency-3, valency-4, valency-5. The pipelined method can be a concern to the advancement of the specified concept for improvement of the featured specifications.

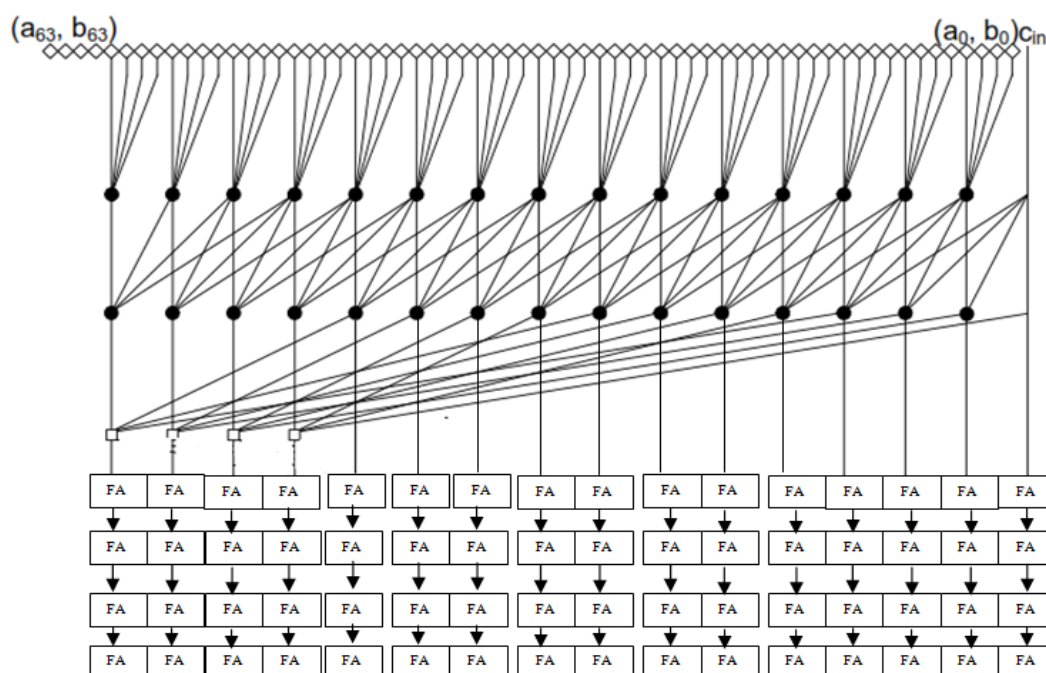
The valency will be design by equal segmentation and by this method we generate all the carry's for the summation. This proposed design will give pipelined parallel design with different values. This shows less gate count as compared to the different structures of carry tree adder gate count, more over our proposed method operation time also less when compared to the other carry tree designs [8],[11],[12].



**Fig8.** Proposed 20-bit adder carry tree with valency  $2*5*2$

In our proposed adder method the concept is similar to that of the normal Sparse Kogge-Stone adder but we utilized the different valency black cell as the elements.

From the design we know the carry's generated for four bits. This carry is given to the block of four full adders which will propagate carry through it. This design consumes less gate count compared with the present existing different adder methods.



**Fig9.** Proposed 64 bit Sparse Kogge-Stone adder

#### 4. RESULTS AND DISCUSSION

	Msgs			
+ /hybridadder64/a	100121	100121727		
+ /hybridadder64/b	549487	549487465955		
+ /hybridadder64/s	549587	549587583586		
+ /hybridadder64/g	000000	000000000000000000000000000000000000...		
+ /hybridadder64/p	000000	00000000000000000000000000000000000011111111110...		
+ /hybridadder64/mp	000000	0000000110100000		
+ /hybridadder64/mg	000000	00000000000001111		
+ /hybridadder64/ng	000101	000101		
+ /hybridadder64/np	000000	000000		
+ /hybridadder64/og	00000	00000		
+ /hybridadder64/op	00000	00000		
+ /hybridadder64/wg	0000	0000		
+ /hybridadder64/wp	0000	0000		
+ /hybridadder64/xp	0000	0000		
+ /hybridadder64/xg	0000	0000		
+ /hybridadder64/temp_c	000000	0000000000001011		
+ /hybridadder64/temp_ca	000000	0000000000001111		

**Fig10.** Proposed 64 bit simulation result

There are five Pipelined stages in our design and each stage operates simultaneously. In our design we give a=100121727 as input in binary format, and give b=549487465955 also in binary format, when added together sum 54958758356 is produced. Different adders were designed and Maximum combinational path delay compared with XILINX 14.3.

**Table1.** Maximum combinational path delay of different adders

Type of adder	Maximum combinational path delay
16 bit Kogge_Stone[11]	14.041ns
16 bit Ripple_Carry[11]	24.686ns
16 bit Spanning_Tree[11]	21.720ns
16 bit Sparse_Kogge[11]	17.527ns
16 bit carry_select_adder[12]	17.501ns
16 bit carry_skip[8]	24.841ns
Proposed method	15.450ns

Our design shows the better speed performance compared with the different adder structures that are listed in the table. Our design shows the 1 ns delay is greater than the kogge-stone adder but kogge-stone adder takes 5% of LUT utilization but our design consumes only 3% of LUT.

#### 5. CONCLUSION

In our proposal we designed different adders like kogge-stone, brunt kung, spanning tree, sparse kogge-stone and proposed hybrid adder of same bit width and the characteristic evaluation can be done, which shows our proposal gives good records. Moreover we designed the 64-bit sparse kogge-stone adder also designed. For designing Verilog HDL can be used, simulated using Model-Sim and synthesized by XILINX configured to Spartan-6 FPGA. By this, we can conclude that this method not only give the better result for the smaller numbers but also for the higher bit lengths. This will be well suited for the faster applications which demanding less power and area.

#### REFERENCES

- [1] N. H. E. Weste and D. Harris, *CMOS VLSI Design*, 4edition, Pearson–Addison-Wesley, 2011.
- [2] R. P. Brent and H. T. Kung, "A regular layout for parallel adders," *IEEE Trans. Comput.*, vol. C-31, pp.260-264, 1982.

- [3] D. Harris, "A Taxonomy of Parallel Prefix Networks," in *Proc. 37th Asilomar Conf. Signals Systems and Computers*, pp. 2213–7, 2003.
- [4] P. M. Kogge and H. S. Stone, "A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations," *IEEE Trans. on Computers*, Vol. C-22, No 8, August 1973.
- [5] P. Ndai, S. Lu, D. Somesekhar, and K. Roy, "Fine Grained Redundancy In Adders," *Int. Symp. On Quality Electronic Design*, pp. 317-321, March 2007.
- [6] T. Lynch and E. E. Swartzlander, "A Spanning Tree Carry Look ahead Adder," *IEEE Trans. on Computers*, vol. 41, no. 8, pp. 931-939, Aug. 1992.
- [7] D. Gizopoulos, M. Psarakis, A. Paschalis, and Y. Zorian, "Easily Testable Cellular Carry Lookahead Adders," *Journal of Electronic Testing: Theory and Applications* 19, 285-298, 2003.
- [8] S. Xing and W. W. H. Yu, "FPGA Adders: Performance Evaluation and Optimal Design," *IEEE Design & Test of Computers*, vol. 15, no. 1, pp. 24-29, Jan. 1998.
- [9] M. Becvár and P. Štukjunger, "Fixed-Point Arithmetic in FPGA," *Acta Polytechnica*, vol. 45, no. 2, pp. 67-72, 2005.
- [10] K. Vitoroulis and A. J. Al-Khalili, "Performance of Parallel Prefix Adders Implemented with FPGA Technology," *IEEE Northeast Workshop on Circuits and Systems*, pp. 498-501, Aug. 2007.
- [11] David H. K. Hoe, Chris Martinez and Sri Jyothsna Vundavalli "Design and Characterization of Parallel Prefix Adders using FPGAs" 2011 IEEE.
- [12] O. J. Bedrij, "Carry-select adder," *IRE Trans. Electron. Comput*, pp. 340–344, 1962.

## Design and Implementation of Digit Serial Fir Filter

Mohammed Arif<sup>1</sup>, K.S.Indrani<sup>2</sup>

<sup>1</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (PG Scholar)

<sup>2</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (Associate Professor)

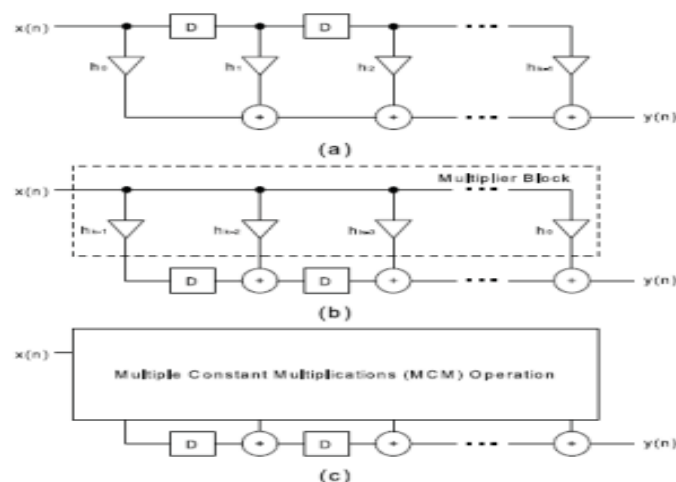
### ABSTRACT

In many digital signal processing systems, the multiple constant multiplication (MCM) operation has significant impact on the complexity and performance of the design because a large number of constant multiplications are required and is also a performance bottleneck in many other DSP systems. For this purpose many efficient algorithms and architectures such as shift-add and common sub-expression elimination (CSE) have been introduced for the design of low complexity MCM operations. These algorithms are very accurate but it contains complexity if suppose we increase the bit length. Therefore to overcome this disadvantage and for low complexity MCM operations we introduce an digit serial fir filter whose operators occupy less area and are independent of data word length. Thus the proposed method reduces the delay, area and power consumption. Hence the proposed method simulation results are shown as less delay and area utilization compared to existing common sub-expression elimination (CSE) method.

**Keywords:** Digit serial arithmetic, multiple constant multiplication (MCM), common sub-expression elimination, finite impulse response (FIR) filters.

### INTRODUCTION

In digital signal processing (DSP) systems such as fast Fourier transforms, discrete cosine transforms (DCT's), and error correcting codes, Finite impulse response (FIR) filters have great importance because of their linear- phase characteristics and feed forward implementations. The two forms of FIR filters i.e, direct and transposed form FIR filter implementations are illustrated in fig 1(a) and (b).



**Figure1.** FIR filters implementations. (a) Direct form. (b) Transposed form with generic multipliers. (c) Transposed form with an MCM block.

Although both architecture have similar complexity in hardware, the transposed form is generally preferred because of its higher performance and power efficiency [1]. The multiplier block of the digital FIR filter in its transposed form, where the multiplication of filter coefficients with the filter

*\*Address for correspondence:*

arif.mohd29691@gmail.com

input is realized, has significant impact on the complexity and performance of the design because a large number of constant multiplications are required. This is generally known as multiple constant multiplications (MCM) operation. Although area, delay, and power- efficient multiplier architectures, such as Wallace [2] and modified booth [3] multipliers, have been proposed, the full flexibility of a multiplier is not necessary for the constant multiplications, since filter coefficients are fixed and determined beforehand by the DSP algorithms [4].

### Shift- Add Architecture

In this method the multiplication of filter coefficients with the input data is generally implemented under a shift adds architecture [5], where each constant multiplication is realized using addition/subtraction and shift operations in an MCM operation [Fig 1(c)]. For the shift adds Implementation of constant multiplications, a straightforward method, generally known as digit based recoding [6], initially defines the constant in binary. Then, for each “1” in the binary representation of the constant, according to its bit position, it shifts the variable and adds up the shifted variables to obtain the result.

As a simple example, consider the constant multiplications  $29x$  and  $43x$ . Their decompositions in binary are listed as follows: Which requires six addition operations as illustrated in Fig. 2(a).

$$29x = (11101) b_{in}^x = x \ll 4 + x \ll 3 + x \ll 2 + x \quad (1)$$

$$43x = (101011) b_{in}^x = x \ll 4 + x \ll 3 + x \ll 2 + x \quad (2)$$

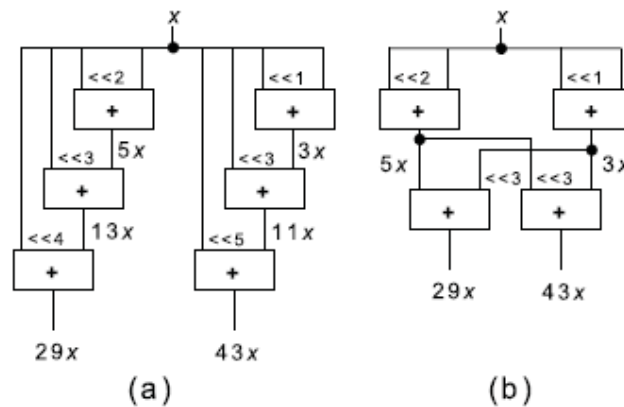
### EXISTING SYSTEM

MCM is involved to produce constant multiplication in Digital Signal Processing (DSP) systems, MIMO (Multiple Input Multiple Output) systems, Error correcting codes, Frequency multiplication, Graphics and Control applications. In such applications full fledge of multipliers are not needed. Since coefficients are constant to produce constant multiplication. Once the MCM architecture is constructed, it can be called as many times it required. Constant multiplication either can be done by digit parallel design or digit serial design. Digit parallel design of constant multiplier needs external wire for shifting. It requires more area while implementation takes place in FPGA or any other ASIC. Hence digit serial design overcomes area constrain with acceptable delay timing. Multiplication with constant is called constant multiplication.

#### Common Sub-Expression Elimination Method

In existing, in order to design MCM architecture without partial product sharing algorithm (Digit based recoding), Common Sub-expression Elimination (CSE) algorithm [7]-[9] is used in existing methods. According to MCM principle constant multiplication is performed by number of shifting and addition operation. For the pair,  $29x$  and  $43x$  without partial product sharing algorithm requires six addition and six shifting operations and CSE algorithm requires four additions and four shifting operations. As shown in below figure 2(a) and 2(b).

However, the digit-based recoding technique does not exploit the sharing of common partial products, which allows great reductions in the number of operations and consequently and consequently, in area and power dissipation of the MCM design at the gate level. Hence, the fundamental optimization problem, called the MCM problem, is defined as finding the minimum number of addition and subtraction operations that implement the constant multiplications. Note that, in parallel design of constant multiplications, shifts can be realized using only wires in hardware without representing any area cost.

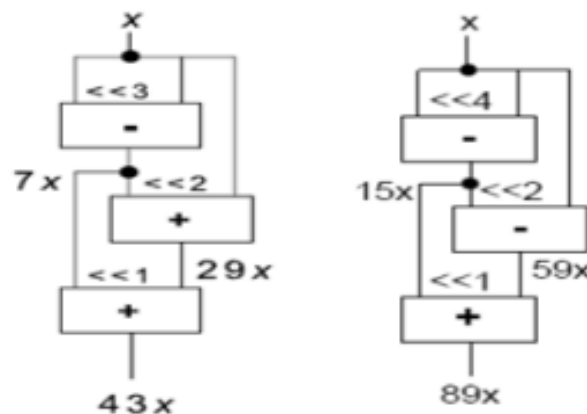


**Figure2.** Shift-adds implementations of  $29x$  and  $43x$ . (a) Without partial product sharing [7] and with partial product sharing. (b) Exact CSE algorithm [10]

The algorithm design for the MCM problem is common sub expression elimination(CSE) algorithm initially extract all possible sub expressions from the representation of the constants when they are defined under binary ,canonical signed digit(CSD) [7], or minimal signed digit(MSD) [8]. Then they find the “best” sub expression, generally the most common, to be shared among the constant multiplications. Returning to our example in fig.2 the exact CSE algorithm gives a solution with four operations by finding the most common partial products  $3x=(11)$ binary  $x$  and  $5x=(101)$ binary  $x$  when constants are defined under binary, as illustrated in fig.2. Compare to the digit based recoding technique CSE MCM algorithm is very accurate but it contains complexity if suppose we increase the bit-length. We know that whenever we increase the bit length the architecture of multiplier will gets changes so that changing the bit length will increases the architecture in MCM. So, because of this disadvantage what happens was the efficiency of the system will gets reduced because with the increased architecture delay will be more so that area, power will gets increases. In order to overcome these constraints we proposed a new method which is a Digit-series FIR Filter

## PROPOSED SYSTEM

Here in proposed method which is a Digit-Series FIR Filter, here we are increasing the bit length so that what happens means even though we are increasing the bit length , architecture will not gets increases it just divides the bit length [13] so that there is no such a complexity exists in this new method. We implemented for both  $29x, 43x, 59x, 89x$ , so in this we are going to compare the delay area, efficiency and power between the existing method and proposed method. So compare to existing method proposed method must gives the less delay compare to previous method.



**Figure3.** 29,43,59,89 coefficient for Digit-series FIR Filters

However, in existed method the algorithms assume that the input data x is processed in parallel. On the other hand, in digit serial arithmetic, the data words are divided into digit sets, consisting of d bits, that are processed one at a time. Since digit-serial architectures offer alternative low complexity designs when compared to the bit-parallel MCM design.

$$29x = x \ll 3 + x \ll 2 + x.$$

$$43x = x \ll 2 + x \ll 1 + x.$$

$$59x = x \ll 4 - x \ll 2 - x.$$

$$89 = x \ll 2 - x \ll 1 - x.$$

In existing method the output for each and every coefficients will needs only one clock cycle. In proposed method the output which are coefficients will needs atleast two or more than two clock cycles i.e, the output will comes after some clock cycles but the complexity will be less so that even though it needs more clock cycles to retrieve the output delay will be very less and efficiency will be more.

## RESULTS AND DISSCUSION

The proposed design has been simulated using Xilinx and Modelsim, the wave form obtained after simulating is as shown in below figs.

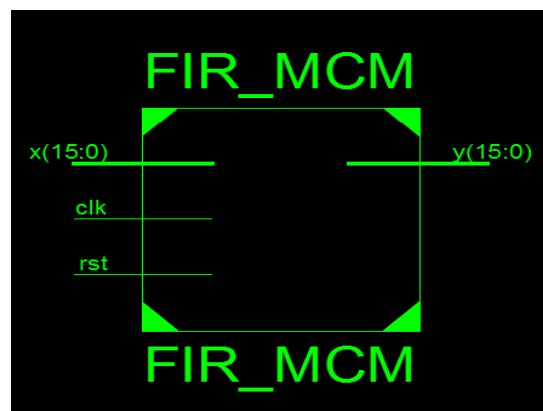


Figure4. Block diagram of FIR\_MCM

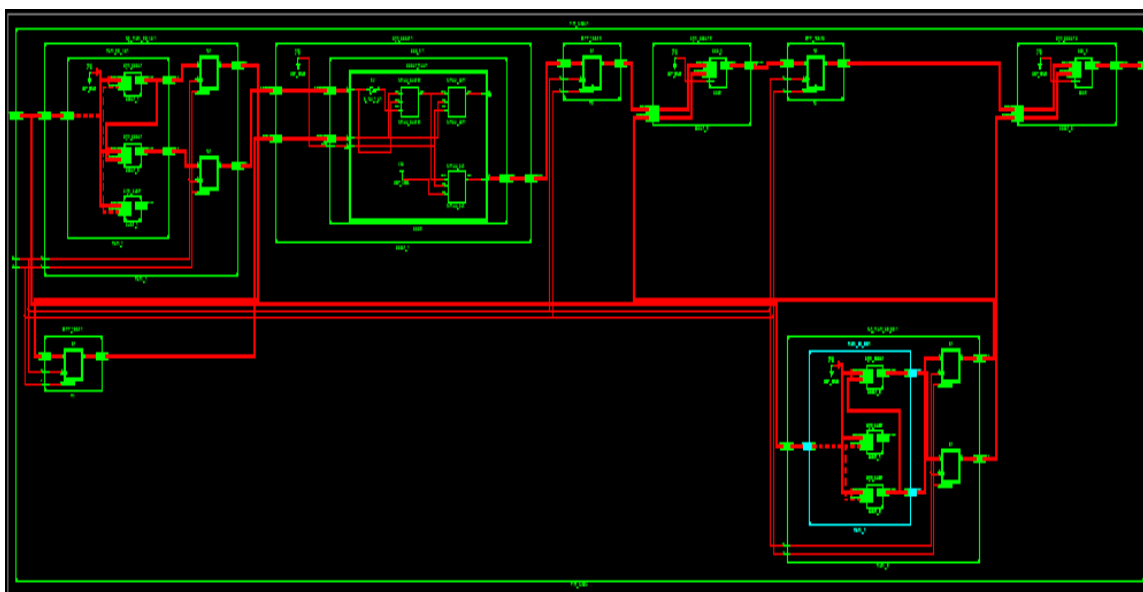


Figure5. RTL ARCHITECTURE of FIR MCM

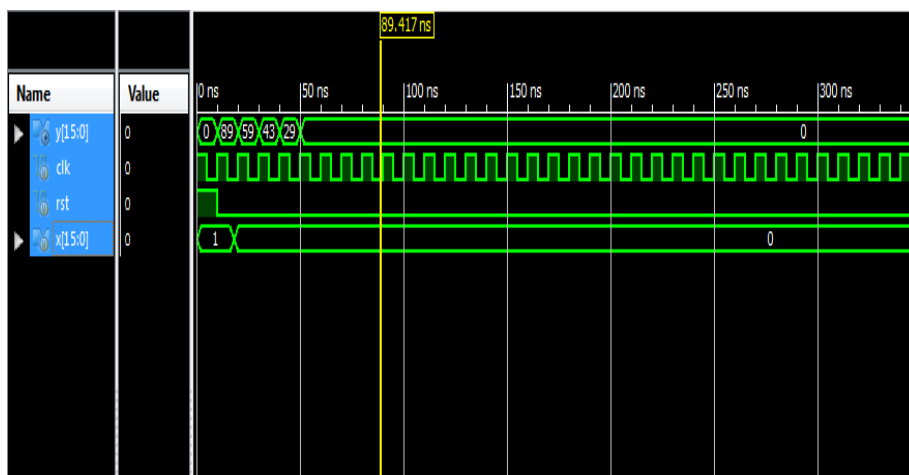


Figure6. Output Waveform of FIR MCM

Table1. Delay for Existing FIR MCM

Timing constraint: Default OFFSET OUT AFTER for Clock 'clk'	
Total number of paths / destination ports: 272 / 16	
Offset:	10.889ns (Levels of Logic = 9)
Source:	mcm_1/m_43_0 (FF)
Destination:	y<14> (PAD)
Source Clock:	clk rising

Table2. Area of Existing FIR MCM

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slice Registers	105	18224	0%
Number of Slice LUTs	271	9112	2%
Number of fully used LUT-FF pairs	84	292	28%
Number of bonded IOBs	34	232	14%
Number of BUFG/BUFGCTRLs	1	16	6%

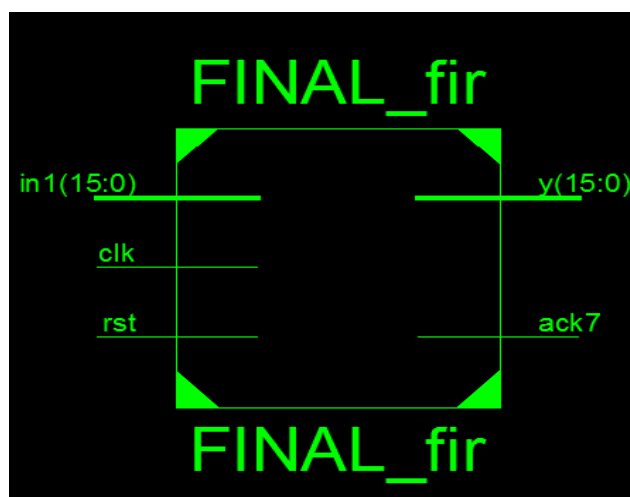


Figure7. Block Diagram for Digit Serial FIR FILTER

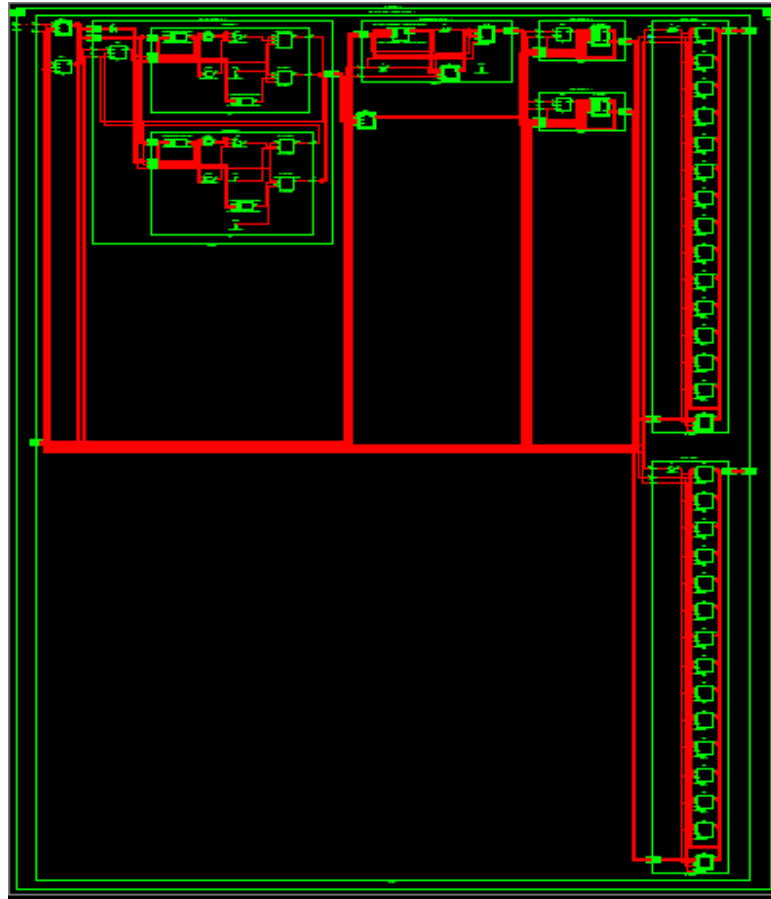


Figure8. RTL schematic architecture for DIGIT SERIAL FIR FILTER

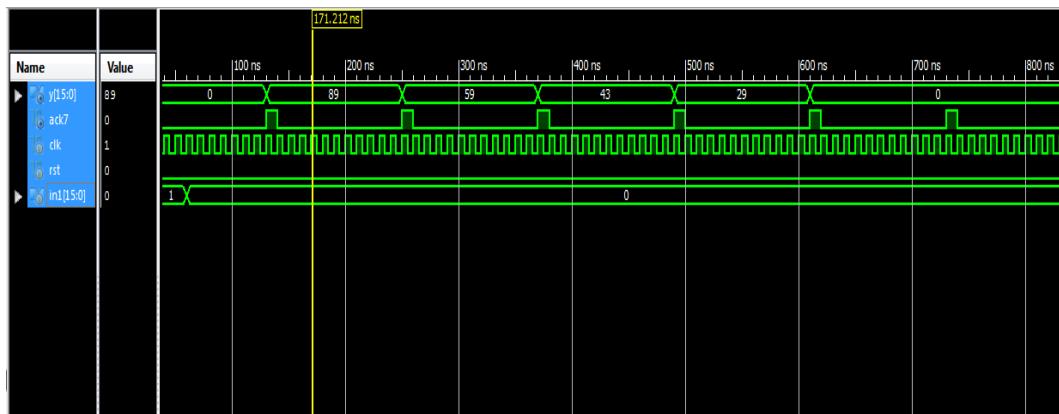


Figure9. Output Waveform for Digit Serial FIR FILTER

Table3. Delay for DIGIR SERIAL FIR FILTER

Timing constraint: Default OFFSET OUT AFTER for Clock 'clk'					
Total number of paths / destination ports: 17 / 17					
-----					
Offset:	4.323ns (Levels of Logic = 1)				
Source:	adder_3/ack (FF)				
Destination:	ack7 (PAD)				
Source Clock:	clk rising				
Data Path: adder_3/ack to ack7					
Cell:in->out	fanout	Gate Delay	Net Delay	Logical Name (Net Name)	
-----					
FDR:C->Q	33	0.447	1.305	adder_3/ack (adder_3/ack)	
OBUF:I->O		2.571		ack7_OBUF (ack7)	
-----					
Total		4.323ns (3.018ns logic, 1.305ns route)			
		(69.8% logic, 30.2% route)			

**Table4.** Area of SERIAL FIR FILTER

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slice Registers	264	18224	1%
Number of Slice LUTs	182	9112	1%
Number of fully used LUT-FF pairs	90	356	25%
Number of bonded IOBs	35	232	15%
Number of BUFG/BUFGCTRLs	1	16	6%

## CONCLUSION

In this paper, we introduce the Digit Serial FIR Filter it divides the input bit-length and then by using shifting, addition, subtraction it process its operation and gives the output at different clock cycles. But in previous method it was designed by using MCM, in this method it has high complexity but it gives the output at single clock cycle only. The main difference between the existed and proposed method was about complexity, area, delay and power .In, existing the output need only single clock cycle but in proposed the output achieves at different clock cycle which is one of the constraint in this proposed method. Here, we are comparing Delay, Area, Power and Efficiency with the previous method which is a MCM method. So, based on Experimental results the proposed method got less delay and area and efficiency got increased when compared to the previous method.

## ACKNOWLEDGEMENT

I am very much thankful to my project guide Mrs. K.S.INDRANI, Associate Professor, for his extensive patience and guidance throughout my project work.

I am grateful to our Project Coordinator, Mr.Kesavan Gopal, Associate Professor, for extending his support and cooperation throughout my project work.

I would like to thank Dr. M. Ch. P. Jagadesh, Head of the ECE Department, for providing the freedom to use all the facilities available in the department, for successful completion of project.

I whole-heartedly express my deep sense of gratitude to our beloved Principal, Dr. S.SUDHAKAR REDDY, Principal, MALLA REDDY ENGINEERING COLLEGE (Autonomous) for providing me adequate facilities for my successful completion of both course work and project work.

I sincerely thank all the staff of the Department, for their timely suggestions, healthy criticism and motivation during the course of my project work.

I would also like to thank all my friends for providing help and moral support at the right timing. With great affection and love, I thank my parents who were the backbone behind my deeds.

## REFERENCES

- [1] L. Wanhammar, DSP Integrated Circuits. New York: Academic, 1999.
- [2] C. Wallace, “A suggestion for a fast multiplier,” IEEE Trans. Electron. Comput., vol. 13, no. 1, pp. 14–17, Feb. 1964.
- [3] W. Gallagher and E. Swartzlander, “High radix booth multipliers using reduced area adder trees,” in Proc. Asilomar Conf. Signals, Syst. Comput., vol. 1. Pacific Grove, CA, Oct.–Nov. 1994, pp. 545–549.
- [4] J. McClellan, T. Parks, and L. Rabiner, “A computer program for designing optimum FIR linear phase digital filters,” IEEE Trans. Audio Electroacoust., vol. 21, no. 6, pp. 506–526, Dec. 1973.

- [5] H. Nguyen and A. Chatterjee, “Number-splitting with shift-and-add decomposition for power and hardware optimization in linear DSP synthesis,” IEEE Trans. Very Large Scale Integer. (VLSI) Syst., vol.8, no. 4, pp. 419–424, Aug. 2000.
- [6] M. Ercegovac and T. Lang, Digital Arithmetic. San Mateo, CA: Morgan Kaufmann, 2003.
- [7] R. Hartley, “Subexpression sharing in filters using canonic signed digit multipliers,” IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 43, no. 10, pp. 677–688, Oct. 1996.
- [8] I.-C. Park and H.-J. Kang, “Digital filter synthesis based on minimal signed digit representation,” in Proc. DAC, 2001, pp. 468–473
- [9] L. Aksoy, E. Costa, P. Flores, and J. Monteiro, “Exact and approximate algorithms for the optimization of area and delay in multiple constant multiplications,” IEEE Trans. Comput.-Aided Design Integer. Circuits Syst., vol. 27, no. 6, pp. 1013–1026, Jun. 2008.
- [10] A. Dempster and M. Macleod, “Use of minimum-adder multiplier blocks in FIR digital filters,” IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 42, no. 9, pp. 569–577, Sep. 1995.
- [11] Y. Voronenko and M. Püschel, “Multiplierless multiple constant multiplication,” ACM Trans. Algor., vol. 3, no. 2, pp. 1–39, May 2007.
- [12] L. Aksoy, E. Gunes, and P. Flores, “Search algorithms for the multiple constant multiplications problem: Exact and approximate,” J. Microprocess. Microsyst., vol. 34, no. 5, pp. 151–162, Aug. 2010.
- [13] R. Hartley and k. Parhi, Digit-serial computation. Norwell, ma: kluwer, 1995.

## **AUTHOR’S BIOGRAPHY**



### **MOHAMMED ARIF**

B.Tech Passed in 2013 in ECE branch from Shadan College of engineering and Technology, JNTUH, Ranga Reddy.

M.Tech from Malla Reddy Engineering College (AUTONOMOUS)

*Branch:* Digital Systems and Computer Electronics, Dept of ECE

## Interfacing OCP on A On-Chip Bus

Mohammad Yousufuddin<sup>1</sup>, G. Bharathi Subhashini<sup>2</sup>

<sup>1</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (PG Scholar)

<sup>2</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (Associate Professor)

### ABSTRACT

This paper proposes a multi modes AHB on-chip bus tracer for versatile System on Chip (SOC) debugging and monitoring named AHB multi resolution bus tracer. This bus tracer is capable of capturing the bus trace with different resolutions, with all the efficient built-in compression mechanism to meet different range of needs. In addition, it also allows the users to switch the tracer resolution dynamically so that the appropriate resolutions levels can be applied to different segments of the trace. In this paper for this work we are using the well-defined interface standard, the Open Core Protocol (OCP) and will be focusing on the design of the internal bus architecture. We are going to built an efficient bus architecture that support the advanced bus functionalities which are defined in Open core Protocol (OCP), which also includes different types of transactions using modes. Here we are going to compare the proposed method with the existing method in order to prove the proposed method is more efficient than the existing. We also prove some constraints like area, delay, power, efficiency. It is done in XILINX 14.2 version using Verilog HDL language.

**Keywords:** soc, ocp, tracing, resolution, modes.

### INTRODUCTION

The On-chip bus is an important System on Chip (SOC) infrastructure. In other words, grouping lot of protocols or any other blocks on a single chip is said to be System on Chip (SOC) which connects other major components. Monitoring the On-chip bus is crucial to the SoC debugging and performance analysis/optimization. Because of this, such signals are difficult to observe since they are embedded in a SoC and there are often no sufficient I/O pins to access these signals. Therefore, a straight forward approach is taken where a bus tracer in SoC will capture the bus signal trace and store the trace in an on-chip storage in such a way that the trace memory can be offloaded to outside world for analysis. SoC chip mainly contains a large number of IP cores that communicate with each other through on-chip buses. In recent days, the growth of SoC chips and reusable IP cores were given higher priority because of its less cost and reduction in the period of time to market. These interfaces play an important role in SoC and should be taken care because of the communication between the IP cores. As the VLSI technology is increasing continuously, the data communication between the IP cores also increasing substantially. Due to which the ability of the on-chip bus to deal with large amount of data is becoming a dominant factor for all the over-all performance. The on-chip bus design can be mainly divided into two parts bus architecture and bus interface. The bus interface refers to the set of their interface signals and corresponding timing relationship. Whereas the bus architecture contains the internal components of buses and interconnections among the IP cores. The widely accepted on-chip bus, AMBA AHB [1], defines a set of bus interface to facilitate basic (single) and burst read/write transactions. AHB also defines the internal bus architecture, which is mainly a shared bus component of multiplexors. The multiplexer based bus architecture works-well with a design which has less/small number of IP cores. When the number of integrated IP cores increases, it is quite obvious that the data communication between the IP cores also increases. And it will become quite frequent that two or more master IP's

*\*Address for Correspondence:*

Yousufuddin425@gmail.com

would request the data from different slaves at the same time. Here, in this work the bus tracer is capable of tracing signals before/after the event triggering, named pre-T/post-T tracing respectively. This factor provides more flexible tracing to focus on the interesting points.

## OCP INTERFACE

Open Core Protocol (OCP) is an interface aiming to standardize and simplify the system integration problem. The Open Core Protocol (OCP) will focus mainly on the IP core requirements instead of the on-chip bus requirements. The Open Core Protocol (OCP) protocol is a communication protocol which defines point to point interface between two communicating entities, such as IP cores and bus interface modules. Where one entity acts as the Master and other as the Slave. If we want to communicate multiple masters and multiple slaves at a time with a grand signal with address and the data lines, there will be a miss communication occurs. There is data over riding miss communication occurs. To avoid data over riding and for the efficient communication we are making use of a protocol the OCP protocol. This will be deciding ‘what master has to make use of the properties of what slave, and what slave has to make use of the properties of what master’ at a time.

Most of the bus functionalities defined in both AXI and OCP are quite similar. The main difference between them is that divides the AXI address channel independent channel write address and read address channel such that both the write and read transactions can be processed at the same time. However the additional area of the channels is the punishment. Some previous work has examined on chip buses of different aspects. The work in [3] and [4] shows high-level AMBA bus models with fast simulation speed and high accuracy of the timing.

In this work we present a powerful on-chip bus design with OCP as the bus interface. We are choosing this OCP because this OCP will focus mainly on the IP core requirements instead of on-chip bus requirements. It is also be chosen because it is open to public and provide some free tools to verify this protocol. Moreover, the proposed bus is flexible in such a way that the bus architecture can adapt to system requirements.

## ON CHIP BUS

The different bus functionalities are Burst, Lock, Pipelined and Out of order transactions.

### Burst Transactions

The burst transactions are comprised of a set of transfer linked together having a defined address sequence and number of transfers. Different combinations of protocol phases are used by different types of transfers. It can be classified into multi-request burst and single request burst. Figure 1 shows the two types of burst transactions. In Figure 1(a) shows the multi burst transaction where the address information must be issued for each command of a burst transaction. Due to which unnecessary overhead occurs. Whereas Figure 1(b) shows the single burst transaction where the address information is issued only once for each burst transaction.

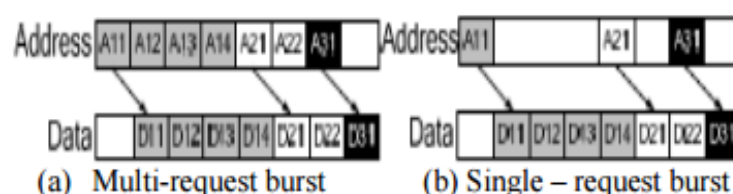


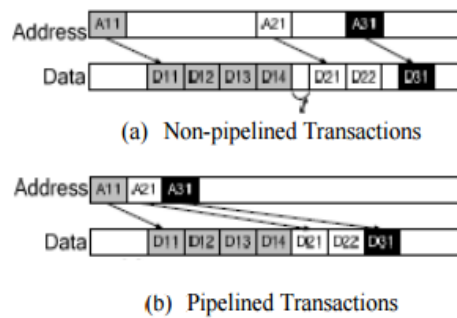
Fig1. Burst transactions

## Lock Transactions

Lock Transaction is a protection mechanism for masters that have low bus priorities. Whenever a higher priority master issues a request, the read/write transactions of the master with the lower priority would be interrupted without this mechanism. It prevents an arbiter from performing arbitration and assures that the lower priority masters can complete its granted transactions without being interrupted.

## Pipelined Transactions (Outstanding Transaction)

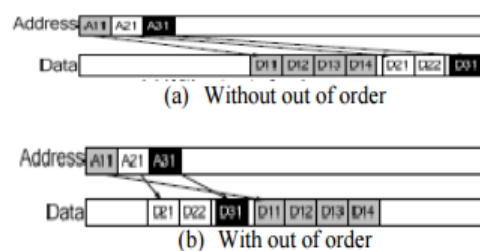
The below Figure 2(a) and (b) shows the difference between the Pipelined and Non-pipelined (also called Outstanding transaction). In Fig 2(a) for a Non-pipelined transaction a read data must be returned after its corresponding address is issued plus a period of latency. For E.g.; D21 is sent right after A21 is issued plus  $t$ . On other side for a pipelined transaction shown in Fig 2(b) this hard link is not required. Therefore, A21 can be issued right after A11 is issued without waiting for the return of data requested by A11(i.e., D11-D14).



**Fig2.** Pipelined transactions

## Out-of-Order Transaction

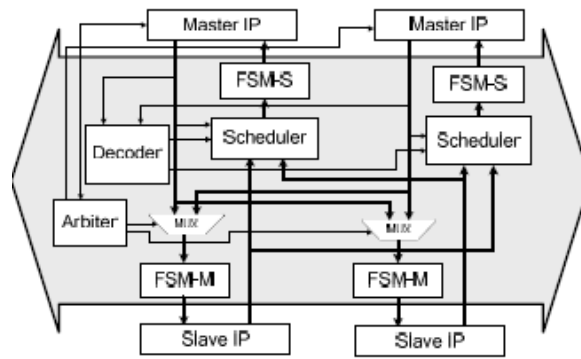
The out-of-order transactions allow the return order of responses to be different from the order of their requests. These types of transaction can definitely improves or increases the communication efficiency of the SoC systems containing IP cores with various access latencies as shown in the below Figure3. The Fig 3(a) shows the transaction which does not allow out-of-order transaction the corresponding responses of A21 and A31 must be returned after the response of A11. Where on the other hand Fig 3(b) support out-of-order transaction where the response with the shorter access latency ( D21, D22 and D31) can be returned before those with longer latency (D11-D14) and thus the transaction can be completed in much less cycles.



**Fig3.** Out-of-order transactions

## ON CHIP BUS DESIGN

The architecture of the proposed on-chip bus is illustrated in Figure 4, where an example with two masters and two slaves is shown. A crossbar architecture is employed such that more than one master can communicate with more than one slave simultaneously.



**Fig4.** Block diagram of OCP bus architecture

### Arbiter

At a time multiple master is been granting or taking the service of a single slave then we can prioritize these masters with the help of a block called Arbiter block. In shared bus architecture, resource contention happens whenever more than one master request the bus at the same time. For crossbar/partial crossbar architecture, resource contention occurs when more than one master is to access the same slave simultaneously. Therefore, here each slave IP is associated with an arbiter that determines which master can access the slave.

### Decoder

In the OCP protocol decoder selects the one slave among the slaves which are more than one exists in the system and decodes the address and decides which slave return response to the target master. Decoder also checks whether the transactions address is illegal or nonexistent and responses with an error message.

### Multiplexer

A multiplexer solves the problem of resource contention whenever more than one slave returns the responses to the same master. It selects the response from the slave that has the highest priority.

### FSM-M & FSM-S

The request and response processes is purely depends on the fact that whether a transaction is a read or a write operation. For a write transaction, the data which is to be written is sent out along with the address of the target slave, the transaction is considered to be completed only when the target slave accepts the data and acknowledges the reception of the data. For the read operation, the address of the target slave is first sent out and the target slave will issue an accept signal when it receives the message.

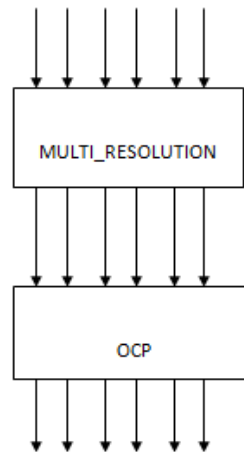
### Scheduler

Out-of-order transactions in both OCP and AXI allow the order of the returned responses to be different from the order of the requests. In particularly OCP protocol, each out-of-order transaction is tagged with a tag ID by a master. For those transactions with the same Tag ID, they must be returned in the same order as requested.

### EXISTING METHOD

Here, the existing method was about the Resolution concept. In this method we are interfacing OCP with the Multi- Resolution based AHB Bus. Resolution architecture mainly contains Master, arbiters, slave and their corresponding responses. This method is a continuous process i.e, here there exists a response signal between master to arbiter, arbiter to slave and from slave to arbiter which is a bi-directional response. But there is no such a response exists between slave to master.

**Operation:** Whenever we transmit the data to the master it will send that data to the arbiter based on acknowledgement and then the arbiter send response first to the slave in order to know whether slave is ready to take data or not. Slave receives the signal from the arbiter and it again resends the response to the arbiter whether empty or not. so based on slave's response the arbiter will send the data to slave. But because of lack of response signal between the master and slave there is no such a response reply's exist between slave and master so that master doesn't know whether the data received or not so master waits for some amount of time and it again sends the new data like this it continuously repeats its process without any Break in the signal so, because of this type of continuous process whenever we need a particular data from that address there will be data loss occurs and delay will get increases and efficiency get decreases.

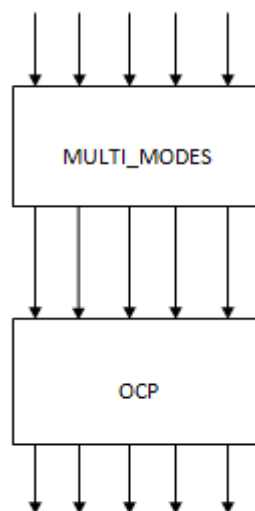


**Fig.** Block Diagram for Multi-Resolution OCP

So, in order to overcome all the above constraints we implemented a new method which is a Modes based OCP.

## PROPOSED METHOD

In order to overcome the constraints in existing method we implemented Modes AHB Bus using OCP. Here, we are interfacing OCP with the Mode based architecture.



**Fig.** Block Diagram for Multi-Mode OCP

Combining the abstract level in the signal dimension and in the timing dimension, we are using five modes in different granularities as shown in the below figures. These five modes are Mode FC (full

signal, cycle level), Mode FT (full signal, transaction level), Mode BC (bus state, cycle level), Mode BT (bus state, transaction level) and Mode MT (master state, transaction level). Now, we will discuss the usage of these modes in the following.

**At Mode FC** (full signal, cycle level), the tracer traces all the bus signals cycle-by-cycle by which the designer can observe the most detailed bus activities. This Mode FC is very useful. By looking at the detail signals the cause of the error can be diagnosed with the help of this Mode. However, since the traced data size of this mode is huge, the trace depth is the shortest among the five modes.

**At Mode FT** (full signal, transaction level), in this Mode the tracer traces all the signals only when their values are changed. In other words, this mode traces the untimed data transaction on the bus. When compare to the Mode FC, the timing granularity is abstracted in the Mode FT. Hence the trace depth in the Mode FT increases.

**At Mode BC** (bus state, cycle level), to represent bus transfer activities in the cycle accurate level, the tracer uses the BSM such as NORMAL, IDLE, ERROR and so on. Compare to Mode FC, although this Mode BC still captures the signals cycle-by-cycle, the signal is abstracted. Thus, the designer can observe the bus handshaking states without analyzing the detail signals.

**At Mode BT** (bus state, transaction level), the tracer uses the bus state to represent bus transfer activities in the transaction level. The traced data is abstracted in both the timing level and signal level. This Mode BT is a combination of Mode BC and Mode BT. In this mode, designer can easily understand the transactions without analyzing at cycle level.

**At Mode MT** (master state, transaction level), here the tracer records only the master's behaviors such as read, write or burst transfer. It is the highest abstraction level. This feature is very suitable for analyzing the master's transaction. The major difference of the Mode MT compared with Mode BT is that, this Mode does not record the transfer handshaking activities and does not capture signals when the state of the bus is IDLE, WAIT and BUSY. Thus, designers can concentrate only on the master's behavior.

## RESULTS & DISCUSSION

The proposed design is coded in VERILOG language and stimulated by using Xilinx ISE tool. The following figures show the stimulated waveforms for BC, BT, FC, FT and MT.

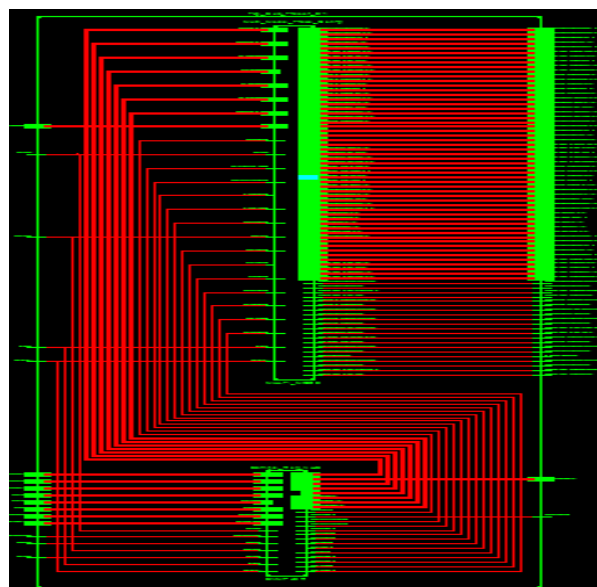
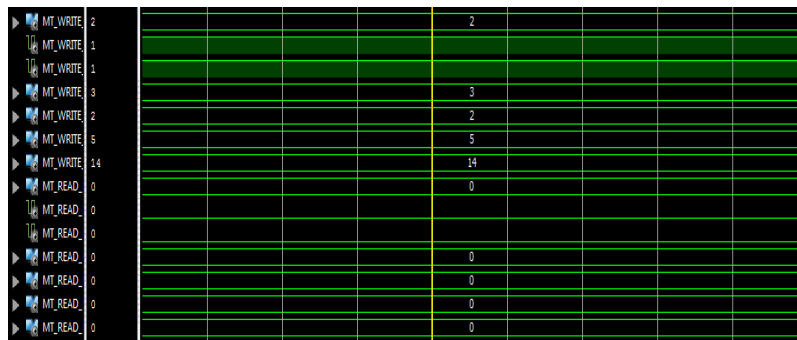
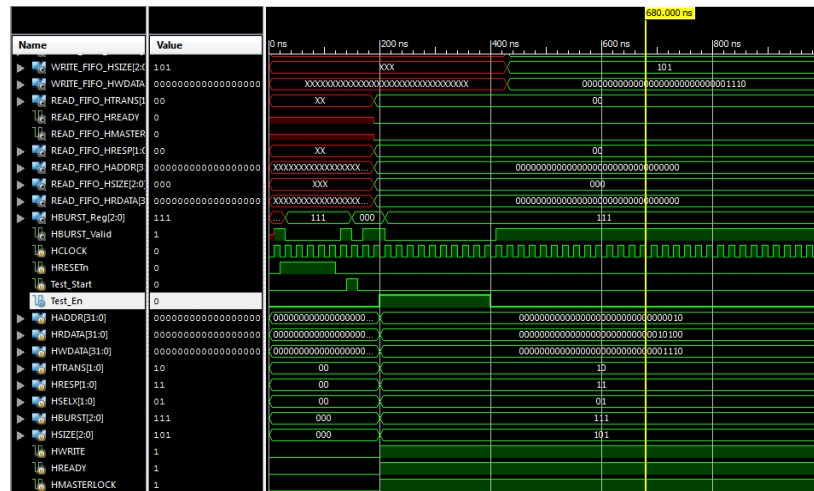


Fig. rtl schematic

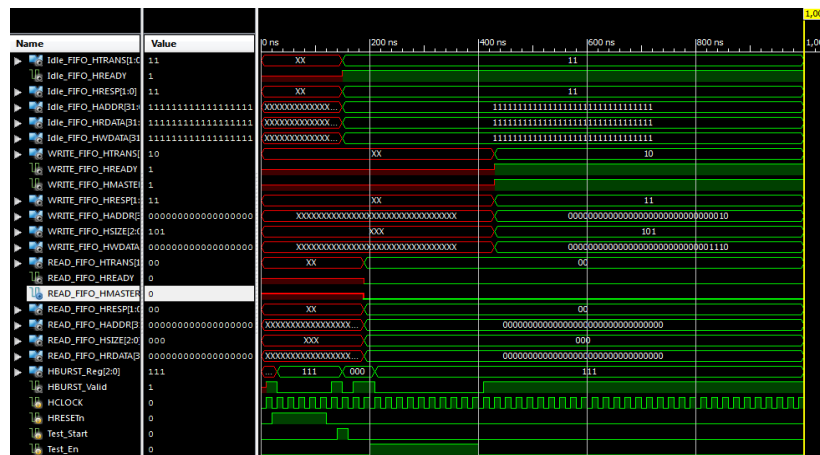




**Fig.** *MT\_Mode*



(a)



(b)

**Fig. (a) & b) Waveforms for MODE-OCP**

## AREA & DELAY

The following figures shows the area and delay reports for existed and proposed methods

Device Utilization Summary (estimated values)			<a href="#">[-]</a>
Logic Utilization	Used	Available	Utilization
Number of Slices	203	768	26%
Number of Slice Flip Flops	354	1536	23%
Number of 4 input LUTs	117	1536	7%
Number of bonded IOBs	364	124	293%
Number of GCLKs	1	8	12%

**Fig.** *exist\_area*

```

=====
Timing constraint: Default OFFSET OUT AFTER for Clock 'HCLOCK'
Total number of paths / destination ports: 251 / 251
=====
Offset:          6.216ns (Levels of Logic = 1)
Source:          TRACE_CONFIG/WRITE_FIFO_HREADY (FF)
Destination:     WRITE_FIFO_HREADY (PAD)
Source Clock:    HCLOCK rising

Data Path: TRACE_CONFIG/WRITE_FIFO_HREADY to WRITE_FIFO_HREADY
      Gate      Net
Cell:in->out    fanout  Delay  Delay  Logical Name (Net Name)
-----
FDE:C->Q        1    0.626  0.681  TRACE_CONFIG/WRITE_FIFO_HREADY (TRACE_C
OBUF:I->O        4.909                WRITE_FIFO_HREADY_OBUF (WRITE_FIFO_HREA
-----
Total          6.216ns (5.535ns logic, 0.681ns route)
                  (89.0% logic, 11.0% route)
=====

```

Fig. delay\_exist

Device Utilization Summary (estimated values)				[1]
Logic Utilization	Used	Available	Utilization	
Number of Slices	554	768		72%
Number of Slice Flip Flops	938	1536		61%
Number of 4 input LUTs	253	1536		16%
Number of bonded IOBs	951	124		766%
Number of GCLKs	1	8		12%

Fig. area\_proposed

```

=====
Timing constraint: Default period analysis for Clock 'HCLOCK'
Clock period: 4.853ns (frequency: 206.063MHz)
Total number of paths / destination ports: 2291 / 1561
=====
Delay:          4.853ns (Levels of Logic = 1)
Source:          SAMPLER/HWRITE_SAMP/HData_out_0 (FF)
Destination:     MULTI_MODE/BT_READ_FIFO_HSIZE_2 (FF)
Source Clock:    HCLOCK rising
Destination Clock: HCLOCK rising

Data Path: SAMPLER/HWRITE_SAMP/HData_out_0 to MULTI_MODE/BT_READ_FIFO_HSIZE_2
      Gate      Net
Cell:in->out    fanout  Delay  Delay  Logical Name (Net Name)
-----
FDSE:C->Q        30    0.626  1.624  SAMPLER/HWRITE_SAMP/HData_out_0 (SAMPLE
LUT4:I2->O        37    0.479  1.599  MULTI_MODE/FC_WRITE_FIFO_HTRANS_not0001
FDE:CE           0.524                MULTI_MODE/FC_WRITE_FIFO_HREADY
-----
Total          4.853ns (1.629ns logic, 3.224ns route)
                  (33.6% logic, 66.4% route)
=====

```

Fig. delay\_propose

## CONCLUSION

In proposed method we implemented OCP Bus using Multi resolution modes, so that based on modes the system will transfer the data properly from Master to Slave. The Real-time Compression and Dynamic Multi-Resolution AHB bus tracer in SoC is designed successfully and the coding is done in VERILOG. The synthesis is done by using the Xilinx ISE tool. The Designed Tracer works properly for all the modes such as Mode FC, Mode FT, Mode BC, Mode BT and Mode MT. Tracer design is verified for all the test cases. So based on the experimental results the proposed method is proved to be very efficient way in data communication when compared to that of the existing method.

## ACKNOWLEDGEMENT

I am very much thankful to my project guide Mrs. G. BHARATHI SUBHASHINI, Associate Professor, for her extensive patience and guidance throughout my project work.

I am grateful to our Project Coordinator, Mr. Kesavan Gopal, Associate Professor, for extending his support and cooperation throughout my project work.

I would like to thank Dr. M. Ch. P. Jagadesh, Head of the ECE Department, for providing the freedom to use all the facilities available in the department, for successful completion of project.

I whole-heartedly express my deep sense of gratitude to our beloved Principal, Dr. S.SUDHAKAR REDDY, Principal, MALLA REDDY ENGINEERING COLLEGE (Autonomous) for providing me adequate facilities for my successful completion of both course work and project work.

I sincerely thank all the staff of the Department, for their timely suggestions, healthy criticism and motivation during the course of my project work.

I would also like to thank all my friends for providing help and moral support at the right timing. With great affection and love, I thank my parents who were the backbone behind my deeds.

## REFERENCES

- [1] Advanced Microcontroller Bus Architecture (AMBA) Specification Rev 2.0 & 3.0, <http://www.arm.com>.
- [2] Open Core Protocol (OCP) Specification, <http://www.ocpip.org/home>.
- [3] Y.-T. Kim, T. Kim, Y. Kim, C. Shin, E.-Y. Chung, K.-M. Choi, J.-T. Kong, S.-K. Eo, “Fast and Accurate Transaction Level Modeling of an Extended AMBA2.0 Bus Architecture,” Design, Automation, and Test in Europe, pages 138-139, 2005.
- [4] G. Schirner and R. Dömer, “Quantitative Analysis of Transaction Level Models for the AMBA Bus,” Design, Automation, and Test in Europe, 6 pages, 2006.
- [5] C.-K. Lo and R.-S. Tsay, “Automatic Generation of Cycle Accurate and Cycle Count Accurate Transaction Level Bus Models from a Formal Model,” Asia and South Pacific Design Automation Conference, pages 558-563, 2009.
- [6] N.Y.-C. Chang, Y.-Z. Liao and T.-S. Chang, “Analysis of Shared-link AXI,” IET Computers & Digital Techniques, Volume 3, Issue 4, pages 373-383, 2009.
- [7] IBM Corporation, “Prioritization of Out-of-Order Data Transfers on Shared Data Bus,” US Patent No. 7,392,353 2008.
- [8] David C.-W. Chang, I.-T. Liao, J.-K. Lee, W.-F. Chen, S.-Y. Tseng and C.-W. Jen, “PAC DSP Core and Application Processors,” International Conference on Multimedia and Expo, pages 289-292, 2006.
- [9] CoWare website, <http://www.coware.com>
- [10] ARM Ltd., San Jose, CA, “AMBA Specification (REV 2.0) ARMIHI0011A,” 1999.
- [11] E. Anis and N. Nicolici, “Low cost debug architecture using lossy compression for silicon debug,” in Proc. IEEE Des., Autom. Test Eur. Conf., Apr. 16–20, 2007, pp. 1–6.
- [12] ARM Ltd., San Jose, CA, “ARM. AMBA AHB Trace Macrocell (HTM) technical reference manual ARM DDI 0328D,” 2007.
- [13] First Silicon Solutions (FS2) Inc., Sunnyvale, CA, “AMBA navigator spec sheet,” 2005.
- [14] J. Gaisler, E. Catovic, M. Isomaki, K. Glembo, and S. Habinc, “GRLIBIP core user’s manual, gaisler research,” 2009.
- [15] Infineon Technologies, Milpitas, CA, “TC1775 TriCore users manual system units,” 2001.
- [16] ARM Ltd., San Jose, CA, “Embedded trace macro cell architecture specification,” 2006.
- [17] E. Rotenberg, S. Bennett, and J. E. Smith, “A trace cache microarchitecture and evaluation,” IEEE Trans. Comput., vol. 48, no. 1, pp. 111–120, Feb. 1999.
- [18] A. B. T. Hopkins and K. D. McDonald-Maier, “Debug support strategy for systems-on-chips with multiple processor cores,” IEEE Trans. Comput., vol. 55, no. 1, pp. 174–184, Feb. 2006.
- [19] B. Tabara and K. Hashmi, “Transaction-level modeling and debug of SoCs,” presented at the IP SoC Conf., France, 2004.

## AUTHOR’S BIOGRAPHY



**Mohammad Yousufuddin**

Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (PG Scholar)

## Thermocouple Interfacing With MSP430F5529 for Furnace Management System

**B.Kiranmai, B.Srinivas**

<sup>1</sup>Department of ECE, Malla Reddy Engineering College (Autonomous), Hyderabad, India  
<sup>2</sup>Assistant Professor, Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India

### ABSTRACT

The thermocouple is standardized by the National Institute of Standards and Technology. This project shows how to implement a thermocouple interface with MSP430F5529. The thermocouple interfaces with the MSP430F5529's integrated 12-bit analog/digital converter (ADC12\_A) through an operational amplifier circuit (LM358). The MSP430 encodes the thermocouple readings into a digital value, converts them to temperature, and stores them in memory and then displays on LCD display. The thermocouple circuit interfaces with the MSP430F5529 microcontroller from Texas Instruments. The MSP430F5529 is a 16-bit ultra-low-power microcontroller with 20-bit addressing and an integrated high performance 12-bit analog-to-digital converter (ADC). The integrated ADC is used to convert the digital values into temperature and stores them in memory. A complete code set accompanies this document. For this document, a Type K thermocouple is used, and the measured temperature range is limited to 0°C to 100°C. A thermocouple is a temperature sensor that consists of two dissimilar metals welded at one end. In industry, certain combinations of alloys have been standardized in this application i.e. (chromel-alumel) is used, a commonly used general purpose Type K thermocouple is used given their low cost and large temperature range (-200°C to +1350°C). Type K thermocouples are the most commonly used general purpose thermocouples. Compared to thermistors, thermocouples sacrifice precision and accuracy for an extremely wide temperature range. Because of this, thermocouples tend to be used in industrial applications where very high temperature may be encountered.

**Keywords:** MSP430F5529, Thermocouple Sensor (Type K), LCD, ADC12, LM358

### INTRODUCTION

As the world grows ever smarter, the use of electronics systems in industrial, consumer, home automation and other areas have become progressively commonplace. With this trend comes the increasing need for the electronics system to be aware of elements of the operating system, whether it be a refrigerator making sure to keep the milk cold, a water heater ensuring the shower is not too hot or too cold or a glucose knowing that the ambient temperature is within the specified range for correct test strip operation, temperature sensing is everywhere. There are many different methods and devices that may be used to measure temperature within an electronic device, and this paper will discuss some considerations and available when selecting a temperature-measuring device. Due to its ultra-low-power nature, the MSP430 microcontroller is well-suited for monitoring temperature within many different devices. The objectives of this paper are understanding the MSP430F5529 controller, Thermocouple sensor for detecting temperature in industries, and code composer studio v6 code development and dumping it on MSP430F5529 controller board.

### LITERATURE REVIEW

This section explains about the block diagram and specifications of the paper. Before going to start this paper, identified the modules and software's which are required to do this proposed system and after this paper, identified the modules and software's which are required to do this proposed system and after that combining all these modules finally design the proposed system of block diagram is done.

#### Existing System

In this existing system, the thermocouple finds the temperature in the educational institutions, furnace regions, metro cities, industrial areas, universities and metro and other locations selected for establishment of such energy centres where the waste heat can be easily available and can be recycled

*\*Address for correspondence:*

bollamkiranmai799@gmail.com

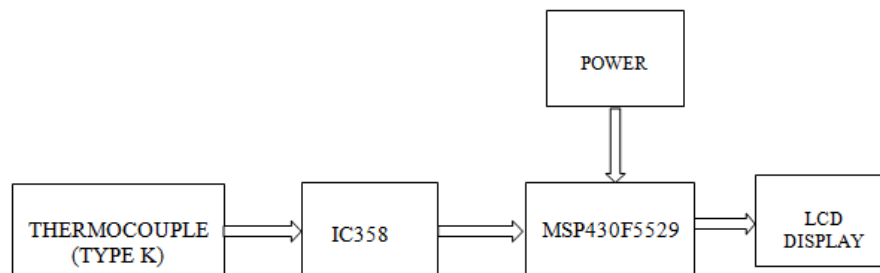
after conversion to the same system. In reality, thermocouples are affected by such issues such as alloy manufacturing uncertainties, aging effects and circuit design mistakes/understandings.

### Proposed System

In this proposed system, we are doing this project by using Furnace management system by using candle light, by soldering rod, water heater. So that we can find out the accurate room temperature we are using this proposed new method by MSP430F5529 microcontroller, so the microcontroller will take consume less power and it saves the power

### HARDWARE IMPLEMENTATION

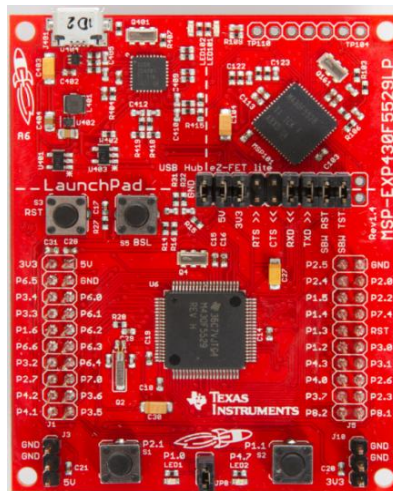
The block diagram of the design is as shown in Fig.1. It consists of power supply unit, microcontroller, Thermocouple sensor, ADC12 and LCD .the brief description of each unit is explained as follows.



**Figure1.** Block diagram Thermocouple Interfacing for MSP430F5529 Microcontroller

### MSP430F5529

The microcontroller MSP430F5529 series is an ultra-low-power mixed signal microcontrollers with built-in 4x16-bit timers with PWM capability, up to 40 I/O pins, a 12 channel comparator, and built-in communication capability using the universal serial communication interface. In addition, it also has a 12-bit analog-to-digital (A/D) converter. It is the heart of the circuit, holds the logic to run each of the peripherals.



**Figure2.** MSP430F5529 launch pad

### Thermocouple Sensor (Type K)

The thermocouple sensor is a sensing device used to measure the temperature range Type K (chromel –alumel) is the most common thermocouple purpose general thermocouple with a sensitivity of approximately  $41\mu\text{V}/^\circ\text{C}$  approximately (chromel positive relative to alumel when the junction temperature is higher than the reference temperature).It is expensive, and a wide variety of probes are available in its  $-200\text{ }^\circ\text{C}$  to  $+1350\text{ }^\circ\text{C}$  /  $-330\text{ }^\circ\text{F}$  to  $+2460\text{ }^\circ\text{F}$  range. Type K was specified at a time when metallurgy was less ad Type K was specified at a time when metallurgy was less advanced than it is today, samples and consequently characteristics may vary considerably bet and consequently characteristics may vary considerably between samples. One of the constituent metals, nickel, is

magnetic; a characteristic of thermocouples made with magnetic material is that they undergo a deviation in output when the material reaches its Curie point; this occurs for type K thermocouples at around 185 °C. Type K thermocouples may be used up to 1260 °C in non-oxidizing or inert atmospheres without rapid aging. In marginally oxidizing atmospheres (such as carbon dioxide) between 800 °C–1050 °C, the chromel wire rapidly corrodes and becomes magnetic in a phenomenon known as "green rot"; this induces a large and permanent degradation of the thermocouple, causing the thermocouple to read too low if the corroded area is exposed to thermal gradient. Another source of drift in type K thermocouples is that near 400 °C, a slow reordering in the chromel wire occurs; this is reversible and leads to hysteresis between heating and cooling.



**Fig3.** Thermocouple sensor(Type K)

### **LM358 IC**

The LM358 is a low power dual operational amplifier integrated circuit originally introduced by National Semiconductor, it is used in detector circuits. The LM358 datasheet that it consist of two independent, high gain internally frequency compensated operational amplifiers which were designed specially to operate from a single power supply over a wide range of voltages designed. Operation from spilt supplies is also possible and the low power supply current drain is independent of the magnitude of the power supply voltage. The LM358 are available in a chip sized package (8-Bumo micro SMD) using National's micro SMD package technology.



**Fig4.** LM358 IC

## **SOFTWARE IMPLEMENTATION**

### **Code Composer Studio**

CCS is the integrated development environment for TI's DSP, microcontroller and application processors. It includes compilers for each of TI's device families, source code editor, project build environment, debugger, profiler, simulators and many other features. Following are the steps for implementing application.

- Open CCS and select a workspace directory.
- Select project > import existing CCS/CCE eclipse project.
- Make sure the project is selected and click finish.
- Build and Debug the code on MSP430F5529.
- Connect "eZ-FET" USB to the PC.1

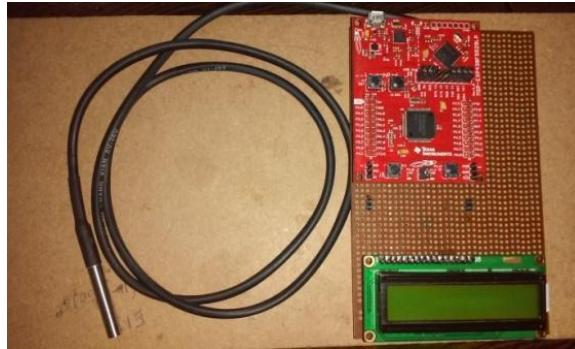
## **PROCEDURE AND RESULTS**

### **Procedure**

The thermocouple will detect the temperature from the furnace and it is being sent to the amplifier IC of type (LM358) and it is used to maximize precision ,i.e it converts micro volts(uv) to mill volts(mv).

The MSP430F5529 collects the information from amplifier and sends the data to the ADC12 which is in built in the microcontroller and its takes the analog signal and converts that signal into digital values. The MSP430 encodes the thermocouple readings into a digital value, converts them to temperature, and stores them in memory and then display on LCD display. The thermocouple circuit interfaces with the MSP430F5529 microcontroller from Texas Instrument.

### Results



**Fig5.** Thermocouple interfacing with MSP430F5529 when Power is off



**Fig6.** Thermocouple interfacing with MSP430F5529 when Power is on and displays the room temperature

### ADVANTAGES

- Temperature range is -200C to +1260C
- Robust
- Rapid response
- No-self heating

### DISADVANTAGES

- Complex signal conditioning
- Accuracy
- Susceptibility to corrosion
- Susceptibility to noise

### APPLICATIONS

- Steel industry
- Gas Appliance safety
- Manufacturing
- Power production

## CONCLUSION AND FUTURE SCOPE

### Conclusion

This proposed model of thermocouple interfacing with MSP430F5529 microcontroller using furnace management system is achieved successfully whenever the temperature is displayed on the LCD, in this we can detect the high temperature and control the temperature whenever it is high, or whenever it is low, so mainly this is used in real time in boilers in industries for measuring the temperature. So for further indication of recognising the temperature values whenever it is high or low, we can add a buzzer, to know when it is danger of exceeding the values.

### Future Scope

The future scope is in these days the society face the energy crisis but also harmful effects of pollution. The thermoelectricity is a “Green Technology” to generate without any harmful effect. The educational institutions, furnace regions, metro cities, industrial areas, universities and other locations can be selected for the establishment of such energy centres where the waste heat can be easily available and can be recycled after conversion to the same system.

## REFERENCES

- [1] KesterWalt, JamesBryant,and Walt Jung –Temperature Sensors
- [2] Implementing a Single-Chip Thermocouple Interface with the MSP430x42x(SLAA216)
- [3] MSP430x5xx/MSP430x6xx Family User’s Guide (SLAAU208)
- [4] MSP430F551x,MSP430F552x Mixed Signal Microcontroller Data Sheet(SLAS590)

## AUTHORS’ BIOGRAPHY



**Mr. B. Srinivas** obtained B.E Degree in 2006 from Osmania University, M.E (Digital System Engineering) in 2008 Osmania University. He published seven papers in journals and presently he is working as Assoc. Professor, Department of ECE, Malla Reddy Engineering College, Dhullapally, Hyderabad, Telangana (state) India.



**Miss. B.Kiranmai** is presently pursuing final semester M.Tech in Embedded Systems at Malla Reddy Engineering College, Hyderabad. She received B.TECH degree in Jayamukhi Institute Of Technological Sciences, Narsampet, Warangal, Telangana(state). Her areas of interest are Embedded Systems, Micro Controllers and Microprocessors.

## Hardware Design Considerations for a Wireless LED Based Display Design

<sup>1</sup>Latha Dasari, <sup>2</sup>Srinivas.B

<sup>1</sup> Department of ECE, Malla Reddy Engineering College (Autonomous), Hyderabad, India  
<sup>2</sup> Associate Professor, Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India

### ABSTRACT

LED's are a better way to illuminate any kind of sign. Advances in processor or design, display technology and lithium batteries have LED to a new generation of powerful, compact and cost efficient. One of the applications of LED's is scoreboards. A scoreboard is a board for publicly displaying the score in a game played between two teams. Most levels of the sport use at least one scoreboard for keeping score, measuring time and displaying statistics. In these, digits are often composed of large dot-matrix or seven segment displays made of incandescent bulbs or light emitting diodes. An official or neutral person will operate the scoreboard using a control panel i.e., IR Remote control. The main objective of this project to display LED's values which are transmitted wireless by using MSP430 microcontroller device. The MSP430 device is selected to interface between the IR communication module and the display module. This application report is only for displaying numerals on the scoreboard but the same concept can be applied to display alphabets. A scoreboard that is used as an inside display for inside sporting events is selected for this project.

**Keywords:** MSP430F5529 microcontroller, IR remote, TSOP Decoder, SMD LED module.

### INTRODUCTION

SMD LED's used in the scoreboard are quite small since it has no leads or surrounding packaging that comes with a standard LED. Two main factors determine whether a display can be used inside or outside. The first factor is weather durability and secondly, outside displays use different LEDs than inside displays. This means it's best handled, not by a human, but by automated assembly equipment. For instance, here I am using for indoor display screens. SMD LEDs work well because you can arrange a large number of diodes in tight groups of red, blue and green, generating a large variety of colors. This LED gives high brightness while it has lower power consumption than of a normal LED light tubes (e.g. dip or cluster led) and traditional light bulbs. The purpose of this project is to provide generic information by displaying numerical on the scoreboard which is transmitted wirelessly with the help of IR Remote module by using MSP430F5529 microcontroller. RC5 is an encoding standard used in infrared remote control signal transmission. The TSOP 1738 decoder is a member of IR remote control receiver series, supporting all major transmission codes.

### OBJECTIVE

The main objective of this project is to display LED's values which are transmitted wireless by using MSP430 microcontroller device. The recommended design provides information regarding LED display which is specifically designed for either inside or outside use. The MSP430 device is selected to interface between the IR communication module and the display module.

*\*Address for correspondence:*

Lathananda.d@gmail.com

Objectives of this paper is understanding MSP430 controller and all specifications like power requirements I/O ports availability, IR Remote control module for communication purpose, SMD LED and Embedded C code development is in CCS4 IDE. Finally understanding the interfacing of IR remote control module and display module is done with MSP430F5529 controller.

The rest of this paper is written as follows. In section III, operation principle is introduced. Section IV originates the hardware implementation for SMD led display. Section V clarifies about firmware implementation. Section VI explains about working procedure and results. Future scope and conclusion are presented in section VII.

## **OPERATING PRINCIPLE**

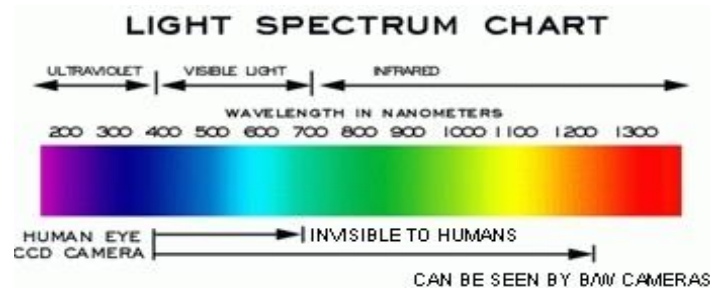
In this project, the IR remote control module is used to operate consumer electronic devices. The hardware design consideration for a wireless SMD LED based display design consists of two modules: IR transmitter section and IR receiver section.

### **A. Infrared Remote Control**

High performance infrared remote control systems are perfect for industrial and home based applications where accidental operations are undesirable. It increases operator safety and productivity. The scoreboard is controlled by IR remote, even when if there is no witness for the players, they still can add their scores by the remote control without going up to the board as other regular scoreboards.

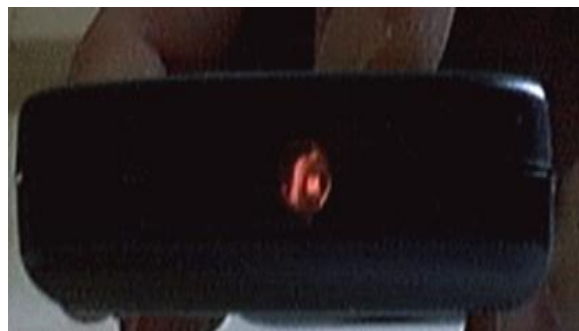
### **B. Infrared**

Infrared(IR) is invisible radiant energy, electromagnetic radiation with longer wavelengths than those of visible light, extending from the normal red edge of the visible spectrum at 950 nm. The main technology used in LED remote controls is infrared (IR) light. There must be a line of sight between the transmitter (light source) and the receiver (light detector).which is invisible to the human eye. The human eye can see in the visible light portion of the spectrum chart as shown in below fig.1



**Figure2.** *Light Spectrum chart*

Take a TV remote and press a button while pointing it at camera for a simple demonstration of how a camera can see IR as shown in fig.2



**Figure2.** *Infrared Light*

### C. IR Modulation

The infrared light is emitted from an IR led controlled by the modulated signal from the transmitter's MCU. Modulation can help the receiver distinguish the desired signals from all other infrared noise sources. The modulation is done by RC5 modulation protocol

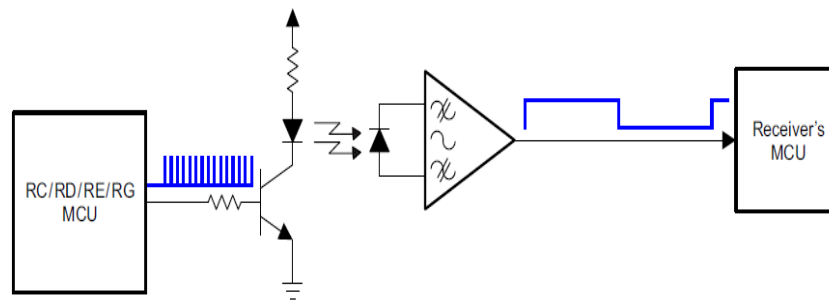


Figure3. Signal modulation

In fig.3, IR modulation is a technique that involves PWM an infrared led at certain 36 KHz frequency. The demodulated signal can be directly connected to the receiver's MCU for decoding.

### D. RC5 Protocol

A feature of bi-phase encoding (Manchester coding) is that the encoded signal will consists entirely of pulses signal high and signal low either a full duration. The encoder accepts data and control signals, encodes commands and outputs the commands to a suitable LED transmission circuit.

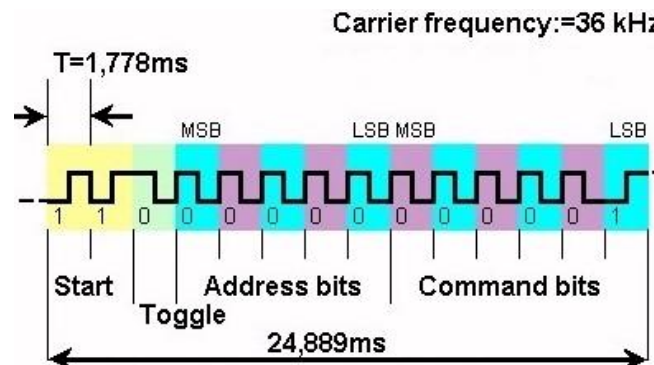


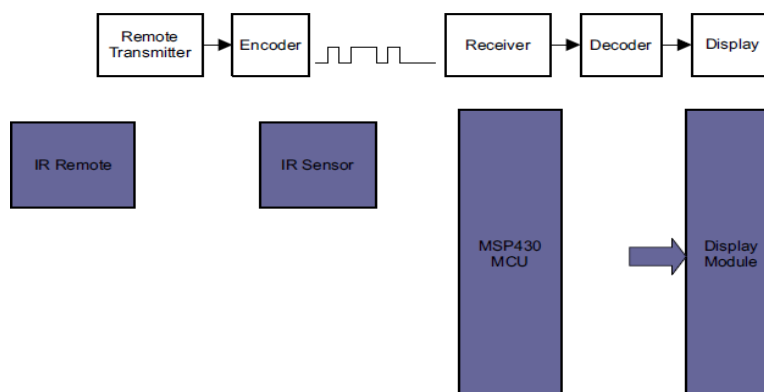
Figure4. RC5 Protocol

### E. Communication

Wireless communication, as the term implies, allows information to be exchanged between two devices without the use of wire or cables. A wireless keypad sends information to the computer without the use of a keypad cable. Information is being transmitted and received using electromagnetic radiation. By controlling the spacing between the transmitted modulated signals, the waveform can be read by an input pin on a microcontroller and decoded as serial bit stream. The distance between the remote and the display is less than 20ft and the front of the sign are easily accessed. The message is changed infrequently and therefore storing all of the possible messages is not necessary.

### HARDWARE IMPLEMENTATION

In this project, hardware design considerations for a wireless LED design consists of two modules. Firstly IR transmitter in which systems is trained by the individuals who will be using the system which consists of IR Remote.



**Figure5.** Block Diagram

Secondly receiver section in which systems is trained to respond to a transmitter signal which consists of microcontroller, led scoreboard and making programming features possible and operate properly.

### A. IR Transmitter

As we know remote control devices uses IR light. This is invisible light about 950nm wavelength. In order to exclude other sources, IR signal is modulated by some frequency and receiver has to be tuned for this frequency. Mostly remote controls transmits IR signal using 36 KHz frequency signals. Transmitting and coding is one part which can be done easily than receiving and decoding. When a key is pressed in remote control the processor wakes up to transmit the appropriate IR command otherwise it remain in low power sleep mode.

### B. IR Receiver

The easiest way to receive the pulse sent from transmitter is to use an integrated IR receiver module like TSOP1738. This TSOP1738 receiver simply removes 36 KHz carrier signal and gives clean pulses that are used for device control. Members of TSOP17XX series are sensitive to different center frequencies of the IR spectrum. This can be used for remote controls burglar alarms etc.

### C. MSP430F5529 Microcontroller

MSP430F5529 Micro controller series is an ultra-low power mixed signal microcontroller. This series are able to run up to 25 MHz, have up to 256 KB flash memory and up to 18 KB RAM. This flash based family features low active power consumption with up to 25 MIPS at 1.8V-3.6 V operation and fully integrated LDO with programmable regulated core supply voltage. The controller is the heart of the hardware circuit. The MSP430 microcontroller device is selected to interface between the IR communication module and the display module.

### D. Light Emitting Diode

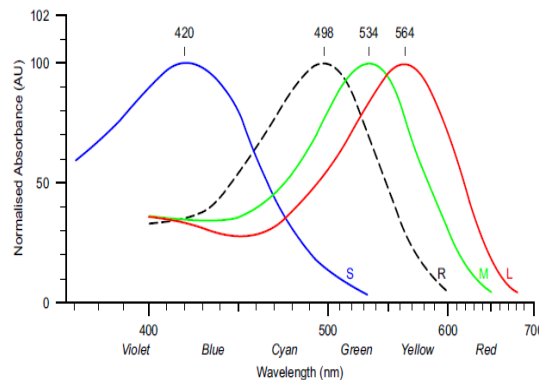
SMD LED stands for surface mount LED. It's a light emitting diode that is mounted onto and soldered onto a circuit board and used in indoor display screens.



**Figure6.** SMD LED

The above fig.6 is an SMD LED with seven segment technique. It works with low current operation, excellent numerical appearance and easy mounting on P.C. boards. Location of use; either inside or outside, viewing angle and time, communication method used and character or numerical size should be considered when selecting an LED scoreboard.

Different techniques can be used in this project for display scoreboards such as dot matrix, seven segment display, TLC5940 (16 channel) device or TLC5947 (24 channel) device. In this project I am using SMD LED with seven segment display. The LEDs are controlled by three SN74LS47N shift registers. This SN74LS47N is an open-collector outputs drive indicators directly and leading/trailing zero suppression.



**Figure7.** Absorbency

A broad range of wavelengths of infrared light spectrum is at 750nm to 1mm is used in this project. The broadest range of visible LED light is red color. Because of the cones in the human retina, the red and yellow colors are best seen by the human eyes. The scoreboard in this application displays red numbers and is clearly visible at a minimum of 30ft.

## **FIRMWARE IMPLEMENTATION**

The development environment used in this project is Code Composer Studio4 (CCS4). Code Composer Studio4 (CCS4) is an integrated development environment for TI's DSPs, microcontrollers and application processors. Code Composer Studio includes a suite of tools used to develop and debug embedded applications, project build environment and many other features. The following are the steps for the firmware implementation

1. Open CCS and select a workspace directory.
2. Select project > Import Existing CCS/CCE Eclipse Project.
3. Browse to the extracted project directory.
4. Make sure the project is selected and click Finish.
5. Connect the MSP-EXP430F5529 Experimenter Board “eZ-FET” USB to the PC.

After the program is compiled and the build is complete a .hex file is generated in .exe folder in the project directory. Then the .hex is programmed to the MSP430 device.

## **WORKING PROCEDURE AND RESULT**

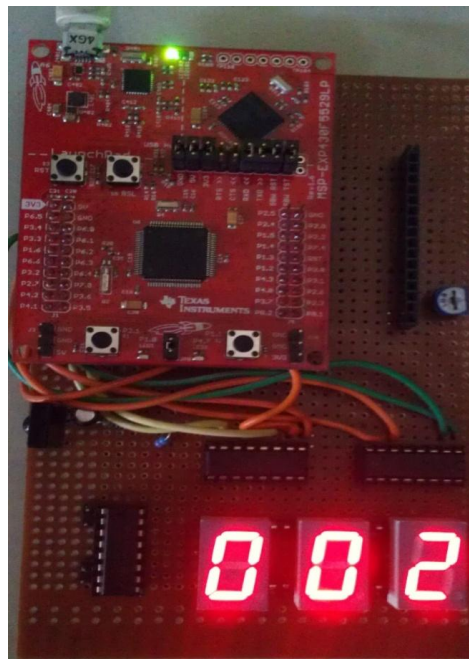
### **A. Working Principle**

The main objective of this project is to use Infrared frequency signal to control the SMD LED scoreboard display. IR remote control consists of keypad is used to send information to receiver side.

The microcontroller device is selected between the IR remote and display module. The microcontroller receives the IR frequency signal from the receiver and it decodes and display information on the scoreboard display. All this process is controlled by MSP430F5529 microcontroller. The distance between the IR remote and display is 20ft.

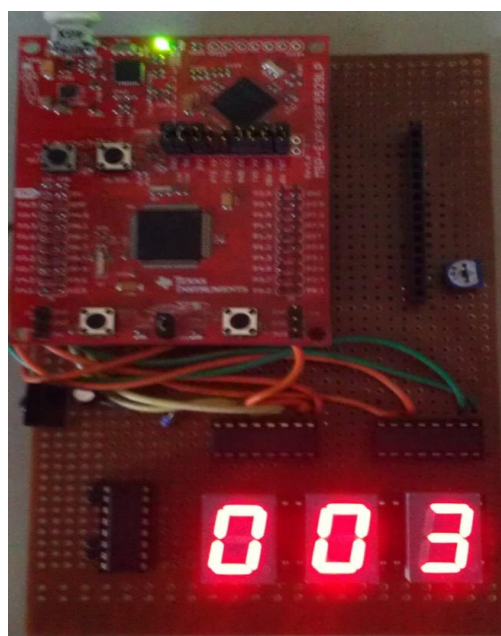
## **B. Result**

The implementation of “Hardware Design Considerations for a Wireless Based LED Display Design” is done successfully. Communication between peripheral modules and microcontroller is properly done.



**Figure8.** *Scoreboard before Increment*

The above figure shows scoreboard before increment and the following figure shows scoreboard after increment.



**Figure9.** *Scoreboard after Increment*

## **CONCLUSION AND FUTURE SCOPE**

### **A. Conclusion**

In the end, there's no question in the digital age for sport stadiums that brightness, high definition scoreboards are an integral part of a stadium. In sports, this is a win-win for everybody the owners get a better stadium and fans are ecstatic.

### **B. Future Scope**

SMD LED technology has immense scope for improvement. The displays of the future will be stretchable, deformable into any shape, twistable and perhaps more importantly durable, efficient and cheap. This is the promise of a new approach for manufacturing inorganic light emitting devices.

### **C. Advantages**

- Safe under all conditions of normal use and health.
- Compared to other LED's, the SMD LED's has long life span.
- Faster response times.

### **D. Applications**

The applications of SMD LED's are:

- Aviation
- Automotive lighting applications
- Traffic signals
- Transmitting information from remote control units and in digital clocks.

## **REFERENCES**

- [1] [www.ti.com/product/msp430f5529](http://www.ti.com/product/msp430f5529)
- [2] [www.ti.com/lit/an/slava655/slava655.pdf](http://www.ti.com/lit/an/slava655/slava655.pdf)
- [3] <https://en.wikipedia.org/wiki/Scoreboard>
- [4] <http://electronics.howstuffworks.com/led>.
- [5] Remote Control System RC-5 Including Command Tables, Philips Semiconductors, December 1992, Publication No. 938870623011
- [6] [Techdoc.altium.com/display/FPGA/Phillips+RC5+Infrared+Transmission+Protocol](http://Techdoc.altium.com/display/FPGA/Phillips+RC5+Infrared+Transmission+Protocol)
- [7] Mayank Bhati, Ankit Gupta, Vikrant, S.K.Dubey. “INFRARED REMOTE CONTROL SWITCH “, International Journal of Advanced Technology in Engineering and Science Volume No.02, Issue No. 05, May 2014
- [8] RC5 Codec Feb 5, 2003 - This Application Note discusses *codec* for the popular *RC5*-based remote.
- [9] [www.vishay.com/docs/82489/tsop322.pdf](http://www.vishay.com/docs/82489/tsop322.pdf) IR Receiver Modules for Remote Control Systems.
- [10] MSP430 Microcontroller Basics by John H. Davies.
- [11] [www.instructables.com/id/remote-controlled-arduino-scoreboards-using-LED-Str](http://www.instructables.com/id/remote-controlled-arduino-scoreboards-using-LED-Str)
- [12] Sang-II Park, Yujie Xiong, Rak-Hwan Kim, e al., Printed Assemblies of Inorganic Light-Emitting Diodes for Deformable and Semitransparent Displays, Science (2009) 325, 977- 981.

**AUTHORS' BIOGRAPHY**



**Mr. B. Srinivas** obtained B.E Degree in 2006 from Osmania University, M.E (Digital System Engineering) in 2008 Osmania University. He published seven papers in journals and presently he is working as Assoc. Professor, Department of ECE, Malla Reddy Engineering College, Dhullapally, Hyderabad, Telangana (state) India.



**Ms. Latha Dasari** is presently pursuing final semester M.Tech in Electronics and Communication Engineering at Malla Reddy Engineering College, Secunderabad. She received her AMIE degree in Electronics and Communication from Siddhartha institute of Engineering and Technology, Hyderabad. Her areas of interest are Microcontrollers and microprocessors, digital system design and analog circuits.

## Automated Utility Meter Reading Using Wireless System Bluetooth with MSP430 Microcontroller

<sup>1</sup>Keerthi M, <sup>2</sup>Raju N

<sup>1</sup>Department of ECE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

<sup>2</sup>Assistant Professor, Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India

### ABSTRACT

This paper is focused on the description of the possible benefits for the electric utilities and residential customers from the automatic meter reading system usage. AMR is always seen as away to read consumption without visiting every meter. Major benefits of the AMR, mentioned in this paper are power quality monitoring, distribution network management, theft detection and so on. The present system in India is obsolete and time consuming and it has major drawbacks. So we are proposing a system is that Automatic Meter Reading System using MSP430 Microcontroller which becomes fully automated and communication is made possible via Bluetooth network.

The electricity meter present in each house is connected by wireless network Bluetooth with the Electricity Billing Office which periodically gets updates from the meter. The Electricity Billing Office using a backend database calculates the amount to be paid according to the number of units consumed and sends it back to the meter for display and also we can see the bill details in the mobile Bluetooth terminal. The advantages of the proposed system make the existing system incompetent. The new system is user friendly, easy to access and far more efficient than the existing system.

**Keywords:** MSP430F5529 Board, Bluetooth Module, Mobile Phone, LCD Display.

### INTRODUCTION

This paper is to make use of new modern technologies and implement them into more practical fields. Our paper deals with the implementation of wireless networks in the field of electricity billing. We can make use of this technology to such an extent such that even complex problems can be handled in a easier way. Wireless networks are the eminent futuristic replacement of cables and power lines that connect every household in a particular area. This type of networks can also be used for creating emergency response networks. Our paper eliminates the need for employing electricity billing meter readers and this set of employers can be used elsewhere.

AMR is the technology of automatically collecting consumption, diagnostic, and status data from energy metering devices and transferring that data to a central database for billing, troubleshooting, and analysing. This technology mainly saves utility providers the expense of periodic trips to each physical location to read a meter. Another advantage is that billing can be based on near real time consumption rather than on estimates based on past or predicted consumption. This timely information coupled with analysis can help both utility providers and customer's better control the use and production of electric energy consumption.

The AMR system is also able to provide a set of different services, which are useful for the utility companies in their operation and planning and maintenance, they are load management, outage and fault reporting, customer services, power quality monitoring, network management, theft detection, billing, balance settling, energy settling, assest management, energy usage information, interruption reporting etc.

*\*Address for correspondence:*

modala.keerthi471@gmail.com

## **LITERATURE REVIEW**

This method can eliminate the problems such as manpower requirement for billing and errors during calculation etc., and can provide necessary information such as tariff variation and due date for payment etc. to the consumer through the wireless medium. The wireless technology can be implemented by having a Bluetooth enabled transceiver interfaced with the electricity billing section server as well as in the consumer side.

### **Existing System**

A Bluetooth Advanced Metering infrastructure is a existing one for automatic meter data collection and energy auditing and management. In this the system operates with multiple channels and frequency hopping and coexists with potential interferers. In this method if any tariff variation occurs, the new tariff rate will be changed only through reprogramming the controller otherwise the previous tariff rate will be displayed on the LCD display. This may cause major problems in billing. Here, there are no intimations given by the electricity board to the consumers about the status of energy consumption. There are no details about the previous month's consumption and the amount of bill paid. We are focusing about these draw backs and overcome in our project Automatic Energy Meter through wireless smart meter using Bluetooth.

### **Proposed System**

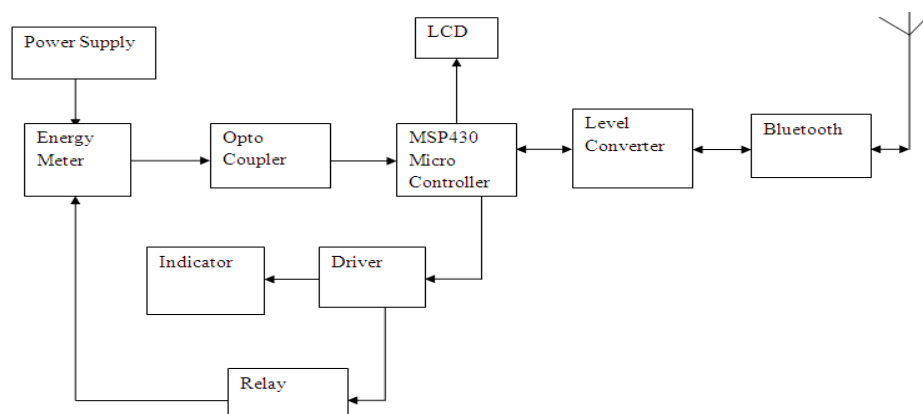
In order to overcome the above mentioned drawbacks, we are proposing a new method automatic energy meter through wireless system Bluetooth with MSP430 Microcontroller which is having the advantages such as no need of manpower, errorless tariff calculation, tripping can be done from the electricity billing side in case of not paying the bill and intimation about tariff variation, amount to be paid and due date for payment.

This system is designed with two modules which are as follows:

- Consumer module in the system
- Electricity Billing Office module in the system

### **Consumer Module**

The consumer side is equipped with energy meter, MSP430 controller, LCD display, indication unit (can be an indication lamp). The controller continuously monitors the energy meter reading and calculates the amount till last usage. These details can be viewed in the LCD display and also it will be sent to the Electricity Billing Office server during each month through the Bluetooth transmitter. The indication unit is provided for the attention of the consumer in case of exceeding normal usage, delaying the payment and in case of any tariff variation by the Electricity Billing Office.



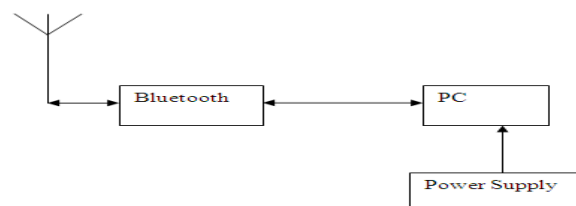
**Figure1.** Consumer Module

### **Working of Consumer Module**

The home module is designed to work as follows. The LED, which indicates the consumption of energy, is replaced by an opto coupler which produces pulses as energy is being used and transfers to the MSP430 microcontroller. The counter which counts the number of times the LED blinks and sends the data to the MSP430 microcontroller. This is received by the controller as an external interrupt. The controller is so programmed that it calculates the amount based on the number of units consumed. These details will be displayed on the LCD panel attached to the home module too. A key button is provided in the LCD display with EEPROM memory which can provide the details of previous month's payment and energy consumption to the consumer. The indicator lamp attached nearer to the LCD display will indicate the overset limit range. A relay is attached to each and every home module which plays the major role of tripping the connection, if the bill is not paid in time. Thus the relay acts as a switching device. The Bluetooth transceiver, which is a wireless communication module, transmits the details calculated by the microcontroller regarding the usage of energy to the electricity board office once in a month. Thus this methodology reduces the manual effort to a great extent. Electricity Billing Office can give information about the tariff variation to the consumer, as notification on LCD display with lamp indication.

### **Electricity Billing Office Side Module**

The Electricity Billing Office module consists of a database at the back end for storing values which are got from the home module via Bluetooth. After the values are got from the home units the cost is calculated and the values are sent back to the home unit and they are displayed in the LCD display for the user to make note of it. Also the cost and amount of units are sent as a SMS to the customer's mobile phone. The data's are transmitted and received using Bluetooth. There is an encoder and decoder both at the central office and also at the individual home units. This facilitates secure transmission and reception in the system.



**Figure2.** Electricity Billing Office side Module

### **Working of Electricity Billing Office Side Module**

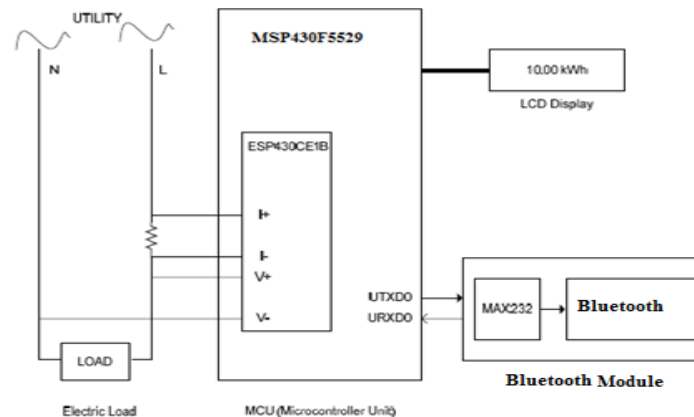
The Electricity Billing Office side module is designed to work as follows. The receiver module is interfaced with a system which is monitored by the officials in the electricity board. The Bluetooth transceiver on the Electricity Billing Office module receives the data and displays it in the terminal C window of the system interfaced with it. The home module will be reset by the Electricity Billing Office officials as the bill is paid. The thing to be noted here is that only when the command is given, board will supply power to the particular customer. Else the supply will be disconnected until the payment of the bill. There by we contribute a small part to prevent the power crises as well.

## **HARDWARE IMPLEMENTATION**

### **Wireless Smart Metering Module**

Two parts of the wireless smart metering system are the smart metering module and the Monitoring module. The smart metering module consists of a MSP430 core and a Bluetooth module. The

MSP430 core collects data from the electrical appliance with the help of suitable integrated sensors. Then it uses the Bluetooth module to send the collected data to the monitoring module. There is a Bluetooth module connected to the computer via a serial communication port in the monitoring module. Bluetooth module uses chip antenna to send and/or receive data. On the other hand, RS-232 serial communication interface is established between the Bluetooth module and the computer. It is also possible to use the USB port with the help of a serial port to USB converter.



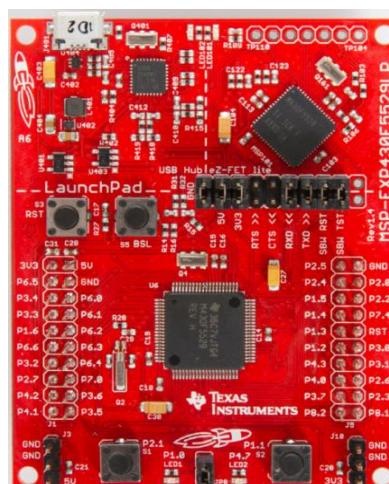
**Figure3.** The hardware diagram of wireless smart meter

## Monitoring Module

The Bluetooth device in monitoring module is responsible for receiving the energy data from wireless smart meters and transmitting the request command from control centre to wireless smart meter. Communication is performed between Bluetooth module and control centre by RS232 serial communication. Also the collected energy data can be sent to the utility companies via internet by a network gateway in monitoring module.

## MSP430F5529

The microcontroller MSP430F5529 series is an ultra-low-power mixed signal microcontrollers with built-in 4x16-bit timers with PWM capability, up to 40 I/O pins, a 12 channel comparator, and built-in communication capability using the universal serial communication interface. In addition, it also has a 12-bit analog-to-digital (A/D) converter. It is the heart of the circuit, holds the logic to run each of the peripherals.



**Figure4.** MSP430F5529 launch pad

## **SOFTWARE IMPLEMENTATION**

### **Code Composer Studio**

CCS is the integrated development environment for TI's DSP, microcontroller and application processors. It includes compilers for each of TI's device families, source code editor, project build environment, debugger, profiler, simulators and many other features. Following are the steps for implementing application.

- Open CCS and select a workspace directory.
- Select project > import existing CCS/CCE eclipse project.
- Make sure the project is selected and click finish.
- Build and Deugg the code on MSP430F5529.
- Connect “eZ-FET” USB to the PC.

## **PROCEDURE AND RESULTS**

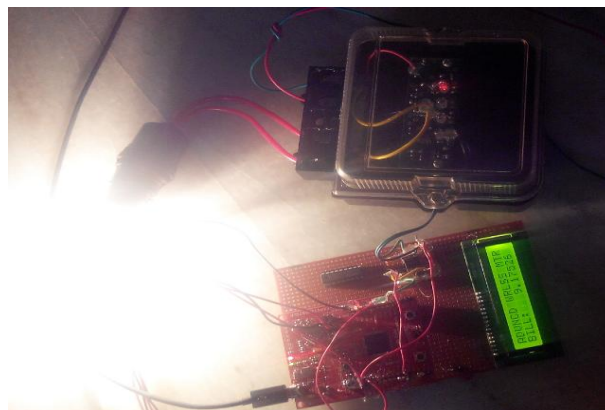
### **Procedure**

MSP430 collects the data from electrical appliances with the help of particular sensors. Bluetooth network will send all this data to the central office with the help of antenna chip. At the central office there is one more Bluetooth network to receive the collected data after receiving the data will be modified and the modified data will be sent back to the customers house which will be displayed on the LCD and customers mobiles.

### **Results**



**Figure5.** *Wireless Metering System When Power is Off*



**Figure6.** *Wireless Metering System when Power is ON*

## **ADVANTAGES**

- Accurate meter reading, no more estimates.
- Improved billing.
- Accurate profile classes and measurement classes, true costs applied.
- Improved security and tamper detection for equipment.
- Less financial burden correcting mistakes.
- Less accrued expenditure.
- Improved procurement power through more accurate data-“de-risking” price.
- In case of shortages, utility will be able to manage/allocate supply.

## **CONCLUSION AND FUTURE SCOPE**

### **Conclusion**

This proposed automated metering system includes an office module which has a PC with its back end connected to a database. The other module is the customer home module which is present at the home. This module is used to make note of the amount of power consumed by the customer and it sends the data to the PC which is present in the billing office. This Electricity Billing Office module calculates the data and sends it to the customer along with the due date. The customer also gets details of the bill on his mobile phone through which he can pay the bill.

### **Future Scope**

In future phase of this project, you will be able to view your daily energy usage, at your convenience, by logging in to your account on website. Knowing both how you use energy, and when, will allow you to decide what energy-saving changes you would like to make. You can always check our website for current information, and the Texas power magazine will continue to provide information on when new options are available.

## **REFERENCES**

- [1] [www.BryanTexasUtilities.com](http://www.BryanTexasUtilities.com)
- [2] [http://www.academia.edu/9135949/bluetooth data transmission using MSP430](http://www.academia.edu/9135949/bluetooth_data_transmission_using_MSP430)
- [3] [http://en.wikipedia.org/wiki/automatic meter reading](http://en.wikipedia.org/wiki/automatic_meter_reading)
- [4] <http://www.itnews.com.au/news/auditor-general-slams-victorian-smart-meters-160398>
- [5] <http://scialert.net/fulltext/?doi=ajsr.2013.88.97&org=11>
- [6] <http://www.seminorprojects.com/thread-automatic-meter-reading-amr#ixzzle7T3vMcj>
- [7] [http://sujeteeee.blogspot.in/2015/02/bluetooth controlled electronic home](http://sujeteeee.blogspot.in/2015/02/bluetooth_controlled_electronic_home)
- [8] <http://researchdesignlab.com/gprs-based-automatic-meter-reading.html>
- [9] <http://www.seminarprojects.com/Thread-automatic-meter-reading-amr#xzzle7T3VMcj>
- [10] N.Baker, "Zigbee and Bluetooth: strengths and weakness for industrial applications," The IEEE computing and control engineering, Vol.16, No.2, April/may 2005, pp.20-25.
- [11] Richa Shrivastava and Nipun Kumar Mishra, "An embedded system for wireless prepaid billing of digital energy meter," international journal of advances in electronics engineering, pp.322-324

## **AUTHORS' BIOGRAPHY**



**Mr. Raju N.** obtained B.E Degree in 2009 from Mona College of Engineering & Technology, M.E (Embedded Systems) in 2013 from CVR College of Engineering (Autonomous). Presently he is working as Assistant Professor, Department of ECE, Malla Reddy Engineering College, Dhullapally, Hyderabad, Telangana, India.



**Mrs. Keerthi M.** is presently pursuing final semester M.Tech in Embedded Systems at Malla Reddy Engineering College, Hyderabad. She received B.E degree in Vidya Vikas Institute of Technology, Chevella, Telangana. Her areas of interest are Embedded Systems, Micro Controllers and Microprocessors.

## Understanding an Fram Technology Using MSP430 Microcontroller

Rajyalakshmi G, Ravitheja T

<sup>1</sup>Dept of ECE, MREC (Autonomous), JNTUH, Hyderabad, India.

<sup>2</sup>Assistant Professor, Dept of ECE, MREC (Autonomous), JNTUH, Hyderabad, India.

### ABSTRACT

FRAM is a non volatile random access memory that uses a ferroelectric capacitor to store the data, which retains stored data even when power is turned off. It offers higher write speed, data retention and endurance over Flash/EEPROM. This technology has been applied to manufacture a family of flow density memories that combine the fast write and reads of SRAM and non volatility of EEPROM with very high read/write endurance. FRAMs have been expressed in many applications such as smartcards, low density memories. FRAMs are forefront of non volatile memories. Future technologies include high density products with high endurance, data retention and the ability to work at lower voltage. Finally, FRAM technology will become the superior non volatile memory technology by the end of decade.

**Keywords:** Ferroelectric materials, FRAM technology, PZT crystal, MSP430microcontroller, MSP430FR5969.

### INTRODUCTION

Established semiconductor memory technologies are classified into different categories. RAMs are random access memories, which have symmetric read and write access time. RAM is volatile memory, it can hold the temporarily. Non volatile memories are traditionally ROM (Read only memory) i.e., floating – gate technology produces electrically erasable memories like flash and EEPROM. These products allow system programming's but write and access times are dissimilar. FRAM is a true non-volatile RAM because it combines the advantages of both non volatile and RAM memories. The write advantages of Flash or EEPROM and non-volatility make it quite suitable for storing data while power is turned off.

### FRAM STRUCTURE

FRAM is also known as FeRAM. A single FRAM cell can be considered a capacitor it consists of a ferroelectric material between two electrode plates. This structure is made up of PZT crystal complex. Here we have to apply an electric field across the crystal, it causes the mobile retained when power is turned off, which makes FRAM is non-volatile. Here data can be stored in the form of "0" or "1".



Here Polarization occurs when an electric field is applied. Electric field polarization remains present even in the absence of an electric field. Here two stabilized states are stored in the form of 1 or 0 data.

*\*Address for correspondence:*

lakshmi.gonuguntla@gmail.com

## FRAM OPERATION

A Ferroelectric crystal has a mobile atom in the centre of the crystal. Whenever applying an electric field across a face of the crystal, it causes this atom to move in the direction of the given field. While reversing the field causes the atom to move in the opposite direction. Here atom positions at the top and bottom of the crystals are stable. Therefore removing an electric field leaves the atom in a stable position, while turning of the power. In case, a memory element, the ferroelectric crystal creates an digital memory.

### Write Operation

All FRAM accesses are limited to 125 nanoseconds or 8MHs access frequency. However, the MSP430FRxx family supports system speeds up to 16 MHz or 24 MHz for the MSP430FR59xx devices, the wait –state generator can be controlled automatically or manually. As with a read, the change of state occurs in under 1ns with a full access taking under 700ns. As with a read operation , a pre charge operation follows a write access memory.

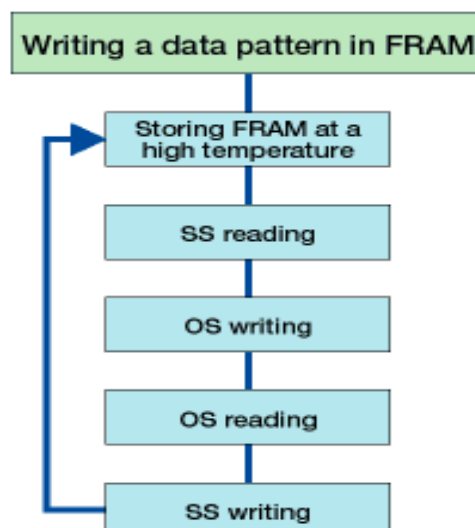
### Read Operation

An electric field is applied across the capacitor. Then the mobile atoms will be moved across the crystals in the direction of the field if they are not already in the appropriate positions. The state which occurs in under one nano sec, with the total circuit access taking less than 70ns. Since read operation memory involves a change of state, the circuit will automatically restore the memory state. Therefore each read access is accompanied by a pre charge operation that restores the memory state. Although the read is destructive, the time during which the memory cell is invalid is under 50 ns.

## RELIABILITY OF FRAM

Same-state refers to the logic state of FRAM; that is, the state of polarization of the Fram ferroelectric crystal prior to the high-temperature bake whenever testing for imprint. • Opposite-state refers to the polarization of the given crystal in a direction of opposite state to that in which it was imprinted. To test for imprint, a data configuration with a set logic state is written onto FRAM, and then the device is exposed to a high-temperature i.e., 125°C. This temperature bakes the bit-cells in one logic state referred to as similar-state. This temperature bake is followed by a read-restore to further strengthen the same-state data. Opposite-state data is written into an FRAM. This state is followed by a thermal depolarization bake to stress the opposite-state data. This temperature bake is performed at the maximum operating temperature for the device (85°C for the MSP430). After the depolarization is completed, then the data is read out to ensure the integrity of the reverse-state data. Then the test is repeated until the testing time is reached. The opposite-state identifies potential imprint issues, while the read operation protects that the FRAM cells have maintained the ability to be polarized with opposite-state data as well maintained enough polarization to be read access without data loss. MSP430 FRAM data retention is tested for a cumulative bake time of 1000 hours at 125°C.

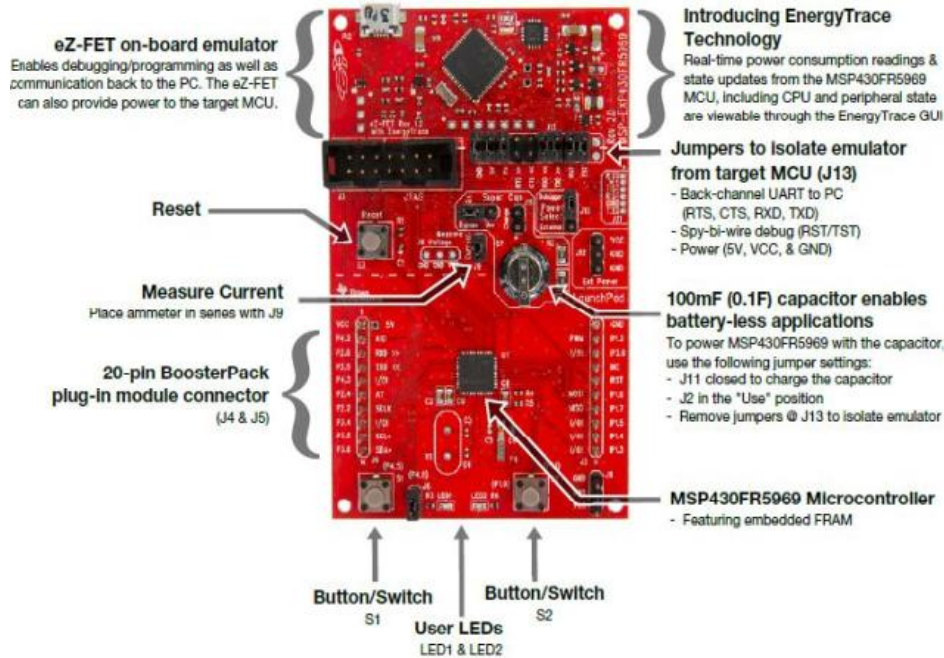
### Test Flow



SS: Same state OS: Opposite state

## HARDWARE IMPLEMENTATION

The MSP430FR5969 launch pad is an easy to use consideration module for the MSP 430FR5969 microcontroller. This MSP430FR5969 micro controller contains everything needed to start developing on MSP430's ultra low power FRAM technology; it includes on-board emulation for pre-programming, debugging and energy measurements. The board features on-board buttons and LED's for high speed integration of a simple user interface as well as a capacitor that allows standalone applications without applying an external power supply. The MSP430FR5969 microcontroller device features embedded FRAM a non volatile memory known for its, maximum high endurance high speed and ultra low power write and read access.



MSP-EXP430FR5969 Overview

The Texas instruments MSP430FR59xx family of ultra low power microcontroller consist of several devices featuring different sets of peripherals. The architecture, joined with seven low power modes are efficient to achieve extended battery life for example in portable measurement applications. The device features a powerful 16 bit registers and constant generators that contribute to maximum code efficiency. The MSP430FR5969 devices are mixed signal microcontroller configurations with up to five 16 bit timers, comparator, universal serial communication interfaces supporting UART, SPI, I2C and hardware multiplier, 12 bit ADC. The FRAM memory can be programmed through the JTAG port, spy-bi-wire, the BSL, or in system by the CPU. Features of MSP430FR5969 FRAM memory include: Ultra low power and ultra fast write non volatile memory, Byte and word access capability and Programmable wait state generation.

## SOFTWARE IMPLEMENTATION

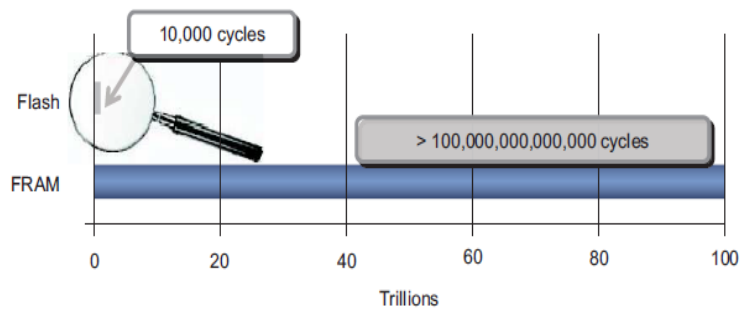
### Code Composer Studio

CCS is the integrated development environment for TI's DSP, microcontroller and application processors. It includes compilers for each of TI's device families, source code editor, project build environment, debugger, profiler, simulators and many other features. Following are the steps for implementing application.

- Open CCS and select a workspace directory.
- Select project > import existing CCS/CCE eclipse project.
- Make sure the project is selected and click finish.
- Build and Debug the code on MSP430F5529.
- Connect “eZ-FET” USB to the PC.

## FRAM FEATURES

### Write Endurance



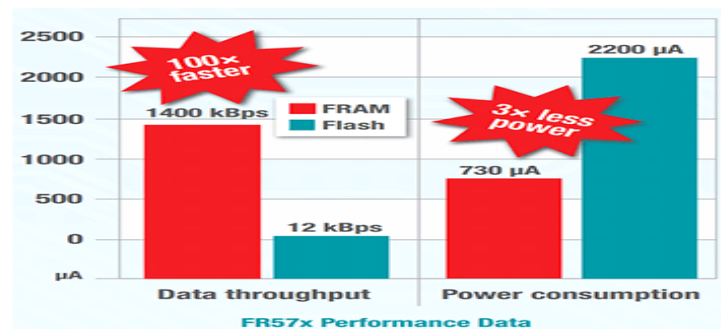
The endurance of maximum  $10^{13}=10$  trillion write/read cycles. FRAM Endurance is more than 10 million times of EEPROM.

### Data Performance

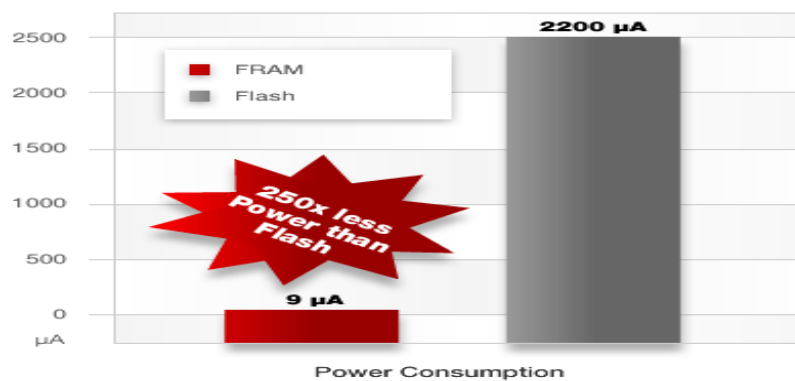
FRAM maximum throughput is greater than 100times faster than flash maximum throughput, while still consuming 3times lower power than flash.

*FRAM maximum throughput=1400kbps@730μA*

*Flash maximum throughput=12kbps@2200μA.*



### Power Consumption:



*FRAM memory consumes 9μA @12 kbps*

*Flash memory consumes 2200μA @12kbps.*

## 8. APPLICATIONS

FRAM memory technologies are used in different applications, those applications are

### Metering:

Power meters, gas meters water meters and smart meters.

### Industrial machines:

Control units, PLC, motors, process monitors robots, solar power, ATM, vending machines, elevators, breakers.

### **Consumer Electronics:**

Car navigation systems, drive recorders, games, battery and ID cards for authentication.

### **Medical field:**

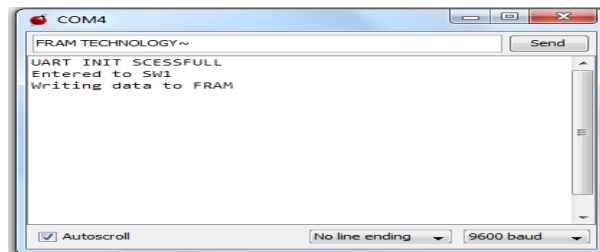
CT scan, Insulin pump, vital meters and oxygen saturation meters.

## **RESULT**

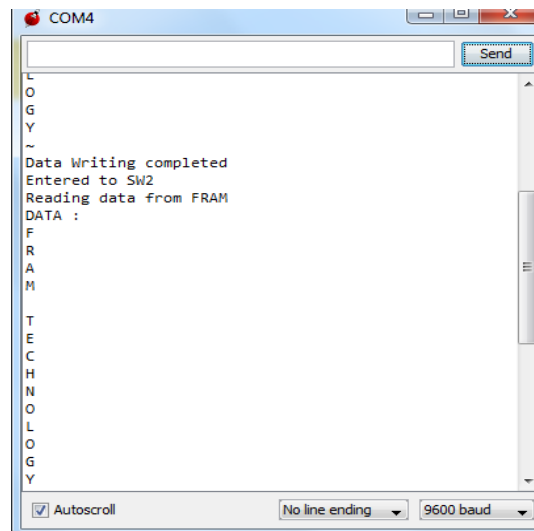
The implementation of UNDERSTANDING AN FRAM TECHNOLOGY USING MSP4305969 MICROCONTROLLER is done successfully. Communication between microcontroller and system work is properly done. In our project writing to FRAM and reading from FRAM is successfully done.



### **Writing to an FRAM**



### **Reading from an FRAM**



Writing to FRAM and reading from an FRAM successfully done. In our project, the writing and reading data is retained whenever the power is turned off Since FRAM is non-volatile memory technology.

## **FUTURE SCOPE AND CONCLUSION**

FRAM is undoubtedly the fastest non volatile embedded memory option available today. Being embedded with the ultra-low power MSP430 architecture makes it a perfect choice for applications needing extremely fast write speeds, low power and high endurance. Some of these applications include data logging using remote sensors, energy harvesting applications, and critical response time

applications. The key factors that influence the maximum possible FRAM write throughput were discussed and the tradeoffs presented. In analyzing the bench test results for the MSP430FR5969, it is seen that the achievable practical write speed is very close to the theoretical maximum. It is up to the user to determine the available resources and design their application in a way that can achieve the fastest possible write speed for FRAM.

## REFERENCES

- [1] "FRAM – Ultra-Low-Power Embedded Memory". Texas Instruments.
- [2] [www.radio-electronics.com/.../fram-ferroelectric-random-access-memory](http://www.radio-electronics.com/.../fram-ferroelectric-random-access-memory)
- [3] <http://www.ti.com/tool/mspexp430fr5969?keyMatch=fr5969&tisearch=Search-EN>.
- [4] <https://store.ti.com/msp-exp430fr5969.aspx>.
- [5] [http://www.researchgate.net/publication/3639286\\_FRAM-the\\_ultimate\\_memory](http://www.researchgate.net/publication/3639286_FRAM-the_ultimate_memory).
- [6] <http://www.ti.com/tool/MSPEXP430FR5969?hootPostID=9fdadf559414e5b5298a528421983bca>
- [7] <http://www.electropages.com/2014/08/farnell-element14-ultra-low-power-texas-instruments-fram-launchpad-development-kit>.
- [8] [http://www.ti.com/product/MSP430FR59691/datasheet/detailed\\_description](http://www.ti.com/product/MSP430FR59691/datasheet/detailed_description).
- [9] [http://en.wikipedia.org/wiki/Ferroelectric\\_RAM](http://en.wikipedia.org/wiki/Ferroelectric_RAM).
- [10] [http://www.physikinstrumente.com/en/products/piezo\\_tutorial.php](http://www.physikinstrumente.com/en/products/piezo_tutorial.php).
- [11] IAR Embedded Workbench for MSP430 C/C++ Compiler User's Guide (IAR): [http://www. iar. com/Products/IAR-Embedded-Workbench/TI-MSP430/User-guides](http://www.iar.com/Products/IAR-Embedded-Workbench/TI-MSP430/User-guides).
- [12] Michael Zweig, Adolf Baumann, et al., An 82  $\mu$  A/MHz Microcontroller with Embedded FeRAM for Energy Harvesting applications(<http://ieeexplore.ieee.org>).
- [13] J. Rodriguez, K. Remack, J. Gertas, L. Wang, C. Zhou, K. Boku, J. Rodriguez-Latorre, K. R. Udayakumar, S. Summerfelt, T. Moise, D. Kim, J. Groat, J. Eliason, M. Depner, F. Chu, Reliability of Ferroelectric Random Access Memory Embedded within 130nm CMOS (<http://ieeexplore.ieee.org>)

## AUTHORS' BIOGRAPHY



**Mr. Thota Ravi Theja** obtained B.Tech Degree in 2011 from JNT University Hyderabad, M.Tech (Embedded Systems) in 2013 JNT University Hyderabad. Presently he is working as Assistant Professor, Department of ECE, Malla Reddy Engineering College, Dhullapally, Hyderabad, Telangana (state) India.



**Ms. Rajyalakshmi Gonuguntla** is presently pursuing final semester M.Tech in Electronics and Communication Engineering at Malla Reddy Engineering College, Secunderabad. She received her AMIE degree in Electronics and Communication from Bomma Institute of Technology and Science, Khammam, Her areas of interest are Microcontrollers and microprocessors, digital system design and analog circuits.

## Video Watermarking by using the Wavelet Transform with Perform SVD Technique

T.Harish Kumar<sup>1</sup>, G.Kumaraswamy<sup>2</sup>

<sup>1</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (PG Scholar)

<sup>2</sup>Department of ECE, MREC (Autonomous), JNTUH, Hyderabad, India (Assistant Professor)

### ABSTRACT

Digital technology allows unauthorized reproduction of digital videos, the protection of the copyrights of digital video is a very important issue. Video watermarking schemes are used to protect the digital videos. It is the process of embedding an imperceptible data (watermark) into cover video. The video watermarking schemes have been widely used to solve the copyright protection problems of digital video related to illegal usage or distribution. To resolve the copyright protection problem, it proposes an effective, robust and imperceptible video watermarking scheme. The combination of Discrete Wavelet Transformation (DWT) and Singular Value Decomposition (SVD) of Blue channel is used to embed the watermark. The copy of the watermark is embedded into high frequency sub band coefficients which are very difficult to remove or destroy. The combination of DWT and SVD increases the security, robustness and imperceptibility of the scheme. The extracted watermark image will be matched with input logo image for authentication to access (play) the video

**Keywords:** Digital Video, Video Watermarking, DWT-SVD Process, Embedding Process, Robustness and Imperceptibility

### INTRODUCTION

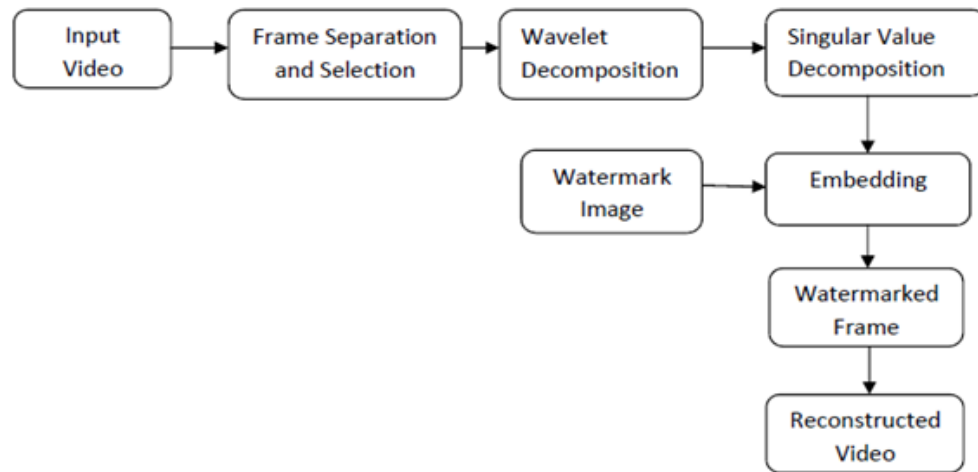
The identification of objects in an image and this process would probably start with image processing techniques such as noise removal, followed by (low-level) feature extraction to locate lines, regions and possibly areas with certain textures. The clever bit is to interpret collections of these shapes as single objects, e.g. cars on a road, boxes on a conveyor belt or cancerous cells on a microscope slide. One reason this is an AI problem is that an object can appear very different when viewed from different angles or under different lighting. Another problem is deciding what features belong to what object and which are background or shadows etc. The human visual system performs these tasks mostly unconsciously but a computer requires skilful programming and lots of processing power to approach human performance. Manipulation of data in the form of an image through several possible techniques. An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen, or movie screen. An image can be processed optically or digitally with a computer. Digital information revolution and the thriving progress in network communication are the major driving forces of this change. The perfect reproduction, the ease of editing, and the Internet distribution of digital Multimedia data have brought about concerns of copyright infringement, illegal distribution, and unauthorized tampering. Techniques of associating some imperceptible data with multimedia sources via embedding started to come out to alleviate these concerns. Interestingly, while most such techniques embed data imperceptibly to retain the perceptual quality and value of the host multimedia source, many of them were referred as digital watermarking whose traditional counterpart is not necessarily imperceptible.

*\*Address for correspondence:*

harishtornala@gmail.com

## **DIGITAL WATERMARKING**

We would normally like to increase the energy of the watermark (or payload of the watermark) in order to increase its robustness. However, increasing the payload of the watermark degrades the visual quality of the image such that human eye will notice the degradation. A dual reasoning leads us to think that it might be better to increase the payload of the watermark by embedding the watermark bits into places where human eye will not detect the changes to the image. Several watermarking schemes were proposed by researchers that aim to exploit the characteristics of the human visual system. For example, [8] suggests to make the gain factor luminance dependent. This is because of the fact that Human Visual System (HVS) is less sensitive to changes in regions of high luminance.



**Figure1.** Block Diagram for Proposed System

We can exploit this property by increasing the payload (energy) of the watermark in those specific areas. We can create a mask image that consists of those areas that are less sensitive to distortions and modulate the watermark bits using this mask image.

$$WI(i,j) = I(i,j) + \text{Mask}(i,j).k.W(i,j)$$

W is the watermark pattern (image), k is the gain factor, and Mask is the mask image as mentioned above. In my implementation, I generate the Mask image using an edge detection algorithm. I convert the edge image into a binary image. I amplify the effect of watermark bits by k on pixels where edge image is '1' and keep the effect of the watermark bits minimal on pixels where edge image is '0'. This increases the energy of the watermark along the edges in the image. I use the canny edge detector to extract the edge information out of the image.

## **VIDEO**

Digital video refers to the capturing, manipulation, and storage of moving images that can be displaced on computer screens. This requires that the moving images be digitally handled by the computer. The word digital refers to a system based on discontinuous events, as opposed to analog, a continuous event. Computers are digital systems; they do not process images the way the human eye does. Stands for Audio Video Interlaced. It is one of the oldest formats. It was

Created by Microsoft to go with Windows 3.1 and it's "Video for Windows" application. Even though it is widely used due to the number of editing systems and software that use AVI by default, this format has many restrictions, specially the compatibility with operations systems and other interface boards (Fisher & Schroeder).



Figure2. Input Videos

Frame processing is the first step in the background subtraction algorithm, the purpose of this step is to prepare the modified video frames by removing noise and unwanted object's in the frame in order to increase the amount of information gained from the frame and the sensitivity of the algorithm. Pre-processing is a process of collecting simple image processing tasks that change the raw input video into a format. This can be processed by subsequent steps.

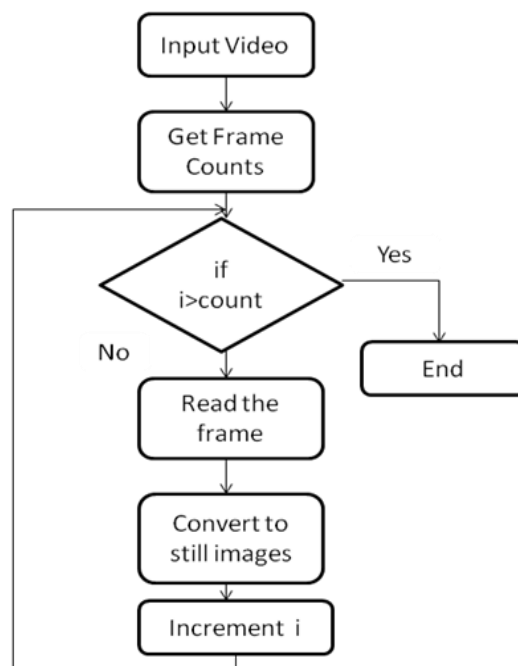
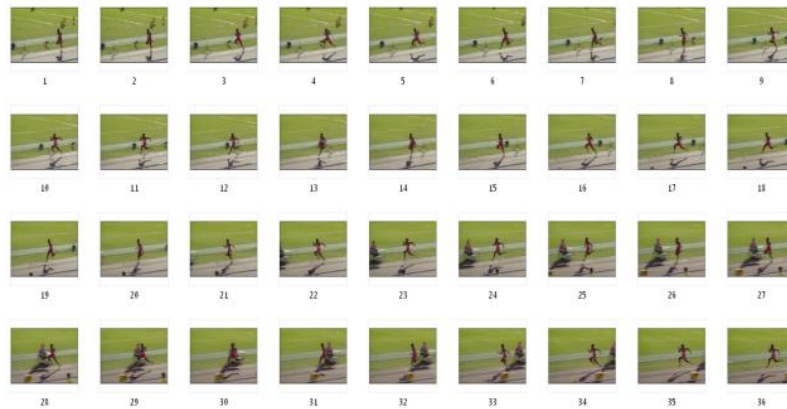


Figure3. Frame Separation Process

An Input Video (**.avi files**) is converted into still images for processing it and to detect the moving objects. These sequences of images gathered from video files by finding the information about it through '**aviinfo**' command. These frames are converted into images with help of the command '**frame2im**' Create the name to each images and this process will be continued for all the video frames.



**Figure4.** *Frame Separations for Input Video*

## **SINGLE FRAME**

According to ADPS a picture is a foresaid as array of numbers that represents lightweight intensities at pixels, which ends in information. Image consists of eight bits per picture element i.e.256 colors. Frame could be a image that has been created or derived and keep in electronic sort of Image Format. A picture is delineate in terms of vector graphics or formation graphics. a picture keep in formation type is typically known as a picture.

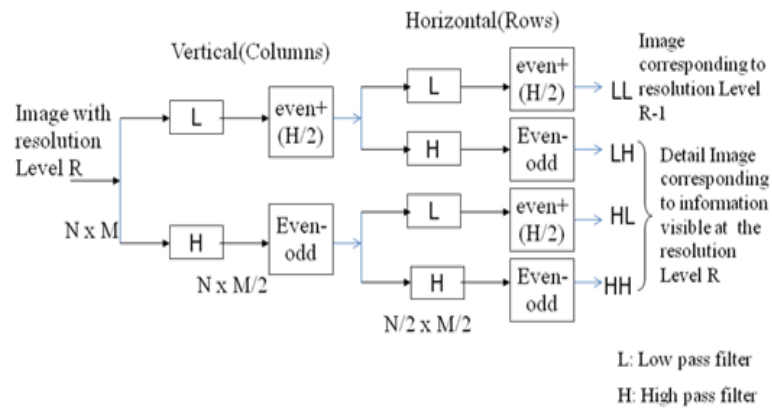


**Figure5.** *(a) RGB Plane Image (b) Single (Blue) Plane Image*

The colors square measure generated from 3 primary colours as red, inexperienced and blue (RGB) [28][11-13]. varied approaches has been designed for image steganography a number of common approaches square measure LSB(Least important Bit) substitution that is that the straightforward and commonest approach of concealment information within pictures. Masking is another technique of embedding messages in important areas. The DWT supported image transformation involves the mathematical relation for concealment information within the pictures.

## **DISCRETE WAVELET TRANSFORM**

These functions contain the direct and inverse lifting riffle remodel (LWT) M-files for each 1-D and 2-D signals. LWT reduces to the poly part version of the DWT algorithmic rule with zero-padding extension mode and while not extra-coefficients. Coming up with new riffles that square measure similar temperament for the distinct wavelet remodel (DWT) is additional delicate and, till recently, was completely a subject for riffle specialists. The 1-D DWT is extended to 2-D remodel exploitation divisible riffle filters. With divisible filters, applying a 1-D remodel to any or all the rows of the input and so continuation on all of the columns will cypher the 2-D remodel. Once one-level 2-D DWT is applied to a picture, four remodel constant sets square measure created.



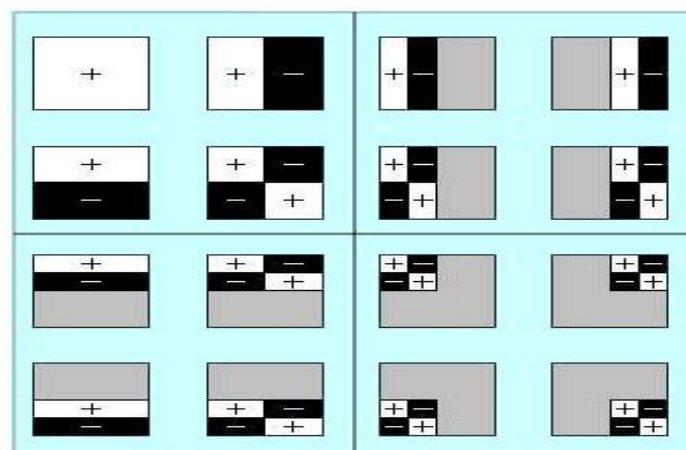
**Figure6.** Wavelet Decomposition Process

An image that undergoes Haar riffle remodel are divided into four bands at every of the remodel level. The primary band represents the input image filtered with an occasional pass filter and compressed to 0.5. This band is additionally known as ‘approximation’. the opposite 3 bands square measure known as ‘details’ wherever the high pass filter is applied. These bands contain directional characteristics. The dimensions of every of the bands is additionally compressed to 0.5.

Specifically, the second band contains vertical characteristics, the third band shows characteristics within the horizontal direction and also the last band represents diagonal characteristics of the input image. Conceptually, Haar riffle is extremely straightforward as a result of it's created from a sq. wave. Moreover, the Haar riffle computation is quick since it solely contains 2 coefficients and it doesn't would like a short lived array for multi-level transformation. Thus, every picture element in a picture that may undergo the riffle remodel computation are used just one occasion and no picture element overlapping throughout the computation.

## HAAR WAVELET PROCESS

The first DWT was fictitious by Hungarian man of science Alfred Haar. For Associate in Nursing input drawn by an inventory of numbers, the Haar riffle remodel could also be thought-about to combine up input values, storing the distinction and spending the add. This method is perennial recursively, pairing up the sums to supply succeeding scale, that results in variations and a final add. The Haar riffle is additionally the only potential riffle. The technical advantage of the Haar riffle is of signals with fast transitions, like watching of tool failure in machines.



**Figure7.** Sub band Representation in Haar wavelet Transform

The Haar wavelet's mother wavelet function  $\psi(t)$  can be described as

$$\psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2}, \\ -1 & \frac{1}{2} \leq t < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Its scaling function  $\phi(t)$  can be described as

$$\phi(t) = \begin{cases} 1 & 0 \leq t < 1, \\ 0 & \text{otherwise.} \end{cases}$$



**Figure8.** Sub band image for Haar Wavelet Transform

## SVD PROCESS

The singular value decomposition (SVD) is a factorization of a real or complex matrix, with many useful applications in signal processing and statistics. Formally, the singular value decomposition of an  $m \times n$  real or complex matrix  $M$  is a factorization of the form follow in this equation.

$$M = U \Sigma V'$$



**Figure9.** SVD Process

Where  $U$  is an  $m \times m$  real or complex unitary matrix,  $\Sigma$  is an  $m \times n$  rectangular diagonal matrix with nonnegative real numbers on the diagonal, and  $V^*$  is an  $n \times n$  real or complex unitary matrix. A non-negative real number  $\sigma$  is a singular value for  $M$  if and only if there exist unit-length vectors  $u$  in  $K^m$  and  $v$  in  $K^n$  such that show as equation

$$Mv = \sigma u$$

The vectors  $u$  and  $v$  are called left-singular and right singular vectors for  $\sigma$ , respectively.

## EMBEDDING PROCESS

The secret image will be decomposed into singular and two orthogonal matrixes. These values are concealing into singular values of high frequency sub bands by modifying it through key value. The key should be selected as least value to reduce the embedding error. The singular value of sub band will be modified by,

$$M_s = C_s + (W_s * K)$$

Where,  $C_s$  – Singular value of cover image sub bands

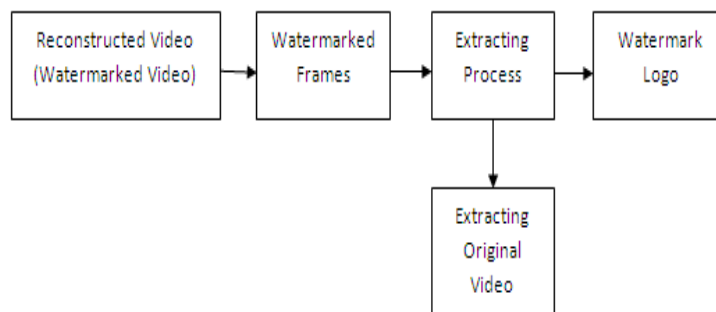
$W_s$  – Singular value of Watermark Image

$M_s$  – Modified Singular matrix

$K$  – Least Key Value.

## WATERMARK EXTRACTION PROCESS

The recognition will be included for accessing the video by person who is having same logo which is already embedded. Before recognition, the watermark image will be extracted from corresponding frame of particular video.



**Figure10.** Block Diagram for Extraction Process

The extracted logo will be matched with query image to check authentication by extracting the statistical features. The features are extracted and its matched with query features by Euclidean distance. If the query image will be matched then corresponding video is accessible otherwise is not opened.

## RESULT ANALYSIS

### Image Quality

Although there are several metrics that tend to be indicative of image quality, each of them has situations in which it fails to coincide with an observer's opinion. However, since running human trials is generally prohibitively expensive, a number of metrics are often computed to help judge image quality; the metrics that see the widest usage are generally quite simple to compute. And once a metric has been used in a seminal article that presents test results, , other researchers will continue using that metric so that their data can be compared to the previous work. This last reason was the determining factor in the use of peak signal-to-noise ratio (PSNR).

### Mean Square Error

Two other quantities that appear frequently when comparing original and reconstructed or approximated data are (root) mean square error. These measures will not be seriously skewed by a single anomaly, since they are measuring average behaviour. RMSE produces the same units as the

original image data, so its results are easy to interpret. Again it should be remembered that these metrics attempt to measure an inverse to image quality.

$$MSE = \frac{\sum \sum [A(i, j) - B(i, j)]^2}{M \times N}$$

Here, A(i,j) = Cover Image (Frame).

B(i,j) = Watermarked Image (Frame).

M X N=row and column of image intensity of pixel vales (255 255) image size.

### **Peak Signal Noise Ratio**

Peak signal-to-noise ratio has two definitions, the original more precise definition, and the second easier to compute and more commonly used. It is this second definition that we use throughout this report. This is the first metric mentioned so far where the results generally run proportional to image quality rather than the inverse.

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right)$$

Generally when PSNR is 20 dB or greater, then the original and the reconstructed images are virtually in-distinguishable by human eyes.

### **CONCLUSION**

The Project presented an effective, robust and imperceptible video watermarking scheme for logo matching based on chaotic crypto system with SVD based data concealment. Here, discrete wavelet transform was used to reserve space for concealing data effectively and chaos encryption was used as to protect image contents. Watermark recognition is used to recognize the input water mark for verification to access the video. This system was generated the Watermark image with less error under maximum data hiding capacity. Finally, the performance of system was evaluated with quality metrics such as error and PSNR factor. It is widely used for copy right protection of image or videos during internet sharing. It was better compatible approach and flexibility with better efficiency rather than prior methods.

### **ACKNOWLEDGEMENT**

I am very much thankful to my project guide Mr. G. Kumaraswamy, Assistant Professor, for his extensive patience and guidance throughout my project work.

I am grateful to our Project Coordinator, Mrs.K.S. Indrani, Associate Professor, for extending his support and cooperation throughout my project work.

I would like to thank Dr. M. Ch. P. Jagadeesh, Head of the ECE Department, for providing the freedom to use all the facilities available in the department, for successful completion of project.

I whole-heartedly express my deep sense of gratitude to our beloved Principal, Dr. S.SUDHAKARA REDDY, Principal, MALLA REDDY ENGINEERING COLLEGE (Autonomous) for providing me adequate facilities for my successful completion of both course work and project work.

I sincerely thank all the staff of the Department, for their timely suggestions, healthy criticism and motivation during the course of my project work.

I would also like to thank all my friends for providing help and moral support at the right timing. With great affection and love, I thank my parents who were the backbone behind my deeds.

## REFERENCES

- [1] Ray-Shine Run, Shi-Jinn Horng, Jui-Lin Lai, Tzong-Wang Kao, Rong-Jian Chen, "An improved SVD-based watermarking technique for copyright protection", Expert Systems with Applications 39, 2012, pp-673–689.
- [2] Shanjun, Zhang; Kazuyoshi, Yoshino, "DWT-Based Watermarking Using QR Code" Science Journal of Kanagawa University, pp3-6, 2008.
- [3] Thitapa Poomvichid, Pantida Patirupanusara and Mahasak Ketcham, "The QR Code for Audio Watermarking using Genetic Algorithm", IMLCS'2012, pp 11-12, 2012.
- [4] Ahmad A. Mohammad, Ali Alhaj, Sameer Shaltaf, "An improved SVD-based watermarking scheme for protecting righ.
- [5] I.J. Cox, M. L. Miller, J. A. Bloom, "Digital Watermarking", Academic Press 2002.
- [6] Bai Ying Lei n, IngYannSoon, ZhenLi, "Blind and robust audio watermarking schemes based on SVD–DCT" Signal Processing, Vol- 91, 2011, pp-1973–1984.
- [7] G.C. Langelaar, I. Setyawan and R.L. Lagendijk, "Watermarking Digital Image and Video Data; A state of The Art Overview," IEEE Signal Processing Magazine, pp. 20-46, Sept. 2003.
- [8] Suppat Rungraungsilp, Mahasak Ketcham, Tanee Wiputtikul, Kanchana Phonphak, and Sartid Vongpradhip, "Data Hiding Method for QR Code Based on Watermark by comparing DFT with DWT Domain" ICCCT', May 26-27, 2012.
- [9] S. J. Sangwine, "Colour in Image Processing", Electronics & Communication Engineering Journal, pp. 211-219, Oct. 2000.
- [10] Veysel Aslantas, "An optimal robust digital image watermarking based on SVD using the differential evolution algorithm.

## AUTHOR’S BIOGRAPHY



### **T. HARISH KUMAR**

B.Tech Passed in 2013 in ECE branch from Anurag Engineering College, JNTUH, Nalgonda.

M.Tech from Malla Reddy Engineering College (AUTONOMOUS)

*Branch:* VLSI System Design, Dept of ECE

## VLSI Implementation of Signed Multiplier using Quaternary Signed Digit Number System

G. RAJKUMAR<sup>1</sup>, V. SRINIVAS<sup>2</sup>

<sup>1</sup>PG Scholar, Dept of ECE (VLSI), Malla Reddy Engineering College, Hyderabad, India.

<sup>2</sup>Assistant Professor, Dept of ECE, Malla Reddy Engineering College, Hyderabad, India.

**Abstract:** A carry-free arithmetic operation can be achieved using a higher Radix number system such as Quaternary Signed Digit (QSD). In QSD, each digit can be represented by a number from -3 to 3. This number system allows multiple representations of any integer. By exploiting this feature, we can design an adder without ripple carry. The implementation of quaternary addition and multiplication results in a fix delay independent of the number of digits. Operations on a large number of digits such as 64, 128, or more, can be implemented with constant delay and less complexity. The QSD multiplier which is Signed multiplier can be designed by employing the QSD adder. The QSD is constant delay adder i.e. for any number of input bits it produces constant propagation delay therefore when we use this adder in multiplier designing the propagation delay will be reduces. The QSD multiplier can multiply both positive and negative numbers without any extra delay and complexity.

**Keywords:** Quaternary Signed Digit (QSD), Logic Circuits, BSD.

### I. INTRODUCTION

Arithmetic operations play an important role in various digital systems such as computers, process controllers, signal processors computer graphics and image processing. The high performance adders and Multipliers are essential since the speed of the digital processor depends heavily on the speed of the adders used is the system. In this paper, we propose a high speed QSD arithmetic logic unit which is capable of carry free addition. The QSD addition operation employs a fixed number of min terms for any operand size. Signed digit number system offers the possibility of carry free addition. QSD Adder / QSD Multiplier circuits are logic circuits designed to perform high-speed arithmetic operations. In QSD number system carry propagation chain are eliminated which reduce the computation time substantially, thus enhancing the speed of the machine. VLSI implementation of Fast Addition using Quaternary Signed Digit Number System” It offers the advantage of reduced circuit complexity both in terms of transistor count and interconnections. QSD number uses 25% less space than BSD to store number and it can be verified by the theorem described as under.

#### A. QSD Numbers Save 25% Storage Compared to BSD

To represent a numeric value  $N$   $\log_4 N$  number of QSD digits and  $3 \log_4 N$  binary bits are required while for the same  $\log_2 N$  BSD digits and  $2 \log_2 N$  binary bits are required in BSD representation. Ratio of number of bits in QSD to BSD representation for an arbitrary number  $N$  is,

$$\frac{3 \lceil \log_4 N \rceil}{2 \lceil \log_2 N \rceil} \quad (1)$$

Which roughly equal to  $\frac{3}{4}$ . Therefore QSD saves  $\frac{1}{4}$  of the storage used by BSD? So the proposed QSD adder is better than RBSD adder in terms of number of gates, input connections and delay though both perform addition within constant time. Proposed design has the advantages of both parallelisms as well as reduced gate complexity. The computation speed and circuit complexity increases as the number of computation steps decreases. A two step schemes appear to be a prudent choice in terms of computation speed and storage complexity. Quaternary is the base 4 redundant number system. The degree of redundancy usually increases with the increase of the radix. The signed digit number system allows us to implement parallel arithmetic by using redundancy. QSD numbers are the SD numbers with the digit set as:  $\{\bar{3}, \bar{2}, \bar{1}, 0, 1, 2, 3\}$  where  $\bar{3}, \bar{2}, \bar{1}$  represent -3, -2, and -1 respectively.

In general, a signed-digit decimal number  $D$  can be represented in terms of an  $n$  digit quaternary signed digit number as

$$D = \sum_{i=0}^{n-1} X_i 4^i \quad (2)$$

Where  $X_i$  can be any value from the set  $\{\bar{3}, \bar{2}, \bar{1}, 0, 1, 2, 3\}$  for producing an appropriate decimal representation. For digital implementation, QSD numbers are represented using 3-bit 2's complement notation. A QSD negative number is the QSD complement of the QSD positive number. Conversion of decimal number to QSD number

$$(107)_{10} = 2 \times 4^3 + \bar{2} \times 4^2 + 3 \times 4^1 + \bar{1} \times 4^0 = (2\bar{2}3\bar{1})_{\text{QSD}}$$

$$(233)_{10} = 3 \times 4^3 + 3 \times 4^2 + \bar{2} \times 4^1 + 1 \times 4^0 = (33\bar{2}1)_{\text{QSD}}$$

$$\text{Hence, } (-233)_{10} = (\bar{3}\bar{3}2\bar{1})_{\text{QSD}}$$

## II. DESIGN ALGORITHM OF QSD ADDER

In QSD number system carry propagation chain are eliminated which reduce the computation time substantially, thus enhancing the speed of the machine [31]. As range of QSD number is from -3 to 3, the addition result of two QSD numbers varies from -6 to +6. Table 3.3 depicts the output for all possible combinations of two numbers. The decimal numbers in the range of -3 to +3 are represented by one digit QSD number. As the decimal number exceeds from this range, more than one digit of QSD number is required. For the addition result, which is in the range of -6 to +6, two QSD digits are needed. In the two digits QSD result the LSB digit represents the sum bit and the MSB digit represents the carry bit. To prevent this carry bit to propagate from lower digit position to higher digit position QSD number representation is used. QSD numbers allow redundancy in the number representations. The same decimal number can be represented in more than one QSD representations. So we choose such QSD represented number which prevents further rippling of carry. To perform carry free addition, the addition of two QSD numbers can be done in two steps:

- **Step 1:** First step generates an intermediate carry and intermediate sum from the input QSD digits i.e., addend and augends.
- **Step 2:** Second step combines intermediate sum of current digit with the intermediate carry of the lower significant digit.

So the addition of two QSD numbers is done in two stages. First stage of adder generates intermediate carry and intermediate sum from the input digits. Second stage of adder adds the intermediate sum of current digit with the intermediate carry of lower significant digit. To remove the further rippling of carry there are two rules to perform QSD addition in two steps:

- **Rule 1:** First rule states that the magnitude of the intermediate sum must be less than or equal to 2 i.e., it should be in the range of -2 to +2.
- **Rule 2:** Second rule states that the magnitude of the intermediate carry must be less than or equal to 1 i.e., it should be in the range of -1 to +1.

According to these two rules the intermediate sum and intermediate carry from the first step QSD adder can have the range of -6 to +6. But by exploiting the redundancy feature of QSD numbers we choose such QSD represented number which satisfies the above mentioned two rules. When the second step QSD adder adds the intermediate sum of current digit, which is in the range of -2 to +2, with the intermediate carry of lower significant digit, which is in the range of -1 to +1, the addition result cannot be greater than 3 i.e., it will be

in the range of -3 to +3. The addition result in this range can be represented by a single digit QSD number; hence no further carry is required. In the step 1 QSD adder, the range of output is from -6 to +6 which can be represented in the intermediate carry and sum in QSD format as shown in table 3.3. We can see in the first column of Table 3.3 that some numbers have multiple representations, but only those that meet the above defined two rules are chosen. The chosen intermediate carry and intermediate sum are listed in the last column of Table 3.3 as the QSD coded number.

TABLE I: Intermediate Carry and Sum Between -6&6

Sum	QSD represented number	QSD coded number
-6	$\bar{2}2, \bar{1}\bar{2}$	$\bar{1}\bar{2}$
-5	$\bar{2}3, \bar{1}\bar{1}$	$\bar{1}\bar{1}$
-4	$\bar{1}0$	$\bar{1}0$
-3	$\bar{1}1, 0\bar{3}$	$\bar{1}1$
-2	$\bar{1}2, 0\bar{2}$	$0\bar{2}$
-1	$\bar{1}3, 0\bar{1}$	$0\bar{1}$
0	$00$	$00$
1	$01, 1\bar{3}$	$01$
2	$02, 1\bar{2}$	$02$
3	$03, 1\bar{1}$	$1\bar{1}$
4	$10$	$10$
5	$11, 2\bar{3}$	$11$
6	$12, 2\bar{2}$	$12$

This addition process can be well understood by following examples:

**Example:** To perform QSD addition of two numbers A=107 and B=-233 (One number is positive and one number is negative). First convert the decimal number to their equivalent QSD representation:

$$(107)_{10} = 2 \times 4^3 + \bar{2} \times 4^2 + 3 \times 4^1 + \bar{1} \times 4^0 = (2\bar{2}3\bar{1})_{\text{QSD}}$$

$$(233)_{10} = 3 \times 4^3 + 3 \times 4^2 + \bar{2} \times 4^1 + 1 \times 4^0 = (33\bar{2}1)_{\text{QSD}}$$

$$\text{Hence, } (-233)_{10} = (\bar{3}\bar{3}2\bar{1})_{\text{QSD}}$$

Now the addition of two QSD numbers can be done as follows:

A = 107	2	$\bar{2}$	3 $\bar{1}$
B = -233	$\bar{3}$	$\bar{3}$	2 $\bar{1}$
Decimal Sum	-1	-5	5 -2
IC	$\bar{1}$	1	0
IS	$\bar{1}$	$\bar{1}$	1 $\bar{2}$
S	$\bar{2}$	0	1 $\bar{2}$

## VLSI Implementation of Signed Multiplier using Quaternary Signed Digit Number System

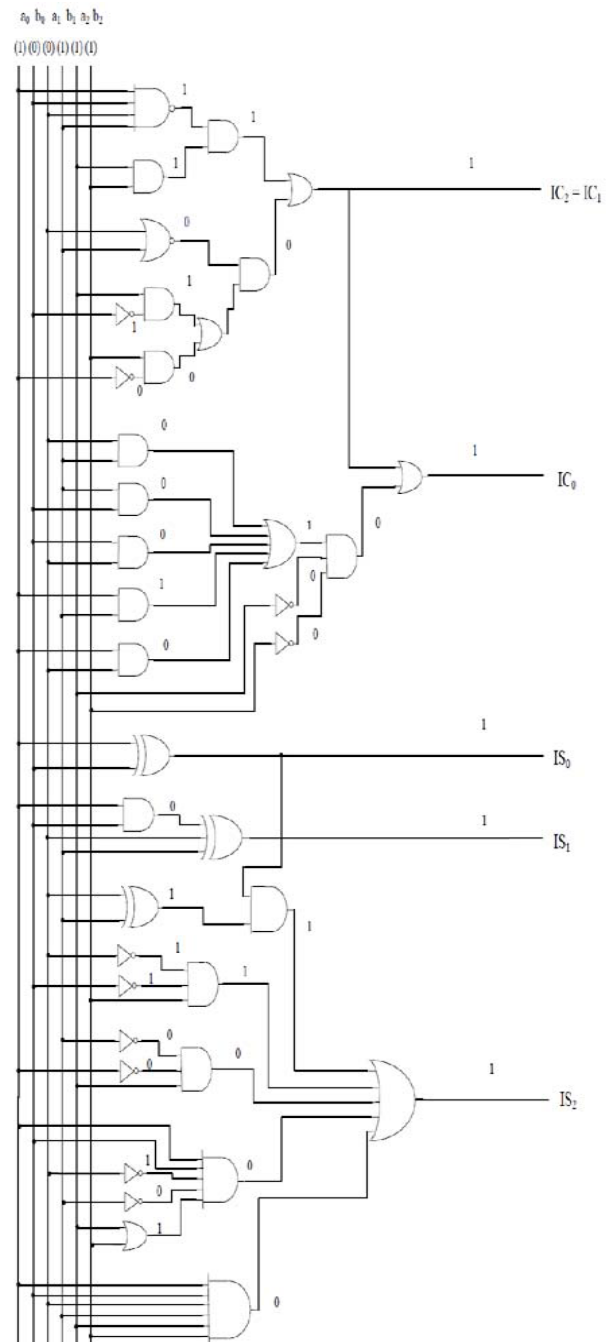
The sum output is  $(\bar{2}01\bar{2})_{QSD}$  which is equivalent to  $(-126)_{10}$  and carry output is 0. From these examples it is clear that the QSD adder design process will carry two stages for addition. The first stage generates intermediate carry and sum according to the defined rules. In the second stage the intermediate carry from the lower significant digit is added to the intermediate sum of current digit which results in carry free output. In this step the current digit can always absorb the carry-in from the lower digit.

**TABLE II: Outputs of All Possible Combinations of a Pair of Intermediate Carry (A) And Sum (B)**

INPUT				OUTPUT		
QSD		Binary		Decimal	QSD	Binary
$A_i$	$B_i$	$A_i$	$B_i$	Sum	$S_i$	$S_i$
1	2	01	010	3	3	111
1	1	01	001	2	2	010
0	2	00	010	2	2	010
0	1	00	001	1	1	001
1	0	01	000	1	1	001
-1	2	11	010	1	1	001
0	-1	00	111	-1	-1	111
-1	0	11	000	-1	-1	111
1	-2	01	110	-1	-1	111
-1	-1	11	111	-2	-2	110
0	-2	00	110	-2	-2	110
-1	-2	11	110	-3	-3	001

### III. LOGIC DESIGN AND IMPLEMENTATION USING OF SINGLE DIGIT QSD ADDER

There are two steps involved in the carry-free addition. The first step generates an intermediate carry and sum from the addend and augends. The second step combines the intermediate sum of the current digit with the carry of the lower significant digit. To prevent carry from further rippling, two rules are defined. The first rule states that the magnitude of the intermediate sum must be less than or equal to 2. The second rule states that the magnitude of the carry must be less than or equal to 1. Consequently, the magnitude of the second step output cannot be greater than 3 which can be represented by a single-digit QSD number; hence no further carry is required. In step 1, all possible input pairs of the addend and augend are considered. The lower significant digit serves as sum and most significant digit serves as carry. The generation of the carry can be avoided by mapping the two digits into a pair of intermediate sum and intermediate carry such that the  $n$ th intermediate sum and the  $(n-1)$ th intermediate carry never form any carry generating pair (3,3), (3,2), (3,1), (3,3), (3,2), (3,1). If we restrict the representation such that the intermediate carry is limited to a maximum of 1, and the intermediate sum is restricted to be less than 2, then the final addition will become carry free. Both inputs and outputs can be encoded in 3-bit 2's complement binary number. The mapping between the inputs, addend and augend, and the outputs, the intermediate carry and sum are shown in binary.



**Fig.1.**

#### A. Intermediate Carry and Sum Are Shown In Binary Format in Table II

To remove the further carry propagation the redundancy feature of QSD numbers is used. We restrict the representation such that all the intermediate carries are limited to a maximum of 1, and the intermediate sums are restricted to be less than 3, then the final addition will become carry free. The QSD representations according to these rules are shown in Table 4.3 for the range of -6 to +6. As the range of intermediate carry is from -1 to +1, it can be represented in 2 bit binary number but we take the 3 bit representation for the bit compatibility with the intermediate sum. At the input side,

the addend  $A_i$  is represented by 3 variable input as  $A_2, A_1, A_0$  and the augend  $B_i$  is represented by 3 variable input as  $B_2, B_1, B_0$ . At the output side, the intermediate carry  $IC$  is represented by  $IC_2, IC_1, IC_0$  and the intermediate sum  $IS$  is represented by  $IS_2, IS_1, IS_0$ . The six variable expressions for intermediate carry and intermediate sum in terms of inputs ( $A_2, A_1, A_0, B_2, B_1$  and  $B_0$ ) can be derived from Table 4.3. So we get the six output expressions for  $IC_2, IC_1, IC_0, IS_2, IS_1$  and  $IS_0$ . As the intermediate carry can be represented by only 2 bits, the third appended bit  $IC_2$  is equal to  $IC_1$  so the expression for both outputs will be the same.

**TABLE III: The Intermediate Carry and Sum Between -6 To 6**

INPUT				OUTPUT					
QSD		Binary		Decimal	QSD		Binary		
$A_i$	$B_i$	$A_i$	$B_i$	Sum	$C_i$	$S_i$	$C_i$	$S_i$	
3	3	011	011	6	1	2	01	010	
3	2	011	010	5	1	1	01	001	
2	3	010	011	5	1	1	01	001	
3	1	011	001	4	1	0	01	000	
1	3	001	011	4	1	0	01	000	
2	2	010	010	4	1	0	01	000	
1	2	001	010	3	1	-1	01	111	
2	1	010	001	3	1	-1	01	111	
3	0	011	000	3	1	-1	01	111	
0	3	000	011	3	1	-1	01	111	
1	1	001	001	2	0	2	00	010	
0	2	000	010	2	0	2	00	010	
2	0	010	000	2	0	2	00	010	
3	-1	011	111	2	0	2	00	010	
-1	3	111	011	2	0	2	00	010	
0	1	000	001	1	0	1	00	001	
1	0	001	000	1	0	1	00	001	
2	-1	010	111	1	0	1	00	001	
-1	2	111	010	1	0	1	00	001	
3	-2	011	110	1	0	1	00	001	
2	3	010	011	1	0	1	00	001	
0	0	000	000	0	0	0	00	000	
1	-1	001	111	0	0	0	00	000	
-1	1	111	001	0	0	0	00	000	
2	-2	010	110	0	0	0	00	000	
-2	2	110	010	0	0	0	00	000	
-3	3	101	011	0	0	0	00	000	
3	-3	011	101	0	0	0	00	000	
0	-1	000	111	-1	0	-1	00	111	
-1	0	111	000	-1	0	-1	00	111	
-2	1	110	001	-1	0	-1	00	111	
1	-2	001	110	-1	0	-1	00	111	
-3	2	101	010	-1	0	-1	00	111	
2	-3	010	101	-1	0	-1	00	111	
-1	-1	111	111	-2	0	-2	00	110	
0	-2	000	110	-2	0	-2	00	110	
-2	0	110	000	-2	0	-2	00	110	
-3	1	101	001	-2	0	-2	00	110	
1	-3	001	101	-2	0	-2	00	110	
-1	-2	111	110	-3	-1	1	11	001	
-2	-1	110	111	-3	-1	1	11	001	
-3	0	101	000	-3	-1	1	11	001	
0	-3	000	101	-3	-1	1	11	001	
-3	-1	101	111	-4	-1	0	11	000	
-1	-3	111	101	-4	-1	0	11	000	
-2	-2	110	110	-4	-1	0	11	000	
-3	-2	101	110	-5	-1	-1	11	111	
-2	-3	110	101	-5	-1	-1	11	111	
-3	-3	101	101	-6	-1	-2	11	110	

## B. Arrey Multiplier

Array multiplier is well known due to its regular structure. Multiplier circuit is based on add and shift algorithm. Each partial product is generated by the multiplication of the multiplicand with one multiplier bit. The partial product are shifted according to their bit orders and then added. The addition can be performed with normal carry propagate adder.

## C. Modified Baugh-Wooley Multiplier

One important complication in the development of the efficient multiplier implementations is the multiplication of two's complement signed numbers. The Modified Baugh-Wooley Two's Complement Signed Multiplier is the best known algorithm for signed multiplication because it maximizes the regularity of the multiplier logic and allows all the partial products to have positive sign bits. Baugh-Woolley technique was developed to design direct multipliers for two's complement numbers. When multiplying two's complement numbers directly, each of the partial products to be added is assigned number. Thus, each partial product has to be sign-extended to the width of the final product in order to form the correct sum by the Carry Save Adder tree. According to the Baugh-woolley approach, an efficient method of adding extra entries to the bit matrix is suggested to avoid having to deal with the negatively weighted bits in the partial product matrix.

## D. QSD Multiplier

There are generally two methods for a multiplication operation: parallel and iterative. QSD multiplication can be implemented in both ways, requiring a QSD partial product generator and QSD adder as basic components. A partial product,  $M_i$ , is a result of multiplication between an  $n$ -digit input,  $A_{n-1}A_0$ , with a single digit input,  $B_i$ , where  $i = 0..n-1$  as shown in Fig.2. The primitive component of the partial product generator is a single-digit multiplication unit whose functionality can be expressed as shown in Table IV.

**TABLE IV: The QSD Representation of A Single-Digit Multiplication Output.4.3**

Multiplication	QSD represented number	QSD coded number
-9	$\bar{2}1, \bar{3}3$	$\bar{2}1$
-6	$\bar{2}2, \bar{1}2$	$\bar{1}2$
-4	$\bar{1}0$	$\bar{1}0$
-3	$\bar{1}1, 0\bar{3}$	$\bar{1}1$
-2	$\bar{1}2, 0\bar{2}$	$0\bar{2}$
-1	$\bar{1}3, 0\bar{1}$	$0\bar{1}$
0	00	00
1	01, $1\bar{3}$	01
2	02, $1\bar{2}$	02
3	03, $1\bar{1}$	$1\bar{1}$
4	10	10
6	12, $2\bar{2}$	12
9	21, $3\bar{3}$	21

## VLSI Implementation of Signed Multiplier using Quaternary Signed Digit Number System

The outputs of all possible combinations of a pair of multiplicand (A) and multiplier (B).

B \ A	-3	-2	-1	0	1	2	3
-3	9	6	3	0	-3	-6	-9
-2	6	4	2	0	-2	-4	-6
-1	3	2	1	0	-1	-2	-3
0	0	0	0	0	0	0	0
1	-3	-2	-1	0	1	2	3
2	-6	-4	-2	0	2	4	6
3	-9	-6	-3	0	3	6	9

Fig.2. Single digit QSD multiplier.

The implementation of an n-digit partial product generator uses n units of the single-digit QSD multiplier as shown in Fig.3. Gathering all the outputs to produce a partial product result presents a small challenge. The QSD representation of a single digit multiplication output, shown in Table 5, contains a carry-out of magnitude 2 when the output is either -9 or 9. This prohibits the use of the second step QSD adder alone as a gatherer. In fact, we can use the complete QSD adder from the previous section as the gatherer. Furthermore, the intermediate carry and sum circuit can be optimized by not considering the input of magnitude 3. The QSD partial product generator implementation is shown in Fig.4.

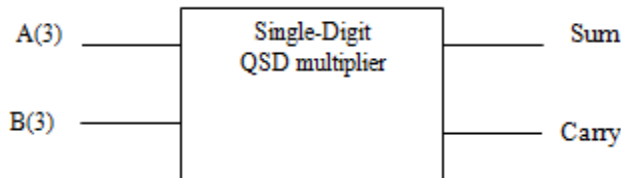


Fig.3. Single digit QSD multiplier.

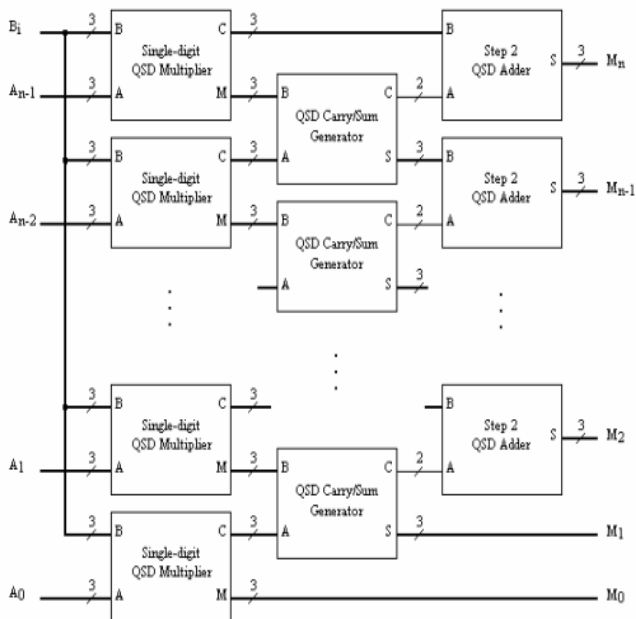


Fig.4. The n-digit QSD partial product generator.

An  $n \times n$ -digit QSD multiplication requires  $n$  partial product terms. In an iterative implementation, a 2ndigit QSD adder is used to perform add-shift operations between the partial product generator and the accumulator. After  $n$  iterations, the multiplication is complete. In contrast, a parallel implementation requires  $n$  partial product circuits and  $n-1$  QSD adder units. A binary reduction sum is applied to reduce the propagation delay to  $O(\log n)$ . The schematic of a  $4 \times 4$  parallel QSD multiplication

TABLE V: Mapping Between Multiplier's Inputs and Output

Input				Output					
QSD		Binary		Decimal	QSD		Binary		
A <sub>i</sub>	B <sub>i</sub>	A <sub>i</sub>	B <sub>i</sub>	M	C <sub>i</sub>	M <sub>i</sub>	C <sub>i</sub>	M <sub>i</sub>	
3	-3	011	101	-9	-2	-1	110	111	
-3	3	101	011						
-2	3	110	011	-6	-1	-2	111	110	
3	-2	011	110						
2	-3	010	101						
-3	2	101	010						
2	-2	010	110	-4	-1	0	111	000	
-2	2	110	010						
-1	3	111	011	-3	-1	1	111	001	
3	-1	011	111						
1	-3	001	101						
-3	1	101	001						
-1	2	111	010	-2	0	-2	000	110	
2	-1	010	111						
1	-2	001	110						
-2	1	110	001						
1	-1	001	111	-1	0	-1	000	111	
-1	1	111	001						
0	0	000	000	0	0	0	000	000	
1	1	001	001	1	0	1	000	001	
-1	-1	111	111						
1	2	001	010	2	0	2	000	010	
2	1	010	001						
-1	-2	111	110						
2	-1	110	111						
1	3	001	011	3	1	-1	001	111	
3	1	011	001						
-1	-3	111	101						
-3	-1	101	111						
2	-2	010	010	4	1	0	001	000	
-2	2	110	110						
2	3	010	011	6	1	2	001	010	
3	2	011	010						
-2	-3	110	101						
-3	-2	101	110						
3	3	011	011	9	2	1	010	001	
-3	-3	101	101						

## IV. SIMULATION AND SYTHESIS RESULTS

In this chapter the simulation results of 5-digit QSD adder, 7-digit QSD adder, Array multiplier, modified baughwooley multiplier and Signed QSD multiplier are presented as shown in Figs.5 and 6. And the results shows that the both 5-digit and 7-digit have the same propagation delay for their operation and the delay associated with Signed QSD multiplier is almost equal to the half that of the array and baughwooley multipliers. These multipliers are simulated in Ishim Simulator of XILINX ISE 13.1 Version and synthesized by

using XST Synthesizer in XILINX ISE 13.1 Version in the family of Virtex5 and selected Device is vlx330.

#### A. Waveform of Signed QSD Multiplier Table



Fig.5. Simulation result of Signed QSD multiplier.

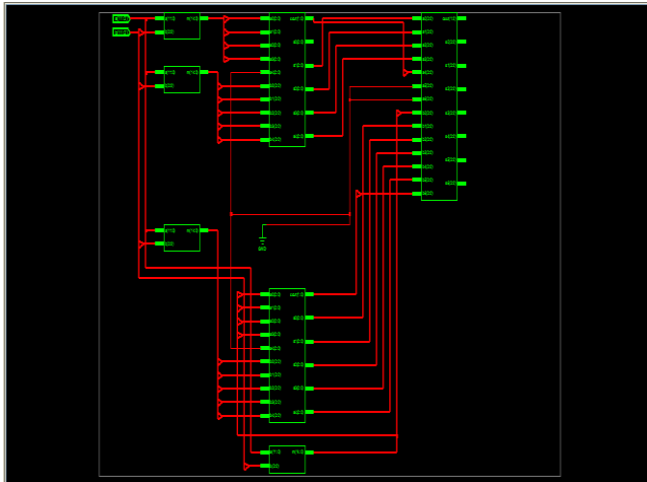


Fig.6. RTL Schematic of QSD multiplier.

#### V. RESULT

TABLE VI: Comparison Between 5-Digit and 7-Digit QSD Adder

	Delay(ns)	LUT's
5-digit QSD adder	5.033	37
7-digit QSD adder	5.033	53

TABLE VII: Comparison Between 12-Bit Multipliers

	Delay(ns)	LUT's
Array multiplier	24.165	248
Modified Baugh-Wooley multiplier	23.840	259
QSD multiplier	9.799	309

#### VI. CONCLUSION

In this project 5-digit and 7-digit QSD adders are designed and synthesized and it showed that both the adders produce equal delay i.e. 5.033ns. By using these QSD adders Signed QSD multiplier is designed and compared with existing multipliers. Which shows that it requires less time compared to other multipliers for same number of input bits, but it occupies more area compared to other multipliers.

#### VII. REFERENCES

- [1] SachinDubey, ReenaRani, SarojKumari "VLSI Implementation of Fast Addition using Quaternary Signed Digit Number System" 2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN 2013)
- [2] A. Avizienis, "Signed-Digit Number Representation for Fast Parallel Arithmetic," IRE Transaction Electron. Comp., EC-10, pp. 389-400, 1961.
- [3] A.A.S. AWOL and J.U. Ahmed, "fast carry free adder design using QSD number system" proceedings of the IEEE 1993 national aerospace and electronic conference, vol 2, pp 1085-1090, 1993.
- [4] F. Kharbash and G. M. Chaudhry, "Reliable Binary Signed Digit Number Adder Design", IEEE Computer Society Annual Symposium on VLSI, pp 479-484, 2007.
- [5] John Moskal, ErdalOruklu and JafarSaniie, "Design and Synthesis of a Carry-Free Signed-Digit Decimal Adder", IEEE International symposium on Circuits and Systems, pp 1089-1092, 2007.
- [6] J. U. Ahmed, A. A. S. AWOL, "Multiplier design using RBSD number system", Proceedings of the 1993 National Aerospace and Electronics Conference, pp. 180-184, Vol. 1, 1993.
- [7] O. Ishizuka, A. Ohta, K. Tannno, Z. Tang, D. Handoko, "VLSI design of a quaternary multiplier with direct generation of partial products," Proceedings of the 27th International Symposium on Multiple-Valued Logic, pp. 169-174, 1997.

## Monitoring and Faults Diagnosis in Wind Turbine by Using Zigbee Wireless Network

Spandana K<sup>1</sup>, Raju N<sup>2</sup>

<sup>1</sup>Department of ECE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

<sup>2</sup>Assistant professor, Malla Reddy Engineering College (Autonomous), Hyderabad, India

<sup>1</sup>Spandana.kommu@gmail.com, <sup>2</sup>razucontact@gmail.com

**Abstract:** The Wind Turbine blast is identified as a major problem against green energy. It is not only the threatening factor for the people but also causes dangerous hazard for human life. To overcome this problem and to increase the green energy, a simple system is introduced to monitor and prevent the fault occurrence in Small Wind Turbine. It collects all the parameters like temperature, vibration and oil level from main components of the turbine and sends it to the control room via wireless Zigbee. At Control room through LCD it is possible to view the current status of the Wind Turbine. In Case of fault occurrence indication automatically command will send to wind turbine section to overcome the fault than the fault occurrences can be prevented.

**Keywords:** wind turbine, monitoring and fault diagnosis, Temperature, Vibration, Oil, LCD Display.

### 1. INTRODUCTION

The wind power industry has experienced large growth in the past years. The growth mainly focuses on a growing market, wind power's better economic conditions (because of political decisions), and the development of large wind turbines and offshore farms. One of the goals is to increase reliability for turbines. The issue is even more important for offshore farms, where service is difficult and expensive. The industry has incentive to make maintenance more efficient. In whatever way, Wind turbines are fault prone, that is they are given in harsh environment such as desert, plains apart from that they are difficult electromechanical system that are situated far away from the control centre. So the chance of fault happens and the side cause will be more, even it leads to power off. It is important to develop the remote controlling and defect diagnosis system to monitor the run time status and the identification of fault to improve the effort and the life time service of the wind turbine.

Wind turbine monitoring systems could be the answer for better wind power industry maintenance management and increased reliability. Such systems are commonly used in other industries. This System continuously monitors the performance of wind turbine parts, e.g., temperature, weather, and transformer, and helps determine the optimal time for specific maintenance. This paper investigates these systems' support for the wind power user. The Health Monitoring of wind turbines, with the aim of diagnosing degradation in different structural parts in wind turbines, has seen its relevance boosted in the operation and maintenance plans for wind turbine industry. As a consequence many researchers are studying the relationship between structural deterioration and the change in different structural dynamics properties associated to some resonances.

In this context, wind turbine manufacturers are clearly increasing efforts in wind farm maintenance in order to achieve more robust and reliable wind turbines under any condition and under any operating regime. Needless to say, this requisite is even more demanding in offshore sites where the maintenance operations are challenging practices involving, in many cases, logistics or access difficulties. A defect is detected, the less impact it has in the wind turbine OPEX (operation expenditure) rate. Therefore, any tool devoted to the task of predicting a relevant structural change at early stage contributes to improve the market competitiveness for a particular wind turbine. Science of early detection of warranty and reliability issues - a topic that is

This paper provides an open-source literature survey on the emerging science of early detection of warranty and reliability issues - a topic that is becoming critical to product life management for the automotive, aerospace and energy industries. An open-source literature survey from a variety of disciplines like reliability engineering, operations research, prognostic-health management, systems engineering, biostatistics, public health and epidemiology.

These are areas that have seen significant recent research activity in the development of early warning systems and algorithms. This paper provides an overview of current and promising techniques with a focus on their underlying statistical theory and related system architecture issues. The concept of the project is to develop a system used to diagnose the fault and monitor the parameters in wind turbine. The project is a special one considering fact that we can monitor parameters and diagnose the fault in a wind turbine from the control station, this project is based on a Zigbee protocol (i.e. message based protocol). Zigbee connects the wind turbine with the control station using wires. This system requires minimal power and provides reliable delivery of data between wind turbine and control station.

### 2. INTRODUCTION TO ZIGBEE

ZigBee is a wireless communication used in Wind Turbine to send data to control room and to receive data from control room. It is based on a typical network architecture using an OSI model over an IEEE 802.15.4-2006. The ZigBee signals function like network signals and strictly similar to Bluetooth and Wi-Fi. ZigBee devices are meant for low power consumption these Devices put themselves to sleep when not in use, thereby conserving power. Hence this makes these devices ideal for battery-operated applications because they can last for a number of years before wanting fresh batteries. ZigBee is top in lucrative applications it scheme lends itself to detecting and monitoring applications and its practice in large scale wireless monitoring [4]. Hence it will be very much suitable for Wind turbine communication.

### 3. RELATED WORKS

The maintenance of offshore Wind Turbine and its failure rate is very high by using the wireless sensor networks for environment monitoring and key components cluster based distribution of wireless sensor network topology the failure rate scan be reduced. The necessity of monitoring and making automation of Wind Turbine with fault diagnosis system and the CAN Bus which is used as Interface module to communicate the monitored parameters between the Wind Turbine and the control centre with the help of RS-232 Serial Communication. Here GPRS is used to transfer the data from wind turbine section to control room. The importance renewable energy and its monitoring technologies of wind turbines facilitate the prevention of accidental and component or structure failure. The acquisition of data under different conditions can be time and cost consuming. Low-cost, low power wireless sensors have materialized as a window for these applications. Economical and flexible wireless sensors networks can be installed within a large structure to evaluate the response and performing monitoring algorithms. In this paper, wireless sensors are deployed into a wind turbine farm to monitor the structures to present models of wind turbine behavior and response to loading and controlling the generator parameters by getting even to rewrite the control algorithm from the system. Methodology to manage the mechanical vibration and temperature from Small Wind Turbine (SWT) with a protection tool developed for preventive maintenance and to avoid catastrophic failure. The triple axis accelerometer.

### 4. PROPOSED SYSTEM

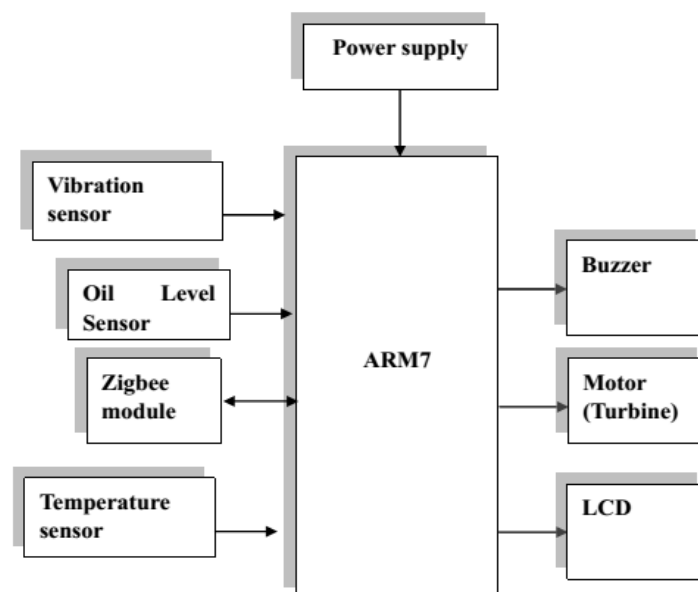
The core objectives of this proposed management of small wind turbine system are to detect the present health condition of the machine, to prevent catastrophic failures caused by the various

components of the wind farm, to improve the power quality before the problem is corrected, to predict the severity level of fault, and to estimate the useful life of the machine. The use of this type of health monitoring system helps to reduce the failure frequency and amount of downtime, maximize the utilization of the wind turbine, and minimize the maintenance overhead and cost due to production lost. Moreover, under the wireless sensor network system, there is no need to install wiring for data collecting and monitoring, thus eliminating the cost of installation and maintenance that would be required by communication cables. The overall performance and reliability of the wind form is improved dramatically.

#### 4.1. Wind Turbine Section

At Wind Turbine Section the System is developed to prevent Wind Turbine blasting hazards. The block diagram Fig explains the hardware components which are placed in the Wind Turbine. In this system the sensor sense the parameters values and send the data to the ARM Controller. The ARM is based on RISC architecture it simple design enables more efficient multi-core CPUs and higher core counts at lower cost, providing higher processing power and improved energy efficiency for servers and it reduces costs, heat and power usage this makes data acquisition and processing from the sensor and send it to the Control room via wireless ZigBee. The temperature sensor senses the temperature and send the data to control room at control room through LCD it is possible to view the current status of the Wind Turbine.

The speed sensor senses the rotational speed from the Generator shaft; the level sensor senses the lubricant oil level from the Wind Turbine components and vibration sensor sense the occurrence of higher vibration in any part of the Wind Turbine components. All this parameter data are send to the control room Section via wireless ZigBee. The display unit is placed is in the Wind Turbine section to show the parameters details which acts as the reference for operator in case of checking the working condition, are for any other revamping parts in the Turbine.



**Fig1.** Block diagram of proposed system

##### 4.1.1. ARM Microcontroller

Microcontroller is a heart of this project. ARM 7 is suitable microcontroller for this proposed embedded system. LPC2148 is ARM 7 controller used in this project. The main feature of LPC2148 are as follows,

- 16-bit/32-bit ARM7TDMI microcontroller.
- 8 kB to 40 kB of on-chip static RAM and 32 kB to 512 kB of on-chip flash memory.
- Two 10-bit ADCs provide a total of 14 analog inputs, with conversion times as low as 2.44 ms per channel.
- CPU operating voltage range of 3.0 V to 3.6 V ( $3.3\text{ V} \pm 10\%$ ) with 5 V tolerant I/O.

### 4.1.2. Power Supply

A device or system that supplies electrical or other types of energy to an output load or group of loads is called a power supply unit or PSU. The term is most commonly applied to electrical energy supplies, less often to mechanical ones and rarely to others.

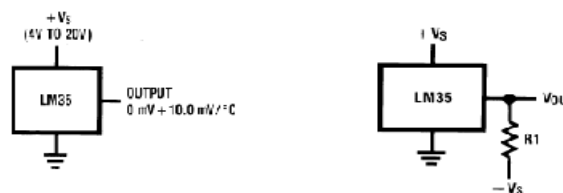
### 4.1.3. Temperature Sensor

LM35 converts temperature value into electrical signals. LM35 series sensors are precision integrated-circuit temperature sensors whose output voltage is linearly proportional to the Celsius temperature. The LM35 requires no external calibration since it is internally calibrated. . The LM35 does not require any external calibration or trimming to provide typical accuracies of  $\pm 1/4^\circ\text{C}$  at room temperature and  $\pm 3/4^\circ\text{C}$  over a full  $-55$  to  $+150^\circ\text{C}$  temperature range.

The LM35's low output impedance, linear output, and precise inherent calibration make interfacing to readout or control circuitry especially easy. It can be used with single power supplies, or with plus and minus supplies. As it draws only  $60\text{ }\mu\text{A}$  from its supply, it has very low self-heating, less than  $0.1^\circ\text{C}$  in still air.

### Features

- Calibrated directly in  $^\circ\text{Celsius}$  (Centigrade)
- Linear  $+10.0\text{ mV}/^\circ\text{C}$  scale factor
- $0.5^\circ\text{C}$  accuracy guaranteed (at  $+25^\circ\text{C}$ )



**Fig2.** The characteristic of this LM35 sensor is: For each degree of centigrade temperature it outputs 10milli volts.

### 4.1.4. Vibration Sensor

Vibration sensors allow you to detect orientation or inclination. They are small, inexpensive, low-power and easy-to-use. If used properly, they will not wear out. Their simplicity makes them popular for toys, gadgets and appliances. Sometimes they are referred to as "mercury switches", "tilt switches" or "rolling ball sensors" for obvious reasons.

They are usually made by a cavity of some sort (cylindrical is popular, although not always) and a conductive free mass inside, such as a blob of mercury or rolling ball. One end of the cavity has two conductive elements (poles). When the sensor is oriented so that that end is downwards, the mass rolls onto the poles and shorts them, acting as a switch throw.

Tilt switches used to be made exclusively of mercury, but are rarer now since they are recognized as being extremely toxic. The benefits of mercury is that the blob is dense enough that it doesn't bounce and so the switch isn't susceptible to vibrations. On the other hand, ball-type sensors are easy to make, won't shatter, and pose no risk of pollution.



**Fig3.** Vibration sensor

#### 4.1.5. Oil Level Sensor

A tank unit (sheet metal / aluminum die cast type) is a instrument to indicate level of Fuel in tank. It is used in all kind of vehicles / stationary engines which have tank of any kind.

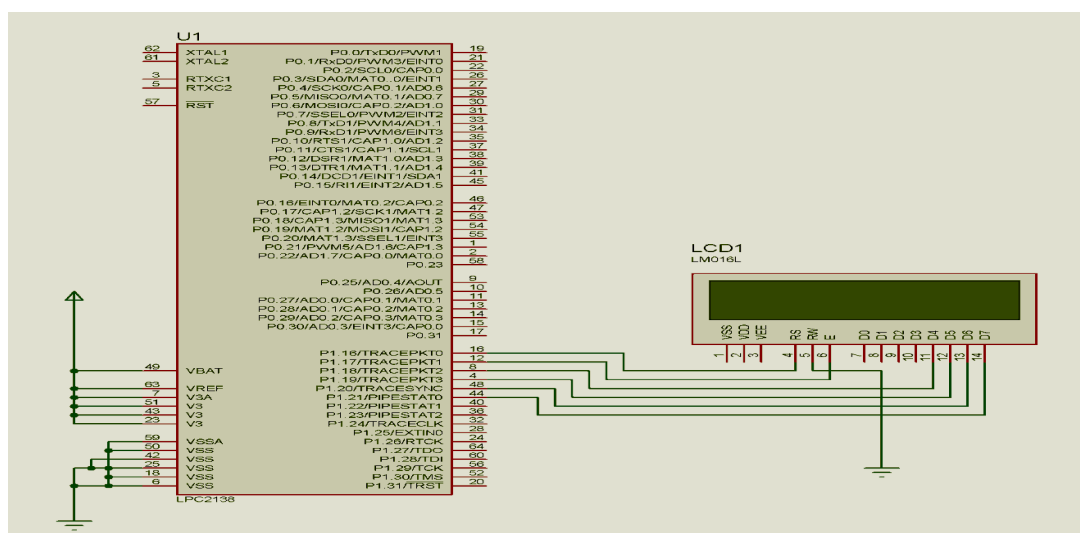
It normally consists of:

- Float (NBR type/PU type) with level arm
- Potentiometer

#### 4.1.6. Liquid Crystal Display

LCD stands for **Liquid Crystal Display**. LCD is finding wide spread use replacing LEDs (seven segment LEDs or other multi segment LEDs) because of the following reasons:

- The declining prices of LCDs.
- The ability to display numbers, characters and graphics. This is in contrast to LEDs, which are limited to numbers and a few characters.
- Incorporation of a refreshing controller into the LCD, thereby relieving the CPU of the task of refreshing the LCD. In contrast, the LED must be refreshed by the CPU to keep displaying the data.



**Fig4.** LPC2148 controller interfacing with LCD diagram

### 4.1.7. Buzzer

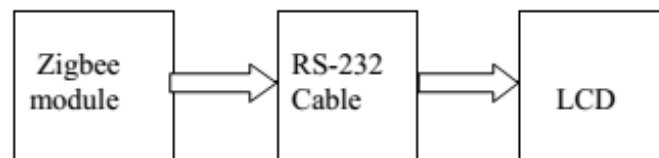
Piezo buzzer is an electronic device commonly used to produce sound. Light weight, simple construction and low price make it usable in various applications like car/truck reversing indicator, computers, call bells etc. Piezo buzzer is based on the inverse principle of piezo electricity discovered in 1880 by Jacques and Pierre Curie. It is the phenomena of generating electricity when mechanical pressure is applied to certain materials and the vice versa is also true. Such materials are called piezo electric materials

### 4.1.8. DC-Motor

Electric motors are used to efficiently used to convert electrical in to mechanical energy. magnetism is the basis of their principles operation. They use permanent magnets, electromagnets, and exploit the magnetic properties of materials in order to create these amazing machines. DC Motor has two leads. It has bidirectional motion If we apply +ve to one lead and ground to another motor will rotate in one direction, if we reverse the connection the motor will rotate in opposite direction. If we keep both leads open or both leads ground it will not rotate (but some inertia will be there). If we apply +ve voltage to both leads then braking will occurs.

## 4.2. Control Room Section

At control room section via LCD it is possible to view the current status of the Wind Turbine. The receiver ZigBee receives the data from the Wind Turbine and it is connected to PC using Serial Communication RS232. Whenever the fault occurs in the Wind Turbine according to the type of fault the comment will be sent from the PC to Wind Turbine to overcome the fault and to prevent blasting hazard.



**Fig5.** Control room section

In case of any fault occurrence like Temperature hike it is identified easily and immediately a comment will be sent from the control room to Wind turbine and the cooler will be switched on to reduce the temperature. Else if there is higher vibration immediately Turbine will be shut down and manual operation will be conducted to recover that fault. If the rotational speed is higher due to storm than immediately turbine speed will be reduced by applying electric breaks and system will be shut down to save the wind turbine blades from damages. For the case of lesser lubrication the auto lubrication process is implemented here. In this the DC motor will be switched on automatically which pumps the oil and fills the lubricant in appropriate parts of the Wind turbine.

## 5. RESULT AND CONCLUSION

The fault identification is done and the parameters are measured through the CAN interface module the monitored data is analyzed and send to LCD through RS232 cable. The location and the type of faults are analyzed before it occurs and are transmitted from wind turbine to the control centre through zigbee network . The proposed system enables the monitoring and prevention of higher vibration, rise in temperature as well as lesser lubrication of the Wind Turbine using the developed methodology to avoid blasting hazard. Thus the ZigBee Wireless communication enables the remote controlling system of all these parameters from control room through monitoring in LCD.

The effect of harsh condition and the nature of large electromechanical system are the causes of fault to be occurred in the wind turbine. It is very important perform the monitoring and fault diagnosis of wind turbine parameters.

### **REFERENCES**

- [1] FU Zhixin and Yuan Yue( 2012), ‘Condition Health Monitoring of Offshore Wind Turbine based on Wireless Sensor Network’ IEEE Transaction and IPEC.
- [2] I. Ribrant and L.B ertling, Survey of failures in Wind power systems with focus on Swedish wind power plants during 1997 - 2005, IEEE transactions on Energy conversion, vol.22, no, 1, pp. 167-173, March 2007.
- [3] Pieter, Jan-willem, Paul, Fleming and Alan (2013),’ LPV Identification of Wind Turbine Rotor Vibrational Dynamics Using Periodic Disturbance Basis Functions’ IEEE TRANSACTIONS ON control systems technology, VOL. 21, NO. 4, JULY 2013.
- [4] <http://compnetworking.about.com/od/zigbee/a/zigbee-wireless-network-standard.html>
- [5] ARM system architecture by Stephen Bo Furber
- [6] <http://www.ohioelectricmotors.com/dc-motors-in-chemical-injection-pump-applications-1107>
- [7] Spacek, Ando junior, Neto, Oliveria, Gruber (2013) ,’Management of Mechanical Vibration and Temperature in Small Wind Turbines Using Wireless Network’ IEEE Transaction on control system.

### **AUTHOR’S BIOGRAPHY**



**Ms. Spandana.K** is presently pursuing final semester M.Tech in Embedded Systems at Malla Reddy Engineering College, Hyderabad. She received B.E degree in Anurag engineering college kodad, Telangana. Her areas of interest are Embedded Systems, Micro Controllers and Microprocessors.



**Mr. Raju N.** obtained B.E Degree in 2009 from Mona College of Engineering & Technology, M.E (Embedded Systems) in 2013 from CVR College of Engineering (Autonomous). Presently he is working as Assoc. Professor, Department of ECE, Malla Reddy Engineering College, Dhullapally, Hyderabad, Telangana, India.

# VLSI DESIGN OF DATA ENCODING SCHEMES FOR LOW POWER

Badugu Chaitanya<sup>1</sup>, V.Srinivas<sup>2</sup>

<sup>1</sup>P.G Scholar, VLSI, Malla Reddy Engineering College

<sup>2</sup>Assistant Professor, Dept. of ECE, Malla Reddy Engineering College

**Abstract:** As technology improves, the power dissipated by the links of a network-on-chip (NoC) starts to compete with the power dissipated by the other elements of the communication subsystem, namely, the routers and the network interfaces (NIs). Here, we present a set of data encoding schemes to reduce the power dissipated by the links of an NoC. In this paper, the encoder in LDPC is replaced with our data encoding schemes in order to reduce the power consumption in Low Density Parity Check Techniques. Experiments carried out on both synthetic and real traffic scenarios show the effectiveness of the proposed schemes.

**Index Terms:** Data encoding, interconnection on chip, low density parity check, majority logic decoding, power analysis.

## I. DATA ENCODING TECHNIQUES

The data encoding techniques are developed to reduce the power consumption caused by the transitions in the interconnect on the chip. The data encoding techniques are based on reducing the number of transitions by considering the types of transitions in the interconnects and by considering them as discussed in the table below and also consider the transitions as different types of inversions available for us. The different types of inversions available for us are odd inversion, full inversion and even inversions. By reducing these inversions we can control the number of transitions in the interconnects which reduces the power consumption caused by these transitions in the links.

Time	Normal			Odd Inverted		
	Type I			Types II, III, and IV		
$t-1$	00, 11	00, 11, 01, 10	01, 10	00, 11	00, 11, 01, 10	01, 10
$t$	10, 01	01, 10, 00, 11	11, 00	11, 00	00, 11, 01, 10	10, 01
	T1*	T1**	T1***	Type III	Type IV	Type II
$t-1$	Type II			Type I		
$t$	01, 10			01, 10		
	10, 01			11, 00		
$t-1$	Type III			Type I		
$t$	00, 11			00, 11		
	11, 00			10, 01		
$t-1$	Type IV			Type I		
$t$	00, 11, 01, 10			00, 11, 01, 10		
	00, 11, 01, 10			01, 10, 00, 11		

By assuming these we have designed three schemes which are as follows:

### A. Scheme I

In scheme I, we focus on reducing the numbers of Type I transitions (by converting them to Types III and IV transitions) and Type II transitions (by converting them to Type I transition). The scheme compares the current data

with the previous one to decide whether odd inversion or no inversion of the current data can lead to the link power reduction.

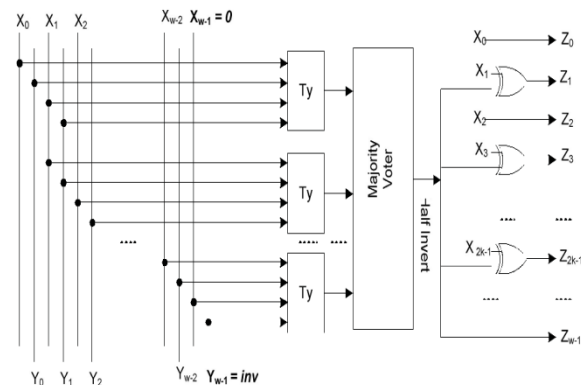
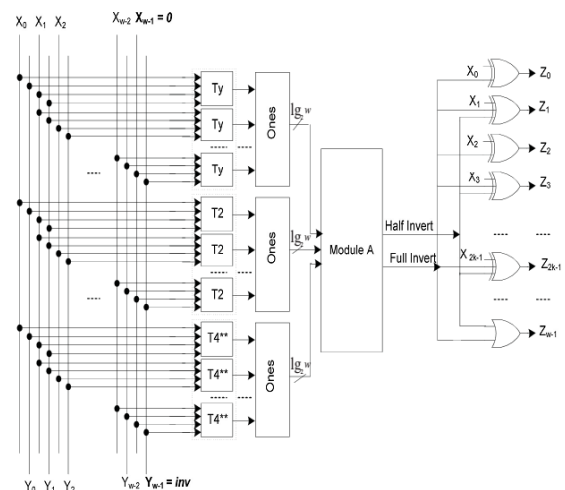


Fig. 1. Encoder architecture scheme I. (a) Circuit diagram [27]. (b) Internal view of the encoder block

### B. Scheme II

In the proposed encoding scheme II, we make use of both odd (as discussed previously) and full inversion. The full inversion operation converts Type II transitions to Type IV transitions. The scheme compares the current data with the previous one to decide whether the odd, full, or no inversion of the current data can give rise to the link power reduction.



### C. Scheme III

In the proposed encoding Scheme III, we add even inversion to Scheme II. The reason is that odd inversion converts some of Type I (T1\*\*\*) transitions to Type II transitions. As can be observed from Table II, if the flit is even inverted, the transitions indicated as T\*\* 1 / T1\*\*\* in the table are converted to Type IV/Type III transitions. Therefore, the

even inversion may reduce the link power dissipation as well. The scheme compares the current data with the previous one to decide whether odd, even, full, or no inversion of the current data can give rise to the link power reduction.

#### Low Density Parity Check Code:

The LDPC code is based on a set of one or more fundamental LDPC codes. Each of the fundamental codes is a systematic linear block code. The fundamental codes can accommodate various code rates and packet sizes.

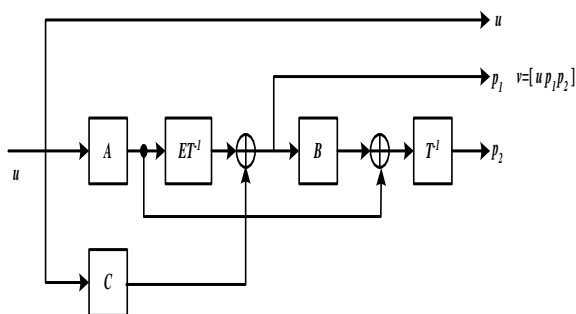
Each LDPC code in the set of LDPC codes is defined by a matrix  $H$  of size  $m$ -by- $n$ , where  $n$  is the length of the code and  $m$  is the number of parity check bits in the code. The number of systematic bits is  $k=n-m$ .

The matrix  $H$  is defined as

$$H = \begin{bmatrix} P_{0,0} & P_{0,1} & P_{0,2} & \cdots & P_{0,n_b-2} & P_{0,n_b-1} \\ P_{1,0} & P_{1,1} & P_{1,2} & \cdots & P_{1,n_b-2} & P_{1,n_b-1} \\ P_{2,0} & P_{2,1} & P_{2,2} & \cdots & P_{2,n_b-2} & P_{2,n_b-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ P_{m_b-1,0} & P_{m_b-1,1} & P_{m_b-1,2} & \cdots & P_{m_b-1,n_b-2} & P_{m_b-1,n_b-1} \end{bmatrix} = P^{H_b}$$

where  $P_{i,j}$  is one of a set of  $z$ -by- $z$  permutation matrices or a  $z$ -by- $z$  zero matrix.

The encoding of a packet at the transmitter generates parity-check bits  $p=(p_0, \dots, p_{m-1})$  based on an information block  $s=(s_0, \dots, s_{k-1})$ , and transmits the parity-check bits along with the information block. Because the current symbol set to be encoded and transmitted is contained in the transmitted codeword, the information block is also known as systematic bits. The encoder receives the information block  $s=(s_0, \dots, s_{k-1})$  and uses the matrix  $H_{bm}$  to determine the parity-check bits. The expanded matrix  $H$  is determined from the model matrix  $H_{bm}$ . Since the expanded matrix  $H$  is a binary matrix, encoding of a packet can be performed with vector or matrix operations conducted over GF.



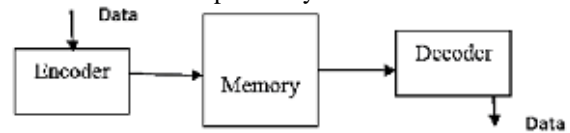
Block diagram of the encoder architecture for the block LDPC code

Using two decoding techniques majority logic decoding and majority logic decoder/detector in LDPC to reduce the delay occurring in these techniques.

#### Majority Logic Decoding:

Majority-logic decoding is a simple and effective scheme for

decoding certain classes of block codes, especially for decoding certain classes of cyclic codes. Majority logic decoding is a method to decode repetition codes, based on the assumption that the largest number of occurrences of a symbol was the transmitted symbol. It will increase the power consumption. Syndrome vector is oldest technology, which is used to detect the error in the code word. Hamming code is one of the examples of syndrome decoder.



Simple memory system schematic

#### Majority Logic detector/decoder:

The ML detector/decoder (MLDD) has been implemented using the Euclidean Geometry LDPC. G-LDPC codes there is an subclass of codes that is one step majority logic decodable (MLD) This method is very practical to generate and check all possible error combinations for codes with small words and affected by a small number of bit flips. When the size of code and the number of bit flip increases, it is difficult to exhaustively test all possible combinations. Therefore the simulations are done in two ways, the error combinations are exhaustively checked when it is feasible and in the rest of the causes the combinations are checked randomly. Since it is convenient to first describe the chosen design and also for simplicity.

## II. DATA ENCODING TECHNIQUES USING LDPC

These data encoding schemes which are discussed in the above sections are replaced by these encoding schemes instead of the encoder in the LDPC block:

- Scheme1 encoding technique in LDPC.
- Scheme2 encoding technique in LDPC.
- Scheme3 encoding technique in LDPC.

#### Scheme1 encoding technique in LDPC:

The proposed encoding architecture, which is based on the odd invert condition defined is shown in Fig. 1. We consider a link width of  $w$  bits. If no encoding is used, the body flits are grouped in  $w$  bits by the NI and are transmitted via the link. In our approach, one bit of the link is used for the inversion bit, which indicates if the flit traversing the link has been inverted or not.

This encoding technique which is scheme 1 of our encoding techniques in used in low density parity check of majority logic decoding are replaced with our encoding scheme1 technique.

#### Scheme2 encoding technique in LDPC:

The principles of this encoder are similar to those of the encoder implementing Scheme I. The proposed encoding architecture, which is based on the odd invert condition and the full invert condition, is shown in Fig. 2. Here again, the  $w$ th bit of the previously and the full invert operating

condition is shown in Fig. 2. Here again, the with bit of the previously encoded body flit is indicated with inv which defines if it was odd or full inverted (inv = 1) or left as it was (inv = 0). In this encoder, in addition to the Ty block in the Scheme I encoder, we have the T2 and T4\*\* blocks which determine if the inversion based on the transition types T2 and T4\*\* should be taken place for the link power reduction. The second stage is formed by a set of 1s blocks which count the number of 1s in their inputs. The output of these blocks has the width of log2 w. The output of the top 1s block determines the number of transitions that odd inverting of pair bits leads to the link power reduction. The middle 1s block identifies the number of transitions whose full inverting of pair bits leads to the link power reduction. Finally, the bottom 1s block specifies the number of transitions whose full inverting of pair bits leads to the increased link power.

The encoding technique discussed above which is scheme2 which is based on the logic of full inversion which is the proposed method of the odd inversion which is scheme1. The technique which is mentioned above is ldpc using shem1 now is replaced with scheme2 in the ldpc and thereby reducing some amount of power compared from scheme1.

#### Scheme3 encoding technique in LDPC:

The operating principles of this encoder are similar to those of the encoders implementing Schemes I and II. The proposed encoding architecture, which is based on the even invert condition of (28), the full invert condition of (29), and the odd invert condition of (30), is shown in Fig. 4.

The (inv = 0). The first stage of the encoder determines the transition types while the second stage is formed by a set of 1s blocks which count the number of ones in their inputs. In the first stage, we have added the Te blocks which determine if any of the transition types of T2, T1\*\*, and T1\*\*\* is detected for each pair bits of their inputs. For these transition types, the even invert action yields link power reduction. Again, we have four Ones blocks to determine the number of detected transitions for each Ty, Te, T2, T4\*\* blocks. The output of the Ones blocks are inputs for Module C. This module determines if odd, even, full, or no invert action corresponding to the outputs "10," "01," "11," or "00," respectively, should be performed.

The encoding technique discussed above which is scheme3 which is based on the logic of even inversion which is the proposed method of the full inversion which is scheme2. The technique which is mentioned above is ldpc using shem1 now is replaced with scheme2 in the ldpc and thereby reducing some amount of power compared from scheme2. Hence, we analyzed that scheme3 is the one which is used in LDPC technique.

### III. RESULTS

Power report for scheme1:

Name	Power (w)	Used	Total Available	Utilization (%)
Clocks	0.001	1	—	—
Logic	0.000	96	3312	1.0
Signals	0.000	149	—	—
I/Os	0.011	102	232	44.0
Total Quiescent Power	0.001			
Total Dynamic Power	0.012			
Total Power	0.013			

Area report for scheme1:

```
Total REAL time to Xst completion: 7.00 secs
Total CPU time to Xst completion: 6.99 secs

-->

Total memory usage is 161352 kilobytes

Number of errors : 0 ( 0 filtered)
Number of warnings : 0 ( 0 filtered)
Number of info : 0 ( 0 filtered)
```

Power report for scheme2:

Name	Power (w)	Used	Total Available	Utilization (%)
Clocks	0.001	1	—	—
Logic	0.000	96	3312	1.0
Signals	0.000	149	—	—
I/Os	0.002	102	232	44.0
Total Quiescent Power	0.001			
Total Dynamic Power	0.003			
Total Power	0.004			

Area report for scheme2:

```
Total REAL time to Xst completion: 7.00 secs
Total CPU time to Xst completion: 6.39 secs

-->

Total memory usage is 161352 kilobytes

Number of errors : 0 ( 0 filtered)
Number of warnings : 0 ( 0 filtered)
Number of info : 0 ( 0 filtered)
```

Power report for scheme3:

Name	Power (w)	Used	Total Available	Utilization (%)
Clocks	0.001	1	—	—
Logic	0.000	96	3312	1.0
Signals	0.000	149	—	—
I/Os	0.002	102	232	44.0
Total Quiescent Power	0.001			
Total Dynamic Power	0.003			
Total Power	0.004			

Area report for scheme3:

```
Total REAL time to Xet completion: 6.00 secs
Total CPU time to Xet completion: 6.01 secs

-->

Total memory usage is 161352 kilobytes

Number of errors   :    0 (    0 filtered)
Number of warnings :    0 (    0 filtered)
Number of infos    :    0 (    0 filtered)
```

#### IV. CONCLUSION

In this paper, we have presented data encoding techniques which are used in the place of encoders in LDPC which reduces the power consumption by eliminating the transitions as discussed before. Here, we analyzed the power consumption for these three schemes and compared their power and area performances.

#### REFERENCES

- [1] International Technology Roadmap for Semiconductors. (2011) .
- [2] M. S. Rahaman and M. H. Chowdhury, "Crosstalk avoidance and error correction coding for coupled RLC interconnects," in Proc. IEEE Int. Symp. Circuits Syst., May 2009, pp. 141–144.
- [3] W. Wolf, A. A. Jerraya, and G. Martin, "Multiprocessor system-on-chip MPSoC technology," IEEE Trans. Comput.-Aided Design Integr. Circuits Syst., vol. 27, no. 10, pp. 1701–1713, Oct. 2008.
- [4] Pedro Reviriego, Juan A. Maestro, and Mark F. Flanagan, "Error Detection in Majority Logic Decoding of Euclidean Geometry Low Density Parity Check (EG-LDPC) Codes," IEEE Trans. Very Large Scale Integra. (VLSI) syst., vol. 21, no. 1, pp.156-159, Jan. 2013.
- [5] Youn Sung Park, Yaoyu Tao, Zhengya Zhang, "A 1.15Gb/s Fully Parallel Nonbinary LDPC Decoder with Fine-Grained Dynamic Clock Gating," 2013 IEEE International Solid-State Circuits Conference.

## RFID Positioning Robot: an Indoor Navigation System

**Malothu mohan**

M.Tech(E.C.E)student,M.R.E.College  
mohannanda690@gmail.com

**B.Sagar**

Asst.Prof(E.C.E),M.R.E.College  
boinapallysagar@gmail.com

**Abstract:** Short form of radio frequency identification also called as dedicated short range communication (DSRC) technology. RFID is used to explain many technologies that are used in radio waves for automatic identifications. RFID works as bar code identification systems however the difference between RFID and bar code technology is that RFID does not depends on the line-of-sight reading that bar code scanning requires to work. Radio-frequency identification makes use of an RFID, tag applied to a product, animal, or person for the purpose of identification and tracking by using radio waves. Almost all RFID tags have two parts. IC for storing and processing information and antenna for receiving and transmitting the wave. This mobile car unit system using RFID technology is supported by using ARM 7 based microcontroller.

**Keywords:** ARM7, LPC2148, RFID Reader, IR Sensor, RFID Tag, 16\*2 LCD, DC Geared Motor, L293D, Kiel Development Tool, Flash magic software

---

### 1. INTRODUCTION

ARM7 (LPC2148) with 512 KB flash, 64 KB of Random Access Memory and several I/O devices. It has control commands like turn right, left speed up, down .It has to be written into the RFID tags beforehand and tags are then stuck on tracks. In this system we are interfaced IR sensor for obstacle detection. Whenever it detects any obstacle it will take the alternate path and moves to the forward direction.

RFID reader can be used to communicate with tags and sending the moving control instructions to Micro Controllable Unit, (MCU) devices. The Micro Controllable Unit module receives moving control commands and then controls mobile car. The proposed car has applications in payment systems, retailing, supply chain management, and manufacturing etc.

ARM/KEIL software can be used for developing the applications.LPC2148 development board also used for testing the built application sensors and wireless nodes are used to prove the complete setup.

### 2. BLOCK DIAGRAM

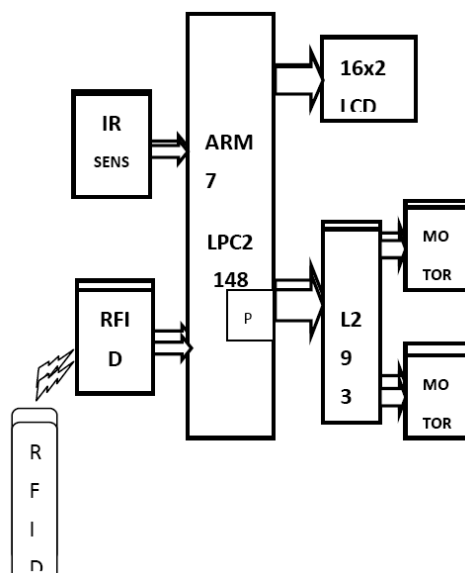


figure.1

### 3. MAIN CONCEPT

The objective of project is to control the robot movement from remote areas, so wireless communication is used to fulfill our application. There are various wireless communications exist. For that we prefer RFID Module.

The purpose of system is to enable data to transmit signal with portable device, called a tag, which is received by an RFID reader and used according to the needs of a required application. The transmitted signal through tag may provide identification or location information of the product tagged, as price, color, of purchase date etc. This technology has been used by various companies for more than a decade.

The following technology gained attention because of its ability to track maneuver object.

### 4. ALGORITHM

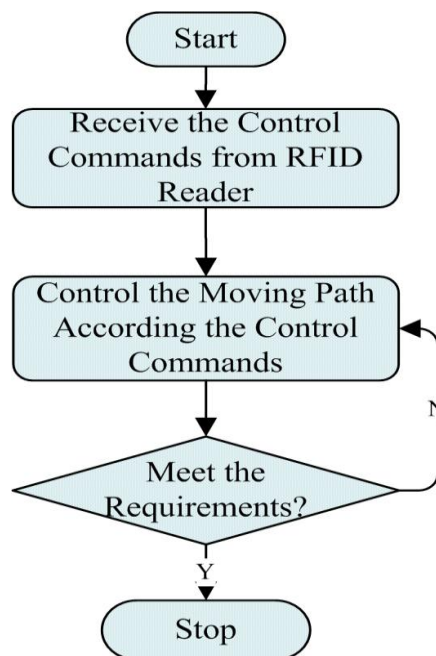


Fig 2

### 5. RFID

Radio frequency identification is well known for dedicated short range communication technology. The module used by RFID is used to describe different technologies that use radio waves to identify people or objects. This identification technology is same as the bar-code recognition system, moreover only difference among RF identification and bar code system is RF identification do not depend on the line-of-sight receiving that bar code scanning process required to work.

RF identification system has three components:

- Antenna
- Transceiver (with decoder)
- Transponder (RF tag)

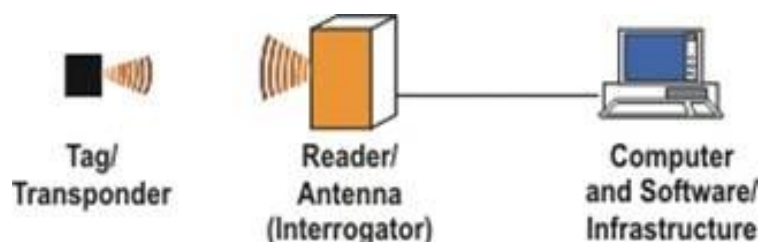


Figure.3

- The provided antenna generates RF signals to activate the tag and to read and write information to it.
- The reader produce radio frequency waves in ranges between one inch to 100 feet or more, depending upon its power produced and the radio frequency used. As RFID tag passes through the electromagnetic zone, it detects the reader's activation signal.
- The reader decodes data encoded in the provided tag's integrated circuit (silicon chip) and then signal is passed on the host computer for processing of selected data.
- The purpose of system is to enable data to transmit signal with portable device, called a tag, which is received by an RFID reader and used according to the needs of a required application. The transmitted signal through tag may provide identification or location information of the product tagged, as price, color, of purchase date etc. This technology has been used by various companies for more than a decade.
- The following technology gained attention because of its ability to track maneuver objects. As the technology is far advanced even pervasive as well as invasive used technology

### 6. IR SENSOR

Infrared radiation is an electromagnetic spectrum having wavelengths higher than visible light wavelengths, but not greater than microwaves, i.e., from  $0.75\mu\text{m}$  to  $1000\mu\text{m}$  is the infrared region of range. Infrared waves are invisible to human. The wavelength provided between  $0.75\mu\text{m}$  to  $3\mu\text{m}$  is known near infrared, and region from  $3\mu\text{m}$  to  $6\mu\text{m}$  is called mid IR and the region more than  $6\mu\text{m}$  is called far IR. (The demarcations are not rigid; regions are defined differently by many).

### 7. CONCLUSION

The design and development of RF identification based robot is successfully designed, fabricated and tested. With implementation of very low cost and flexibility in design, this kit can able reduce human introversion to greater extent. This project shows RF identification signal along with embedded system. This demonstration includes many examples provides how embedded system can make our life simple and stress free. This project includes robot that can be operated by RF identification tags, each and every tag is indicates such a direction. Its use is not limited to industries, with such modification the module can be applied to different purposes like in defense, coal mines and military applications etc.

### 8. FUTURE SCOPE

The main purpose of this project was to be able to create RFID MOBILE ROBOT.

RFID reader can collect the data from the tags and takes the direction.

- Extension of the project is we can interface GPS module can certainly be used to help with global navigation and for robot localization.
- Here we can interface the Camera to take the pictures with in a field and also interface one wireless communication module and send the data to the receiver part.
- The RFID tags can serve as a “sanity check” of sorts to check the work of the particle.

### REFERENCES

- [1] W. Gueaieb and M. S. Miah, “Experiments on a novel modular cost-effective RFID-based mobile robot navigation system,” in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Montral, Canada, October 7–10 2007, pp. 1658–1663.
- [2] J. Borenstein, H. R. Everett, L. Feng, and D. Wehe, “Mobile robot positioning: Sensors and techniques,” *Journal of Robotic Systems*, vol. 14, no. 4, pp. 231–249, April 1997.
- [3] L. R. Ojeda, G. D. Cruz, and J. Borenstein, “Currentbased slippage detection and odometry correction for mobile robots and planetary rovers,” *IEEE Transactions on Robotics*, vol. 22, no. 2, pp. 366–378, April 2006.
- [4] H. Makela and K. Koskinen, “Navigation of outdoor mobile robots using dead reckoning and visually detected landmarks,” in *Fifth International Conference on Advanced Robotics*, 1991.

- [5] T. D’Orazio, M. Ianigro, E. Stella, F. P. Lovergine, and 94 JOURNALS OF COMMUNICATIONS, VOL. 4, NO. 2, MARCH 2009 © 2009
- A. Distanto, “Mobile robot navigation by multi-sensory integration,” in *IEEE International Conference on Robotics and Automation*, vol. 2, May 1993, pp. 373–379.
- [6] S.-Y. Yi and B.-W. Choi, “Autonomous navigation of indoor mobile robots using a global ultrasonic system,” *Robotica Archive*, vol. 22, no. 4, pp. 369–374, August
- [7] 2004.
- [8] Voorhoorst FA, Overbeeke KJ, Smets GJ. Using movement parallax for 3D laparoscopy. *Med Prog Technol* 1996; 21: 211–218.
- [9] Becker H, Melzer A, Schurr MO, Buess G. 3-D video techniques in endoscopic surgery. *Endosc Surg Allied Technol* 1993; 1: 40 – 46.
- [10] Van Bergen P, Kunert W, Bessell J, Buess GF. Comparative study of two-dimensional and three-dimensional vision systems for minimally invasive surgery. *Surg Endosc* 1998; 12: 948–954.
- [11] Birkett DH, Josephs LG, Este-McDonald J. A new 3-D laparoscope in gastrointestinal surgery. *Surg Endosc* 1994; 8: 1448–1451.

#### **AUTHOR BIOGRAPHY**



**Malothu Mohan**, pursuing M.Tech in Embedded Systems from Malla reddy Engineering College [Autonomous], Mysammaguda, Hyderabad Telangana, India-500100.

## Design and Power Analysis of 8T SRAM Cell Using Charge Sharing Technique

**Lingam Rajesh**

M.TECH Mallareaddy Eng College (VLSISD) Hyderabad  
ls.rajeshce@gmail.com

**Santosh J**

Assistant Professor in ECE at MREC, Hyderabad

---

**Abstract:** *The aim of the paper is to design and analyze 8T SRAM Cell using Charge Sharing Technique where a standard 8T SRAM cell performance degrades with low power supplies. In the design the SRAM cell uses a charge sharing technique between the transistors to make SRAM more rigid against noises that occur due to low power supplies. Apart from noise reduction the read discharge power is reused. The comparison between standard 6T, 8T and 8T with charge sharing is made. It shows that less power is consumed by 8T with charge sharing than others.*

**Keywords:** 6T SRAM, 8T SRAM, CMOS, dynamic power budge, SNM

---

### 1. INTRODUCTION

Due to growing demand of portable battery operated embedded systems made a necessity for energy efficient design. As predicted 90 % of systems will be made up of memory and the memory management is need of the time. The SRAM is popular choice for embedded systems for its high speed, robustness and ability easy to manufacture. The larger SRAM cell the larger the power consumed.

A basic low power SRAM cell is designed by using cross-coupled CMOS inverters with 6 transistors giving basic 6T SRAM cell. However with technology scaling below nanometer the power dissipation of 6T SRAM becomes significant with low power supplies as due to this the gate delay is increased which reduces the frequency of operations.

An 8T SRAM cell with charge sharing technique which used at architecture level is implemented at the cell level of design.

### 2. EXISTING WORKS

#### 2.1. Basic SRAM

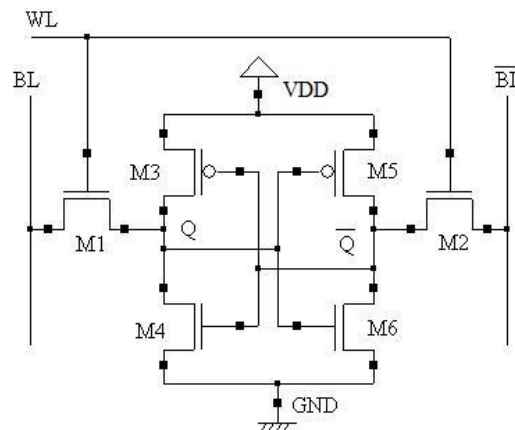
A number of SRAM cell topologies are reportable within the past years. Among these many design architectures, resistive load four-transistor (4T) SRAM bit cell, that load less 4T cell and 6 transistor(6T) SRAM cell have received attention in use, attributable to their symmetry in storing logic 'one' and logic 'zero'. The information storage within the 4T SRAM cells is ensured by the leakage current of the access NMOS transistors. Hence, they're not correct candidates for low-power applications. On the opposite hand, the information stability during a 6T SRAM cell is free lance of the outflow current. Moreover, 6T configuration exhibits a significantly higher tolerance against noise that is a very important benefit particularly within the scaled technologies wherever the noise margins area unit lowering. That's the most reason for the recognition of the 6T SRAM cell in low-power SRAM units rather than the 4T configurations in usage.

The Six Transistor SRAM cell is most generally utilized in embedded memory attributable to its quick time interval and relatively little space. 6T cell style involves complex tradeoffs between varied factors specifically abrupt scaling done on area, most sensible soft error immunity ability, high cell on current, low leakage current through off transistors, has good stability with minimum voltage &

minimum word line voltage pulse. The total CMOS 6T SRAM bit cell configuration is shown in Fig.1. Full CMOS SRAM configuration provides greater noise margin, low static or leakage power dissipation, high change speeds suitability for high density SRAM arrays. Every 6T cell contains a capability of storing one little bit of information.

The 6T SRAM cell consists of two inverters connected back to back. M5, M6 are access transistors that are controlled by the word line (WL) pulse. The cell preserves one in all its 2 possible states denoted as 0 and 1 as long as power is obtainable to the 6T bit cell.

There are 3 operations in SRAM memory cell particularly write, browse and storage operations. Read and write operations are initiated by enabling the word-line (WL). To do write operation the value to be written is applied to the bit lines and for browse operation each BL and BLB is recharged to VDD. Whenever the browse operation the zero storing node is flustered & this might flip the keep data, but the palm write operation desires that information ought to be flipped terribly simply. Historically device size has been adopted to balance the browse versus write style needs.



**Fig1. Conventional 6T SRAM**

The 6T SRAM cell design consists of two access transistors and two cross coupled CMOS inverters. Bit lines are the input/output ports of the cell with high capacitive loading. The operations READ and WRITE are conducted by these bit lines only; we will see how these are carried out. Write style needs.

A 6T SRAM cell consists of 2 cross-coupled CMOS inverters and 2 NMOS access transistors. The output (input) of the inverters construct the internal nodes of the cell A and B. Once active, the access transistors facilitate the communication of the cell internal nodes with the input/output ports of the cell. The input/output ports of the cell unit of measurement called bit lines (BL and BLB.) Bit lines unit of measurement a shared data communications medium among the cells on an {analogous the same} column in an array of cells. Consequently, they need high physical phenomenon loading. The browse and write operations unit of measurement conducted through the bit lines as we tend to are reaching to see inside the long run sections.

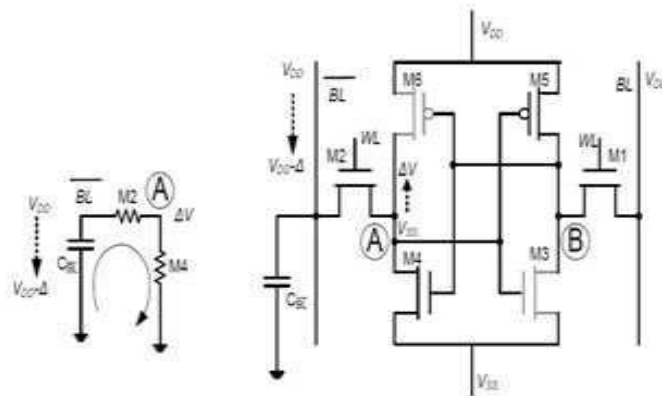
## 2.2. Read Operation

Figure 1 illustrates the operation of the cell throughout a browse access. During this design, node A carries a logic 'zero' and node B carries a logic 'one' before the cell is accessed. Thus, the transistors, M3 and M6, measure 'off' whereas M4 and M5 live 'on' and catch up on the escape current of power supply and M6. In customary vogue, the bit lines are precharged to VDD before the browse operation begins. An SRAM cell during read operation

Enabling of the word lines (WL), i.e., the gate of the access transistors starts the read operation. Because the word lines go high, power provide goes to saturation region whereas M4 operates in thermionic vacuum tube region. Attributable to the short-channel impact, this associated with power provide encompasses a linear relationship with the voltage of the node 'A'. Hence, these transistors behave sort of a resistance throughout this operation. Therefore, power provide and M4 A resistance and elevate node 'A' voltage by  $\Delta V$ .

This voltage drives the input of the convertor M5-M3. to confirm CMOS SRAM: a outline sixteen a non-destructive browse operation  $\Delta V$  is chosen such it does not trigger the M5-M3 electrical

converter and node B remains at VDD over the whole cell interval. Having a seamless voltage of VDD at the gate of M4 warrants the constant resistance assumption for M4 over the interval.



**Fig2.** An SRAM Cell during Read operation

Figure 2 shows the linear model of the bit line discharge path. Throughout this model the bit line capacitance of CBL is pre-charged to VDD. Upon the activation of power offer, CBL discharges through power offer and M4 and causes a free fall of  $\phi$  on BL. Since the gate supply voltage of M1 remains at zero volts (i.e.,  $V_{gs1} = 0V$ ), CBL cannot discharge and remains at VDD. The regardful voltage between BL and BL,  $\phi$ , is amplified using a way electronic equipment to supply the regular logic levels.

Clearly, a faster bit line discharge is achieved by reducing the resistance at intervals the discharge path. However, such enhancements return at the price of larger cell semiconductor device sizes that may not urged for prime density SRAMs.

DC analysis of the operation of the cell transistors is conventionally adopted to substantiate the soundness of the cell throughout the browse operation. as a result of it had been mentioned before, an occasional enough  $\Delta V$  ensures that the output of inverters M5-M3 remains constant at node B. to make sure a non-de structive browse operation, the voltage level  $\Delta V$  is controlled by the resistive magnitude relation of power provide and M4.

## 2.3. Write Operation

Figure 3.3 illustrates the operation of the cell inside the write operation. Throughout this figure the initial conditions of nodes A and B unit VSS and VDD, severally. Re-writing the recent data to the cell is trivial thus we have a tendency to tend to focus on propellant the knowledge of the cell.

In different words, the write operation is complete providing the voltage level on node A and B becomes VDD and VSS, severally. The activation of the word line cannot cause a spare voltage increase on node A to trigger the CMOS inverter M5-M3 if every bit lines unit of measurement pre-charged to VDD.

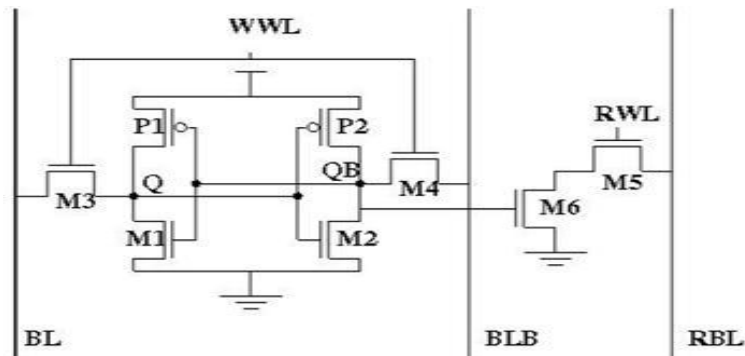
Therefore, the write operation is conducted by reducing the bit line associated with node B, BL, to a sufficiently low voltage (e.g., VSS.) This operation forms a possible divider comprising of M5 and power offer at the beginning of the operation.

To assess the voltage that appears at node B upon activation of the word lines in write operation,  $\Delta V$ . A sufficiently low  $\Delta V$  triggers the convertor M6-M4 that lands up in charging up node A to VDD. Since node a drives the convertor M5-M3, node B is force all the approach right down to VSS through power offer and M5 turns off. Hence, the logic state of the cell is changed. The word line becomes inactive once the completion of the operation. A in write operation ar usually bonded by choosing an accurate PR. A lower PR lands up during a lower  $\Delta V$ , and a lower  $\Delta V$  is expounded to higher drive at the input of convertor M6-M4.

## 3. 8T SRAM CELL

It is noteworthy that for associate SRAM cell, the specified form of operation is commonly set with the correct choice of the bit line voltage. However, this involves additional edge circuits like bit line precharge circuits and writes drivers to create positive correct bit line voltage setting before any

operation. At low provide voltages due the soundness limitations of 6T SRAM cell we tend to use 8T SRAM cell for quick transmission applications. it's like 6T SRAM cell with a scan decoupled path that consists of M5 and M6 transistors. Allow us to see the operating of 8T SRAM style.



**Fig3. 8T SRAM Cell**

### 3.1. Write Operation

The write operation of 8T SRAM cell is same as to the conventional 6T SRAM. The write operation in 8T SRAM is carried out as shown and discussed below

### 3.2. Write '0' Operation

For writing '0', the bit line has to give zero volts and VDD to the bit line (BLbar). And write word line is asserted which makes both the transistors M3 and M4 ON. Hence the value in the bit line is stored at Q. Hence '0' is stored at Q.

### 3.3. Write '1' Operation

Likewise writing '1' is also carried in the likely same. The bit line has to give a value VDD and bit line bar is given a value 0 volts. As WWL is enabled for write operation, the values in bit lines are store at respective nodes that is at Q will have value logical '1' and logical '0' at Qbar. There is no change in the write operation when compared with the basic SRAM operation.

### 3.4. Read Operation

The read operation is initiated by pre-charging the read bit line to VDD which is required in the conventional one.

### 3.5. Read '0' Operation

Read word line (RWL) drives the access transistor M5 ON. If the value stored at Q is '0' then transistor M6 will be ON and RBL is connected to ground directly through M5&M6 transistors discharges. This implies that the value stored at Q in the SRAM is zero.

### 3.6. Read '1' Operation

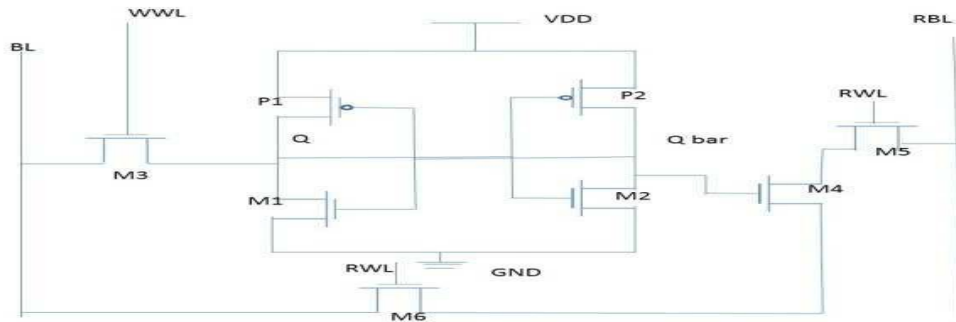
If the value stored at Q is '1', due to M6 transistor will be OFF and there is no discharge path for RBL, and the value in RBL is VDD which shows that value stored at Q is '1'. The circuit diagram of 8T SRAM shown in the figure 4.

The disadvantages in 6T SRAM are minimized in 8T SRAM, even though the transistor count increased the power consumption. The circuit diagram and operations of conventional and 8T SRAM are discussed in this chapter. The SRAM with charge sharing concept is discussed next.

## 4. PROPOSED SRAM DESIGN

Here the proposed SRAM design is using the concept of charge sharing of 10T SRAM design. But the difference is that the design is done with less number of transistors when compared to the above 10T SRAM which also decreases the area of the design and in the proposed design we also reduced the power consumption when compared with the previous design.

The proposed SRAM consists of a single ended 7T bit cell which has one bit line (BL) for write operation and one Read Bit cell for read operation.



During the write operation WWL was enabled and RWL was disabled (i.e  $RWL=0$ ) so M4 M5 M6 are in the off state. The cell acts like single ended 5T SRAM Cell and writes the Bit line data into the cross coupled inverter pair P1 M3 and P2 M2.

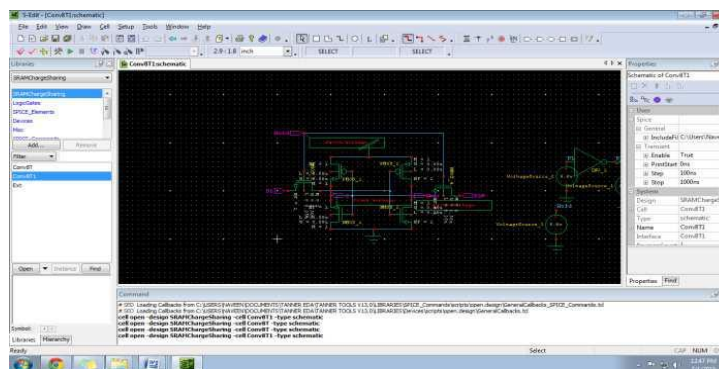
During the read operation the bit line disconnected from inverter pair because of  $WWL=0$  during read phase and RWL was enabled so M6 M5 will be in the ON state. For read operation here we are using separate bit line called RBL instead of using same BL. SO during read operation RBL was pre-recharged.

**Read '0':** In reading '0' M4 was ON state, so RBL has a discharging path from M4 M5 and M6, the M6 will acts like a charge sharing network, instead of discharging the charge to the ground M6 will charge the bit line (BL) so there will no loss of power to the ground.

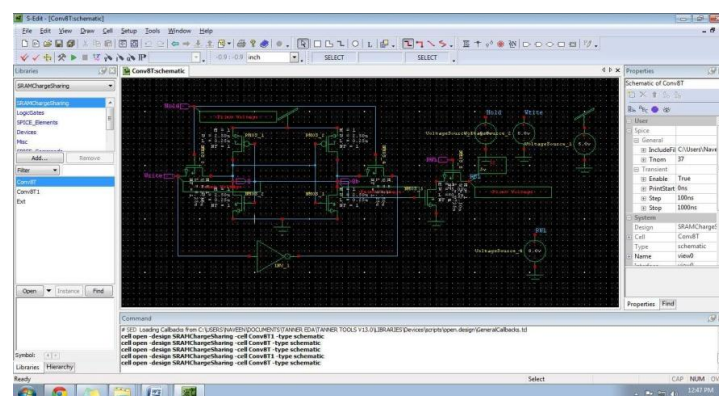
**Read '1':** In reading '1' M4 was OFF state so there will be no discharging path for RBL to discharge maintains the charge and reads the '1'

## 5. SIMULATION AND RESULTS

These SRAM designs are designed and simulated using S-edit and T-Spice using TSMC018 technology in Tanner Tools 13.0. The Conventional 8T SRAM, and Proposed SRAM are simulated there power dissipations are compared and shown in below:



**Fig5. S-edit Design of 6T SRAM Cell**



**Fig6. S-edit Design of 8T SRAM Cell**

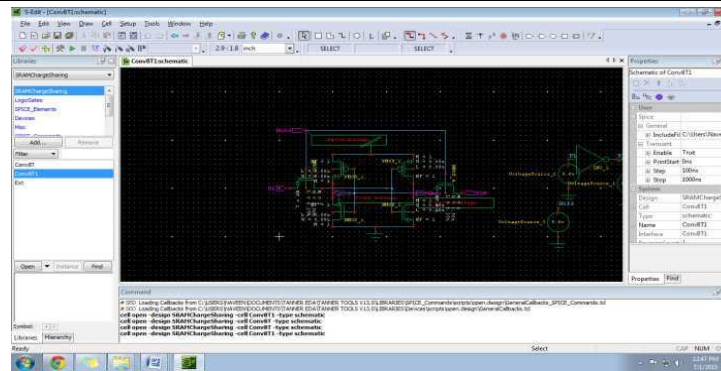


Fig7. S-edit Design of Proposed SRAM Cell

Circuit	Power Consumption
Conventional 6T SRAM	5.517330e-004 watts
8T SRAM Cell	9.461892e-010 watts
Proposed SRAM Cell	5.806951e-010 watts

## 6. CONCLUSION

Here we have used the designs of SRAM of charge sharing and analyzed the design for minimum area and power as well. The design technique will reduce power and maintain at good stability levels.

## 7. FUTURE SCOPE

In future we can extend this multi bit SRAM Cells and overall bit line discharge losses can be reduced by using adiabatic logics.

## REFERENCES

- [1] Kevin Z. Embedded Memories for Nano-Scale VLSIs. Springer Publishing Company, Incorporated. 2009; 400.
- [2] Brown AR, Roy G, Asenov A. Poly-Si-Gate- Related Variability in Decananometer MOSFETs with Conventional Architecture. Electron Devices. IEEE Transactions on 2007; 54(11): 3056-3063.
- [3] Bo Z et al. A Sub-200mV 6T SRAM in 0.13um CMOS.inSolid-State Circuits Conference. ISSCC 2007. Digest of Technical Papers. IEEE International. 2007.
- [4] Cheng B, Roy S, Roy G, Brown A, Asenov A. Impact of Random Dopant Fluctuation on Bulk CMOS 6-T SRAM Scaling. inSolid-State Device Research Conference. ESSDERC 2006.Proceeding of the 36<sup>th</sup> European. 2006.
- [5] Singh J, Simon Hollis DKR, Mohanty SP. A single ended 6T SRAM cell design for ultra-low-voltage applications. IEICE Electronics Express 2008; 5(18): 750-755.
- [6] Mizuno H, Nagano T. Driving source-line cell architecture for sub-1-V high-speed low-power applications. Solid-State Circuits. IEEE Journal of 1996; 31(4): 552-557.
- [7] Takeda K. et al. A read-static-noise-margin-free SRAM cell for low-VDD and high-speed applications. Solid-State Circuits, IEEE Journal of 2006; 41(1): 113-121.
- [8] Chang L. et al. An 8T-SRAM for Variability Tolerance and Low-Voltage Operation in High-Performance Caches. Solid-State Circuits, IEEE Journal of 2008; 43(4): 956-963.
- [9] Tae-Hyoung K. et al. A High-Density Subthreshold SRAM with Data-Independent Bitline Leakage and Virtual Ground Replica Scheme.inSolid-State Circuits Conference. ISSCC 2007. Digest of Technical Papers. IEEE International. 2007.
- [10] Wang X, Roy S, Asenov A. Impact of Strain on the Performance of high-k/metal replacement gate MOSFETs. Proc. 10<sup>th</sup> Ultimate Integration on Silicon (ULIS 2009), 2009.

**AUTHOR'S BIOGRAPHY**



**Lingam Rajesh**, Completed B.TECH in SREE Rama Eng College (ECE) (JNTU-A), M.TECH Mallareaddy Eng College (VLSISD) Hyderabad, Area of Interest: VLSI System Design.



**Santosh J**, Assistant Professor in ECE At MREC, Completed B.TECH: CVR College Of Eng. JNTU-H, M.TECH: VLSI System Design, SDES, JNTU-H Area of Interest VLSI System design Reconfigurable Architecture & Design of Fault Tolerant Systems.



**Department of  
Computer Science and  
Engineering**



## Incorporation of Third Party Auditor for Safe and Secure Data Storage in Cloud Computing



**Sanjeeva Polepaka**  
Associate Professor,  
Department of CSE,  
Malla Reddy Engineering College.



**Mr. Rajeshwarrao kodipaka**  
Assistant Professor,  
Department of CSE,  
Malla Reddy Engineering College.

### Abstract:

Cloud Computing is the long dreamed vision of computing as a utility, where users can remotely store their data into the cloud so as to enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources. By data outsourcing, users can be relieved from the burden of local data storage and maintenance. However, the fact that users no longer have physical possession of the possibly large size of outsourced data makes the data integrity protection in Cloud Computing a very challenging and potentially formidable task, especially for users with constrained computing resources and capabilities.

Thus, enabling public auditability for cloud data storage security is of critical importance so that users can resort to an external audit party to check the integrity of outsourced data when needed. To securely introduce an effective third party auditor (TPA), the following two fundamental requirements have to be met: 1) TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data, and introduce no additional on-line burden to the cloud user; 2) The third party auditing process should bring in no new vulnerabilities towards user data privacy.

### Introduction of domain:

Cloud Computing is one of the best choice for Small and Medium Sized Entrepreneurs' in the world. This has been used widely to cater the needs of organizations with different software applications through SAAS operations of cloud. Cloud has gifted the organizations the platform as a service at most economical rates. The economical rates and swift services has given rise to the cloud computing fame.

But the recent privacy issues and security challenges have degraded the cloud computing marketing. North Bridge (2013)<sup>1</sup>. Cloud market has fallen down because of the data storage security problems, quality of services and privacy preserving issues in cloud computing. The research scholars have done enough research to admeasure the problems but still lot of loop holes and missing links are notified in cloud computing to arrest the problems. There is a great need to revive the situation by doing the research work to ensure data storage security in cloud computing. There are many research works tried to ensure the data security in cloud computing servers with different techniques like digital encryption, fuzzy key word search etc. The ultimate solution for data security and data storages should be given to arrest the problems due to Byzantine failure, malicious data modification attack and server colluding attacks. Ricardo Puttini (2013)

### Aim:

The aim of the project is to utilize the public key based homomorphic authenticator and uniquely integrate it with random mask technique to achieve a privacy-preserving public auditing system for cloud data storage security. The aim is to support efficient handling of multiple auditing tasks in cloud computing. The aim is to explore the technique of bilinear aggregate signature to extend the main result into a multi-user setting, where TPA can perform multiple auditing tasks simultaneously. The aim includes to provide an extensive secured and highly performing system in cloud computing with an effective third party auditor.

### Objectives:

- The main objective of the project is to develop a highly secured and good performance system with third party auditor for data storage in cloud computing.

- The objective is to design and develop a cloud computing application with extensive public data storage with third party auditing system.
- The objective is to demonstrate the third party auditing system is not effecting the performance of the cloud servers and storage operations.
- The implementation of third party auditing with high performance for the public storage in cloud computing is one of the objective.
- The objective is to develop a cloud computing application with high security in storing public data.

## Background:

Cloud computing is one of the most popular technological aspects in computational world. The most important thing in cloud computing is data storage with proper security measures. At the same time cloud computing is facing the data leakage and insecurity. To avoid these challenges in cloud computing a strong secured auditing system is needed to identify the users and their data. The increased security has to be implemented, at the same time the performance should not be degraded. To give proper solution the present project has been developed to provide the higher data security and increased performance while implementing the public data auditing in the cloud computing.

## Existing System:

Though Cloud computing is enriched with the special features like agile, reliable, cost effective and measurable delivery of data. Cloud computing has excellent delivery models with identification, Authentication, Authorization, Confidentiality, Integrity, Non-repudiation and availability as information security requirements. It is blamed by the provision of untrustworthy servers located at remote and un-known locations. This feature has become an issue to store sensitive data and confidential data at untrusted servers at unknown remote locations and caused the heavy computation overhead. Frequently the cloud computing has encountered the security issues such as SQL Injection Attacks, Cross Site Scripting Attacks, Man in the Middle Attacks, Network Level Security Issues, DNS attacks, Sniffer Attacks, BGP Prefix Hijacking and issue of reused IP Address. Apart from these attacks Application level security issues with security concerns with the Hypervisor, Denial of Service Attacks, Cookie Poisoning, Hidden Field Manipulation.

Captcha Breaking, Google Hacking and Distributed DOS Attacks are traced in cloud computing. To overcome all these attacks and computation overhead, Priyadarshini has suggested a cloud storage mechanism with Kerberos authentication and utilization of multi-clouds. Kerberos secure mechanism is very strong to protect weak link from opponent. It is very reliable with a system to back up. The Kerberos mechanism is transparent which can be revealed with the password authentication. It is measurable to support large number of clients and service providers. Kerberos authentication mechanism is rich with Authentication Servers and Ticket Granting Servers which can be incorporated in six steps. Above all these measures Priyadarshini has suggested multi-clouds environment to implement Kerberos Realms and Multiple Kerberos to obtain increased security in cloud computing storage. - A. Priyadarshini [2013].

## Drawbacks of Existing system:

The Department of Education, The United States of America has incorporated a Privacy Technical Assistance Center for learning data privacy, confidentiality and security. Data security threats are classified into two types. These are technical and non-technical. A comprehensive privacy and data security plan has been initiated by PTAC to reduce the vulnerability to security threats. The technical data security threats are regarded as non-existent security architecture, Un-Patched client side Software and Application, Phishing and targeted Attacks which is known as Spear Phishing, Internet Websites, Poor configuration Management, data storage in mobile devices and transfer of data from mobile devices, Cloud Computing, Data from removable media, Botnets [ the series of networks of compromised computers] and Zero-day attacks. The non technical cyber security threats are regarded as Insider attacks, Poor passwords, Physical Security, Insufficient backup and recovery, improper destruction, Social Engineering and social media. PTAC has successfully inculcate the mitigation for all above mentioned threats and suggested the students to follow consistent implementation of the security plan drastically eradicate the cyber threats and establish the security. – PTAC – IB [2011].

## Proposed System:

The proposed system is developed with the public key based homomorphic authenticator and uniquely integrate it with random mask technique to achieve a privacy-preserving public auditing system for cloud data storage.

To support efficient handling of multiple auditing tasks, we further explore the technique of bilinear aggregate signature to extend our main result into a multi-user setting, where Third Party Auditing can perform multiple auditing tasks simultaneously.

Extensive security and performance analysis shows the proposed schemes are provably secure and highly efficient. We also show how to extend our main scheme to support batch auditing for Third Party Auditing upon delegations from multi-users.

### Third Party Auditing:

Third party auditing system is one security mechanism to store the data into the cloud servers. Third party auditing system is arranged by the cloud consumers as well as the cloud service providers. The cloud service providers would safe guard the server with the help of Third party auditing.

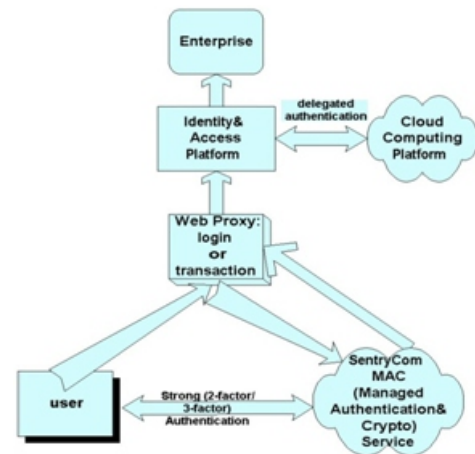
Third party auditing will scan the data whatever stored by the cloud consumer then it will decide whether the data can be stored in the server or not. If the data consists of malicious data then the TPA will delete the data from the server. Predominantly the data storage will be audited by the TPA system and allow the data to place in the server. It can be accessed by the authorized user of the Cloud Consumers.

### Algorithm:

A public auditing scheme consists of four algorithms. These are KeyGen, SigGen, GenProof and VerifyProof.

- KeyGen: key generation algorithm that is run by the user to setup the scheme
- SigGen: used by the user to generate verification metadata, which may consist of MAC, signatures or other information used for auditing
- GenProof: run by the cloud server to generate a proof of data storage correctness
- VerifyProof: run by the TPA to audit the proof from the cloud server.

### Flowchart:



### Modules:

#### 1. Privacy-Preserving Public Auditing Module:

Homomorphic authenticators are unforgeable verification metadata generated from individual data blocks, which can be securely aggregated in such a way to assure an auditor that a linear combination of data blocks is correctly computed by verifying only the aggregated authenticator.

Overview to achieve privacy-preserving public auditing, we propose to uniquely integrate the homomorphic authenticator with random mask technique. In our protocol, the linear combination of sampled blocks in the server's response is masked with randomness generated by a pseudo random function (PRF).

The proposed scheme is as follows:

- Setup Phase
- Audit Phase

#### 2. Batch Auditing Module:

With the establishment of privacy-preserving public auditing in Cloud Computing, TPA may concurrently handle multiple auditing delegations upon different users' requests. The individual auditing of these tasks for TPA can be tedious and very inefficient. Batch auditing not only allows TPA to perform the multiple auditing tasks simultaneously, but also greatly reduces the computation cost on the TPA side.

### 3.Data DynamicsModule:

Hence, supporting data dynamics for privacy-preserving public risk auditing is also of paramount importance. Now we show how our main scheme can be adapted to build upon the existing work to support data dynamics, including block level operations of modification, deletion and insertion. We can adopt this technique in our design to achieve privacy-preserving public risk auditing with support of data dynamics.

### The Functionality of the Project:

The project is designed and developed to demonstrate the cloud computing data storage security with the help of Third Party Auditing system. The cloud computing is declining with the security loopholes. These loop holes of security system can be arrested by employing the third party auditing system. The cloud computer is occupied by the cloud consumers. Sometimes the cloud servers are losing the data integrity and confidentiality because of the cloud consumer's data stealing mechanism. The data stored by the cloud consumers should be streamlined by employing a third party auditor who is most amicable and trustworthy for cloud consumers as well as cloud service providers. The cloud consumer's data should be scanned and verified by the Third Party Auditor, then it should be stored in the cloud servers. This data streamlining mechanism will avoid the malicious data storage in the cloud computing servers. When the data is accessed by the cloud users who are created and permitted by the consumers should also approved by the third party auditor. In this present project the Third Party Auditing system will keep an eye on the data accessing items and keeps the track of them. In this way the data storage mechanism will be audited by the Third Party auditor and safely enable the cloud consumer to store the data into the cloud servers.

### Conclusion :

The project is designed with a novel mechanism to audit the data storage with third party auditing system to filter the data. The data storage options have to be streamlined with proper monitoring and auditing system. To ensure the safe and secure data storage mechanism in cloud computing the proposed third party auditing system is developed in .Net technologies in simulation mode. This has been tested with proper test cases to ensure the mechanism working properly. The test results have been evaluated.

Public auditability also allows clients to delegate the integrity verification tasks to TPA while they themselves can be unreliable or not be able to commit necessary computation resources performing continuous verifications. Another major concern is how to construct verification protocols that can accommodate dynamic data files. In this paper, we explored the problem of providing simultaneous public auditability and data dynamics for remote data integrity check in Cloud Computing.

### References:

- 1.P. Mell and T. Grance, "Draft NIST working definition of cloud computing," Referenced on June. 3rd, 2009 Online at <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>, 2009.
- 2.M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A Berkeley view of cloud computing," University of California, Berkeley, Tech. Rep. UCB-EECS-2009-28, Feb 2009.
- 3.N. Gohring, "Amazon's s3 down for several hours," Online at <http://www.pcworld.com/businesscenter/article/142549/amazons-s3-down-for-several-hours.html>, 2008.
- 4.Amazon.com, "Amazon s3 availability event: July 20, 2008," Online at <http://status.aws.amazon.com/s3-20080720.html>, July 2008.
- 5.S. Wilson, "Appengine outage," Online at <http://www.cio-weblog.com/50226711/appengine-outage.php>, June 2008.
- 6.B. Krebs, "Payment Processor Breach May Be Largest Ever," Online at <http://voices.washingtonpost.com/securityfix/2009/01/payment-processor-breach-may-b.html>, Jan. 2009.
- 7.G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," Cryptology ePrintArchive, Report 2007/202, 2007, <http://eprint.iacr.org/>.
- 8.M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," Cryptology ePrint Archive, Report 2008/186, 2008, <http://eprint.iacr.org/>.

- 9.Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in Proc. of ESORICS'09, Saint Malo, France, Sep. 2009.
- 10.Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, <http://www.cloudsecurityalliance.org>.
- 11.H. Shacham and B. Waters, "Compact proofs of retrievability," in Proc. of Asiacrypt 2008, vol. 5350, Dec 2008, pp. 90–107.
- 12.Juels and J. Burton S. Kaliski, "Pors: Proofs of retrievability for large files," in Proc. of CCS'07, Alexandria, VA, October 2007, pp. 584–597.
- 13.M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in Proc. of HotOS'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 1–6.
- 14.104th United States Congress, "Health Insurance Portability and Accountability Act of 1996 (HIPPA)," Online at <http://aspe.hhs.gov/admsimp/pl104191.htm>, 1996, last access: July 16, 2009.
- 15.D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and verifiably encrypted signatures from bilinear maps," in Proc. of Eurocrypt 2003, volume 2656 of LNCS. Springer-Verlag, 2003, pp. 416–432.
- 16.T. S. J. Schwarz and E. L. Miller, "Store, Forget, and Check: Using Algebraic Signatures to Check Remotely Administered Storage," Proc. of ICDCS '06, pp. 12–12, 2006.
- 17.M. Lillibridge, S. Elnikety, A. Birrell, M. Burrows, and M. Isard, "A Cooperative Internet Backup Scheme," Proc. of the 2003 USENIX Annual Technical Conference (General Track), pp. 29–41, 2003.
- 18.K. D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," Cryptology ePrint Archive, Report 2008/489, 2008, <http://eprint.iacr.org/>.
- 19.L. Carter and M. Wegman, "Universal Hash Functions," Journal of Computer and System Sciences, vol. 18, no. 2, pp. 143–154, 1979.
- 20.J. Hendricks, G. Ganger, and M. Reiter, "Verifying Distributed Erasure-coded Data," Proc. 26th ACM Symposium on Principles of Distributed Computing, pp. 139–146, 2007.
- 21.J. S. Plank and Y. Ding, "Note: Correction to the 1997 Tutorial on Reed-Solomon Coding," University of Tennessee, Tech. Rep. CS-03-504, 2003.
- 22.Q. Wang, K. Ren, W. Lou, and Y. Zhang, "Dependable and Secure Sensor Data Storage with Dynamic Integrity Assurance," Proc. of IEEE INFOCOM, 2009.
- 23.R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession," Proc. of ICDCS '08, pp. 411–420, 2008.
- 24.D. L. G. Filho and P. S. L. M. Barreto, "Demonstrating Data Possession and Uncheatable Data Transfer," Cryptology ePrint Archive, Report 2006/150, 2006, <http://eprint.iacr.org/>.
- 25.M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to Keep Online Storage Services Honest," Proc. 11th USENIX Workshop on Hot Topics in Operating Systems (HOTOS '07), pp. 1–6, 2007.

## Measurable, Safe and Secure Data Management for Sensitive User in Cloud Computing



**Mr. Rajeshwarrao kodipaka**

Assistant Professor,  
Department of CSE,  
Malla Reddy Engineering College.



**Sanjeeva Polepaka**

Associate Professor,  
Department of CSE,  
Malla Reddy Engineering College.

### Abstract:

Several Universities have adopted the paper based system where the students have to submit their work in the paper only during the specific office hours. Even though this was effective at the beginning, this paper based system faced a lot of problems as the number of students started increasingly. Marketing of the students work has become very difficult to the supervisors in this paper based system. To solve this problem online project assessment system is uploaded in cloud computing servers with attribute based algorithms. This online project assessment will not only be useful to the student and the supervisor but also will be useful for the administrator as the maintenance of the system is very easy and also flexible. This online project assessment system will also be very scalable and therefore a new student can be added very easily.

### Keyword:

cloud computing, attribute based access rights, Secure Computing.

### Introduction:

Many Universities are still using paper based method for their students to submit their assignment and their project. In this paper based system the students will make a paper copy of their work and they will submit it before the deadline in the concerned department. This process of paper based submission takes a lot of time for the students and also for the supervisors to mark them. The process of assessment submission, feedback process will be fastened when these assignments are submitted online. In this online submission the assignments will be submitted in the electronic format and the various feedback methods will be used to provide the feedback through online itself. Various tools are present in the market to provide the online submission facilities however the selection of

the best tool will provide various benefits to the online submission of the assessments. This online submission will help the students to submit their work to their tutors with the help of the internet.

### Objectives:

- \* To investigate the alternative solutions for online submission without any time delay and duplicated work.
- \* To design and develop a web based application to facilitate the students project work and assignments submission online without duplication of data.
- \* To explore the possible software application methodology to attain the requirement specification of the concept.
- \* To examine the application utility with user interaction i.e student assignment submission and university verification and marking with relevant comments.
- \* To compare the performance of the application functionality according to the functional requirements of the concept.

### Proposed Solution:

As the number of students is increasing the supervisor will face problem in marking their assessments if it is a paper based submission. To overcome this, an efficient project assessment system like online submissions system is need which can solve this problem. Both the students and the supervisor can feel it easy in working with this system. The time required for submitting the assessment will be less and the time required for marking the student assessment will also be less. The developed online project assessment system will also be reliable, scalable, flexible, and efficient to all its users. This online project assessment system website will be having three users namely

- \* Admin
- \* Supervisor
- \* Students

Based on each user there will be given certain roles and responsibilities. The student will have the chance to upload his work and can also check his result. The supervisor will be having the chance to view the submitted work of the student and also to give the feedback to the student work. The administrator on the other hand will be having the right to add the students and the supervisor. The administrator will allocate a certain number of students to a supervisor and he will also be having the rights to manage all the users present in the project assessment system. The admin must act as the main person who is responsible for maintaining the other users of the online project assessment system. The admin should assign the access rights to view his own work submitted to the guide. And also restrict him to access other works submitted to the guide. The supervisor should check the assessments that are submitted by the students and must mark those works. The student on the other hand should be capable of submitting their work and check his result. In addition to all these requirements the administrator is required to instruct the supervisor to mark the assessments of the students by sending those emails. This proposed system will be developed with the help of visual studio .net and sql server 2008.

## Literature Survey:

### Online college portals:

Universities are rich with large number of students. The assignment, projects, dissertation reports submission in a specific date and time is become very difficult for the administrative staff as well as to the students also -ict support [2011]. The article from Center for Teaching Excellence[2009] describes the specific web applications developed for online learning have attained the goal of formative and summative assessment of the students and their learning capabilities. The online web application for learning has developed to achieve the formative3 assessment techniques online concept mapping with the inbuilt tool. The concept mapping has been developed from the ability to draw reasonable inferences from observations, synthesize and integrate information and ideas and learning concepts and theories in the subject area.

### Attribute Based Access Rights:

Susan Hohenberger (2014) described the attribute is a term used in database. The attribute can be distinguished as the schema of the database. It may be a column or row of the table.

Attribute based concepts are generally used in the data mining concepts. When the data is huge and searching time is remarkably high, this attribute based concepts can be used.

### Online Security:

Richard N. Katz and Associates [2011] states that online security is one of the most important aspects for the university portals. Portals are basically divided into three types. These are vertical portal horizontal portals and University Enterprise portals. Portals are basically interactive.

### HTML Injection:

HTML Injection is an attack for web based applications. This is normally affecting different web sites where security loop holes are there. This will be called as Hypertext Markup Language injection. –Microsoft Corporation has identified that this is sometimes called virtual defacement attack.

## PRIVACY ISSUES AND CHALLENGES:

The benefits extracted from cloud computing are dominating the other fields from 2008. The recent privacy issues and challenges found in the cloud computing have constructed the barrier to the exponential growth of cloud computing Jaydip Sen [2013].

### Research methodology:

Methodology will guide the student to expertize the topics related to the research work. The relevant topics identification, exploration and research study will be possible through the methodology. The following research methods are followed to develop the research report. The present project is a web based application to facilitate the students to upload the submissions of their academic fulfillment within the prescribed time. The present application development has been done on the basis of different knowledge material acquisition from different sources. The research work has been done by doing different research methods and interacting with the people and for knowledge exploration.

## Functional Requirements:

The functional requirements can be defined with the functionality of the project as well as the module description. The main modules of the project are

- \* Administrator
- \* The supervisor
- \* The user

## Module Description:

**Administrator:** This module is designed to describe the details of the university programs. This module is responsible for defining the start and end dates of the assignments, dissertations and other academic programs. **The supervisor:** This module is designed to view the details of the assignments, dissertations and other academic works of the student. This module is designed to insert the feedback in the prescribed window of the module. The feedback will be sent to the student as well as it will be displayed against the students hall ticket no. The supervisor should be facilitated to write the feedback on the uploaded work and store it in the database. The feedback should also be displayed to the user. **The user will be considered as student.** Every student will be identified with the hall ticket no. The student will send the project work, assignment work or dissertation work to the university through the upload. The system will stop the duplicate uploads from the students. The user will have to view the status of the uploaded work. The user should be facilitated to view the details of the feedback of the project or uploaded work.

## Design of Data Flow diagram the project:



## Design Patterns:

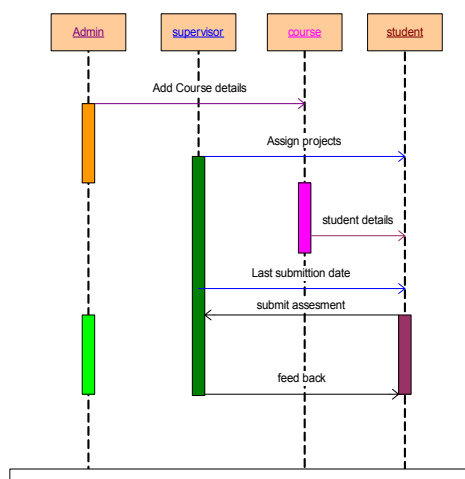
Design patterns in .Net technologies are reusable solution to common software design problems - Christopher Alexander[2013]. Design patterns will be used when a problem is encountered in any application again and again with specific context. Every pattern illustrates the specific problem and its solution in specific environment. Design patterns are used by the developers according to the design of the application. The design patterns will decide the class or object of the application programming. In this application the design patterns are defined and used in connecting the database with the user interface screens with the help of ADO.Net classes and statements.

The application is rich with 'N' tire architecture. The code written for attribute based access rights for each user is illustrated in C#.Net. The project is rich with creational design patterns, structural design patterns. The creational design pattern is used for connecting the user interface screens with the specific tables with factory method, abstract method and builder pattern. N tire architecture is configured with adapter design patterns and composite design patterns. In this system Behavioral design patterns are also used to illustrate the attribute based access rights. In this the chain of responsibility, interpreter, memento, observer and template method design patterns are used. Christopher Alexander[2013]

## Code behind the technique:

Code behind the technique is done with C#.Net and ADO. Net. In this project Stored procedures are used for well connectivity of the tables. The user interface screens are developed using ASP.Net and the tables are created in SQL Server 2005. To connect the tables with the relevant user interface screens ADO.Net classes and statements have been used. The attribute based access rights have been designed in C#.Net. The code behind the technique is predominantly used with the help of ADO.net and C#.Net to give full functionality of the project.

Sequence Diagram



## TEST REPORTS:

The test cases mentioned above are properly executed and given pass results. The consolidated test results are said to be good and passed all the tests.

## CONSOLIDATED TEST RESULTS:

The test cases are given by the test engineer and the test reports are generated as soon as all the test cases are executed positively the summary of the test reports are consolidated module wise. The consolidated test results have revealed that the project has been developed according to the goals and objectives of the project. All the test cases have been passed and the application is proven to be correct.

## Critical Points :

The application should enable the administrator to give access rights to authorized faculty to define the assignments, projects and other academic works. The application should enable the faculty to define the assignments, projects and other works with date of submission. The application should fix the deadline for project submission. The project submitted after submission data and time should not be allowed. The project work or assignment with the duplicate name should not be allowed. The system should compare the submitted work with the previous submitted works and check the duplication. The duplicated works should not be allowed to upload into the system. The feedback on the submitted projects should be viewed by the students.

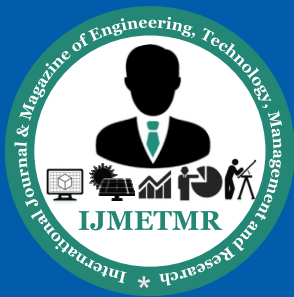
## Conclusion:

Safe and Secure Project assessment using attribute based access rights is a web based application. This application is designed and developed with proper security in ASP. Net and SQL Server technologies. The project is means for all universities which are facilitating the online submission of project work or any assignment work given to the students. The project has been designed and developed according to the software development life cycle model. The specific application development model used for this project is Rapid Application Development model. The database tables are designed and developed with relational database management system. The project is designed with appropriate security for uploading the files from student end and configured the attribute based access rights to the uploaded files to view and give feedback by the faculty members. The project facilitate the student to view the feedback on the submitted project file after submitting and approved by the faculty members. The project functionality is appropriate for online submission system to any university. The critical evaluation results have been tested and found the project is fulfilling the objectives of the project.

## REFERENCES:

- [1]David Jones, Sandy Behrens [2013] Online Assignment Management: An Evolutionary Tale
- [2]Matthew Prineas and Marie Cini [2011] National Institute for Learning Outcomes Assessment published in October 2011 as Assessing Learning in Online Education: The Role of Technology in Improving Student Outcomes
- [3]ICT Support [2011] Online Applicant Portal Project Student Self Access Portal -Project Page 1 of 11 ICT Support 2011
- [4]Mollie McGill[2010]HIGHER EDUCATION WEB PORTALS:SERVING STATE AND STUDENT TRANSFER NEEDS.
- [5]Susan Hohenberger and Brent Waters [2014]Online/ Oine Attribute-Based Encryption supported by National Science Foundation CNS-1154035 and CNS-1228443; the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory under contract FA8750-11-2-0211, DARPA N11AP20006,

- [6]VCU Center for Teaching Excellence [2009] Online Teaching and Learning Resource Guide published As Palloff and Pratt (2009) suggested, trust develops when a community of learners is created.
- [7]ZhongyuanXu, Scott D. Stoller [2013] Mining Attribute-based Access Control Policies (Submitted on 11 Jun 2013 (v1), last revised 5 Jan 2014 (this version, v3)).
- [8]B. Lohman, J.R. van der Hoeven [2013] User Requirements Document v.2 (URD) Computer emulator for digital preservation
- [9]Mark Saunders and Paul Josey [2013]The Layers of Research Design By Mark Saunders and Paul Josey
- [10]Aitchison, A. (2009) Beginning Spatial with SQL Server 2008, Apress.
- [11]Knight, B., Patel, K., (2011) Professional Microsoft SQL Server 2008 Administration, John Wiley & Sons
- [12]Michael R. Harwell[2011]Research Design in Qualitative/Quantitative/Mixed Methods
- [13]Lerman J., (2010) Programming Entity Framework, 2nd Edition,
- [14]Mackey, A. (2010) Introducing .NET 4.0: with Visual Studio 2010, Apress
- [15]Spaanjaars, I., (2010) Beginning ASP.NET 4: in C# and VB, John Wiley & Sons
- [16]Vieria, R. (2011) Beginning Microsoft SQL Server 2008 Programming, John Wiley & Sons
- [17]Booch .Jacobson (2004) "UML Distilled, A brief guide to the standard object modeling language, Third Edition, pp: 25-56.
- [18]Prof. Dr. Knut Hinkelmann Dr. Hans Friedrich Witschel[2013]How to choose a research methodology.
- [19]CemKaner and Jamer Bach (2011) "Lessons Learned in software testing: A context-driven approach", A content driven approach.
- [20]Richard N. Katz and Associates Web Portals and Higher Education Technologies to Make IT Personal
- [21]JaydipSen [2013]Security and Security and Privacy PrivacyIssues in Cloud Computing Computing
- [22]vince\_kornacki'spicturevince\_kornacki[2013] Web Application Penetration Test Tricks Part I – Virus Upload published in cyber readiness and response.
- [23]Michael Dalton, Christos Kozyrakis, and Nickolai-Zeldovich [2009] Nemesis: Preventing Authentication & Access Control Vulnerabilities in Web Applications.
- [24]Barbour, R. (2013). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? British Medical Journal, 322, 1115–1117.
- [25]ShanmugasundaramHariharan [2012] Automatic Plagiarism Detection Using Similarity Analysis published in The International Arab Journal of Information Technology, Vol. 9, No. 4, July 2012
- [26]Romans Lukashenko, Vita Graudina, Janis Grundspenkis [2007]Computer-Based Plagiarism Detection Methods and Tools: An OverviewInternational Conference on Computer Systems and Technologies - CompSysTech'07
- [27]Commonwealth of Australia 2013Data Dictionary (External) – Version 2 (05/13) Settlement Reporting Facility SRF
- [28]Jessie Chen, MEng. (2009) C# .NET Algorithm for Variable Selection Based on the Mallow's Cp Criterion
- [29]WEB LINKS
- [30]<http://www.codeproject.com/Articles/20640/Creating-PDF-Documents-in-ASP-NET>
- [31]<http://www.beansoftware.com/ASP.NET-Tutorials/>
- [32]<http://msdn.microsoft.com/en-us/library/aa479347.aspx>
- [33]<http://www.microsoft.com/sqlserver/en/us/default.aspx>
- [34]<http://msdn.microsoft.com/en-gb/default.aspx>
- [35]<http://msdn.microsoft.com/en-us/vstudio/dd285474>



[36]<http://www.aspfree.com/c/a/ASP.NET-Code/How-to-Use-the-ListBox-Control-in-ASP-NET-2-0/>

[37]<http://www.microsoft.com/visualstudio/eng/downloads>

[38]<http://www.learn-asp.net/asptutorials/FileUpload.aspx>

[39]<http://www.dotnetspider.com/tutorials/DotNet-Tutorials.aspx>

[40]<http://www.microsoft.com/visualstudio/eng/whats-new>

[41]<http://shop.oreilly.com/category/browse-subjects/programming/dotnet.do>

[42]<http://www.asp.net/web-forms/tutorials/data-access/working-with-binary-files/uploading-files-cs>

[43]<http://www.asp.net/ajax/documentation/live/overview/ScriptManagerOverview.aspx>

[44]<http://research-methodology.Net/research-methodology/research-approach/>

# Design of Sentiment Analysis System using Polarity Classification Technique

Rajeshwar Rao Kodipaka  
Asst.Prof, Dept of CSE,  
MREC-Sec-bad, TS

Sanjeeva Polepaka  
Associate.Prof, Dept of CSE,  
MREC-Sec-bad, TS

Md. Rafeeq  
Associate.Prof, Dept of CSE,  
CMRTC-Kandlakoya-Hyd, TS

## ABSTRACT

Twitter is a medium that we can use for communication. All posted tweets we can store in one location and create archive. Archive contains new and old tweets. Now we can start the analyzation on archive tweets that's we can design effective sentiment analysis system. This paper main aim is to determine parts of speech opinion words using polarity classification technique and support vector machine learning algorithm. Surveys of methods are used in various levels of sentiment analysis. It does analyze the tweets information in limited levels of content only. Now in this paper we design new sentiment analysis tool using polarity classification technique. Polarity classification techniques discover top 20-emoticons, learning different classes of words and other features information. These techniques perform in depth tweets analysis. It does provide better analysis results compare to previous methods.

## Keywords

Sentiment analysis, opinion mining, text documents, support vector machine classifier, polarity classification technique.

## 1. INTRODUCTION

Twitter enables all organizations for communication directly with each other. It's possible to tap the global real time communication important events very easy using twitter analysis. It is used to find out opinions from twitter micro blogging tweets. Many numbers of researchers have done good extensive work presently and previously on this topic environment. Analysis starts on the basis of tweets. Present approaches are missing contexts for some number of conclusions.

In this paper using polarity classification technique analyzes the tweets and fulfills the missing context features. Here we fulfill the features like part of speech opinion words of information, frequent occurrences opinions words discovery with the help of support vector machine classifier. It gives more conclusions compare to all previous approaches and more useful also for any real applications decision making.

## 2. RELATED WORK

Sentiment analysis is also known as Opinion Mining. Sentimental analysis major aims to determine the attitude of writer, judgment and communication based on text documents. Sentiment analysis major task is classify text documents, sentences, aspect based level. Express opinions in documents, sentences, aspect based level information are positive, negative or neutral [1].

Twitter data contains different topics information related to different domains. Classify the sentiment documents information using probabilistic model. Probabilistic model is one of the supervised learning algorithm. Each topic related how many documents are available it's not possible to

recognize or predictable. First choose the class labels and classify each and every topic documents separately [1] [2]. This is we can call as a document level sentiment classification.

In each topic sentiment documents again possible to classify or learning the sentiment sentences, aspect level process, opinion words also. The above steps are possible to implement on news, blogs and other categories [2] [3].

In twitter first collect the different categories of data using hash tags input. Hash tags works like class labels, but here there is [3] [4] no sentiment labels information. Supervised learning is not sufficient for extraction of sentiment analysis information. Here we should use the unsupervised learning [4] [2]. Using unsupervised learning identify the sentiment words measurement in documents. Sentiment words are categorized into positive and negative contexts.

Sentiment words are huge. Reduces the sentiment words information and improve the classification result. Classification improves using co-occurrence technique. Highly occurred features we can display as a output [4] [5] using unsupervised clustering. All high occurred features are not semantic or meaningful.

On co-occurrence words information applies correlation technique finally recognizes or predicts the relationship sentiment words content. This correlation sentiment words procedure is best prediction procedure compare to above all procedures [2] [3] [5].

Again in twitter it's possible to predict the similar opinions of information. Identify the social relationship users from total twitter data. Social relationship users it's possible to display using visualization concept. Finally it's possible to display temporal events relationship information also as a final result. Consider the temporal events relationship and possible to view of sentiment words trends information effectively. These trends changes dynamically [4] [5] [6].

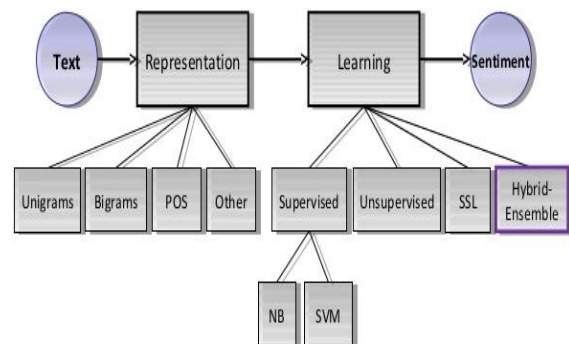


Fig 1: Sentiment Analysis Result with different data mining techniques

### 3. PROBLEM STATEMENT

Face book and Twitter are major resources for sentiment analysis. Nowadays numbers of people are increases for posting the opinions in social media environment. Result is vast amount of unstructured data is available. Manual sentiment analysis is not feasible which generated data is currently. Here in this paper we can propose automated sentiment analysis techniques. Automated sentiment analysis techniques are extracting different characteristics of information. Here we focus on tweet level polarity classification process. Hence it is one of the interesting research topics.

### 4. PROPOSED METHODOLOGY

After posting before analysis of raw tweets we can apply preprocessing operation. Using preprocessing operation removes the unnecessary content and change the format of context information. Next we use the tokeniser operations for dividing the content based on parts of speech. According to parts of speech classify the sentences of information. Classify the sentences depends on hash tags or tokens. Different hash tags are different parts of speech content purpose in our implementation process.

Here we design the different rules for sentiment analysis. Those rules are emoticon, support vector machines with n-grams, Sentic computing and Linguistic rules etc.

#### 4.1 Emoticon Rules

Emoticons are ASCII art. They are formed through creative letters, numbers and symbols. Most of the people represent the facial features information. Here normally for plain text messages we can add the emotional flavor. Next for normal text messages again we can add the smile automatically those messages are converted to happiness and surprise messages.

First choose the twitter data source and collect lakhs or millions of tweets. In millions of tweets we can perform the analyzation operation and we discovered 20 usage patterns information.

#	Emoticon	Usage	Percent	Notes
#1	:)	32,115,789	33.380%	Happy face
#2	:D	10,595,385	11.006%	Laugh
#3	:(	7,613,014	7.908%	Sad face
#4	;)	7,238,295	7.519%	Wink
#5	:-)	4,254,708	4.420%	Happy face (with nose)
#6	:P	3,588,863	3.728%	Tongue out
#7	=)	3,564,080	3.702%	Happy face
#8	(:	2,720,383	2.826%	Happy face (mirror)
#9	;-)	2,085,015	2.166%	Wink (with nose)
#10	:/	1,840,827	1.912%	Uneasy, undecided, skeptical, annoyed?
#11	XD	1,795,792	1.865%	Big grin
#12	=D	1,434,004	1.490%	Laugh
#13	:O	1,077,124	1.119%	Shock, Yawn
#14	=]	1,055,517	1.096%	Happy face
#15	D:	1,048,320	1.089%	Grin (mirror)
#16	;D	1,004,509	1.043%	Wink and grin
#17	:]	954,740	0.992%	Happy face
#18	:-)	816,170	0.848%	Unhappy
#19	=/	809,760	0.841%	Uneasy, undecided, skeptical, annoyed?
#20	=(	760,800	0.790%	Unhappy

Fig2: Top 20 Emoticons

In Above diagrams top-20 emoticons we displayed here in our implementation. These emoticons are occurred frequently in number of tweets messages.

#### 4.2 N-Grams Technique

In posted tweets observe the features and store features into feature vector. All required features are available in sequence or contiguous or not we can check with n-grams concept. Afterwards using TF-IDF identifies and calculates the frequency count. Consider the frequency count generate weighted features information. Weighted features information controls the number of dimensions. In all frequent words we can categorize based on parts of speech. Consider the different number of parts of speech tags information and categorize the words into noun, adverb, verb etc.

#### 4.3 Linguistic Rules

After collection of training tweets information next we can apply support vector machine classifier. It will separate the features of information n different classes. Each and every class contains some data points of information content. Each and every class of words again categorize into two classes. Those two classes are positive and negative words. Calculate the polarity value like positive and negative content. Polarity value is nothing but decision score.

### 5. EXPERIMENTS AND RESULTS

We evaluate our proposed system on available real time datasets. Every and every year tweets we can collect separately and create dataset. Every year dataset contains different categories tweets are available. Those categories are positive, negative and neutral tweets. Here first remove neutral tweets from total number of tweets information.

Evaluate two datasets and calculates efficiently different performance metrics parameters information. Those parameters are precision, recall and f-measure.

Table1: Performance Metrics

Method	Positive			Negative			Average		
	P	R	F	P	R	F	P	R	F
N-grams	89.92	81.90	85.72	61.20	75.66	67.67	75.56	78.78	76.69
N-grams and Emoticon Rules	89.74	83.05	86.27	62.50	74.85	68.11	76.12	78.95	77.19
Modified N-grams	89.39	82.90	86.02	62.00	73.93	67.44	75.69	78.41	76.73
Modified N-grams, and Emoticon Rules	89.25	83.97	86.53	63.29	73.22	67.89	76.27	78.60	77.21
Modified N-grams, Emoticon Rules, and Word-level Unsupervised Rules	90.22	86.24	88.19	67.37	75.25	71.09	78.80	80.75	79.64
Modified N-grams, Emoticon Rules, and Concept-level Unsupervised Rules	90.41	86.20	88.25	67.45	75.76	71.37	78.93	80.98	79.81

### 6. CONCLUSION AND FUTURE WORK

In this paper new twitter sentiment analysis system apply rules and discover more useful text. Useful text contains meaningful features. Those features are discovered using linguistic content and sentic computing rules. Features are like emoticon symbols and n-grams content information. These meaningful features are support for decision making in all real time applications.

In future we plan to improve the performance using other unsupervised classifiers. We plan to develop some more rules

for efficient text predictions and multi model sentiment analysis.

## 7. REFERENCES

- [1] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *J. Am. Soc. Inform.Sci. Technol.*, vol. 60, no. 11, pp. 2169–2188, 2009.
- [2] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Micro-blogging as online word of mouth branding," in *Proc. Extended Abstr. Human Factors Comput. Syst.*, 2009, pp. 3859–3864.
- [3] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *J. Comput. Sci.*, vol. 2, no. 1, pp. 1–8, 2011.
- [4] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welp, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *Proc. 4th Int. AAAI Conf. Weblogs Soc. Media*, 2010, vol. 10, pp. 178–185.
- [5] L. T. Nguyen, P. Wu, W. Chan, W. Peng, and Y. Zhang, "Predicting collective sentiment dynamics from time-series social media," in *Proc. 1st Int. Workshop Issues Sentiment Discovery Opinion Mining*, 2012, p. 6.
- [6] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in twitter events," *J. Am. Soc. Inform. Sci. Technol.*, vol. 62, no. 2, pp. 406–418, 2011.
- [7] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in *Proc. Workshop Lang. Soc. Media*, 2011, pp. 30–38.
- [8] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lect. Human Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.
- [9] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, "User-level sentiment analysis incorporating social networks," in *Proc. 17<sup>th</sup> ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1397–1405.
- [10] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification," in *Proc. 45th Annu. Meeting Assoc. Comput. Linguistics*, 2007, vol. 7, pp. 440–447.
- [11] F. Li, S. J. Pan, O. Jin, Q. Yang, and X. Zhu, "Cross-domain coextraction of sentiment and topic lexicons," in *Proc. 50th Annu. Meeting Assoc. Comput. Linguistics: Long Papers*, 2012, pp. 410–419.
- [12] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, "Cross-domain sentiment classification via spectral feature alignment," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 751–760.
- [13] I. Ounis, C. Macdonald, J. Lin, and I. Soboroff, "Overview of the trec-2011 microblog track," in *Proc. 20th Text Retrieval Conf.*, 2011, <http://trec.nist.gov/pubs/trec20/t20.proceedings.html>
- [14] I. Soboroff, I. Ounis, J. Lin, and I. Soboroff, "Overview of the trec- 2012 microblog track," in *Proc. 21st Text REtrieval Conf.*, 2012.
- [15] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N Project Report*, Computer Science Department, Stanford, USA, pp. 1–12, 2009.
- [16] S. Li, C.-R. Huang, G. Zhou, and S. Y. M. Lee, "Employing personal/ impersonal views in supervised and semi-supervised sentiment classification," in *Proc. 48th Annu. Meeting Assoc. Comput. Linguistics*, 2010, pp. 414–423.

## 8. AUTHOR PROFILE

**Rajeshwarrao. Kodipaka** (ISTE,CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.

**Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer network.

**Md. Rafeeq** Doing external PhD in JNTUH-Hyd,Telangana,India. I would like thank to my guide Dr. Sunil Kumar, and co-guide Dr. Subhash Chandra for constant support for Understanding the subject.

# Security and Privacy as data Service in the Multi-cloud Node Environment

Mr. Rajeshwar rao kodipaka\*1, Mr. Sanjeeva Polepaka\*2, Mr.M.Amarendhar Reddy\*3

*Asst.Prof, Dept of CSE, MREC-Sec-bad, TS.*

*Associate.Prof, Dept of CSE, MREC-Sec-bad, TS*

*Asst.Prof, Dept of CSE, B.V. Raju Institute of Technology (UGC-Autonomous), Dist: Medak, TS*

## ABSTRACT:

Technology and its significance have changed the global world to the next level of the IT Village. In the Context of the revolution of the Information Technology , we have gone to the next level of computing to make the Live easier , smooth , faster and accurate. Hence this terminology is to heard but to achieve is typical. Hence, In this Paper, We have kept for the next generation computing scaling like Cloud Computing with next residual of the Security. Hence, in the current context we have given the glimpse of the time and material to make cloud feasible from the security point of view for the usability of the domains like Banking and financial with high end feasible and scalable solution. In this we have implemented the node based Architectural model with acknowledgement encrypted based on the level of the data to these data moved in the network with the classification of the node based data in the preview model of the datacenter to monitor the traffic of the data and the shortest path Mechanism.

**INDEX TERMS:** cloud, security, privacy, multi-cloud, application partitioning, tier partitioning, data partitioning, multi party computation.

## 1.INTRODUCTION

Technology and its advancement lead us to research for the next level of the advancement. In the context of the Cloud computing which we can tell as the technological advancement plays the vital role in the industry of Information Technology. In the modern age of the cloud computing generation, the idea of cloud computing is almost as old as the computer itself. Its principle is to have the user's computer, Smartphone, tablet or any internet-connected device acting as a front-end displaying an application, as the resources connected in remote servers. Regarding

data confidentiality, all three providers claim that they provide and support encryption. Many cloud providers allow their customers to encrypt data before sending it to the cloud. For instance, Amazon S3 makes it optional to users to encrypt data for additional security, but not recommended in the case of third party or external auditors depending on the model of service they are offering. The first model is the Infrastructure as a service (IaaS), which is the most basic model. The IaaS model consists in providing Virtual Machines to the clients whom will have to install operating systems as well as their applications on top of the vendors providers like Amazon EC2, Google Compute Engine .Today's data centers are already containing thousands of servers, and this number is very likely to increase during the forthcoming years. Furthermore, the customers can request the creation or the removal node to meet their needs.

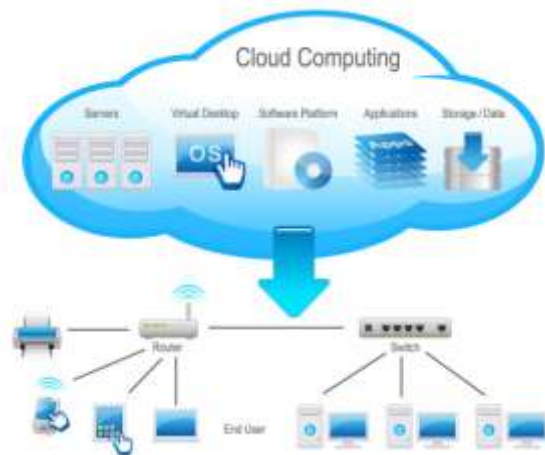


**Fig.1.1. Tiered structure of the cloud**

Therefore, the networks are highly elastic and can reach an immense number of traffic. Because of that, networks architecture has to be scalable in order to ensure a good performance whatever number of VIRTUAL is running in the network.

## II. RELATED WORK

In the IaaS model, the customers to dynamically scale up and down to as many machines as needed inside the cloud in a “pay-as-use” manner. The clients can dynamically provision resources to meet the current demand by adjusting their leasing of resources from the cloud provider. Customers can also then use more efficient hardware without being preoccupied by its maintenance, cooling and storage. In each model, the cloud providers can use multi-tenancy, where virtual machines from multiple customers can share the same sets of physical servers and the network, in order to reduce the waste of resources. However, economies of scale are so important that their users benefit from both aspects like efficient and cheaper solution.



**Fig.2.1. Illustration of the Centralized cloud node**

Due to the of an ever growing number of people connected to the internet, the constant evolution of computer science along with the decreasing of hardware prices, the use of cloud computing instead of classic networks architecture has recently become a reality. In order to optimize the resources in the data centers or for maintenance reasons, cloud providers often have to migrate from the physical server to another. Migrations are used in order to pool the active and then reduce the number of turned on physical servers inside the data center, in order to reduce cost. Migrations also reveal themselves handy when there is a hardware problem on a physical server, for example. The running on this server is moved to other physical servers and the alleged

deficient server can be turned off and replaced. The impact of node migration on the network is important. Since the virtualization is potentially moving from a physical server to another, the routes between them are also impacted, and the topology of the network is thus changing. Most importantly, the end points of the routes are changing. This causes the need to reconfigure the network with each migration.

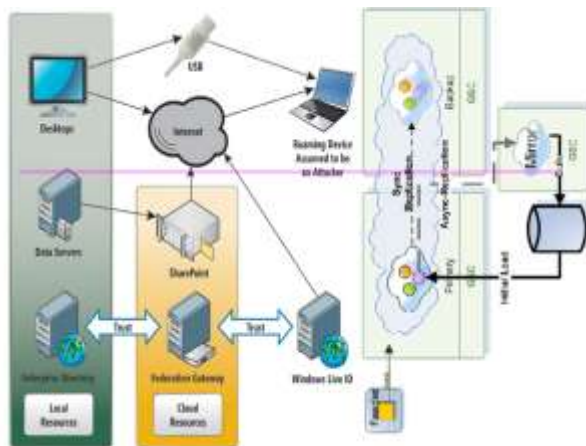
## III. PROPOSED METHODOLOGY

In this paper, we have given emphasis on the Privacy and Security involved in the multi-cloud architecture. The big threat to cloud is the security which has not yet robust and need much more advancement to give the client that is the best solution especially in the industry domain of banking and financial services. The advancement of the technology and its usefulness makes us to research the best of the best service for the automation world more precisely the cloud computing security can be breached by several actors. In the cloud, there are the cloud provider and the tenants. Our project aims to secure the traffic of a tenant. In this situation, this traffic could first be threatened by other tenants. Indeed, malicious tenants could be willing to access data of other tenants or to gain access to their network by using techniques such as ARP cache poisoning or IP spoofing. Furthermore, a configuration of routing appliances could lead to a breach in confidentiality. The other threat comes from inside the tenant's network itself. This threat is the most important one in security today, as malicious software can enter the enterprise network downloaded by the employees browsing the internet. In the enterprise network, the traffic is secured by the traversal of virtual middle boxes along the path. Our project aims to create a framework enabling the traversal of virtual middle boxes despite the elastic nature of the cloud, meanwhile providing isolation of traffic in order to prevent the risk of an attack by another tenant. The switches between the source switch and the first ingress switch will all have the same type of rule. Based on the, the packet will be routed to a particular port.

This port has been calculated during the route calculation. At the ingress switch of a virtual middle box, the rule will consist in matching one

attribute and applying two actions. The matching attribute is obviously the tag corresponding to the virtual middle boxes. The packet must then be sent to the virtual middle boxes, but before that, the tag must be popped, so the virtual middle boxes receive an unmodified packet. The progress switch of the virtual middle boxes will receive an untagged packet. The policy management in a secure enterprise network today can therefore be quite complex. For instance, it may require restricting a machine containing sensitive data to be accessed only by a small group of users, or preventing external traffic from directly reaching internal servers.

The actual realization may involve servers having complex communication patterns governed by network access control, such as the traversal of several virtual middle boxes before being reached. When enterprises decide to move to the cloud, they want to keep the same requirements regarding their policy management. It would be possible for the network manager of the enterprise to implement the virtual middle boxes and the routing policies on in the same way as before moving to the cloud. However, one of the goals of moving to the cloud is to escape the burden of network administration and configuration. Furthermore, the type of security policies in place in enterprises networks is often quite similar as it consists in the traversal of several virtual middle boxes.



**Fig.3.1. Multi- Cloud Architecture design involves the Agility platform**

A tenant could be attacked by the cloud provider or by other clients. The cloud provider could be attacked

by tenants. Our project does not consider the situation where the provider is malicious. Indeed, we suppose that the cloud provider is honest. We further suppose that all network appliances are compliant and secure. Based on the source and destination field, as well as the port from which the packet arrived, the controller will have to push the down to the packet. A second rule consists in matching this tag and to route the packet to the adequate port. The rules set in the forthcoming switches are the same as previous, until we reach the last virtual middle boxes. From the progress switch of the last virtual middle boxes to the destination switch, there are two cases. First, the destination machine is located in the current zone, and the destination switch is connected to the destination machine. In that case, the routing in the last segment is done by matching the destination address. Network security is a key aspect in designing modern applications.

#### IV.EVALUATION AND ANALYSIS

We can see that in both cases, the security policy has been reconfigured automatically after migration. The middle box sequences have been applied accordingly to the Application ID present in the packet, even though the instances traversed have changed. We chose to show both these cases of migration in order to demonstrate that *all or part* of the traversed middle box instances can be modified. Our model allows us to dynamically enforce security policies in a multi-tenant cloud network. In addition, the security policies stay coherent in spite of the node migration. So as to create a prototype of our network controller, we had to implement several modules in order to demonstrate the relevance of our policy enforcement mechanism. Particularly, a routing protocol has been implemented in order to figure out the route from the source to the destination, throughout the virtual middle boxes.

#### V.CONCLUSION AND FUTURE WORK

Apart from the above mentioned domain other domains are fair and fine enough to us the technology. In this one we use the data partitioning and the security of the node and agent based cryptography, which will provide the tiered or layered security. In the context of the security; particularly, we considered the criteria of scalability

and auto in a context where the network is shared by multiple tenants and the migration of nodes are increasing with the number of request. We first considered the different existing solutions in order to differentiate the traffic between the different tenants, as the isolation between tenants is the basis of a secure network. There are many solutions allowing traffic separation. However, the precision of these solutions varies greatly, as the isolation can be made from a complete separation of physical network and automatic centralized control of the network. We then analyzed the way security policies are defined and enforced. The security policies are defined in several ways. Some solutions focus on the isolation only or the routing rules and their consistency, whereas other architectures put more emphasis on the traversal of virtual middle boxes.

## VI. REFERENCES

- [1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing, Version 15," Nat'l Inst. of Standards and Technology, Information Technology Laboratory, vol. 53, p. 50, <http://csrc.nist.gov/groups/SNS/cloud-computing/>, 2010.
- [2] F. Gens, "IT Cloud Services User Survey, pt.2: Top Benefits & Challenges," blog, <http://blogs.idc.com/ie/?p=210>, 2008.
- [3] Gartner, "Gartner Says Cloud Adoption in Europe Will Trail U.S. by at Least Two Years," <http://www.gartner.com/it/page.jsp?id=2032215>, May 2012.
- [4] J.-M. Bohli, M. Jensen, N. Gruschka, J. Schwenk, and L.L.L. Iacono, "Security Prospects through Cloud Computing by Adopting Multiple Clouds," Proc. IEEE Fourth Int'l Conf. Cloud Computing (CLOUD), 2011.
- [5] D. Hubbard and M. Sutton, "Top Threats to Cloud Computing V1.0," Cloud Security Alliance, <http://www.cloudsecurityalliance.org/topthreats>, 2010.
- [6] M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, "On Technical Security Issues in Cloud Computing," Proc. IEEE Int'l Conf. Cloud Computing (CLOUD-II), 2009.
- [7] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds," Proc. 16th ACM Conf. Computer and Comm. Security (CCS '09), pp. 199-212, 2009.
- [8] Y. Zhang, A. Juels, M.K.M. Reiter, and T. Ristenpart, "Cross-VIRTUAL Side Channels and Their Use to Extract Private Keys," Proc. ACM Conf. Computer and Comm. Security (CCS '12), pp. 305-316, 2012.
- [9] N. Gruschka and L. Lo Iacono, "Vulnerable Cloud: SOAP Message Security Validation Revisited," Proc. IEEE Int'l Conf. Web Services (ICWS '09), 2009.
- [10] M. McIntosh and P. Austel, "XML Signature Element Wrapping Attacks and Countermeasures," Proc. Workshop Secure Web Services, pp. 20-27, 2005.
- [11] J. Kincaid, "Google Privacy Blunder Shares Your Docs without Permission," TechCrunch, <http://techcrunch.com/2009/03/07/huge-google-privacy-blunder-shares-your-docs-withoutpermission/>, 2009.
- [12] J. Somorovsky, M. Heiderich, M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, "All Your Clouds Are Belong to Us: Security Analysis of Cloud Management Interfaces," Proc. Third ACM Workshop Cloud Computing Security Workshop (CCSW '11), pp. 3-14, 2011.
- [13] S. Bugiel, S. Nurnberger, T. Poppelmann, A.-R. Sadeghi, and T. Schneider, "AmazonIA: When Elasticity Snaps Back," Proc. 18th ACM Conf. Computer and Comm. Security (CCS '11), pp. 389-400, 2011.
- [14] D. Bernstein, E. Ludvigson, K. Sankar, S. Diamond, and M. Morrow, "Blueprint for the Intercloud—Protocols and Formats for Cloud Computing Interoperability," Proc. Int'l Conf. Internet and Web Applications and Services, pp. 328-336, 2009.
- [15] A. Celesti, F. Tusa, M. Villari, and A. Puliafito, "How to Enhance Cloud Architectures to Enable Cross-Federation," Proc. IEEE Third Int'l Conf. Cloud Computing (CLOUD), pp. 337-345, 2010.
- [16] R. Turpin and B.A. Coan, "Extending Binary Byzantine Agreement to Multivalued Byzantine Agreement," Information Processing Letters, vol. 18, no. 2, pp. 73-76, 1984.
- [17] I. Koren and C.M.C. Krishna, Fault-Tolerant Systems. Morgan Kaufmann, 2007.

## AUTHORS PROFILE:



**1. Rajeshwarrao. Kodipaka** (ISTE, CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.



2. **Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer networks.



3. **M. AMARENDHAR REDDY** is presently associated with B. V. Raju Institute of Technology (UGC - Autonomous), Narsapur, Medak

(Dist), Telangana State as Asst. Professor in Information technology Department. He graduated from JNTU Hyderabad in B.Tech CS & IT and completed M.Tech (SE) from JNTUH-School of Information Technology Campus. Presently with 10 years of rich experience in Teaching, Administration and Research in the area of Object Oriented Modeling, Software Engineering Data Warehousing & Mining and Programming Languages. Prior to his current position with B. V. Raju Institute of Technology he served multiple positions as I/C HOD, Exam Branch In charge and Students' Advisor in various leading organization. His teaching experience includes teaching UG and PG course on Object oriented analysis and design, Java, C, C++, Software Engineering, Formal Languages and Automata Theory and Automata and Compiler Design. As part of Students' Advisor responsibilities he conducted workshops on Data mining and STQA.

# Cloud query Based allocation system in the Dynamic environment of Log Based Personalized Web Mining

Mr. Rajeshwar rao kodipaka\*1, Mr. Sanjeeva Polepaka\*2, Mr.M.Amarendhar Reddy\*3

Asst.Prof, Dept of CSE, MREC-Sec-bad, TS,  
Associate.Prof, Dept of CSE, MREC-Sec-bad, TS,

Asst.Prof, Dept of CSE, B.V. Raju Institute of Technology (UGC-Autonomous), Dist: Medak, TS,

## ABSTRACT:

*Information Technology and its implementation are vast and stretched as the endless ocean. Its beauty lies in the heart of the mother land India, where the main theme gets constructed. For Instance if we keep the Virtual implementation of IT is the cloud which makes us enabling the in the optimization of resource, economy and many more compare to the classical technology. In this paper, we try to put forward the cost reduction system which we call as pay per use and network. In the System of the elasticity which we can as the balancing the data load, making the environment dynamic, number of servers and the optimizing the traffic of the network., where we play an important role in the upgrading of technology which usually we call it as the Technology of change. concept of the caching the first step of reducing the network, making the virtual node which depends on the peak of the data load in order to balance the computing environment which we call as cloud computing. Technologically its meaning remains virtual, but in reality while exposing to a real world of enterprise solution.*

**INDEX TERMS:** Cloud computing, cost efficiency, differential query services, privacy, personalized web search, utility, risk, profile.

## 1.INTRODUCTION

Software as a service (SaaS), which is the most widely known among internet users. In this model, the end-user is using directly the software installed by the cloud provider, which could be an online e-mail service, for instance. we will first take a look at the current state of commercial Web search, since it touches most of our daily lives. Then we discuss related work pertaining to several classes of sanitization mechanisms and corresponding vulnerabilities. For the first part of the related work, we use an informal definition of privacy, namely that

a sanitization mechanism preserves privacy if, given the output of the mechanism, an analyst cannot make a reliable guess as to who a query belongs to. A little later on, we will introduce a more formal definition used in much of the most recent research on the topic. Some of largest US based search engine companies (Google, Yahoo, Microsoft, AOL, and Ask.com) have privacy policies that do little to search data, keeping IP address and browser cookie identifiers associated with searches for as long as 18 months (Schwartz & Cooper, 2007). This last model differs slightly to the previous as the cloud users are the customers of the application, in comparison of the two first models where the cloud users are often businesses. Cloud and data storage service which provides both secure data outsourcing service and efficient data retrieval and repair service, including four different entities: the data owner, the data user, the cloud server, and the third party server. The data owner outsources the encoded fragments of the file M to n cloud servers denoted as storage servers. If the data owner requires keeping the data content confidential, the file can be first encrypted before encoding. Outsourced data are attached by some metadata like verification tags to provide integrity check capability. After the data outsourcing, a data user can select any k storage servers to retrieve encoded segments, and recover the file M, which can be further decrypted in case the file is encrypted. Meanwhile, the third party server periodically checks the integrity of data stored in cloud servers. Failed cloud servers can be repaired with the help of other healthy cloud servers.

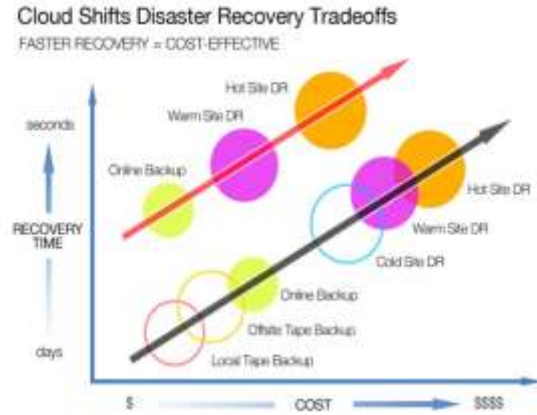


**Fig.1.1 Illustration of the Cloud Cost Model**

Security as malicious software can enter the enterprise network downloaded by the employees browsing the internet. Indeed, malicious tenants could be willing to access data of other tenants or to gain access to their network by using techniques such as ARP cache poisoning or IP spoofing. Furthermore, a configuration of routing appliances could lead to a breach in confidentiality.

## II.RELATED WORK

We can process user logs to build user-specific models, which we call local models. We do not sanitize local models under the assumption that users interacting with their own sensitive information do not constitute breaches of privacy. (We do not consider the case where multiple individuals access the same account, though this may be a common situation with, e.g., shared computers in homes, libraries, or Internet. Regarding data confidentiality, all three providers claim that they provide and support encryption. Many cloud providers allow their customers to encrypt data before sending it to the cloud. For instance, Amazon S3 makes it optional to users to encrypt data for additional security, but not recommended in the case of third party or external auditors. Cloud front is an Amazon web service which provides data confidentiality while being transferred. Attempts to apply functional repair in the LT codes based distributed storage should first solve how to recode packets, because the random linear recoding in the functional repair of network coding-based storage codes cannot satisfy the degree distribution in LT codes.



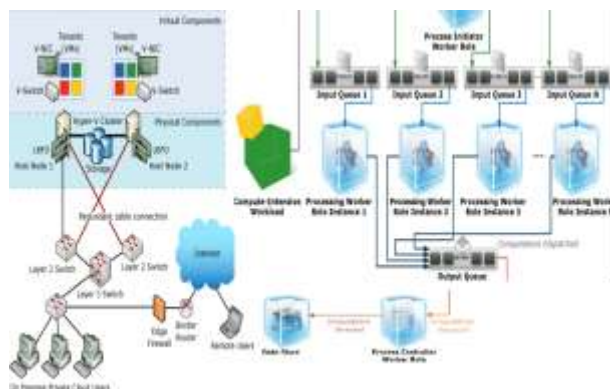
**Fig.2.1 Cloud Balance in the Traffic**

A tenant could be attacked by the cloud provider or by other clients. The cloud provider could be attacked by tenants. Our project does not consider the situation where the provider is malicious. Indeed, we suppose that the cloud provider is honest. We further suppose that all network appliances are compliant and secure. It seems that this problem can be solved by utilizing the recently proposed LT network codes (LTNC) which provides efficient decoding at the cost of slightly more communication in the single-source broadcasting scenario. However, after several rounds of repair with same recoding operations regulated in LT network codes, data users experience decoding failure with high probability. Our project aims to secure the traffic of a tenant. In this situation, this traffic could first be threatened by other tenants.

## III.PROPOSED METHODOLOGY

Search logs, or query logs as they are sometimes called, are databases containing information about user search activity. A typical log might contain a set of entries consisting of an identifier, a time stamp, an event descriptor, and pertinent information about the event. Common identifiers are IP addresses, cookie identifiers, or account numbers. Example event descriptions include query or click. Pertinent information about the event might include the query text or a clicked URL. When we assume that all identical identifiers correspond to a single individual, we refer to a collection of search log entries sharing the same identifier as a user search log or user log for short.

To begin with, the dedicated hardware, even though it may be inactive at a given time, has to be monopolized for the tenants who need the highest security. In other words, the pooling of resources would be limited which may lead to an oversized network. Indeed, the dedicated servers could have hosted other tenants' Virtual Machines. However, these VMs would have to be located on other servers, contributing to the waste of resources, as the servers would not be loaded at their maximal capacity. Furthermore, as the use of the cloud would grow, we can easily imagine that the expansion of such a rigid architecture would become harder and harder to maintain and configure. The Proof Generation process is run by the storage provider in order to generate proof transcript which could be a set of corresponding tags or, in many cases, an aggregation of the perturbed tags and aggregation of data blocks. A proof transcript allows the verifier to derive the proof in order to check the integrity of the challenged blocks. Upon receiving the response from the storage provider, the verifier executes the Proof Verification protocol to verify the validity.



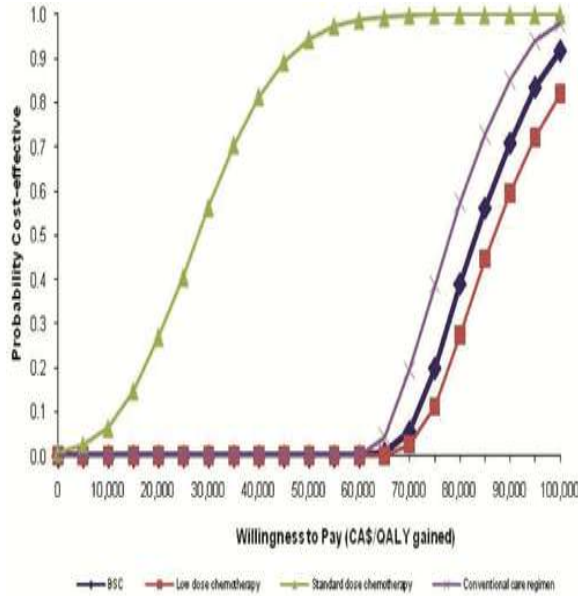
**Fig.3.1 Node Values Cloud Computing Peak Architect of High End Security.**

When enterprises decide to move to the cloud, they want to keep the same requirements regarding their policy management. It would be possible for the network manager of the enterprise to implement the middle boxes and the routing policies on VMs in the same way as before moving to the cloud. However, one of the goal of moving to the cloud is to escape the burden of network administration and configuration. Furthermore, the type of security policies in place in enterprises networks is often quite

similar as it consists in the traversal of several middle boxes. These data centers are located in Northern Virginia (USA) and the company's cluster of cloud computing services in Virginia were "currently experiencing degraded performance" as indicated on the Amazon website. Consideration of both data repair and data retrieval, and design a LT codes-based cloud storage service (LTCS). Multi-keyword ranked search over encrypted cloud data, and establish a variety of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching", i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity measure.

#### IV. ANALYSIS AND INTERPRETATION

Today's data centers are already containing hundreds of thousands of servers, and this number is very likely to increase during the forthcoming years. Furthermore, the customers can request the creation or the removal of VMs in order to meet their similar incident happened about four months earlier due to an electrical storm that caused some disturbance to the same data centers. In the field of IR, personalization is the act of incorporating information about a user into the processing of an IR application. For example, boosting the rank of search results that are more similar to web sites a user has visited in the past. Another common practice is the use of a user's location, often inferred by IP address or GPS coordinates. We concentrate on local-only personalization that can be performed in isolation on an individual's computer. We consider two different types of personalization: task-aware and supervised learning from user-provided annotations. Note that, in general, personalization can also involve a group of users



**Fig.3.1.1 Comparison of Dataset and Security feature with Size and Cost**

Thus, cloud customers are enthusiastic to be allowed to store their data in the cloud and at the same time they would like to be able to check by themselves that their data is protected.

## V.CONCLUSION

In the modeling feature of the cloud computing as pay per use which makes us to in making the system everything global and making the smaller level organization as base to as use. It is distinguished with the traditional multi-domain approaches by the characteristics of the emerging cloud environments. We use a top-down methodology in the research referring the PEI stack. Secure and reliable cloud storage with the efficiency. By mapping users to multiple pseudo users, it is more difficult to make the same links that the AOL data allowed. First, queries issued on two different days are not directly connected (provided the user was inactive for part of the intervening time period). Second, sessions generally consist of a small number of queries and are therefore less likely to cover a diverse set of topics. When these two points are taken together, it becomes more difficult to gather sufficient evidence to identify the individuals behind the queries.

## VI.REFERENCES

- [1] Thomas S. J. Schwarz and Ethan L. Miller. Store, forget, and check: Using algebraic signatures to check remotely administered storage. In ICDCS '06: Proceedings of the 26th IEEE International Conference on Distributed Computing Systems, Washington, DC, USA, 2006.
- [2] F. Sebe', J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient Remote Data Possession Checking in Critical Information Infrastructures," IEEE Trans. Knowledge Data Eng., vol. 20, no. 8, pp. 1034-1038, Aug. 2008.
- [3] G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. Fourth Int'l Conf. Security Privacy Comm. Networks, pp. 1-10, 2008.
- [4] C. Erway, A. Ku' pc,u", C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," Proc. 16th ACM Conf. Computer Comm. Security, pp. 213-222, 2009.
- [5] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," Proc. 14th European Conf. Research Computer Security, pp. 355-370, 2009.
- [6] A.F. Barsoum and M.A. Hasan, "Provable Possession and Replication of Data over Cloud Servers," Technical Report 2010/32, Centre for Applied Cryptographic Research, <http://www.cacr.math.uwaterloo.ca/techreports/2010/cacr2010-32.pdf>. 2010.
- [7] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession," Proc. 28th Int'l Conf. Distributed Computing Systems, pp. 411-420, 2008.
- [8] A.F. Barsoum and M.A. Hasan, "On Verifying Dynamic Multiple Data Copies over Cloud Servers," Technical Report 2011/447, Cryptology Eprint Archive, <http://eprint.iacr.org/>, 2011.
- [9] K.D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," Proc. 16th ACM Conf. Computer Comm. Security, pp. 187-198, 2009.
- [10] Y. Dodis, S. Vadhan, and D. Wichs, "Proofs of Retrievability via Hardness Amplification," Proc. Sixth Theory Cryptography Conf. Theory Cryptography, 2009.
- [11] A. Juels and B.S. Kaliski, "PORs: Proofs of Retrievability for Large Files," Proc. 14th ACM Conf. Computer Comm. Security, pp. 584-597, 2007.
- [12] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. 14th Int'l Conf. Theory Appl. Cryptology Information Security, pp. 90-107, 2008.
- [13] M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu, "Plutus: Scalable Secure File Sharing on Untrusted Storage," Proc. Second USENIX Conf. File Storage Technologies, 2003.
- [14] E.-J. Goh, H. Shacham, N. Modadugu, and D. Boneh, "SiRiUS: Securing Remote entrusted Storage," Proc. Network Distributed System Security Symp., 2003.

## AUTHORS PROFILE:



**1. Rajeshwarrao. Kodipaka** (ISTE,CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.



2. **Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer networks.



3. **M. AMARENDHAR REDDY** is presently associated with B. V. Raju Institute of Technology (UGC - Autonomous), Narsapur, Medak (Dist), Telangana State as Asst. Professor in

Information technology Department. He graduated from JNTU Hyderabad in B.Tech CS & IT and completed M.Tech (SE) from JNTUH-School of Information Technology Campus. Presently with 10 years of rich experience in Teaching, Administration and Research in the area of Object Oriented Modeling, Software Engineering Data Warehousing & Mining and Programming Languages. Prior to his current position with B. V. Raju Institute of Technology he served multiple positions as I/C HOD, Exam Branch In charge and Students' Advisor in various leading organization. His teaching experience includes teaching UG and PG course on Object oriented analysis and design, Java, C, C++, Software Engineering, Formal Languages and Automata Theory and Automata and Compiler Design. As part of Students' Advisor responsibilities he conducted workshops on Data mining and STQA.

# Privacy and Social Security Mechanism in the Era of Social Multi Cloud Environment

Mr. Rajeshwar rao kodipaka\*<sup>1</sup>, Mr. Sanjeeva Polepaka\*<sup>2</sup>

*Asst.Prof, Dept of CSE, MREC-Sec-bad, TS.*

*Associate. Prof, Dept of CSE, MREC-Sec-bad, TS.*

## ABSTRACT:

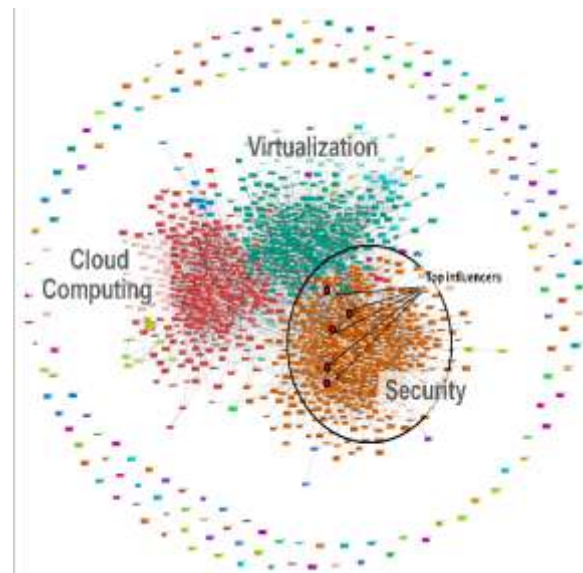
Time and the Constraint are both typically interrelated and lead us to the next level high technological revolution journey where we analyse the Human Automation instead of System Automation. Technological in the View Point cognitive science Information technology will have to make high level of demanding journey; in that aspect we consider the best of the technical revolution in the Social media where we share the thoughts, informational to many more. If we consider the Face book like such popular network sharing side leads us to the best of research, where technological Face book is high level to manage such huge volume of Data effectively and efficiently, But Lacking with the point of loop whole is the Information of Public and Private and Preventive social Inference of the Aspect. Hence; we look forward to put the paper as the most and best to best Information flow in the aspect of content and curative illustrative technological solution. However, the precision of these solutions varies greatly, as the isolation can be made from a complete separation of physical network, recent breakthroughs in Software-Defined Networking have let us to believe that the most suitable technology for our project today is the use of the technology as it provides an isolation per of as well as a automatic centralized control of the network.

**KEYWORDS:** Social network analysis, data mining, social network privacy, security, privacy, multi-cloud, application partitioning.

## 1.INTRODUCTION

In the Revolution n of Social Media, by structurally mapping out relationships between software developers, the presence of power-law distributions was found for project sizes, cluster sizes of connected developers, and the number of projects joined by developers. et al. Conclude Open Source software development can be modeled as self-organizing, collaboration, social

networks. It et al. explored the statistics and topological information of the Open Source software developer's collaboration network further by extracting project evolution parameters, by inspecting the network over a time period of two years. 50,000 projects involving 80,000 developers were investigated. Again, power-law distributions were found for the cluster distribution and degree distribution, which is the amount of ties within a network. Also they found during this time frame the average degree of projects on Source Forge slightly increased, though the network diameter slightly decreased. Face book, which is currently the world's most popular OSN, opened to the general public on 26 Sep 2006.



**Fig.2.1 Illustration of Data Sharing in the Cloud Security**

Prior to that, users were required to have a verifiable e-mail address from a "network" such as a university. Suppose a University of Cambridge student named Alice created a Face book account in August 2006 and immediately limited the sharing of all personal information to direct "friends" only, but left

her “search ability” setting at “Everyone” so that friends could search for her and send friend requests. Without this setting, it was very difficult for real-life friends to find each other on Face book: it was impossible for two users with the most restrictive search ability settings to “friend” each other, which discouraged users from choosing such settings.

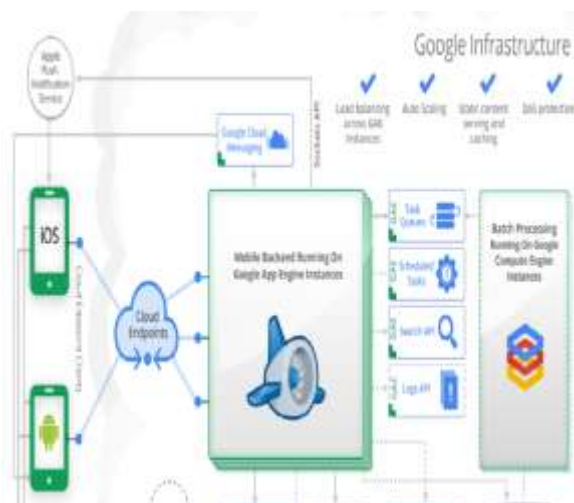
## II. RELATED WORK

Taken social network theory into account the organizational network structure affects team viability and team task performance. Team viability involves the committed of team members to stay together, where team task performance involves attaining the goals of a team. Team viability and team task performance can be conceived as success measures of teams. and Harrison noted although these two dimensions of success are conceptually distinct, in reality there is a close connection and cross-correlation between team task performance and team viability. In addition they found teams with central leaders, and teams central in a network full of other teams, tend to be better performers. Today's OSNs provide rich platforms for third-party social applications.

being given it anyway. Several applications used information in ways that contravened Facebook's Terms of Service, making information visible to users who would normally be unable to view it. People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time. **Mark Zuckerberg, January 2010.**

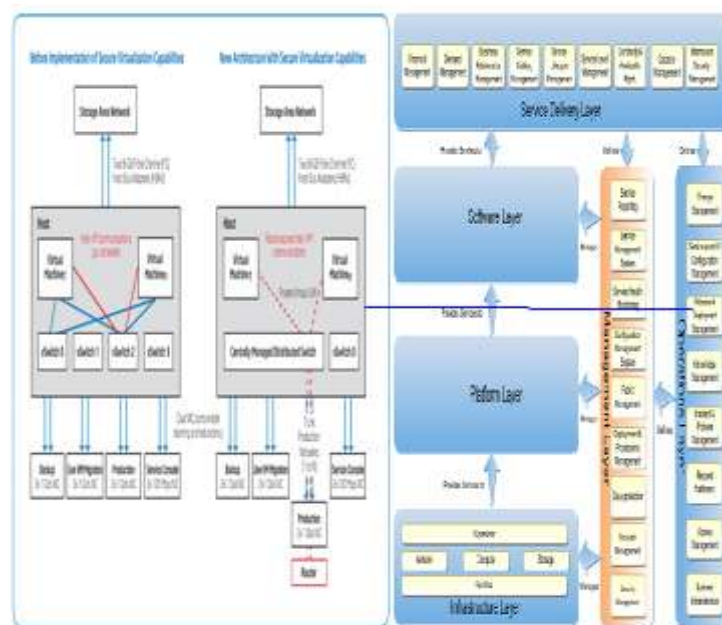
### III. PROPOSED METHODOLOGY

Today's centralized online social networks (OSNs) are trusted to make access control decisions on behalf of users, but those decisions do not always align with users' intentions. In some cases, technical flaws lead to data leakage contrary to the wishes of both users and operators. In other cases, there is a mismatch between user understanding of privacy settings and the actual protections afforded by the system. In many cases, however, the interests of the user and the network are simply not aligned. In these cases, when users' data must be safeguarded by operators without the incentive to do so, private information has a history of leaking out.



**Fig.2.1. Mobile Cloud Social Security Model**

Application functionality is supported by private user information, often served in greater quantities than actually required. Safeguards against misuse rely heavily on legal rather than technical protection. Felt and Evans, in their 2008 study of 150 Facebook applications [103], found that 90% of applications had no need of private user data, but were



**Fig.3.1 High Level Flow of Role, Language and Typically the Privacy Level**

Most of these forms of communication are not archived and therefore difficult to investigate. Next, the nature of Open Source software communities is to be open and transparent. Therefore, the project's public forums are taken into account as the investigated conversational network. On the public forums registered members can post messages and react on each other, resulting in forum threads. Registered members form the nodes of the conversational network. And, when two members of a project group post in the same thread, a relationship between those two members, or a link between those two nodes, exists. Finally, the several orthogonal cognitive considerations in social networks generally, describing them as problems with “decentralized systems”. For instance, the authors state that “more control over personal data almost inevitably translates to more decisions, which leads to cognitive overload”, but this is not a reason to avoid decentralized architectures. Centralized services such as Facebook provide very detailed and fine-grained control over some personal data, capable of causing just as much cognitive overload as an equivalent decision matrix in a decentralized system. What Facebook does not have is an incentive to explore alternative schemes for privacy policy configuration if those schemes hinder a “free flow of information” within the network. In contrast, a distributed system that does not extract monetary value from user data has every incentive to combat cognitive overload by accurately capturing user intent, reducing decision fatigue and improving the user experience of privacy management. While the authors of this critique raise some important issues, their case against decentralized data architectures is greatly overstated.

#### IV.ANALYSIS

In this we used between ness centrality to measure the effectiveness with which an attacker might influence information moving through a social network.

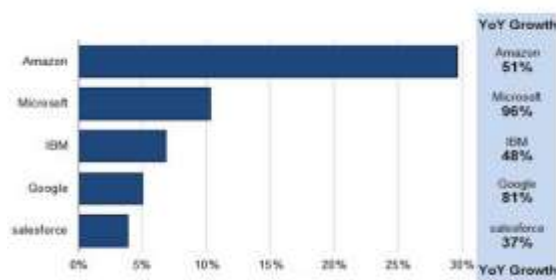


Fig.3.1.1 Content Highlighting to the Privacy Based on the Organisation

Suppose an attacker can compromise  $N$  nodes in a social graph and use them to influence information that passes through them. If the attacker can choose which  $N$  nodes to compromise, the optimal attack strategy is to choose the most central nodes in the network, yielding the highest probability of being able to influence any communication path. If the attacker has imperfect centrality information, the probability of successful attack will be reduced.

#### V.CONCLUSION AND FUTURE WORK

Technologically Each and every Aspect of Information has its own loop whole. Blocks are encrypted with block-specific symmetric keys; their content is only readable by principals which possess the relevant keys. Blocks in the encrypted store contain explicit links to other blocks, which are only visible to principals that possess the block's symmetric key. The system also provides opportunities for covert communication, including a low-bitrate perfectly unobservable communications. Some difficulties can be obviated under the random oracle assumption, but proofs in the standard model would be a natural direction to pursue. It is desirable to prove that, first, convergent encryption provides in distinguishability.

#### VI.REFERENCES

- [1] Facebook Beacon, 2007.
- [2] T. Zeller, "AOL Executive Quits After Posting of Search Data," The New York Times, no. 22, <http://www.nytimes.com/2006/08/22/technology/22iht-aol.2558731.html?pagewanted=all&r=0>, Aug.2006.
- [3] K.M. Heussner, "Gaydar' n Facebook: Can Your Friends Reveal Sexual Orientation?" ABC News, <http://abcnews.go.com/Technology/gaydar-facebook-friends/story?id=8633224#>, UZ939UqheOs, Sept. 2009.
- [4] C. Johnson, "Project Gaydar," The Boston Globe, Sept. 2009.
- [5] L. Backstrom, C. Dwork, and J. Kleinberg, "Wherefore Art Thou r3579x?: Anonymized Social Networks, Hidden Patterns, and Structural Steganography," Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 181-190, 2007.
- [6] M. Hay, G. Miklau, D. Jensen, P. Weis, and S. Srivastava, "Anonymizing Social Networks," Technical Report 07-19, Univ. of Massachusetts Amherst, 2007.
- [7] K. Liu and E. Terzi, "Towards Identity Anonymization on Graphs," Proc. ACM SIGMOD Int'l Conf. Management of Data (SIGMOD '08), pp. 93-106, 2008.
- [8] J. He, W. Chu, and V. Liu, "Inferring Privacy Information from Social Networks," Proc. Intelligence and Security Informatics, 2006.
- [9] E. Zheleva and L. Getoor, "Preserving the Privacy of Sensitive Relationships in Graph Data," Proc. First ACM SIGKDD Int'l Conf. Privacy, Security, and Trust in KDD, pp. 153-171, 2008.
- [10] R. Gross, A. Acquisti, and J.H. Heinz, "Information Revelation and Privacy in Online Social Networks," Proc. ACM Workshop Privacy in the Electronic Soc. (WPES '05), pp. 71-80, <http://dx.doi.org/10.1145/1102199.1102214>, 2005.

- [11] H. Jones and J.H. Soltren, "Facebook: Threats to Privacy," technical report, Massachusetts Inst. of Technology, 2005.
- [12] P. Sen and L. Getoor, "Link-Based Classification," Technical Report CS-TR-4858, Univ. of Maryland, Feb. 2007.
- [13] B. Tasker, P. Abbeel, and K. Daphne, "Discriminative Probabilistic Models for Relational Data," Proc. 18th Ann. Conf. Uncertainty in Artificial Intelligence (UAI '02), pp. 485-492, 2002.
- [14] A. Menon and C. Elkan, "Predicting Labels for Dyadic Data," Data Mining and Knowledge Discovery, vol. 21, pp. 327-343, 2010.
- [15] E. Zheleva and L. Getoor, "To Join or Not to Join: The Illusion of Privacy in Social Networks with Mixed Public and Private user Profiles," Technical Report CS-TR-4926, Univ. of Maryland, College Park, July 2008.
- [16] N. Talukder, M. Ouzzani, A.K. Elmagarmid, H. Elmeleegy, and M. Yakout, "Privometer: Privacy Protection in Social Networks," Proc. IEEE 26th Int'l Conf. Data Eng. Workshops (ICDE '10), pp. 266- 269, 2010.

#### **AUTHORS PROFILE:**



**1. Rajeshwarrao. Kodipaka** (ISTE, CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested

domain is Data mining, cloud computing, Computer Networks.



**2. Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer network.

# Implementation of Cryptography Encryption Algorithm for Plane Text

BV Ramudu<sup>1</sup>, Sanjeeva Polepaka<sup>2</sup>

Assoc Professor, MallaReddy Engineering College (Autonomous) Hyderabad, India<sup>1,2</sup>

**Abstract:** In this paper an efficient security architecture design and implementation of all candidates of AES encryption standards AES-128, AES-192 and AES-256 on the same hardware is proposed. AES algorithm proposed by NIST has been widely accepted as best cryptosystem for wireless communication security. The hardware implementation is useful in wireless security like military and mobile phones. This contribution investigates implementation of AES Encryption with regards to FPGA and VHDL. Optimized and synthesized VHDL code for AES-128, AES-192 and AES-256 for encryption of 128-bit data is implemented. Xilinx ISE 9.2i software is used for simulation. Each algorithm is tested with sample vectors provided by NIST output results are perfect with minimal delay. The proposed design consumes less power and area which is suitable battery driven mobile phones. Throughput reaches the value of 666.67 Mbps for encryption of 128-bit data with AES-128 key.

**Keywords:** Cryptography, Cipher, Reconfiguration, Encryption, Decryption.

## 1. INTRODUCTION

Cryptography is the science of information and communication security. Cryptography is the science of secret codes, enabling the confidentiality of communication through an insecure channel. It protects against unauthorized parties by preventing unauthorized alteration of use. Generally speaking, it uses a cryptographic system to transform a plaintext into a cipher text, using most of the time a key. One has to notice that there exist certain ciphers that don't need a key at all. A famous example is ROT13 (abbreviation from Rotation 13), a simple Caesar-cipher that obscures text by replacing each letter with the letter thirteen places down in the alphabet. Since our alphabet has 26 characters, it is enough to encrypt the cipher text again to retrieve the original message. Let me just mention briefly that there are secure public-key ciphers, like the famous and very secure Rivest-Shamir-Adleman (commonly called RSA) that uses a public key to encrypt a message and a secret key to decrypt it. Cryptography is a very important domain in computer science with many applications. The most famous example of cryptography is certainly the Enigma machine, the legendary cipher machine used by the German Third Reich to encrypt their messages, whose security breach ultimately led to the defeat of their submarine force. Before continuing, please read carefully the legal issues involving cryptography as in several countries even the domestic use of cryptography is prohibited: Cryptography has long been of interest to intelligence gathering agencies and law enforcement agencies. Because of its facilitation of privacy, and the diminution of privacy attendant on its prohibition, cryptography is also of considerable interest to civil rights supporters. Accordingly, there has been a history of controversial legal issues surrounding cryptography, especially since the advent of inexpensive computers has made possible widespread access to high quality cryptography. In some countries, even the domestic use of

cryptography is, or has been, restricted. Until 1999, France significantly restricted the use of cryptography domestically. In China, a license is still required to use cryptography. Many countries have tight restrictions on the use of cryptography. Among the more restrictive are laws in Belarus, Kazakhstan, Mongolia, Pakistan, Russia, Singapore, Tunisia, Venezuela, and Vietnam.[31] In the United States, cryptography is legal for domestic use, but there has been much conflict over legal issues related to cryptography. One particularly important issue has been the export of cryptography and cryptographic software and hardware. Because of the importance of cryptanalysis in World War II and an expectation that cryptography would continue to be important for national security, many western governments have, at some point, strictly regulated export of cryptography. After World War II, it was illegal in the US to sell or distribute encryption technology overseas; in fact, encryption was classified as a munitions, like tanks and nuclear weapons.[32] Until the advent of the personal computer and the Internet, this was not especially problematic. Good cryptography is indistinguishable from bad cryptography for nearly all users, and in any case, most of the cryptographic techniques generally available were slow and error prone whether good or bad. However, as the Internet grew and computers became more widely available, high quality encryption techniques became well-known around the globe. As a result, export controls came to be seen to be an impediment to commerce and to research.

### 1.1 Description of AES Algorithm:

Advanced Encryption Standard is the successor of Data Encryption Standard which was in use during the early 1977 to 1990. In DES encryption is based on a symmetric key algorithm that uses a 56-bit key. However by the mid 1990's, it was clear that the DES with 56-bit is insecure for many applications since the key is very small. Then it

was upgraded to Triple DES which was believed to be practically secure although there are theoretical attacks. Thus in Nov-26-2001 the FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 197(FIPS 197) specifies an algorithm called Advanced Encryption Standard (AES). AES is based on the principle known as Substitution Permutation network (SP-network) which means there will be a series of linked mathematical operations in the block cipher algorithm. AES encrypts a data block of 128-bits which is fixed with three different key sizes 128,192,256 bits.

The operations are based on Rijndael algorithm. The input of AES algorithm is 128-bit or 16 byte data which can be specified as a block. The basic unit of processing in the AES algorithm is a byte. All byte values in the AES algorithm will be presented as the concatenation of its individual bit values (0 or 1) between the braces in the order (b7, b6, b5, b4, b3, b2, b1, b0). These bytes are interpreted as finite field elements using a polynomial representation as follows

Algorithm	Key length (Nk words)	Block Size (Nb words)	Number of rounds (Nr)
AES-128	4	4	10
AES-192	6	4	12
AES-256	8	4	14

$$b_7X^7 + b_6X^6 + b_5X^5 + b_4X^4 + b_3X^3 + b_2X^2 + b_1X + b_0 = \sum_{i=0}^7 b_i X^i$$

Internally in AES algorithm operations are performed on a two-dimensional array of bytes called the state. The state consists of four rows of bytes, each containing Nb bytes, where Nb is the block length divided by 32 (4 for 128-bit key, 6 for 192-bit key, 8 for 256-bit key). Likewise the key length and number of rounds (iterations) differ from key to key as shown in table 1.

Figure1: Different keys and its attributes

### AES Encryption:

Encryption is the process of converting the plain text into a format which is not easily readable and is called as cipher. The cipher is got by doing a series of mathematical operations iteratively.

#### a) Sub Bytes:

In this sub bytes step the data in the plain text is substituted by some pre-defined values from a substitution box. The substitution box is invertible.

#### b) Shift Rows:

In shift rows operation the rows in the 4×4 matrix is shifted to left r bits and r varies with the rows of the matrix (r=0 for row1, r=1 for row2, r=2 for row3, r=3 for row 4). This process is illustrated in fig 2. This has the effect of moving positions of lower positions in the row, while the lowest bytes wrap around to the top of the row.

#### c) Mix Columns:

Mix column is calculated using the below formula.

Here a0, a1, a2, a3 is calculated using the polynomials as below

$$a(x) = \{2\}x^3 + \{3\}x^2 + \{1\}x + \{1\}.$$

The mix column transformation operates on the state column by column, treating each column as a four term polynomial. The columns are considered as polynomials over GF (2<sup>8</sup>) and multiplied modulo x<sup>4</sup> + 1 with a fixed polynomial a(x) which is got from the above formula. This can also written as a matrix multiplication

$$s'(x) = a(x) s(x)$$

#### d) Add Round Key:

In the add round key step the 128 bit data is xored with the sub key of the current round using the key expansion operation. The add round key is used in two different places one during the start that is when round r=0 and then during the other rounds that is when 1 ≤ round ≤ Nr, where Nr is the maximum number of rounds. The formula to perform the add round key is S'(x) = S(x) R(x)

where S'(x) – state after adding round key ,S(x) – state before adding round key and R(x) – round key

#### e) Key Expansion:

The key expansion has three steps: Byte Substitution subword(), Rotation rotword() and Xor with RCON (round constant). The input to key schedule is the cipher key K. Key expansion generates a total of Nb(Nr + 1) words. The algorithm requires an initial set of Nb words, and each of the Nr rounds requires Nb words of key data. The resulting key schedule consists of a linear array of 4-byte words, denoted [wi], with i in the range 0 ≤ i < Nb(Nr + 1). The subword() function takes a four byte input and applies the byte substitution operation and produces an output word. The rotword() takes a word [a0, a1, a2, a3] as input and performs a cyclic permutation to produce [a1, a2, a3, a0] as output word. The round constant word array rcon[i] is calculated using the below formula in rijndale finite field.

$$rcon[i] = x^{(2^i - 1)} \text{ mod } x^4 + x + 1$$

The first Nk words of the expanded key are filled with the cipher key. Every following word w[i] is equal to the xor of previous word w[i-1] and the word Nk positions earlier w[i-Nk]. For words in positions that are a multiple of Nk, a transformation is applied to w[i-1] prior to the XOR, followed by an XOR with a round constant Rcon[i]. This transformation consists of a cyclic shift of the bytes in a word rotword() and byte substitution subword().

But in key expansion of 256 -bit cipher if Nk=8 and i-4 is a multiple of Nk then subword() function is applied to w[i-1] prior to the xor. The algorithm for the key expansion routine is given in table 2. Thus with all the above operations the algorithm for the encryption of the data is as follows. Since it begins and ends with the add round key operation there is no wasted unkeyed step in the beginning or the end. The table 3 shows the algorithm for implementation of all three AES encryption.

### Algorithm for key expansion

```

KeyExpansion(byte key[4*Nk], word w[Nb*(Nr+1)], Nk)
begin
word temp
i=0
while(i<Nk)
{
w[i] = word(key[4*i], key[4*i+1], key[4*i+2], key[4*i+3])
i = i+1
}
end while
I = Nk
while (i < Nb * (Nr+1))
{
temp = w[i-1]
if (i mod Nk = 0)
temp = SubWord(RotWord(temp)) xor Rcon[i/Nk]
else if (Nk > 6 and i mod Nk = 4)
temp = SubWord(temp)
end if

w[i] = w[i-Nk] xor temp
i = i + 1
}
end while
end

```

### Algorithm for encryption

```

byte state[4,Nb]
state = in
AddRoundKey(state, key Schedule[0, Nb-1])
for round = 1 step 1 to Nr-1
{
SubBytes(state)
ShiftRows(state)
MixColumns(state)
AddRoundKey(state, key Schedule[round*Nb, (round+1)*Nb-1])
}
SubBytes(state)
ShiftRows(state)
AddRoundKey(state, key Schedule[Nr*Nb, (Nr+1)*Nb-1])
out = state

```

## 2. HARDWARE IMPLEMENTATION OF ENCRYPTION ALGORITHM

Many hardware implementation of encryption algorithm using VHDL is available. In most of the case hardware implementations of AES uses only the AES-128 candidate. Some software implementation of AES192 and AES-256 are available. In the proposed architecture all candidates of AES i.e. AES-128, AES192 and AES-256 are implemented in the same device. The proposed design is implemented using VHDL coding in Xilinx ISE 9.2. Iterative looping techniques is followed to implement the entire design modules of AES encryption algorithm to the

minimize hardware utilization. The key controller unit, key expansion unit, and round function unit and mix column unit everything are implemented in hardware.

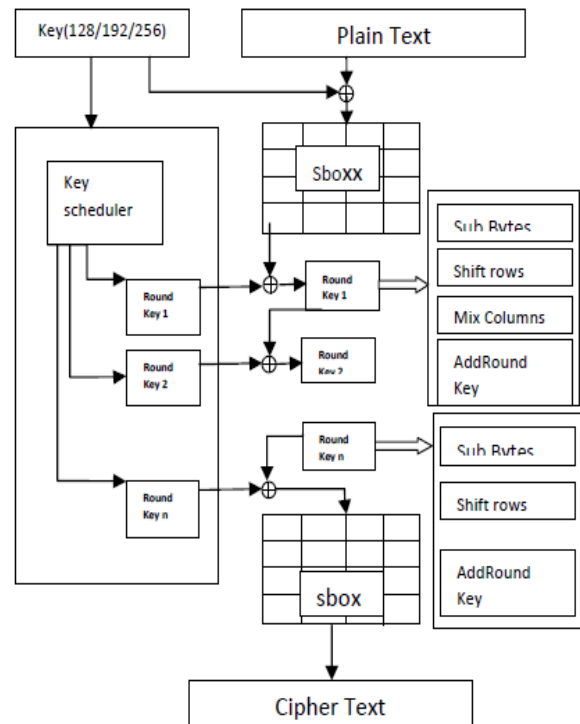


Figure 2. Flow Chart of AES Encryption Implementation

The table 2 Show s that proposed design outperforms all designs based on iterative looping in terms of area and throughput. The performance of AES-192 and AES 256 is also verified and simulation results are given. The novel architecture to implement all AES candidates in same hardware proposed is shown in figure 2. the simulation results of various AES key lengths is shown in figure 3,4 and 5.



Figure 3. Schematic Diagram AES-128/192/256 Architecture

## 3. CONCLUSION

The AES algorithm is an iterative private key symmetric block cipher that can process data block of 128-bits through the use of cipher keys with key length 128,192 and 256 bits. An efficient FPGA implementation of 128 bit block and keys 128, 192 and 256 bits of AES –

Rijndael algorithm has been presented in this paper. Optimized and synthesizable VHDL code is developed for implementation of all AES-128/192/256 bit key encryption and is verified using xilinx ISE 9.2 simulation tool. All the transformations of algorithm are simulated using an iterative design approach in order to minimize the hardware utilization. Thus it can reduce the space by enclosing three different encryption standards in a single architecture and the power consumption can also be reduced which makes it usable in battery operated network devices having Bluetooth and wireless communication devices like software radio. Throughput reaches the value of 666.7Mbps for AES-128 encryption with FPGA device XC2V6000BF957-6.

### REFERENCES

- [1] J. Daemen, V.Rijmen: The Rijndael Block Cipher: AES Proposal : First AES Candidate Conference (AES1) : August 20-22, 1998
- [2] A. Dandalis, V.K. Prasanna, J.D.P. Rolim: A Comparative Study of Performance of AES Candidates Using FPGAs: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [3] T. Ichikawa, T. Kasuya, M. Matsui: Hardware Evaluation of the AES Finalists: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [4] K. Gaj, P. Chodowiec: Comparison of the Hardware Performance of the AES Candidates using Reconfigurable Hardware: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [5] Xilinx VirtexTM-E 1.8V Field Programmable Gate Arrays: URL: <http://www.xilinx.com>: November 2000.
- [6] M.McLoone, J.V. McCanny: Apparatus for Selectably Encrypting and Decrypting Data: UK Patent Application No. 0107592.8: Filed March 2001.
- [7] B. Weeks, M. Bean, T. Rozylowicz, C. Ficke: Hardware Performance Simulations of Round 2 Advanced Encryption Standard Algorithms: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [8] Announcing the ADVANCED ENCRYPTION STANDARD (AES)" Federal Information Processing Standards Publication 197 November 26, 2001

### BIOGRAPHY



**B. Venkata Ramudu** completed B. Tech (CSE) M. Tech (CSE) from JNTU Hyderabad. He is working as associate professor in Malla Reddy Engineering College (Autonomous) Hyderabad. His research includes data mining.



**Sanjeeva Polepaka** completed B. Tech (CSE) from Andhra University and M. Tech (CSE) from Acharya Nagarjuna University. He is pursuing Ph D from JNTU Hyderabad; He is working as associate professor in Malla Reddy Engineering College (Autonomous) Hyderabad. His research includes digital image processing.

# Cloud query Based allocation system in the Dynamic environment of Log Based Personalized Web Mining

Mr. Rajeshwar rao kodipaka\*1, Mr. Sanjeeva Polepaka\*2, Mr.M.Amarendhar Reddy\*3

Asst.Prof, Dept of CSE, MREC-Sec-bad, TS,  
Associate.Prof, Dept of CSE, MREC-Sec-bad, TS,

Asst.Prof, Dept of CSE, B.V. Raju Institute of Technology (UGC-Autonomous), Dist: Medak, TS,

## ABSTRACT:

*Information Technology and its implementation are vast and stretched as the endless ocean. Its beauty lies in the heart of the mother land India, where the main theme gets constructed. For Instance if we keep the Virtual implementation of IT is the cloud which makes us enabling the in the optimization of resource, economy and many more compare to the classical technology. In this paper, we try to put forward the cost reduction system which we call as pay per use and network. In the System of the elasticity which we can as the balancing the data load, making the environment dynamic, number of servers and the optimizing the traffic of the network., where we play an important role in the upgrading of technology which usually we call it as the Technology of change. concept of the caching the first step of reducing the network, making the virtual node which depends on the peak of the data load in order to balance the computing environment which we call as cloud computing. Technologically its meaning remains virtual, but in reality while exposing to a real world of enterprise solution.*

**INDEX TERMS:** Cloud computing, cost efficiency, differential query services, privacy, personalized web search, utility, risk, profile.

## 1.INTRODUCTION

Software as a service (SaaS), which is the most widely known among internet users. In this model, the end-user is using directly the software installed by the cloud provider, which could be an online e-mail service, for instance. we will first take a look at the current state of commercial Web search, since it touches most of our daily lives. Then we discuss related work pertaining to several classes of sanitization mechanisms and corresponding vulnerabilities. For the first part of the related work, we use an informal definition of privacy, namely that

a sanitization mechanism preserves privacy if, given the output of the mechanism, an analyst cannot make a reliable guess as to who a query belongs to. A little later on, we will introduce a more formal definition used in much of the most recent research on the topic. Some of largest US based search engine companies (Google, Yahoo, Microsoft, AOL, and Ask.com) have privacy policies that do little to search data, keeping IP address and browser cookie identifiers associated with searches for as long as 18 months (Schwartz & Cooper, 2007). This last model differs slightly to the previous as the cloud users are the customers of the application, in comparison of the two first models where the cloud users are often businesses. Cloud and data storage service which provides both secure data outsourcing service and efficient data retrieval and repair service, including four different entities: the data owner, the data user, the cloud server, and the third party server. The data owner outsources the encoded fragments of the file M to n cloud servers denoted as storage servers. If the data owner requires keeping the data content confidential, the file can be first encrypted before encoding. Outsourced data are attached by some metadata like verification tags to provide integrity check capability. After the data outsourcing, a data user can select any k storage servers to retrieve encoded segments, and recover the file M, which can be further decrypted in case the file is encrypted. Meanwhile, the third party server periodically checks the integrity of data stored in cloud servers. Failed cloud servers can be repaired with the help of other healthy cloud servers.

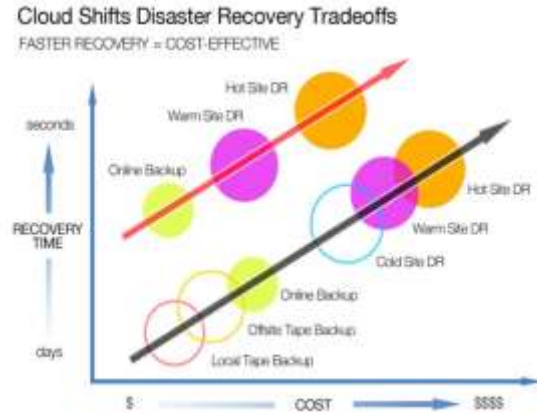


**Fig.1.1 Illustration of the Cloud Cost Model**

Security as malicious software can enter the enterprise network downloaded by the employees browsing the internet. Indeed, malicious tenants could be willing to access data of other tenants or to gain access to their network by using techniques such as ARP cache poisoning or IP spoofing. Furthermore, a configuration of routing appliances could lead to a breach in confidentiality.

## II.RELATED WORK

We can process user logs to build user-specific models, which we call local models. We do not sanitize local models under the assumption that users interacting with their own sensitive information do not constitute breaches of privacy. (We do not consider the case where multiple individuals access the same account, though this may be a common situation with, e.g., shared computers in homes, libraries, or Internet. Regarding data confidentiality, all three providers claim that they provide and support encryption. Many cloud providers allow their customers to encrypt data before sending it to the cloud. For instance, Amazon S3 makes it optional to users to encrypt data for additional security, but not recommended in the case of third party or external auditors. Cloud front is an Amazon web service which provides data confidentiality while being transferred. Attempts to apply functional repair in the LT codes based distributed storage should first solve how to recode packets, because the random linear recoding in the functional repair of network coding-based storage codes cannot satisfy the degree distribution in LT codes.



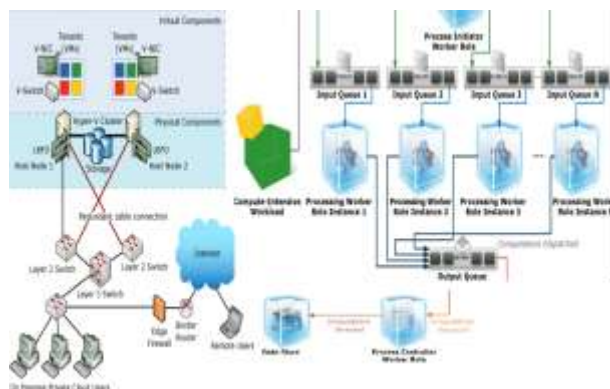
**Fig.2.1 Cloud Balance in the Traffic**

A tenant could be attacked by the cloud provider or by other clients. The cloud provider could be attacked by tenants. Our project does not consider the situation where the provider is malicious. Indeed, we suppose that the cloud provider is honest. We further suppose that all network appliances are compliant and secure. It seems that this problem can be solved by utilizing the recently proposed LT network codes (LTNC) which provides efficient decoding at the cost of slightly more communication in the single-source broadcasting scenario. However, after several rounds of repair with same recoding operations regulated in LT network codes, data users experience decoding failure with high probability. Our project aims to secure the traffic of a tenant. In this situation, this traffic could first be threatened by other tenants.

## III.PROPOSED METHODOLOGY

Search logs, or query logs as they are sometimes called, are databases containing information about user search activity. A typical log might contain a set of entries consisting of an identifier, a time stamp, an event descriptor, and pertinent information about the event. Common identifiers are IP addresses, cookie identifiers, or account numbers. Example event descriptions include query or click. Pertinent information about the event might include the query text or a clicked URL. When we assume that all identical identifiers correspond to a single individual, we refer to a collection of search log entries sharing the same identifier as a user search log or user log for short.

To begin with, the dedicated hardware, even though it may be inactive at a given time, has to be monopolized for the tenants who need the highest security. In other words, the pooling of resources would be limited which may lead to an oversized network. Indeed, the dedicated servers could have hosted other tenants' Virtual Machines. However, these VMs would have to be located on other servers, contributing to the waste of resources, as the servers would not be loaded at their maximal capacity. Furthermore, as the use of the cloud would grow, we can easily imagine that the expansion of such a rigid architecture would become harder and harder to maintain and configure. The Proof Generation process is run by the storage provider in order to generate proof transcript which could be a set of corresponding tags or, in many cases, an aggregation of the perturbed tags and aggregation of data blocks. A proof transcript allows the verifier to derive the proof in order to check the integrity of the challenged blocks. Upon receiving the response from the storage provider, the verifier executes the Proof Verification protocol to verify the validity.



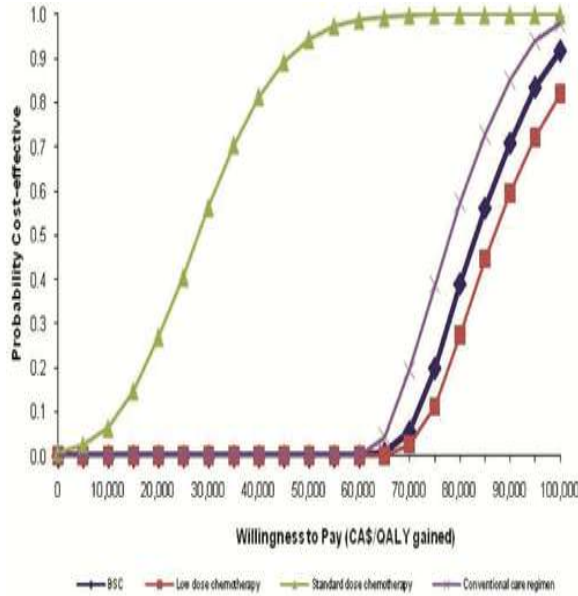
**Fig.3.1 Node Values Cloud Computing Peak Architect of High End Security.**

When enterprises decide to move to the cloud, they want to keep the same requirements regarding their policy management. It would be possible for the network manager of the enterprise to implement the middle boxes and the routing policies on VMs in the same way as before moving to the cloud. However, one of the goal of moving to the cloud is to escape the burden of network administration and configuration. Furthermore, the type of security policies in place in enterprises networks is often quite

similar as it consists in the traversal of several middle boxes. These data centers are located in Northern Virginia (USA) and the company's cluster of cloud computing services in Virginia were "currently experiencing degraded performance" as indicated on the Amazon website. Consideration of both data repair and data retrieval, and design a LT codes-based cloud storage service (LTCS). Multi-keyword ranked search over encrypted cloud data, and establish a variety of privacy requirements. Among various multi-keyword semantics, we choose the efficient similarity measure of "coordinate matching", i.e., as many matches as possible, to effectively capture the relevance of outsourced documents to the query keywords, and use "inner product similarity" to quantitatively evaluate such similarity measure.

#### IV. ANALYSIS AND INTERPRETATION

Today's data centers are already containing hundreds of thousands of servers, and this number is very likely to increase during the forthcoming years. Furthermore, the customers can request the creation or the removal of VMs in order to meet their similar incident happened about four months earlier due to an electrical storm that caused some disturbance to the same data centers. In the field of IR, personalization is the act of incorporating information about a user into the processing of an IR application. For example, boosting the rank of search results that are more similar to web sites a user has visited in the past. Another common practice is the use of a user's location, often inferred by IP address or GPS coordinates. We concentrate on local-only personalization that can be performed in isolation on an individual's computer. We consider two different types of personalization: task-aware and supervised learning from user-provided annotations. Note that, in general, personalization can also involve a group of users



**Fig.3.1.1 Comparison of Dataset and Security feature with Size and Cost**

Thus, cloud customers are enthusiastic to be allowed to store their data in the cloud and at the same time they would like to be able to check by themselves that their data is protected.

## V.CONCLUSION

In the modeling feature of the cloud computing as pay per use which makes us to in making the system everything global and making the smaller level organization as base to as use. It is distinguished with the traditional multi-domain approaches by the characteristics of the emerging cloud environments. We use a top-down methodology in the research referring the PEI stack. Secure and reliable cloud storage with the efficiency. By mapping users to multiple pseudo users, it is more difficult to make the same links that the AOL data allowed. First, queries issued on two different days are not directly connected (provided the user was inactive for part of the intervening time period). Second, sessions generally consist of a small number of queries and are therefore less likely to cover a diverse set of topics. When these two points are taken together, it becomes more difficult to gather sufficient evidence to identify the individuals behind the queries.

## VI.REFERENCES

- [1] Thomas S. J. Schwarz and Ethan L. Miller. Store, forget, and check: Using algebraic signatures to check remotely administered storage. In ICDCS '06: Proceedings of the 26th IEEE International Conference on Distributed Computing Systems, Washington, DC, USA, 2006.
- [2] F. Sebe', J. Domingo-Ferrer, A. Martinez-Balleste, Y. Deswarte, and J.-J. Quisquater, "Efficient Remote Data Possession Checking in Critical Information Infrastructures," IEEE Trans. Knowledge Data Eng., vol. 20, no. 8, pp. 1034-1038, Aug. 2008.
- [3] G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," Proc. Fourth Int'l Conf. Security Privacy Comm. Networks, pp. 1-10, 2008.
- [4] C. Erway, A. Ku' pc,u", C. Papamanthou, and R. Tamassia, "Dynamic Provable Data Possession," Proc. 16th ACM Conf. Computer Comm. Security, pp. 213-222, 2009.
- [5] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling Public Verifiability and Data Dynamics for Storage Security in Cloud Computing," Proc. 14th European Conf. Research Computer Security, pp. 355-370, 2009.
- [6] A.F. Barsoum and M.A. Hasan, "Provable Possession and Replication of Data over Cloud Servers," Technical Report 2010/32, Centre for Applied Cryptographic Research, <http://www.cacr.math.uwaterloo.ca/techreports/2010/cacr2010-32.pdf>, 2010.
- [7] R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession," Proc. 28th Int'l Conf. Distributed Computing Systems, pp. 411-420, 2008.
- [8] A.F. Barsoum and M.A. Hasan, "On Verifying Dynamic Multiple Data Copies over Cloud Servers," Technical Report 2011/447, Cryptology Eprint Archive, <http://eprint.iacr.org/>, 2011.
- [9] K.D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," Proc. 16th ACM Conf. Computer Comm. Security, pp. 187-198, 2009.
- [10] Y. Dodis, S. Vadhan, and D. Wichs, "Proofs of Retrievability via Hardness Amplification," Proc. Sixth Theory Cryptography Conf. Theory Cryptography, 2009.
- [11] A. Juels and B.S. Kaliski, "PORs: Proofs of Retrievability for Large Files," Proc. 14th ACM Conf. Computer Comm. Security, pp. 584-597, 2007.
- [12] H. Shacham and B. Waters, "Compact Proofs of Retrievability," Proc. 14th Int'l Conf. Theory Appl. Cryptology Information Security, pp. 90-107, 2008.
- [13] M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu, "Plutus: Scalable Secure File Sharing on Untrusted Storage," Proc. Second USENIX Conf. File Storage Technologies, 2003.
- [14] E.-J. Goh, H. Shacham, N. Modadugu, and D. Boneh, "SiRiUS: Securing Remote entrusted Storage," Proc. Network Distributed System Security Symp., 2003.

## AUTHORS PROFILE:



**1. Rajeshwarrao. Kodipaka** (ISTE,CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.



2. **Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer networks.



3. **M. AMARENDHAR REDDY** is presently associated with B. V. Raju Institute of Technology (UGC - Autonomous), Narsapur, Medak (Dist), Telangana State as Asst. Professor in

Information technology Department. He graduated from JNTU Hyderabad in B.Tech CS & IT and completed M.Tech (SE) from JNTUH-School of Information Technology Campus. Presently with 10 years of rich experience in Teaching, Administration and Research in the area of Object Oriented Modeling, Software Engineering Data Warehousing & Mining and Programming Languages. Prior to his current position with B. V. Raju Institute of Technology he served multiple positions as I/C HOD, Exam Branch In charge and Students' Advisor in various leading organization. His teaching experience includes teaching UG and PG course on Object oriented analysis and design, Java, C, C++, Software Engineering, Formal Languages and Automata Theory and Automata and Compiler Design. As part of Students' Advisor responsibilities he conducted workshops on Data mining and STQA.

# Privacy and Social Security Mechanism in the Era of Social Multi Cloud Environment

Mr. Rajeshwar rao kodipaka\*<sup>1</sup>, Mr. Sanjeeva Polepaka\*<sup>2</sup>

*Asst.Prof, Dept of CSE, MREC-Sec-bad, TS.*

*Associate. Prof, Dept of CSE, MREC-Sec-bad, TS.*

## ABSTRACT:

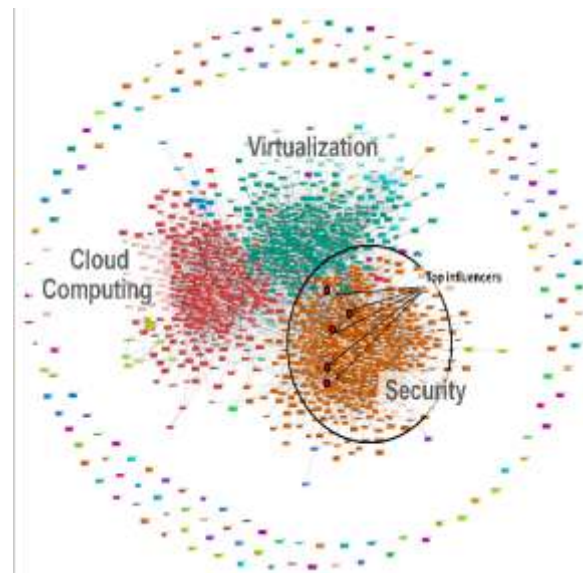
Time and the Constraint are both typically interrelated and lead us to the next level high technological revolution journey where we analyse the Human Automation instead of System Automation. Technological in the View Point cognitive science Information technology will have to make high level of demanding journey; in that aspect we consider the best of the technical revolution in the Social media where we share the thoughts, informational to many more. If we consider the Face book like such popular network sharing side leads us to the best of research, where technological Face book is high level to manage such huge volume of Data effectively and efficiently, But Lacking with the point of loop whole is the Information of Public and Private and Preventive social Inference of the Aspect. Hence; we look forward to put the paper as the most and best to best Information flow in the aspect of content and curative illustrative technological solution. However, the precision of these solutions varies greatly, as the isolation can be made from a complete separation of physical network, recent breakthroughs in Software-Defined Networking have let us to believe that the most suitable technology for our project today is the use of the technology as it provides an isolation per of as well as a automatic centralized control of the network.

**KEYWORDS:** Social network analysis, data mining, social network privacy, security, privacy, multi-cloud, application partitioning.

## 1.INTRODUCTION

In the Revolution n of Social Media, by structurally mapping out relationships between software developers, the presence of power-law distributions was found for project sizes, cluster sizes of connected developers, and the number of projects joined by developers. et al. Conclude Open Source software development can be modeled as self-organizing, collaboration, social

networks. It et al. explored the statistics and topological information of the Open Source software developer's collaboration network further by extracting project evolution parameters, by inspecting the network over a time period of two years. 50,000 projects involving 80,000 developers were investigated. Again, power-law distributions were found for the cluster distribution and degree distribution, which is the amount of ties within a network. Also they found during this time frame the average degree of projects on Source Forge slightly increased, though the network diameter slightly decreased. Face book, which is currently the world's most popular OSN, opened to the general public on 26 Sep 2006.



**Fig.2.1 Illustration of Data Sharing in the Cloud Security**

Prior to that, users were required to have a verifiable e-mail address from a "network" such as a university. Suppose a University of Cambridge student named Alice created a Face book account in August 2006 and immediately limited the sharing of all personal information to direct "friends" only, but left

her “search ability” setting at “Everyone” so that friends could search for her and send friend requests. Without this setting, it was very difficult for real-life friends to find each other on Face book: it was impossible for two users with the most restrictive search ability settings to “friend” each other, which discouraged users from choosing such settings.

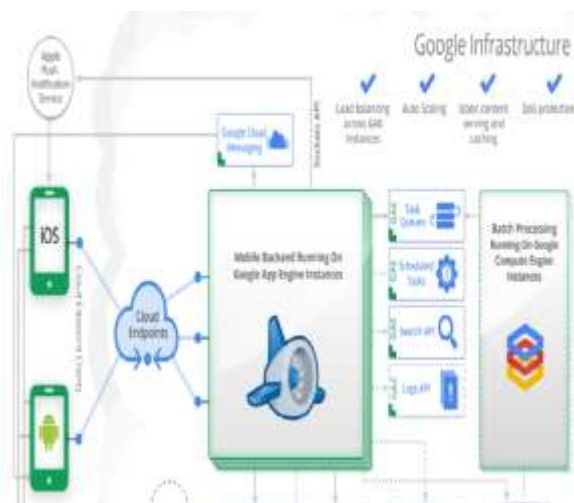
## II. RELATED WORK

Taken social network theory into account the organizational network structure affects team viability and team task performance. Team viability involves the committed of team members to stay together, where team task performance involves attaining the goals of a team. Team viability and team task performance can be conceived as success measures of teams. and Harrison noted although these two dimensions of success are conceptually distinct, in reality there is a close connection and cross-correlation between team task performance and team viability. In addition they found teams with central leaders, and teams central in a network full of other teams, tend to be better performers. Today's OSNs provide rich platforms for third-party social applications.

being given it anyway. Several applications used information in ways that contravened Facebook's Terms of Service, making information visible to users who would normally be unable to view it. People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time. **Mark Zuckerberg, January 2010.**

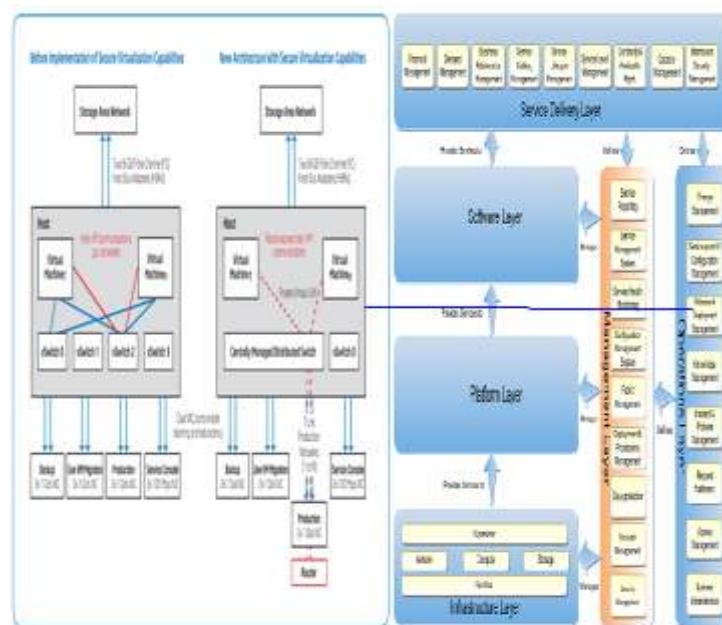
### III. PROPOSED METHODOLOGY

Today's centralized online social networks (OSNs) are trusted to make access control decisions on behalf of users, but those decisions do not always align with users' intentions. In some cases, technical flaws lead to data leakage contrary to the wishes of both users and operators. In other cases, there is a mismatch between user understanding of privacy settings and the actual protections afforded by the system. In many cases, however, the interests of the user and the network are simply not aligned. In these cases, when users' data must be safeguarded by operators without the incentive to do so, private information has a history of leaking out.



**Fig.2.1. Mobile Cloud Social Security Model**

Application functionality is supported by private user information, often served in greater quantities than actually required. Safeguards against misuse rely heavily on legal rather than technical protection. Felt and Evans, in their 2008 study of 150 Facebook applications [103], found that 90% of applications had no need of private user data, but were

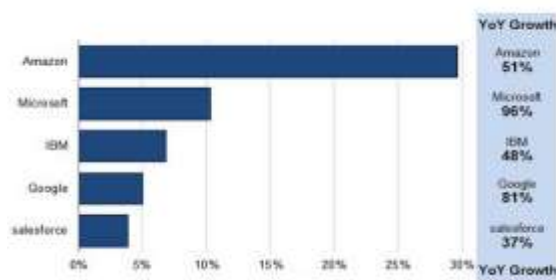


**Fig.3.1 High Level Flow of Role, Language and Typically the Privacy Level**

Most of these forms of communication are not archived and therefore difficult to investigate. Next, the nature of Open Source software communities is to be open and transparent. Therefore, the project's public forums are taken into account as the investigated conversational network. On the public forums registered members can post messages and react on each other, resulting in forum threads. Registered members form the nodes of the conversational network. And, when two members of a project group post in the same thread, a relationship between those two members, or a link between those two nodes, exists. Finally, the several orthogonal cognitive considerations in social networks generally, describing them as problems with “decentralized systems”. For instance, the authors state that “more control over personal data almost inevitably translates to more decisions, which leads to cognitive overload”, but this is not a reason to avoid decentralized architectures. Centralized services such as Facebook provide very detailed and fine-grained control over some personal data, capable of causing just as much cognitive overload as an equivalent decision matrix in a decentralized system. What Facebook does not have is an incentive to explore alternative schemes for privacy policy configuration if those schemes hinder a “free flow of information” within the network. In contrast, a distributed system that does not extract monetary value from user data has every incentive to combat cognitive overload by accurately capturing user intent, reducing decision fatigue and improving the user experience of privacy management. While the authors of this critique raise some important issues, their case against decentralized data architectures is greatly overstated.

#### IV.ANALYSIS

In this we used between ness centrality to measure the effectiveness with which an attacker might influence information moving through a social network.



**Fig.3.1.1 Content Highlighting to the Privacy Based on the Organisation**

Suppose an attacker can compromise  $N$  nodes in a social graph and use them to influence information that passes through them. If the attacker can choose which  $N$  nodes to compromise, the optimal attack strategy is to choose the most central nodes in the network, yielding the highest probability of being able to influence any communication path. If the attacker has imperfect centrality information, the probability of successful attack will be reduced.

#### V.CONCLUSION AND FUTURE WORK

Technologically Each and every Aspect of Information has its own loop whole. Blocks are encrypted with block-specific symmetric keys; their content is only readable by principals which possess the relevant keys. Blocks in the encrypted store contain explicit links to other blocks, which are only visible to principals that possess the block's symmetric key. The system also provides opportunities for covert communication, including a low-bitrate perfectly unobservable communications. Some difficulties can be obviated under the random oracle assumption, but proofs in the standard model would be a natural direction to pursue. It is desirable to prove that, first, convergent encryption provides in distinguishability.

#### VI.REFERENCES

- [1] Facebook Beacon, 2007.
- [2] T. Zeller, "AOL Executive Quits After Posting of Search Data," The New York Times, no. 22, <http://www.nytimes.com/2006/08/22/technology/22iht-aol.2558731.html?pagewanted=all&r=0>, Aug.2006.
- [3] K.M. Heussner, "Gaydar' n Facebook: Can Your Friends Reveal Sexual Orientation?" ABC News, <http://abcnews.go.com/Technology/gaydar-facebook-friends/story?id=8633224#>, UZ939UqheOs, Sept. 2009.
- [4] C. Johnson, "Project Gaydar," The Boston Globe, Sept. 2009.
- [5] L. Backstrom, C. Dwork, and J. Kleinberg, "Wherefore Art Thou r3579x?: Anonymized Social Networks, Hidden Patterns, and Structural Steganography," Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 181-190, 2007.
- [6] M. Hay, G. Miklau, D. Jensen, P. Weis, and S. Srivastava, "Anonymizing Social Networks," Technical Report 07-19, Univ. of Massachusetts Amherst, 2007.
- [7] K. Liu and E. Terzi, "Towards Identity Anonymization on Graphs," Proc. ACM SIGMOD Int'l Conf. Management of Data (SIGMOD '08), pp. 93-106, 2008.
- [8] J. He, W. Chu, and V. Liu, "Inferring Privacy Information from Social Networks," Proc. Intelligence and Security Informatics, 2006.
- [9] E. Zheleva and L. Getoor, "Preserving the Privacy of Sensitive Relationships in Graph Data," Proc. First ACM SIGKDD Int'l Conf. Privacy, Security, and Trust in KDD, pp. 153-171, 2008.
- [10] R. Gross, A. Acquisti, and J.H. Heinz, "Information Revelation and Privacy in Online Social Networks," Proc. ACM Workshop Privacy in the Electronic Soc. (WPES '05), pp. 71-80, <http://dx.doi.org/10.1145/1102199.1102214>, 2005.

- [11] H. Jones and J.H. Soltren, "Facebook: Threats to Privacy," technical report, Massachusetts Inst. of Technology, 2005.
- [12] P. Sen and L. Getoor, "Link-Based Classification," Technical Report CS-TR-4858, Univ. of Maryland, Feb. 2007.
- [13] B. Tasker, P. Abbeel, and K. Daphne, "Discriminative Probabilistic Models for Relational Data," Proc. 18th Ann. Conf. Uncertainty in Artificial Intelligence (UAI '02), pp. 485-492, 2002.
- [14] A. Menon and C. Elkan, "Predicting Labels for Dyadic Data," Data Mining and Knowledge Discovery, vol. 21, pp. 327-343, 2010.
- [15] E. Zheleva and L. Getoor, "To Join or Not to Join: The Illusion of Privacy in Social Networks with Mixed Public and Private user Profiles," Technical Report CS-TR-4926, Univ. of Maryland, College Park, July 2008.
- [16] N. Talukder, M. Ouzzani, A.K. Elmagarmid, H. Elmeleegy, and M. Yakout, "Privometer: Privacy Protection in Social Networks," Proc. IEEE 26th Int'l Conf. Data Eng. Workshops (ICDE '10), pp. 266- 269, 2010.

#### **AUTHORS PROFILE:**



**1. Rajeshwarrao. Kodipaka** (ISTE, CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested

domain is Data mining, cloud computing, Computer Networks.



**2. Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer network.

# An Effective Approach for Maintaining High Accurateness in Cloud Environment

Mr. Rajeshwar Rao Kodipaka<sup>1</sup>, Mr. M. Amarendhar Reddy<sup>2</sup>, B. Venkat Ramudu<sup>3</sup>

Asst. Professor, Department of CSE, MallaReddy Engineering College (Autonomous), Hyderabad, India<sup>1,3</sup>

Asst. Professor, Department of IT, B.V. Raju Institute of Technology, Medak, India<sup>2</sup>

**Abstract:** Generally Software-as-a-service clouds were built on software as a service along with service-oriented architecture. This assists the providers of application service in conveying their applications by means of substantial cloud computing infrastructure. In the traditional efforts while confidentiality as well as privacy protection exertions was broadly studied, the problem of service integrity attestation has not been appropriately addressed. In our work we mainly spotlight on data processing services which have turned out to be popular with applications in numerous usage areas. We present IntTest, which is a scalable and efficient integrated service integrity attestation structure intended for software-as-a-service cloud multitenant cloud systems. The proposed approach offers a novel integrated attestation graph analysis method which can not only identify attackers more resourcefully by taking an integrated approach, but also can hold back aggressive attackers and limit the scope of damage that is caused by colluding attacks.

**Keywords:** Software-as-a-service, Integrated attestation graph analysis, Cloud computing, IntTest, Attackers.

## 1. INTRODUCTION

Infrastructures of cloud computing are usually shared by application service providers from various domains of security which make them exposed to malicious attacks. Service integrity is the most prevailing problem that has to be addressed regardless of processing of public or private data by cloud system [1]. In major multitenant systems of cloud, numerous malicious attackers might commence colluding attacks on definite targeted service functions to nullify the assumption. In our work we focus on services of data streaming applications for clouds with numerous real-world applications. In our work to address this challenge we present IntTest, which is an integrated service integrity attestation structure intended for multitenant cloud systems. IntTest put forward result auto-correction that can automatically restore corrupted data processing results that are produced by malevolent attackers with superior results that are produced by benign service providers. IntTest approach considers a holistic approach by thoroughly examining consistency as well as inconsistency relationships between several service providers within complete cloud system. IntTest offers a realistic service integrity attestation system that does not imagine trustworthy entities on third-party service.

services might attract several service providers for profit. To maintain automatic service composition, we can install a set of portal nodes that serve as the gateway for user to access composed services in Software-as-a-service cloud. The portal node can combine various service components into composite services on the basis of user's needs. Altered from other distributed systems for instance peer-to-peer networks as well as volunteer computing setting, Software-as-a-service cloud systems acquire a set of distinctive features. Application service providers of third-party usually do not want to make known internal functioning details of their software services for protection of intellectual property. Both cloud infrastructure providers as well as third-party providers are independent entities. It is not practical to enforce any special hardware or else protected kernel support on individual service provisioning sites. For privacy fortification, only portal nodes encompass global information regarding which service functions are offered by which service providers in Software-as-a-service cloud. Our work draw attention in the direction of on data processing services which have turned out to be more and more popular with applications in numerous real-world usage areas.

## 2. MODELLING OF CLOUD MODEL OF SAAS

Software-as-a-service clouds mostly were put up on concepts of software as a service along with service-oriented architecture which facilitate application service providers in conveying their applications by means of substantial cloud computing infrastructure. In a comprehensive Software-as-a-service cloud, identical service function can be offered by various application service providers [2][3]. These components of functionally equivalent service exist due to service providers might generate replicated service components in support of load Balancing as well as fault tolerance purposes; and popular

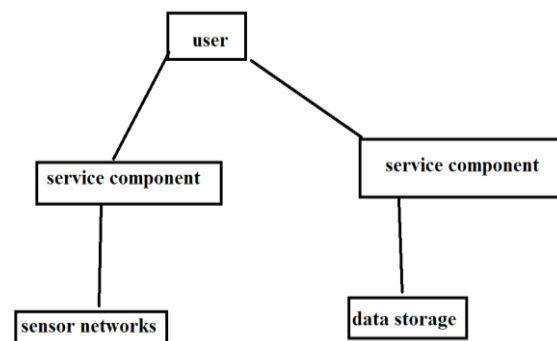


Fig1: An overview of service integrity attacks.

In our work present IntTest, which is a scalable and efficient integrated service integrity attestation structure intended for software-as-a-service cloud multitenant cloud systems. IntTest offers a novel integrated attestation graph analysis method that can make available stronger attacker pinpointing power than earlier schemes. Given a Software-as-a-service cloud system, objective of IntTest is to identify any malevolent service provider that offers a misleading service function [4]. IntTest considers all service components as black boxes, which does not necessitate any special hardware on cloud platform.

### 3. AN OVERVIEW OF PROPOSED IntTest SYSTEM

IntTest put up on earlier work and AdapTest but can grant well-built malicious attacker pinpointing control than RunTest and AdapTest. RunTest and AdapTest as well as conventional majority voting methods need to imagine that providers of benign service receive majority in each service function. IntTest scrutinize per-function consistency graphs as well as global inconsistency graph. The analysis of per-function consistency graph can limit extent of damage that is caused by colluding attackers, while global inconsistency graph analysis can efficiently depict those attackers that attempt to compromise numerous service functions for this reason, IntTest can still identify malicious attackers although they turn out to be majority for several service function. IntTest offer result auto-correction that can automatically restore corrupted data processing results that are produced by malevolent attackers with superior results that are produced by benign service providers. It presents a realistic service integrity attestation system that does not imagine trustworthy entities on third-party service. By taking an integrated approach, IntTest can not only identify attackers more resourcefully but also can hold back aggressive attackers and limit the scope of damage that is caused by colluding attacks. In our work we focus on services of data streaming applications for clouds with numerous real-world applications. Our work spotlight on data processing services which have turned out to be more and more popular with applications in numerous real-world usage areas [5]. IntTest can not only identify malicious service providers but moreover automatically spot on corrupted data processing results to get better result quality of cloud data processing service. The offers a novel integrated attestation graph analysis method that can make available stronger attacker pinpointing power than earlier schemes. To distinguish service integrity attack as well as pinpoint malicious service contributor, our algorithm depends on replay-based consistency check to obtain consistency/inconsistency relationships among service providers. The intuition following our method is that when two service providers differ with each other on processing result of the similar input, not less than one of them has to be malevolent. We do not forward an input data item as well as its duplicates simultaneously. As a substitute, we attestation data on several service providers after receiving processing result original data consequently, malicious attackers cannot avoid risk of being detected when they construct false results on original data [6]. By

means of replay-based consistency check, we can check functionally equivalent service providers and get hold of their consistency as well as inconsistency relationships.

### 4. CONCLUSION

We draw attention towards services of data stream processing that are considered for clouds with abundant applications. Our work put forward an effective and scalable IntTest, which is efficient integrated service integrity attestation structure intended for software-as-a-service cloud multitenant cloud systems. Scalable IntTest approach takes for consideration a holistic approach by thoroughly examining consistency as well as inconsistency relationships between several service providers within complete cloud system. The proposed system offer result auto-correction that can automatically restore corrupted data processing results that are produced by malevolent attackers with superior results that are produced by benign service providers. The system put forward realistic service integrity attestation system that does not imagine trustworthy entities on third-party service. To make a distinction of service integrity attack as well as pinpoint malicious service contributor, our algorithm depends on replay-based consistency check to obtain consistency/inconsistency relationships among providers of service.

### REFERENCES

- [1] L.Alchaal, V.Roca, and M.Habert, "Managing and Securing Web Services with VPNs," Proc. IEEE Int'l Conf. Web Services, pp. 236- 243, June 2004.
- [2] H. Zhang, M.Savoie, S.Campbell, S.Figuerola, G.von Bochmann, and B.S.Arnaud, "Service-Oriented Virtual Private Networks for Grid Applications," Proc. IEEE Int'l Conf. Web Services, pp. 944-951, July 2007.
- [3] M. Burnside and A.D. Keromytis, "F3ildCrypt: End-to-End Protection of Sensitive Information in Web Services," Proc. 12<sup>th</sup> Int'l Conf. Information Security (ISC), pp. 491-506, 2009.
- [4] J.L. Griffin, T. Jaeger, R. Perez, and R. Sailer, "Trusted Virtual Domains: Toward Secure Distributed Services," Proc. First Workshop Hot Topics in System Dependability, June 2005.
- [5] L. Lamport, R. Shostak, and M. Pease, "The Byzantine Generals Problem," ACM Trans. Programming Languages and Systems, vol. 4, no. 3, pp. 382-401, 1982.
- [6] T. Ho et al., "Byzantine Modification Detection in Multicast Networks Using Randomized Network Coding," Proc. IEEE Int'l Symp. Information Theory (ISIT), 2004.

### BIOGRAPHIES

**MR. RAJESHWAR RAO KODIPAKA** (ISTE, CSI LIFE MEMBER), Working as Asst. Professor, in Department of CSE in MallaReddy Engineering College (Autonomous), Hyderabad. He has vast experience in teaching. His research includes data mining; Cloud computing, Software Engineering and Computer networks.

**M. AMARENDHAR REDDY** is presently associated with B. V. Raju Institute of Technology (UGC - Autonomous), Narsapur, Medak (Dist), Telangana State as Asst. Professor in Information technology Department with 10 years of rich experience in Teaching, Administration and Research in the area of Object Oriented Modeling, Software Engineering Data Warehousing & Mining and Programming Languages.

**B VENKATA RAMUDU** is working as Asst. Professor Malla Reddy Engineering College (Autonomous), Hyderabad. His research interests in Computer Networks, Cryptography and security.

# Security and Privacy as data Service in the Multi-cloud Node Environment

Mr. Rajeshwar rao kodipaka\*1, Mr. Sanjeeva Polepaka\*2, Mr.M.Amarendhar Reddy\*3

*Asst.Prof, Dept of CSE, MREC-Sec-bad, TS.*

*Associate.Prof, Dept of CSE, MREC-Sec-bad, TS*

*Asst.Prof, Dept of CSE, B.V. Raju Institute of Technology (UGC-Autonomous), Dist: Medak, TS*

## ABSTRACT:

Technology and its significance have changed the global world to the next level of the IT Village. In the Context of the revolution of the Information Technology , we have gone to the next level of computing to make the Live easier , smooth , faster and accurate. Hence this terminology is to heard but to achieve is typical. Hence, In this Paper, We have kept for the next generation computing scaling like Cloud Computing with next residual of the Security. Hence, in the current context we have given the glimpse of the time and material to make cloud feasible from the security point of view for the usability of the domains like Banking and financial with high end feasible and scalable solution. In this we have implemented the node based Architectural model with acknowledgement encrypted based on the level of the data to these data moved in the network with the classification of the node based data in the preview model of the datacenter to monitor the traffic of the data and the shortest path Mechanism.

**INDEX TERMS:** cloud, security, privacy, multi-cloud, application partitioning, tier partitioning, data partitioning, multi party computation.

## 1.INTRODUCTION

Technology and its advancement lead us to research for the next level of the advancement. In the context of the Cloud computing which we can tell as the technological advancement plays the vital role in the industry of Information Technology. In the modern age of the cloud computing generation, the idea of cloud computing is almost as old as the computer itself. Its principle is to have the user's computer, Smartphone, tablet or any internet-connected device acting as a front-end displaying an application, as the resources connected in remote servers. Regarding

data confidentiality, all three providers claim that they provide and support encryption. Many cloud providers allow their customers to encrypt data before sending it to the cloud. For instance, Amazon S3 makes it optional to users to encrypt data for additional security, but not recommended in the case of third party or external auditors depending on the model of service they are offering. The first model is the Infrastructure as a service (IaaS), which is the most basic model. The IaaS model consists in providing Virtual Machines to the clients whom will have to install operating systems as well as their applications on top of the vendors providers like Amazon EC2, Google Compute Engine .Today's data centers are already containing thousands of servers, and this number is very likely to increase during the forthcoming years. Furthermore, the customers can request the creation or the removal node to meet their needs.

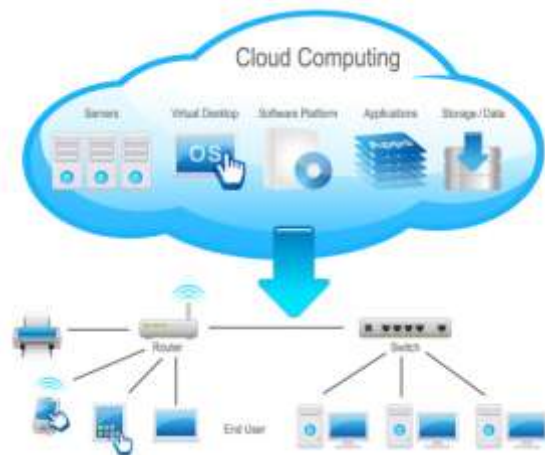


**Fig.1.1. Tiered structure of the cloud**

Therefore, the networks are highly elastic and can reach an immense number of traffic. Because of that, networks architecture has to be scalable in order to ensure a good performance whatever number of VIRTUAL is running in the network.

## II. RELATED WORK

In the IaaS model, the customers dynamically scale up and down to as many machines as needed inside the cloud in a “pay-as-use” manner. The clients can dynamically provision resources to meet the current demand by adjusting their leasing of resources from the cloud provider. Customers can also then use more efficient hardware without being preoccupied by its maintenance, cooling and storage. In each model, the cloud providers can use multi-tenancy, where virtual machines from multiple customers can share the same sets of physical servers and the network, in order to reduce the waste of resources. However, economies of scale are so important that their users benefit from both aspects like efficient and cheaper solution.



**Fig.2.1. Illustration of the Centralized cloud node**

Due to the of an ever growing number of people connected to the internet, the constant evolution of computer science along with the decreasing of hardware prices, the use of cloud computing instead of classic networks architecture has recently become a reality. In order to optimize the resources in the data centers or for maintenance reasons, cloud providers often have to migrate from the physical server to another. Migrations are used in order to pool the active and then reduce the number of turned on physical servers inside the data center, in order to reduce cost. Migrations also reveal themselves handy when there is a hardware problem on a physical server, for example. The running on this server is moved to other physical servers and the alleged

deficient server can be turned off and replaced. The impact of node migration on the network is important. Since the virtualization is potentially moving from a physical server to another, the routes between them are also impacted, and the topology of the network is thus changing. Most importantly, the end points of the routes are changing. This causes the need to reconfigure the network with each migration.

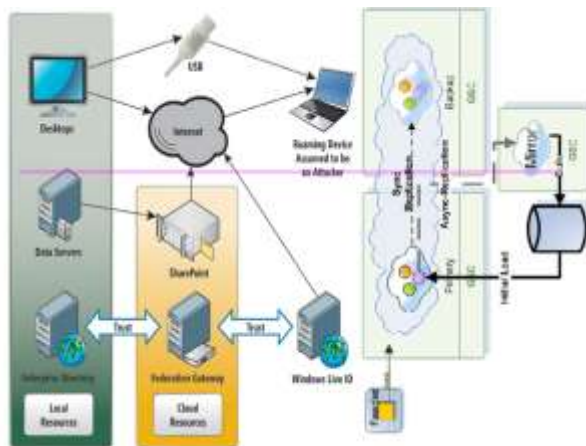
## III. PROPOSED METHODOLOGY

In this paper, we have given emphasis on the Privacy and Security involved in the multi-cloud architecture. The big threat to cloud is the security which has not yet robust and need much more advancement to give the client that is the best solution especially in the industry domain of banking and financial services. The advancement of the technology and its usefulness makes us to research the best of the best service for the automation world more precisely the cloud computing security can be breached by several actors. In the cloud, there are the cloud provider and the tenants. Our project aims to secure the traffic of a tenant. In this situation, this traffic could first be threatened by other tenants. Indeed, malicious tenants could be willing to access data of other tenants or to gain access to their network by using techniques such as ARP cache poisoning or IP spoofing. Furthermore, a configuration of routing appliances could lead to a breach in confidentiality. The other threat comes from inside the tenant's network itself. This threat is the most important one in security today, as malicious software can enter the enterprise network downloaded by the employees browsing the internet. In the enterprise network, the traffic is secured by the traversal of virtual middle boxes along the path. Our project aims to create a framework enabling the traversal of virtual middle boxes despite the elastic nature of the cloud, meanwhile providing isolation of traffic in order to prevent the risk of an attack by another tenant. The switches between the source switch and the first ingress switch will all have the same type of rule. Based on the, the packet will be routed to a particular port.

This port has been calculated during the route calculation. At the ingress switch of a virtual middle box, the rule will consist in matching one

attribute and applying two actions. The matching attribute is obviously the tag corresponding to the virtual middle boxes. The packet must then be sent to the virtual middle boxes, but before that, the tag must be popped, so the virtual middle boxes receive an unmodified packet. The progress switch of the virtual middle boxes will receive an untagged packet. The policy management in a secure enterprise network today can therefore be quite complex. For instance, it may require restricting a machine containing sensitive data to be accessed only by a small group of users, or preventing external traffic from directly reaching internal servers.

The actual realization may involve servers having complex communication patterns governed by network access control, such as the traversal of several virtual middle boxes before being reached. When enterprises decide to move to the cloud, they want to keep the same requirements regarding their policy management. It would be possible for the network manager of the enterprise to implement the virtual middle boxes and the routing policies on in the same way as before moving to the cloud. However, one of the goals of moving to the cloud is to escape the burden of network administration and configuration. Furthermore, the type of security policies in place in enterprises networks is often quite similar as it consists in the traversal of several virtual middle boxes.



**Fig.3.1. Multi- Cloud Architecture design involves the Agility platform**

A tenant could be attacked by the cloud provider or by other clients. The cloud provider could be attacked

by tenants. Our project does not consider the situation where the provider is malicious. Indeed, we suppose that the cloud provider is honest. We further suppose that all network appliances are compliant and secure. Based on the source and destination field, as well as the port from which the packet arrived, the controller will have to push the down to the packet. A second rule consists in matching this tag and to route the packet to the adequate port. The rules set in the forthcoming switches are the same as previous, until we reach the last virtual middle boxes. From the progress switch of the last virtual middle boxes to the destination switch, there are two cases. First, the destination machine is located in the current zone, and the destination switch is connected to the destination machine. In that case, the routing in the last segment is done by matching the destination address. Network security is a key aspect in designing modern applications.

#### IV.EVALUATION AND ANALYSIS

We can see that in both cases, the security policy has been reconfigured automatically after migration. The middle box sequences have been applied accordingly to the Application ID present in the packet, even though the instances traversed have changed. We chose to show both these cases of migration in order to demonstrate that *all or part* of the traversed middle box instances can be modified. Our model allows us to dynamically enforce security policies in a multi-tenant cloud network. In addition, the security policies stay coherent in spite of the node migration. So as to create a prototype of our network controller, we had to implement several modules in order to demonstrate the relevance of our policy enforcement mechanism. Particularly, a routing protocol has been implemented in order to figure out the route from the source to the destination, throughout the virtual middle boxes.

#### V.CONCLUSION AND FUTURE WORK

Apart from the above mentioned domain other domains are fair and fine enough to us the technology. In this one we use the data partitioning and the security of the node and agent based cryptography, which will provide the tiered or layered security. In the context of the security; particularly, we considered the criteria of scalability

and auto in a context where the network is shared by multiple tenants and the migration of nodes are increasing with the number of request. We first considered the different existing solutions in order to differentiate the traffic between the different tenants, as the isolation between tenants is the basis of a secure network. There are many solutions allowing traffic separation. However, the precision of these solutions varies greatly, as the isolation can be made from a complete separation of physical network and automatic centralized control of the network. We then analyzed the way security policies are defined and enforced. The security policies are defined in several ways. Some solutions focus on the isolation only or the routing rules and their consistency, whereas other architectures put more emphasis on the traversal of virtual middle boxes.

## VI. REFERENCES

- [1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing, Version 15," Nat'l Inst. of Standards and Technology, Information Technology Laboratory, vol. 53, p. 50, <http://csrc.nist.gov/groups/SNS/cloud-computing/>, 2010.
- [2] F. Gens, "IT Cloud Services User Survey, pt.2: Top Benefits & Challenges," blog, <http://blogs.idc.com/ie/?p=210>, 2008.
- [3] Gartner, "Gartner Says Cloud Adoption in Europe Will Trail U.S. by at Least Two Years," <http://www.gartner.com/it/page.jsp?id=2032215>, May 2012.
- [4] J.-M. Bohli, M. Jensen, N. Gruschka, J. Schwenk, and L.L.L. Iacono, "Security Prospects through Cloud Computing by Adopting Multiple Clouds," Proc. IEEE Fourth Int'l Conf. Cloud Computing (CLOUD), 2011.
- [5] D. Hubbard and M. Sutton, "Top Threats to Cloud Computing V1.0," Cloud Security Alliance, <http://www.cloudsecurityalliance.org/topthreats>, 2010.
- [6] M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, "On Technical Security Issues in Cloud Computing," Proc. IEEE Int'l Conf. Cloud Computing (CLOUD-II), 2009.
- [7] T. Ristenpart, E. Tromer, H. Shacham, and S. Savage, "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds," Proc. 16th ACM Conf. Computer and Comm. Security (CCS '09), pp. 199-212, 2009.
- [8] Y. Zhang, A. Juels, M.K.M. Reiter, and T. Ristenpart, "Cross-VIRTUAL Side Channels and Their Use to Extract Private Keys," Proc. ACM Conf. Computer and Comm. Security (CCS '12), pp. 305-316, 2012.
- [9] N. Gruschka and L. Lo Iacono, "Vulnerable Cloud: SOAP Message Security Validation Revisited," Proc. IEEE Int'l Conf. Web Services (ICWS '09), 2009.
- [10] M. McIntosh and P. Austel, "XML Signature Element Wrapping Attacks and Countermeasures," Proc. Workshop Secure Web Services, pp. 20-27, 2005.
- [11] J. Kincaid, "Google Privacy Blunder Shares Your Docs without Permission," TechCrunch, <http://techcrunch.com/2009/03/07/huge-google-privacy-blunder-shares-your-docs-withoutpermission/>, 2009.
- [12] J. Somorovsky, M. Heiderich, M. Jensen, J. Schwenk, N. Gruschka, and L. Lo Iacono, "All Your Clouds Are Belong to Us: Security Analysis of Cloud Management Interfaces," Proc. Third ACM Workshop Cloud Computing Security Workshop (CCSW '11), pp. 3-14, 2011.
- [13] S. Bugiel, S. Nurnberger, T. Poppelmann, A.-R. Sadeghi, and T. Schneider, "AmazonIA: When Elasticity Snaps Back," Proc. 18th ACM Conf. Computer and Comm. Security (CCS '11), pp. 389-400, 2011.
- [14] D. Bernstein, E. Ludvigson, K. Sankar, S. Diamond, and M. Morrow, "Blueprint for the Intercloud—Protocols and Formats for Cloud Computing Interoperability," Proc. Int'l Conf. Internet and Web Applications and Services, pp. 328-336, 2009.
- [15] A. Celesti, F. Tusa, M. Villari, and A. Puliafito, "How to Enhance Cloud Architectures to Enable Cross-Federation," Proc. IEEE Third Int'l Conf. Cloud Computing (CLOUD), pp. 337-345, 2010.
- [16] R. Turpin and B.A. Coan, "Extending Binary Byzantine Agreement to Multivalued Byzantine Agreement," Information Processing Letters, vol. 18, no. 2, pp. 73-76, 1984.
- [17] I. Koren and C.M.C. Krishna, Fault-Tolerant Systems. Morgan Kaufmann, 2007.

## AUTHORS PROFILE:



**1. Rajeshwarrao. Kodipaka** (ISTE, CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.



2. **Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer networks.



3. **M. AMARENDHAR REDDY** is presently associated with B. V. Raju Institute of Technology (UGC - Autonomous), Narsapur, Medak

(Dist), Telangana State as Asst. Professor in Information technology Department. He graduated from JNTU Hyderabad in B.Tech CS & IT and completed M.Tech (SE) from JNTUH-School of Information Technology Campus. Presently with 10 years of rich experience in Teaching, Administration and Research in the area of Object Oriented Modeling, Software Engineering Data Warehousing & Mining and Programming Languages. Prior to his current position with B. V. Raju Institute of Technology he served multiple positions as I/C HOD, Exam Branch In charge and Students' Advisor in various leading organization. His teaching experience includes teaching UG and PG course on Object oriented analysis and design, Java, C, C++, Software Engineering, Formal Languages and Automata Theory and Automata and Compiler Design. As part of Students' Advisor responsibilities he conducted workshops on Data mining and STQA.

# Distributed Databases Rules of Mining in the field of Association

**K.Rajeshwar Rao<sup>1</sup>, A.Prakash<sup>2</sup>, B.Venkat Ramudu<sup>3</sup>**

Asst.Professor, Department of CSE, MREC(A), Hyderabad , India<sup>1,3</sup>

Asso.Professor, Department of CSE, St. Martyn's Engineering College, Hyderabad, India<sup>2</sup>

**Abstract:** Association Rule mining is one among the data mining technique employed in distributed databases. This technique disclose some attention grabbing relationship between locally massive associate degreed globally massive itemsets and proposed an algorithm, Fast distributed mining of association(FDM) , unsecured distributed version of Apriori algorithm and generates a low range of candidate sets and well reduces the amount of message to be passed at mining association rules. Two secure multi party algorithms are main ingredients in our protocol. One that computes the union of private subsets that every of the interacting players hold, and another that tests the inclusion of element by one player in a subset held by another. This protocol offers increased privacy with relation to the protocol and additionally, it is less complicated and considerably a lot of efficient in terms of Communication rounds, communication cost and computation cost.

**Index terms:** Privacy Preserving Data Mining, Association Rules, Distributed Databases.

## I. INTRODUCTION

Aim of data mining is to extract vital information from massive datasets, however typically these datasets are split among varied parties. Data mining is defined as the technique for extracting hidden, predictive and knowledge data from large distributed databases. The technology that has emerged as method of identifying patterns and trends from large quantities of knowledge. This paper studies the matter of association rule mining in horizontally distributed databases. In the distributed databases, there are many players thathold same databases that share same schema however hold data totally on different entities. The goal is to search out all association rules with support s and confidence c to attenuate the data disclosed regarding the personal databases command by those players [1].

Kantarcioğlu and Clifton studied the problem whenever a lot of suitable security definitions that permit parties to settle on their desired level of security are required, effective solutions that maintain the required security [2]. So they devised a protocol for its solution. The main part of that protocol is the sub protocol for secure computation of union of private subsets that are controlled by the various players. It makes the protocol expensive and its implementation depends upon cryptographic primitive's strategies, oblivious transfer and hash function also the leakage of data renders the protocol not secured perfectly.

This paper proposed an algorithm privacy preserving distributed data mining algorithm for horizontally distributed data sets and to get most interesting association or correlation relationships among an outsized or large set of data items and to incorporate encryption security techniques to reduce the information that goes to shared with others, while adding little overhead to the mining task [1]. In the proposed association rule that hold within the unified database with support and confidence no smaller than the given threshold s and c, accordingly.

The data that would like to secure during this paper is not only individual transaction within the completely different databases, however more public or global information like what association rules are supported locally in every of these databases. The proposed protocol improves upon that in Kantarcioğlu and Clifton [2] in terms of simplicity efficiency such as privacy and security.

## II. LITERATURE REVIEW

Data mining in centralized model assumes that required data is either available at or can sent to a central site and data mining algorithm applied at central site. A simple approach not to share data is, perform data mining tool at each site independently and combine the results. However, it fails in globally valid results because of data skewness and disparity between local and global results. Individual sites are unable to detect cross-site correlations in data mining. Because the same item may be duplicated at different sites and there will be an overweight in the results.

FP tree is a compact data structure and without candidate set generation it finds the frequent itemset by traversing itemsets through FP tree. Paper [10] uses the FP tree algorithm for association rule mining in distributed environment to find the frequent itemsets. Data encryption standards are used to generate the privacy to database by using two keys. With the key1, first party encrypts the dataset. This encrypted dataset again encrypted with key2. At the receiver side at first, it decrypt the dataset with key2 and then after with key1. It provides high security to the databases compared with other cryptographic techniques and it is also known as Double encryption. With data leakage of zero percentage it provides higher privacy to the database; however it applicable to homogeneous databases only.

The paper [11] uses apriori algorithm with extended Distributed RK sum secure protocol for privacy preserving data mining. At first, on given dataset Apriori algorithm is applied then to find the frequent itemsets of individual parties and then to get global results Extended distributed RK sum secure algorithm is applied. Distributed RK sum secure protocol used for secure multi party protocol but drawback is more than two parties join together, they can know the data of some party. The drawback is reduced in extended distributed RK sum protocol with more privacy to the data, the communication and computation complexity  $O(n)$ .

The paper [13] studied five different algorithms such as Apriori algorithm, MCISI algorithm, MS algorithm, Apriori with systematic rules and HMT (Hash Mapping Table). Apriori algorithms combined with hashing techniques to reduce the time and space complexity. MSApriori (minimum support) method provides different minimum support itemset values for different items. Apriori systematic rules used to save the time by scanning the database only once rather than multiple scans.

### Privacy Preserving Data Mining

In [4] studied the problem in privacy preserving association rule mining in horizontally distributed databases among  $n$  ( $n > 2$ ) sites, with no sites can be treated as trusted. The proposed protocol uses hash based cryptographic techniques to find the global frequent itemsets. Privacy preserving rule mining categorized into three types based on privacy preserving/protection techniques.

1. Heuristic based techniques
2. Reconstruction based techniques
3. Cryptographic based techniques

Heuristic based techniques uses adaptive modification and it only modifies the selected values, utility loss is minimized. Cryptographic based techniques are used for both horizontal distributed databases and vertically distributed databases and it is based on the secure multi party computation. Here no database knows anything except his input database and the result, so that computation is secured. Reconstruction based techniques are used for different data such as numerical, binary and categorical data, its work on the problem of privacy preservation by perturbing the data. Original distribution of data is constructed from randomized data by Reconstruction based techniques.

### Association Rule:

Association rule mining is well known and popular method finding interesting relationships between from large databases. Today association rules are used in many applications including intrusion detection, biometrics, web usage mining and continuous production. The concept was popularized by Agrawal et al. article in 1993. It is a two step process i.e. Minimum support and confidence, minimum support used to find all frequent itemsets from given database and the constraint confidence used to form

the rules. The association rule mining problem is stated/defined as Let  $I = \{i_1, i_2, i_3, \dots, i_n\}$  are the set of items, database  $D$  consists set of transaction  $T$ . Each transaction  $T$  is a set of items from  $I$  such that  $T \subseteq I$ , with a unique transaction id TID. Each transaction  $T$  contains  $X$ , and a set of some items in  $I$ , if  $X \subseteq T$ . An association rule is an implication of the form  $X \Rightarrow Y$ , where  $X \subset I$ ,  $Y \subset I$  and  $X \cap Y = \emptyset$ . The association rule  $X \Rightarrow Y$  holds in the transaction set  $D$ , with confidence  $c$  if  $c\%$  of transaction in  $D$  that contain  $X$  also contain  $Y$ . The rule  $X \Rightarrow Y$  has support  $s$  in the transaction set  $D$  if  $S\%$  of transactions in  $D$  contain  $X \cup Y$ . In shopping centers association rules are used to place the items side by side, so that more items will be sold. In Amazon, association rules mining to recommended you to buy the products based on present item buying or browsing. Google uses association rules for mining, when you type the words for search it will give frequent associated words.

### Apriori Algorithm

In transactional databases, apriori algorithm is used for mining of frequent itemset and association rule learning. Apriori algorithm proceeds by identifying the frequent itemsets and extending them to larger and larger itemsets in database. It identifies the itemset at least  $C$  transaction in the database. Apriori algorithm uses the bottom up approach and proposed by Agarwal and Srikanth in 1994. Its usage is more in market basket analysis. To count candidate itemset efficiently Apriori algorithm uses breadth-first search and hash tree structure. Apriori algorithm works based on Apriori principle and to generate candidate  $k$ -item sets from  $(k-1)$  frequent itemsets, pruning technique used to avoid the measuring certain itemsets. Hash based methods with Apriori algorithms are used to reduce the time and space complexity.

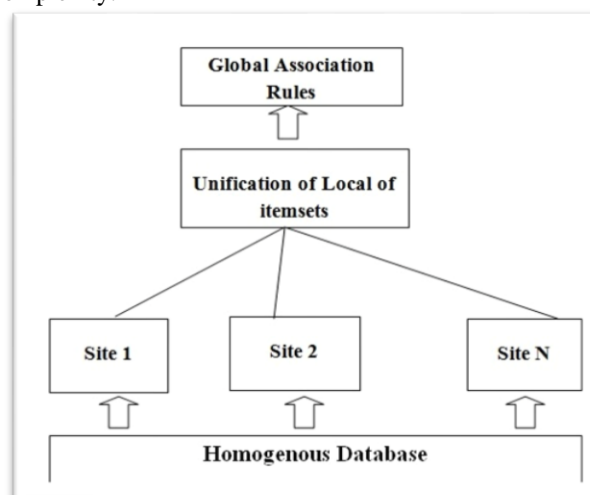


Fig1 Distributed Databases Architecture of the system

### Synthetic Database Generation

To calculate the performance of algorithms over a very large range of data characteristics the synthetic transactions are used and the process involved here is data anonymization. It is used in data mining process. Synthetic data are generated to get certain conditions and

to meet specific needs that may not be found in real, original data. Following are the parameters used in synthetic data generation and these values are similar to used in [2], [5], [12].

### FDM Algorithm

Following are the steps for FDM algorithm

1. Initialization
2. Candidate set generation
3. Local Pruning
4. Unifying the candidate itemset
5. Computing local support
6. Broadcast the mining results

Cheung [4] proposed Fast distributed mining for association rules, for distributed databases, Apriori algorithm for association rule mining for individual databases. In [2] proposed an insecure version of FDM algorithm but algorithm violates privacy in two stages, in step four and step six. In step 4 whenever the players broadcast the itemsets that are locally frequent in their private databases. In step 6 where they broadcast the sizes of the local supports of candidate itemsets. Kantarcioglu & Clifton [2] proposed secure implementations of these steps and also in [1] describe the various implementations and proceed to analyze implementations in terms of privacy, efficiency and compare them. Finally it show that our protocol offers better/enhanced privacy which is less complicated and is considerably additional efficient in terms of communication rounds, communication cost and computation cost.

### III. DISCUSSIONS

Following are the steps to secure computation itemsets using FDM K&C protocol.

1. Selection of Cryptographic primitive cipher.
  - Players choose the required commutative cipher and its corresponding private key and selected hash function applied on all itemsets for encryption.
  - A lookup table with hash values is generated to compute pre image of hash values.
  - Encrypt all the itemsets.
2. Merging of itemsets
  - Odd players send its encrypted set to player1.
  - Even players send its encrypted set to player2.
  - Player1 unifies all itemsets were sent by odd players and duplicates are removed.
  - Players 2 unifies all itemsets were sent by odd players and duplicates are removed.
  - Player 2 sends his permuted list of itemsets to player 1.
  - Player 1 unifies his list of itemsets and the list received from player 2 and from the unified list duplicates are removed.
  - Denote the final list as  $EC_S^k$ .
3. Decryption.

Parameter	Interception
N	No. of transactions in database
L	No. of items
$A_t$	Average size of transaction
$A_f$	Maximal potentially large itemsets average size
$N_f$	No. of maximal potentially large itemsets
CS	Clustering size
PS	Pool size
C or	Correlation level
M F	Multiplying factor

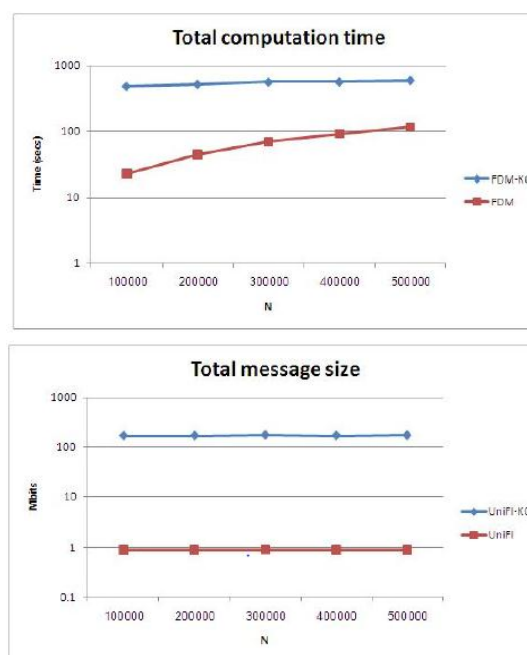


Fig2 Computation and communication cost versus number of transactions N

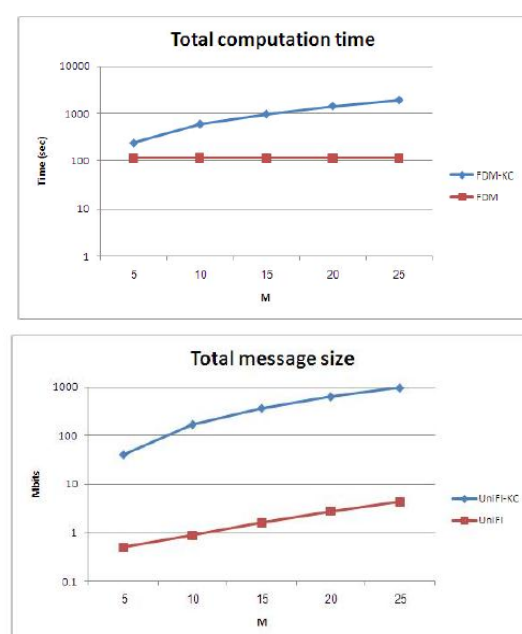


Fig3 Computation and communication costs versus the number of players M

The fig. 3 illustrates the leading protocol K and C. In this protocol main ingredients are two novel secure multiparty algorithms in which union and intersection operations. This protocol works if the number of players is more than two.

#### IV. CONCLUSION

The proposed protocol secure mining of association rules in horizontally distributed databases offers enhanced privacy and security than the leading protocol [2]. Two secure multi party algorithms are main ingredients in our protocol.

One that computes the union of private subsets that every of the interacting players hold, and another that tests the inclusion of element by one player in a subset held by another. Proposed protocol is more significant in communication cost, communication rounds and computation cost

#### REFERENCES

- [1] Tamir Tassa, "secure Mining of Association rules in Horizontally Distributed Databases" IEEE trans. Knowledge and Data Engg. Vol. 26, no. 2, April 2014.
- [2] M. Kantarcioglu & C. Clifton," Privacy preserving distributed mining of association rules on horizontally partitioned data. IEEE Transactions on Knowledge and Data Engineering, 16:1026-1037, 2004.
- [3] Agarwal and R. Srikant "Fast algorithms for mining association rules in large databases. In VLDB, pages 487-499, 1994.
- [4] D.W.L Cheung, V.T.Y. Ng, A.W.C. Fu, and Y. Fu. Efficient mining of association rules in distributed databases. *IEEE Trans. Knowl. Data Eng.*, 8(6):911-922, 1996Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.
- [5] D.W.L. Cheung, J. Han, V.T.Y. Ng, A.W.C. Fu, and Y. Fu. A fast distributed algorithm for mining association rules. In *PDIS*, pages 31-42, 1996.
- [6] A.V. Evfimievski, R. Srikant, R. Agrawal, and J. Gehrke. Privacy preserving mining of association rules. In *KDD*, pages 217-228, 2002.
- [7] Larry A. Dunning, Member, IEEE, and Ray Kresman,"Privacy Preserving Data Sharing With Anonymous ID Assignment", IEEE Transaction On Information Forensics and security, VOL. 8, NO. 2, FEBRUARY 2013.
- [8] J. Vaidya and C. Clifton, "Privacy preserving association rule mining in vertically partitioned data," in *The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton, Alberta, Canada, July 23-26 2002, pp. 639-644.
- [9] H. Grosskreutz, B. Lemmen, and S. R'uping. Secure distributed subgroup discovery in horizontally partitioned data. *Transactions on Data Privacy*, 4:147-165, 2011.
- [10] J Jyotirmayee Rautaray, Raghavendra Kumar, "Privacy Preserving Databases Using Data Encryption (DES) " International Journal of Innovative Research in Science , Engineering and Technology Vol. 2 , Issue 3, March 2013.
- [11] Meers Treesa Mathews, Manju E.V," Extended Distributed RK Secure Sum Protocol in Apriori Algorithm for Privacy Preserving" International Journal of Engineering and Advanced Technology(IJEAT), Volume-3, Issue-4, April 2014.
- [12] J.S. Park, M.S. Chen, and P.S. Yu. An effective hash based algorithm for mining association rules. In *SIGMOD Conference*, pages 175-186, 1995.
- [13] Priyanka Asthana, Anju Singh , Diwakar Singh," A Survey on Association Rule Mining Using Apriori Based Algorithm and Hash Based Methods ", International Journal of Advanced Research in Computer Science and Software Engineering. Volume 3, Issue 7, July 2013.

#### BIOGRAPHIES

**K .RAJESHWAR RAO** ,Working as Asst.professor, in Department of CSE in Malla Reddy Engineering College , Hyderabad . He has vast experience in teaching. His research includes data mining , Web technologies, Artificial intelligence , Software Engineering and Computer networks.

**A.PRAKASH**, currently working as Head of the Department, CSE, St.Martin's Engineering College. He completed his M.Tech in Software Engineering from JNTUH. He had an experience of more than 12 years in teaching and his areas of interest are Education Technologies, Computer Networks, Cryptography and security.

**B VENKATA RAMUDU**, is working as Asst.professor Malla Reddy Engineering College (Auto nous) Hyderabad Education Technologies, Computer Networks, Cryptography and security.

## Incorporation of Third Party Auditor for Safe and Secure Data Storage in Cloud Computing



**Sanjeeva Polepaka**

Associate Professor,  
Department of CSE,

Malla Reddy Engineering College.



**Mr. Rajeshwarrao kodipaka**

Assistant Professor,  
Department of CSE,

Malla Reddy Engineering College.

### Abstract:

Cloud Computing is the long dreamed vision of computing as a utility, where users can remotely store their data into the cloud so as to enjoy the on-demand high quality applications and services from a shared pool of configurable computing resources. By data outsourcing, users can be relieved from the burden of local data storage and maintenance. However, the fact that users no longer have physical possession of the possibly large size of outsourced data makes the data integrity protection in Cloud Computing a very challenging and potentially formidable task, especially for users with constrained computing resources and capabilities.

Thus, enabling public auditability for cloud data storage security is of critical importance so that users can resort to an external audit party to check the integrity of outsourced data when needed. To securely introduce an effective third party auditor (TPA), the following two fundamental requirements have to be met: 1) TPA should be able to efficiently audit the cloud data storage without demanding the local copy of data, and introduce no additional on-line burden to the cloud user; 2) The third party auditing process should bring in no new vulnerabilities towards user data privacy.

### Introduction of domain:

Cloud Computing is one of the best choice for Small and Medium Sized Entrepreneurs' in the world. This has been used widely to cater the needs of organizations with different software applications through SAAS operations of cloud. Cloud has gifted the organizations the platform as a service at most economical rates. The economical rates and swift services has given rise to the cloud computing fame.

But the recent privacy issues and security challenges have degraded the cloud computing marketing. North Bridge (2013)<sup>1</sup>. Cloud market has fallen down because of the data storage security problems, quality of services and privacy preserving issues in cloud computing. The research scholars have done enough research to admeasure the problems but still lot of loop holes and missing links are notified in cloud computing to arrest the problems. There is a great need to revive the situation by doing the research work to ensure data storage security in cloud computing. There are many research works tried to ensure the data security in cloud computing servers with different techniques like digital encryption, fuzzy key word search etc. The ultimate solution for data security and data storages should be given to arrest the problems due to Byzantine failure, malicious data modification attack and server colluding attacks. Ricardo Puttini (2013)

### Aim:

The aim of the project is to utilize the public key based homomorphic authenticator and uniquely integrate it with random mask technique to achieve a privacy-preserving public auditing system for cloud data storage security. The aim is to support efficient handling of multiple auditing tasks in cloud computing. The aim is to explore the technique of bilinear aggregate signature to extend the main result into a multi-user setting, where TPA can perform multiple auditing tasks simultaneously. The aim includes to provide an extensive secured and highly performing system in cloud computing with an effective third party auditor.

### Objectives:

- The main objective of the project is to develop a highly secured and good performance system with third party auditor for data storage in cloud computing.

- The objective is to design and develop a cloud computing application with extensive public data storage with third party auditing system.
- The objective is to demonstrate the third party auditing system is not effecting the performance of the cloud servers and storage operations.
- The implementation of third party auditing with high performance for the public storage in cloud computing is one of the objective.
- The objective is to develop a cloud computing application with high security in storing public data.

## Background:

Cloud computing is one of the most popular technological aspects in computational world. The most important thing in cloud computing is data storage with proper security measures. At the same time cloud computing is facing the data leakage and insecurity. To avoid these challenges in cloud computing a strong secured auditing system is needed to identify the users and their data. The increased security has to be implemented, at the same time the performance should not be degraded. To give proper solution the present project has been developed to provide the higher data security and increased performance while implementing the public data auditing in the cloud computing.

## Existing System:

Though Cloud computing is enriched with the special features like agile, reliable, cost effective and measurable delivery of data. Cloud computing has excellent delivery models with identification, Authentication, Authorization, Confidentiality, Integrity, Non-repudiation and availability as information security requirements. It is blamed by the provision of untrustworthy servers located at remote and un-known locations. This feature has become an issue to store sensitive data and confidential data at untrusted servers at unknown remote locations and caused the heavy computation overhead. Frequently the cloud computing has encountered the security issues such as SQL Injection Attacks, Cross Site Scripting Attacks, Man in the Middle Attacks, Network Level Security Issues, DNS attacks, Sniffer Attacks, BGP Prefix Hijacking and issue of reused IP Address. Apart from these attacks Application level security issues with security concerns with the Hypervisor, Denial of Service Attacks, Cookie Poisoning, Hidden Field Manipulation.

Captcha Breaking, Google Hacking and Distributed DOS Attacks are traced in cloud computing. To overcome all these attacks and computation overhead, Priyadarshini has suggested a cloud storage mechanism with Kerberos authentication and utilization of multi-clouds. Kerberos secure mechanism is very strong to protect weak link from opponent. It is very reliable with a system to back up. The Kerberos mechanism is transparent which can be revealed with the password authentication. It is measurable to support large number of clients and service providers. Kerberos authentication mechanism is rich with Authentication Servers and Ticket Granting Servers which can be incorporated in six steps. Above all these measures Priyadarshini has suggested multi-clouds environment to implement Kerberos Realms and Multiple Kerberos to obtain increased security in cloud computing storage. - A. Priyadarshini [2013].

## Drawbacks of Existing system:

The Department of Education, The United States of America has incorporated a Privacy Technical Assistance Center for learning data privacy, confidentiality and security. Data security threats are classified into two types. These are technical and non-technical. A comprehensive privacy and data security plan has been initiated by PTAC to reduce the vulnerability to security threats. The technical data security threats are regarded as non-existent security architecture, Un-Patched client side Software and Application, Phishing and targeted Attacks which is known as Spear Phishing, Internet Websites, Poor configuration Management, data storage in mobile devices and transfer of data from mobile devices, Cloud Computing, Data from removable media, Botnets [ the series of networks of compromised computers] and Zero-day attacks. The non technical cyber security threats are regarded as Insider attacks, Poor passwords, Physical Security, Insufficient backup and recovery, improper destruction, Social Engineering and social media. PTAC has successfully inculcate the mitigation for all above mentioned threats and suggested the students to follow consistent implementation of the security plan drastically eradicate the cyber threats and establish the security. – PTAC – IB [2011].

## Proposed System:

The proposed system is developed with the public key based homomorphic authenticator and uniquely integrate it with random mask technique to achieve a privacy-preserving public auditing system for cloud data storage.

To support efficient handling of multiple auditing tasks, we further explore the technique of bilinear aggregate signature to extend our main result into a multi-user setting, where Third Party Auditing can perform multiple auditing tasks simultaneously.

Extensive security and performance analysis shows the proposed schemes are provably secure and highly efficient. We also show how to extend our main scheme to support batch auditing for Third Party Auditing upon delegations from multi-users.

### Third Party Auditing:

Third party auditing system is one security mechanism to store the data into the cloud servers. Third party auditing system is arranged by the cloud consumers as well as the cloud service providers. The cloud service providers would safe guard the server with the help of Third party auditing.

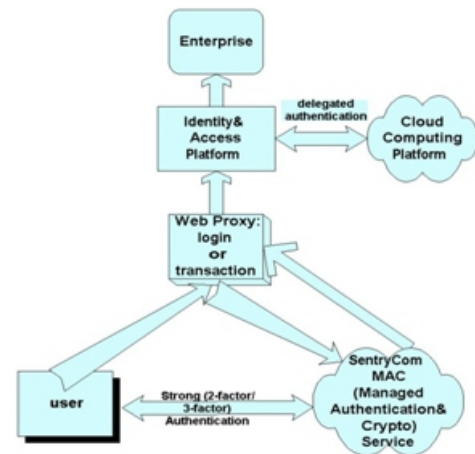
Third party auditing will scan the data whatever stored by the cloud consumer then it will decide whether the data can be stored in the server or not. If the data consists of malicious data then the TPA will delete the data from the server. Predominantly the data storage will be audited by the TPA system and allow the data to place in the server. It can be accessed by the authorized user of the Cloud Consumers.

### Algorithm:

A public auditing scheme consists of four algorithms. These are KeyGen, SigGen, GenProof and VerifyProof.

- KeyGen: key generation algorithm that is run by the user to setup the scheme
- SigGen: used by the user to generate verification metadata, which may consist of MAC, signatures or other information used for auditing
- GenProof: run by the cloud server to generate a proof of data storage correctness
- VerifyProof: run by the TPA to audit the proof from the cloud server.

### Flowchart:



### Modules:

#### 1.Privacy-Preserving Public Auditing Module:

Homomorphic authenticators are unforgeable verification metadata generated from individual data blocks, which can be securely aggregated in such a way to assure an auditor that a linear combination of data blocks is correctly computed by verifying only the aggregated authenticator.

Overview to achieve privacy-preserving public auditing, we propose to uniquely integrate the homomorphic authenticator with random mask technique. In our protocol, the linear combination of sampled blocks in the server's response is masked with randomness generated by a pseudo random function (PRF).

The proposed scheme is as follows:

- Setup Phase
- Audit Phase

#### 2.Batch Auditing Module:

With the establishment of privacy-preserving public auditing in Cloud Computing, TPA may concurrently handle multiple auditing delegations upon different users' requests. The individual auditing of these tasks for TPA can be tedious and very inefficient. Batch auditing not only allows TPA to perform the multiple auditing tasks simultaneously, but also greatly reduces the computation cost on the TPA side.

### 3.Data DynamicsModule:

Hence, supporting data dynamics for privacy-preserving public risk auditing is also of paramount importance. Now we show how our main scheme can be adapted to build upon the existing work to support data dynamics, including block level operations of modification, deletion and insertion. We can adopt this technique in our design to achieve privacy-preserving public risk auditing with support of data dynamics.

### The Functionality of the Project:

The project is designed and developed to demonstrate the cloud computing data storage security with the help of Third Party Auditing system. The cloud computing is declining with the security loopholes. These loop holes of security system can be arrested by employing the third party auditing system. The cloud computer is occupied by the cloud consumers. Sometimes the cloud servers are losing the data integrity and confidentiality because of the cloud consumer's data stealing mechanism. The data stored by the cloud consumers should be streamlined by employing a third party auditor who is most amicable and trustworthy for cloud consumers as well as cloud service providers. The cloud consumer's data should be scanned and verified by the Third Party Auditor, then it should be stored in the cloud servers. This data streamlining mechanism will avoid the malicious data storage in the cloud computing servers. When the data is accessed by the cloud users who are created and permitted by the consumers should also approved by the third party auditor. In this present project the Third Party Auditing system will keep an eye on the data accessing items and keeps the track of them. In this way the data storage mechanism will be audited by the Third Party auditor and safely enable the cloud consumer to store the data into the cloud servers.

### Conclusion :

The project is designed with a novel mechanism to audit the data storage with third party auditing system to filter the data. The data storage options have to be streamlined with proper monitoring and auditing system. To ensure the safe and secure data storage mechanism in cloud computing the proposed third party auditing system is developed in .Net technologies in simulation mode. This has been tested with proper test cases to ensure the mechanism working properly. The test results have been evaluated.

Public auditability also allows clients to delegate the integrity verification tasks to TPA while they themselves can be unreliable or not be able to commit necessary computation resources performing continuous verifications. Another major concern is how to construct verification protocols that can accommodate dynamic data files. In this paper, we explored the problem of providing simultaneous public auditability and data dynamics for remote data integrity check in Cloud Computing.

### References:

- 1.P. Mell and T. Grance, "Draft NIST working definition of cloud computing," Referenced on June. 3rd, 2009 Online at <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>, 2009.
- 2.M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A Berkeley view of cloud computing," University of California, Berkeley, Tech. Rep. UCB-EECS-2009-28, Feb 2009.
- 3.N. Gohring, "Amazon's s3 down for several hours," Online at <http://www.pcworld.com/businesscenter/article/142549/amazons-s3-down-for-several-hours.html>, 2008.
- 4.Amazon.com, "Amazon s3 availability event: July 20, 2008," Online at <http://status.aws.amazon.com/s3-20080720.html>, July 2008.
- 5.S. Wilson, "Appengine outage," Online at <http://www.cio-weblog.com/50226711/appengine-outage.php>, June 2008.
- 6.B. Krebs, "Payment Processor Breach May Be Largest Ever," Online at <http://voices.washingtonpost.com/securityfix/2009/01/payment-processor-breach-may-b.html>, Jan. 2009.
- 7.G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," Cryptology ePrintArchive, Report 2007/202, 2007, <http://eprint.iacr.org/>.
- 8.M. A. Shah, R. Swaminathan, and M. Baker, "Privacy-preserving audit and extraction of digital contents," Cryptology ePrint Archive, Report 2008/186, 2008, <http://eprint.iacr.org/>.

- 9.Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in Proc. of ESORICS'09, Saint Malo, France, Sep. 2009.
- 10.Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, <http://www.cloudsecurityalliance.org>.
- 11.H. Shacham and B. Waters, "Compact proofs of retrievability," in Proc. of Asiacrypt 2008, vol. 5350, Dec 2008, pp. 90–107.
- 12.Juels and J. Burton S. Kaliski, "Pors: Proofs of retrievability for large files," in Proc. of CCS'07, Alexandria, VA, October 2007, pp. 584–597.
- 13.M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to keep online storage services honest," in Proc. of HotOS'07. Berkeley, CA, USA: USENIX Association, 2007, pp. 1–6.
- 14.104th United States Congress, "Health Insurance Portability and Accountability Act of 1996 (HIPPA)," Online at <http://aspe.hhs.gov/admnsimp/pl104191.htm>, 1996, last access: July 16, 2009.
- 15.D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and verifiably encrypted signatures from bilinear maps," in Proc. of Eurocrypt 2003, volume 2656 of LNCS. Springer-Verlag, 2003, pp. 416–432.
- 16.T. S. J. Schwarz and E. L. Miller, "Store, Forget, and Check: Using Algebraic Signatures to Check Remotely Administered Storage," Proc. of ICDCS '06, pp. 12–12, 2006.
- 17.M. Lillibridge, S. Elnikety, A. Birrell, M. Burrows, and M. Isard, "A Cooperative Internet Backup Scheme," Proc. of the 2003 USENIX Annual Technical Conference (General Track), pp. 29–41, 2003.
- 18.K. D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," Cryptology ePrint Archive, Report 2008/489, 2008, <http://eprint.iacr.org/>.
- 19.L. Carter and M. Wegman, "Universal Hash Functions," Journal of Computer and System Sciences, vol. 18, no. 2, pp. 143–154, 1979.
- 20.J. Hendricks, G. Ganger, and M. Reiter, "Verifying Distributed Erasure-coded Data," Proc. 26th ACM Symposium on Principles of Distributed Computing, pp. 139–146, 2007.
- 21.J. S. Plank and Y. Ding, "Note: Correction to the 1997 Tutorial on Reed-Solomon Coding," University of Tennessee, Tech. Rep. CS-03-504, 2003.
- 22.Q. Wang, K. Ren, W. Lou, and Y. Zhang, "Dependable and Secure Sensor Data Storage with Dynamic Integrity Assurance," Proc. of IEEE INFOCOM, 2009.
- 23.R. Curtmola, O. Khan, R. Burns, and G. Ateniese, "MR-PDP: Multiple-Replica Provable Data Possession," Proc. of ICDCS '08, pp. 411–420, 2008.
- 24.D. L. G. Filho and P. S. L. M. Barreto, "Demonstrating Data Possession and Uncheatable Data Transfer," Cryptology ePrint Archive, Report 2006/150, 2006, <http://eprint.iacr.org/>.
- 25.M. A. Shah, M. Baker, J. C. Mogul, and R. Swaminathan, "Auditing to Keep Online Storage Services Honest," Proc. 11th USENIX Workshop on Hot Topics in Operating Systems (HOTOS '07), pp. 1–6, 2007.

# Design of Sentiment Analysis System using Polarity Classification Technique

Rajeshwar Rao Kodipaka  
Asst.Prof, Dept of CSE,  
MREC-Sec-bad, TS

Sanjeeva Polepaka  
Associate.Prof, Dept of CSE,  
MREC-Sec-bad, TS

Md. Rafeeq  
Associate.Prof, Dept of CSE,  
CMRTC-Kandlakoya-Hyd, TS

## ABSTRACT

Twitter is a medium that we can use for communication. All posted tweets we can store in one location and create archive. Archive contains new and old tweets. Now we can start the analyzation on archive tweets that's we can design effective sentiment analysis system. This paper main aim is to determine parts of speech opinion words using polarity classification technique and support vector machine learning algorithm. Surveys of methods are used in various levels of sentiment analysis. It does analyze the tweets information in limited levels of content only. Now in this paper we design new sentiment analysis tool using polarity classification technique. Polarity classification techniques discover top 20-emoticons, learning different classes of words and other features information. These techniques perform in depth tweets analysis. It does provide better analysis results compare to previous methods.

## Keywords

Sentiment analysis, opinion mining, text documents, support vector machine classifier, polarity classification technique.

## 1. INTRODUCTION

Twitter enables all organizations for communication directly with each other. It's possible to tap the global real time communication important events very easy using twitter analysis. It is used to find out opinions from twitter micro blogging tweets. Many numbers of researchers have done good extensive work presently and previously on this topic environment. Analysis starts on the basis of tweets. Present approaches are missing contexts for some number of conclusions.

In this paper using polarity classification technique analyzes the tweets and fulfills the missing context features. Here we fulfill the features like part of speech opinion words of information, frequent occurrences opinions words discovery with the help of support vector machine classifier. It gives more conclusions compare to all previous approaches and more useful also for any real applications decision making.

## 2. RELATED WORK

Sentiment analysis is also known as Opinion Mining. Sentimental analysis major aims to determine the attitude of writer, judgment and communication based on text documents. Sentiment analysis major task is classify text documents, sentences, aspect based level. Express opinions in documents, sentences, aspect based level information are positive, negative or neutral [1].

Twitter data contains different topics information related to different domains. Classify the sentiment documents information using probabilistic model. Probabilistic model is one of the supervised learning algorithm. Each topic related how many documents are available it's not possible to

recognize or predictable. First choose the class labels and classify each and every topic documents separately [1] [2]. This is we can call as a document level sentiment classification.

In each topic sentiment documents again possible to classify or learning the sentiment sentences, aspect level process, opinion words also. The above steps are possible to implement on news, blogs and other categories [2] [3].

In twitter first collect the different categories of data using hash tags input. Hash tags works like class labels, but here there is [3] [4] no sentiment labels information. Supervised learning is not sufficient for extraction of sentiment analysis information. Here we should use the unsupervised learning [4] [2]. Using unsupervised learning identify the sentiment words measurement in documents. Sentiment words are categorized into positive and negative contexts.

Sentiment words are huge. Reduces the sentiment words information and improve the classification result. Classification improves using co-occurrence technique. Highly occurred features we can display as a output [4] [5] using unsupervised clustering. All high occurred features are not semantic or meaningful.

On co-occurrence words information applies correlation technique finally recognizes or predicts the relationship sentiment words content. This correlation sentiment words procedure is best prediction procedure compare to above all procedures [2] [3] [5].

Again in twitter it's possible to predict the similar opinions of information. Identify the social relationship users from total twitter data. Social relationship users it's possible to display using visualization concept. Finally it's possible to display temporal events relationship information also as a final result. Consider the temporal events relationship and possible to view of sentiment words trends information effectively. These trends changes dynamically [4] [5] [6].

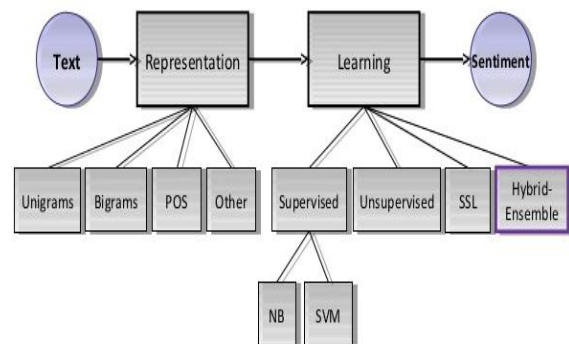


Fig 1: Sentiment Analysis Result with different data mining techniques

### 3. PROBLEM STATEMENT

Face book and Twitter are major resources for sentiment analysis. Nowadays numbers of people are increases for posting the opinions in social media environment. Result is vast amount of unstructured data is available. Manual sentiment analysis is not feasible which generated data is currently. Here in this paper we can propose automated sentiment analysis techniques. Automated sentiment analysis techniques are extracting different characteristics of information. Here we focus on tweet level polarity classification process. Hence it is one of the interesting research topics.

### 4. PROPOSED METHODOLOGY

After posting before analysis of raw tweets we can apply preprocessing operation. Using preprocessing operation removes the unnecessary content and change the format of context information. Next we use the tokeniser operations for dividing the content based on parts of speech. According to parts of speech classify the sentences of information. Classify the sentences depends on hash tags or tokens. Different hash tags are different parts of speech content purpose in our implementation process.

Here we design the different rules for sentiment analysis. Those rules are emoticon, support vector machines with n-grams, Sentic computing and Linguistic rules etc.

#### 4.1 Emoticon Rules

Emoticons are ASCII art. They are formed through creative letters, numbers and symbols. Most of the people represent the facial features information. Here normally for plain text messages we can add the emotional flavor. Next for normal text messages again we can add the smile automatically those messages are converted to happiness and surprise messages.

First choose the twitter data source and collect lakhs or millions of tweets. In millions of tweets we can perform the analyzation operation and we discovered 20 usage patterns information.

#	Emoticon	Usage	Percent	Notes
#1	:)	32,115,789	33.380%	Happy face
#2	:D	10,595,385	11.006%	Laugh
#3	:(	7,613,014	7.908%	Sad face
#4	;)	7,238,295	7.519%	Wink
#5	:-)	4,254,708	4.420%	Happy face (with nose)
#6	:P	3,588,863	3.728%	Tongue out
#7	=)	3,564,080	3.702%	Happy face
#8	(:	2,720,383	2.826%	Happy face (mirror)
#9	;-)	2,085,015	2.166%	Wink (with nose)
#10	:/	1,840,827	1.912%	Uneasy, undecided, skeptical, annoyed?
#11	XD	1,795,792	1.865%	Big grin
#12	=D	1,434,004	1.490%	Laugh
#13	:O	1,077,124	1.119%	Shock, Yawn
#14	=]	1,055,517	1.096%	Happy face
#15	D:	1,048,320	1.089%	Grin (mirror)
#16	;D	1,004,509	1.043%	Wink and grin
#17	:]	954,740	0.992%	Happy face
#18	:-)	816,170	0.848%	Unhappy
#19	=/	809,760	0.841%	Uneasy, undecided, skeptical, annoyed?
#20	=(	760,800	0.790%	Unhappy

Fig2: Top 20 Emoticons

In Above diagrams top-20 emoticons we displayed here in our implementation. These emoticons are occurred frequently in number of tweets messages.

### 4.2 N-Grams Technique

In posted tweets observe the features and store features into feature vector. All required features are available in sequence or contiguous or not we can check with n-grams concept. Afterwards using TF-IDF identifies and calculates the frequency count. Consider the frequency count generate weighted features information. Weighted features information controls the number of dimensions. In all frequent words we can categorize based on parts of speech. Consider the different number of parts of speech tags information and categorize the words into noun, adverb, verb etc.

### 4.3 Linguistic Rules

After collection of training tweets information next we can apply support vector machine classifier. It will separate the features of information n different classes. Each and every class contains some data points of information content. Each and every class of words again categorize into two classes. Those two classes are positive and negative words. Calculate the polarity value like positive and negative content. Polarity value is nothing but decision score.

### 5. EXPERIMENTS AND RESULTS

We evaluate our proposed system on available real time datasets. Every and every year tweets we can collect separately and create dataset. Every year dataset contains different categories tweets are available. Those categories are positive, negative and neutral tweets. Here first remove neutral tweets from total number of tweets information.

Evaluate two datasets and calculates efficiently different performance metrics parameters information. Those parameters are precision, recall and f-measure.

Table1: Performance Metrics

Method	Positive			Negative			Average		
	P	R	F	P	R	F	P	R	F
N-grams	89.92	81.90	85.72	61.20	75.66	67.67	75.56	78.78	76.69
N-grams and Emoticon Rules	89.74	83.05	86.27	62.50	74.85	68.11	76.12	78.95	77.19
Modified N-grams	89.39	82.90	86.02	62.00	73.93	67.44	75.69	78.41	76.73
Modified N-grams, and Emoticon Rules	89.25	83.97	86.53	63.29	73.22	67.89	76.27	78.60	77.21
Modified N-grams, Emoticon Rules, and Word-level Unsupervised Rules	90.22	86.24	88.19	67.37	75.25	71.09	78.80	80.75	79.64
Modified N-grams, Emoticon Rules, and Concept-level Unsupervised Rules	90.41	86.20	88.25	67.45	75.76	71.37	78.93	80.98	79.81

### 6. CONCLUSION AND FUTURE WORK

In this paper new twitter sentiment analysis system apply rules and discover more useful text. Useful text contains meaningful features. Those features are discovered using linguistic content and sentic computing rules. Features are like emoticon symbols and n-grams content information. These meaningful features are support for decision making in all real time applications.

In future we plan to improve the performance using other unsupervised classifiers. We plan to develop some more rules

for efficient text predictions and multi model sentiment analysis.

## 7. REFERENCES

- [1] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *J. Am. Soc. Inform.Sci. Technol.*, vol. 60, no. 11, pp. 2169–2188, 2009.
- [2] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Micro-blogging as online word of mouth branding," in *Proc. Extended Abstr. Human Factors Comput. Syst.*, 2009, pp. 3859–3864.
- [3] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *J. Comput. Sci.*, vol. 2, no. 1, pp. 1–8, 2011.
- [4] A. Tumasjan, T. O. Sprenger, P. G. Sandner, and I. M. Welp, "Predicting elections with twitter: What 140 characters reveal about political sentiment," in *Proc. 4th Int. AAAI Conf. Weblogs Soc. Media*, 2010, vol. 10, pp. 178–185.
- [5] L. T. Nguyen, P. Wu, W. Chan, W. Peng, and Y. Zhang, "Predicting collective sentiment dynamics from time-series social media," in *Proc. 1st Int. Workshop Issues Sentiment Discovery Opinion Mining*, 2012, p. 6.
- [6] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in twitter events," *J. Am. Soc. Inform. Sci. Technol.*, vol. 62, no. 2, pp. 406–418, 2011.
- [7] A. Agarwal, B. Xie, I. Vovsha, O. Rambow, and R. Passonneau, "Sentiment analysis of twitter data," in *Proc. Workshop Lang. Soc. Media*, 2011, pp. 30–38.
- [8] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lect. Human Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.
- [9] C. Tan, L. Lee, J. Tang, L. Jiang, M. Zhou, and P. Li, "User-level sentiment analysis incorporating social networks," in *Proc. 17<sup>th</sup> ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1397–1405.
- [10] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification," in *Proc. 45th Annu. Meeting Assoc. Comput. Linguistics*, 2007, vol. 7, pp. 440–447.
- [11] F. Li, S. J. Pan, O. Jin, Q. Yang, and X. Zhu, "Cross-domain coextraction of sentiment and topic lexicons," in *Proc. 50th Annu. Meeting Assoc. Comput. Linguistics: Long Papers*, 2012, pp. 410–419.
- [12] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, "Cross-domain sentiment classification via spectral feature alignment," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 751–760.
- [13] I. Ounis, C. Macdonald, J. Lin, and I. Soboroff, "Overview of the trec-2011 microblog track," in *Proc. 20th Text Retrieval Conf.*, 2011, <http://trec.nist.gov/pubs/trec20/t20.proceedings.html>
- [14] I. Soboroff, I. Ounis, J. Lin, and I. Soboroff, "Overview of the trec- 2012 microblog track," in *Proc. 21st Text REtrieval Conf.*, 2012.
- [15] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *CS224N Project Report*, Computer Science Department, Stanford, USA, pp. 1–12, 2009.
- [16] S. Li, C.-R. Huang, G. Zhou, and S. Y. M. Lee, "Employing personal/ impersonal views in supervised and semi-supervised sentiment classification," in *Proc. 48th Annu. Meeting Assoc. Comput. Linguistics*, 2010, pp. 414–423.

## 8. AUTHOR PROFILE

**Rajeshwarrao. Kodipaka** (ISTE,CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.

**Sanjeeva Polepaka** is working as associate professor in Malla Reddy Engineering college (Autonomous) Hyderabad ,TS State. He completed B.Tech(CSE) from Andhra University ,M Tech(CSE) from ANU Guntur Perusing Ph D from JNTU Hyderabad. He has 12 years of teaching experience in various engineering colleges. He is life member of CSI and ISTE .His area of interest includes Image processing mobile , grid computing and computer network.

**Md. Rafeeq** Doing external PhD in JNTUH-Hyd,Telangana,India. I would like thank to my guide Dr. Sunil Kumar, and co-guide Dr. Subhash Chandra for constant support for Understanding the subject.

# An Effective Approach for Maintaining High Accurateness in Cloud Environment

Mr. Rajeshwar Rao Kodipaka<sup>1</sup>, Mr. M. Amarendhar Reddy<sup>2</sup>, B. Venkat Ramudu<sup>3</sup>

Asst. Professor, Department of CSE, MallaReddy Engineering College (Autonomous), Hyderabad, India<sup>1,3</sup>

Asst. Professor, Department of IT, B.V. Raju Institute of Technology, Medak, India<sup>2</sup>

**Abstract:** Generally Software-as-a-service clouds were built on software as a service along with service-oriented architecture. This assists the providers of application service in conveying their applications by means of substantial cloud computing infrastructure. In the traditional efforts while confidentiality as well as privacy protection exertions was broadly studied, the problem of service integrity attestation has not been appropriately addressed. In our work we mainly spotlight on data processing services which have turned out to be popular with applications in numerous usage areas. We present IntTest, which is a scalable and efficient integrated service integrity attestation structure intended for software-as-a-service cloud multitenant cloud systems. The proposed approach offers a novel integrated attestation graph analysis method which can not only identify attackers more resourcefully by taking an integrated approach, but also can hold back aggressive attackers and limit the scope of damage that is caused by colluding attacks.

**Keywords:** Software-as-a-service, Integrated attestation graph analysis, Cloud computing, IntTest, Attackers.

## 1. INTRODUCTION

Infrastructures of cloud computing are usually shared by application service providers from various domains of security which make them exposed to malicious attacks. Service integrity is the most prevailing problem that has to be addressed regardless of processing of public or private data by cloud system [1]. In major multitenant systems of cloud, numerous malicious attackers might commence colluding attacks on definite targeted service functions to nullify the assumption. In our work we focus on services of data streaming applications for clouds with numerous real-world applications. In our work to address this challenge we present IntTest, which is an integrated service integrity attestation structure intended for multitenant cloud systems. IntTest put forward result auto-correction that can automatically restore corrupted data processing results that are produced by malevolent attackers with superior results that are produced by benign service providers. IntTest approach considers a holistic approach by thoroughly examining consistency as well as inconsistency relationships between several service providers within complete cloud system. IntTest offers a realistic service integrity attestation system that does not imagine trustworthy entities on third-party service.

services might attract several service providers for profit. To maintain automatic service composition, we can install a set of portal nodes that serve as the gateway for user to access composed services in Software-as-a-service cloud. The portal node can combine various service components into composite services on the basis of user's needs. Altered from other distributed systems for instance peer-to-peer networks as well as volunteer computing setting, Software-as-a-service cloud systems acquire a set of distinctive features. Application service providers of third-party usually do not want to make known internal functioning details of their software services for protection of intellectual property. Both cloud infrastructure providers as well as third-party providers are independent entities. It is not practical to enforce any special hardware or else protected kernel support on individual service provisioning sites. For privacy fortification, only portal nodes encompass global information regarding which service functions are offered by which service providers in Software-as-a-service cloud. Our work draw attention in the direction of on data processing services which have turned out to be more and more popular with applications in numerous real-world usage areas.

## 2. MODELLING OF CLOUD MODEL OF SAAS

Software-as-a-service clouds mostly were put up on concepts of software as a service along with service-oriented architecture which facilitate application service providers in conveying their applications by means of substantial cloud computing infrastructure. In a comprehensive Software-as-a-service cloud, identical service function can be offered by various application service providers [2][3]. These components of functionally equivalent service exist due to service providers might generate replicated service components in support of load Balancing as well as fault tolerance purposes; and popular

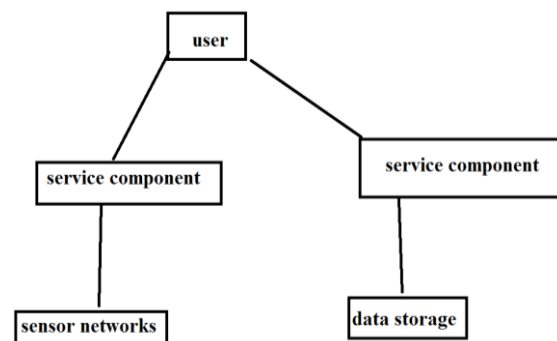


Fig1: An overview of service integrity attacks.

In our work present IntTest, which is a scalable and efficient integrated service integrity attestation structure intended for software-as-a-service cloud multitenant cloud systems. IntTest offers a novel integrated attestation graph analysis method that can make available stronger attacker pinpointing power than earlier schemes. Given a Software-as-a-service cloud system, objective of IntTest is to identify any malevolent service provider that offers a misleading service function [4]. IntTest considers all service components as black boxes, which does not necessitate any special hardware on cloud platform.

### 3. AN OVERVIEW OF PROPOSED IntTest SYSTEM

IntTest put up on earlier work and AdapTest but can grant well-built malicious attacker pinpointing control than RunTest and AdapTest. RunTest and AdapTest as well as conventional majority voting methods need to imagine that providers of benign service receive majority in each service function. IntTest scrutinize per-function consistency graphs as well as global inconsistency graph. The analysis of per-function consistency graph can limit extent of damage that is caused by colluding attackers, while global inconsistency graph analysis can efficiently depict those attackers that attempt to compromise numerous service functions for this reason, IntTest can still identify malicious attackers although they turn out to be majority for several service function. IntTest offer result auto-correction that can automatically restore corrupted data processing results that are produced by malevolent attackers with superior results that are produced by benign service providers. It presents a realistic service integrity attestation system that does not imagine trustworthy entities on third-party service. By taking an integrated approach, IntTest can not only identify attackers more resourcefully but also can hold back aggressive attackers and limit the scope of damage that is caused by colluding attacks. In our work we focus on services of data streaming applications for clouds with numerous real-world applications. Our work spotlight on data processing services which have turned out to be more and more popular with applications in numerous real-world usage areas [5]. IntTest can not only identify malicious service providers but moreover automatically spot on corrupted data processing results to get better result quality of cloud data processing service. The offers a novel integrated attestation graph analysis method that can make available stronger attacker pinpointing power than earlier schemes. To distinguish service integrity attack as well as pinpoint malicious service contributor, our algorithm depends on replay-based consistency check to obtain consistency/inconsistency relationships among service providers. The intuition following our method is that when two service providers differ with each other on processing result of the similar input, not less than one of them has to be malevolent. We do not forward an input data item as well as its duplicates simultaneously. As a substitute, we attestation data on several service providers after receiving processing result original data consequently, malicious attackers cannot avoid risk of being detected when they construct false results on original data [6]. By

means of replay-based consistency check, we can check functionally equivalent service providers and get hold of their consistency as well as inconsistency relationships.

### 4. CONCLUSION

We draw attention towards services of data stream processing that are considered for clouds with abundant applications. Our work put forward an effective and scalable IntTest, which is efficient integrated service integrity attestation structure intended for software-as-a-service cloud multitenant cloud systems. Scalable IntTest approach takes for consideration a holistic approach by thoroughly examining consistency as well as inconsistency relationships between several service providers within complete cloud system. The proposed system offer result auto-correction that can automatically restore corrupted data processing results that are produced by malevolent attackers with superior results that are produced by benign service providers. The system put forward realistic service integrity attestation system that does not imagine trustworthy entities on third-party service. To make a distinction of service integrity attack as well as pinpoint malicious service contributor, our algorithm depends on replay-based consistency check to obtain consistency/inconsistency relationships among providers of service.

### REFERENCES

- [1] L.Alchaal, V.Roca, and M.Habert, "Managing and Securing Web Services with VPNs," Proc. IEEE Int'l Conf. Web Services, pp. 236- 243, June 2004.
- [2] H. Zhang, M.Savoie, S.Campbell, S.Figuerola, G.von Bochmann, and B.S.Arnaud, "Service-Oriented Virtual Private Networks for Grid Applications," Proc. IEEE Int'l Conf. Web Services, pp. 944-951, July 2007.
- [3] M. Burnside and A.D. Keromytis, "F3ildCrypt: End-to-End Protection of Sensitive Information in Web Services," Proc. 12<sup>th</sup> Int'l Conf. Information Security (ISC), pp. 491-506, 2009.
- [4] J.L. Griffin, T. Jaeger, R. Perez, and R. Sailer, "Trusted Virtual Domains: Toward Secure Distributed Services," Proc. First Workshop Hot Topics in System Dependability, June 2005.
- [5] L. Lamport, R. Shostak, and M. Pease, "The Byzantine Generals Problem," ACM Trans. Programming Languages and Systems, vol. 4, no. 3, pp. 382-401, 1982.
- [6] T. Ho et al., "Byzantine Modification Detection in Multicast Networks Using Randomized Network Coding," Proc. IEEE Int'l Symp. Information Theory (ISIT), 2004.

### BIOGRAPHIES

**MR. RAJESHWAR RAO KODIPAKA** (ISTE, CSI LIFE MEMBER), Working as Asst. Professor, in Department of CSE in MallaReddy Engineering College (Autonomous), Hyderabad. He has vast experience in teaching. His research includes data mining; Cloud computing, Software Engineering and Computer networks.

**M. AMARENDHAR REDDY** is presently associated with B. V. Raju Institute of Technology (UGC - Autonomous), Narsapur, Medak (Dist), Telangana State as Asst. Professor in Information technology Department with 10 years of rich experience in Teaching, Administration and Research in the area of Object Oriented Modeling, Software Engineering Data Warehousing & Mining and Programming Languages.

**B VENKATA RAMUDU** is working as Asst. Professor Malla Reddy Engineering College (Autonomous), Hyderabad. His research interests in Computer Networks, Cryptography and security.

# Implementation of Cryptography Encryption Algorithm for Plane Text

BV Ramudu<sup>1</sup>, Sanjeeva Polepaka<sup>2</sup>

Assoc Professor, MallaReddy Engineering College (Autonomous) Hyderabad, India<sup>1,2</sup>

**Abstract:** In this paper an efficient security architecture design and implementation of all candidates of AES encryption standards AES-128, AES-192 and AES-256 on the same hardware is proposed. AES algorithm proposed by NIST has been widely accepted as best cryptosystem for wireless communication security. The hardware implementation is useful in wireless security like military and mobile phones. This contribution investigates implementation of AES Encryption with regards to FPGA and VHDL. Optimized and synthesized VHDL code for AES-128, AES-192 and AES-256 for encryption of 128-bit data is implemented. Xilinx ISE 9.2i software is used for simulation. Each algorithm is tested with sample vectors provided by NIST output results are perfect with minimal delay. The proposed design consumes less power and area which is suitable battery driven mobile phones. Throughput reaches the value of 666.67 Mbps for encryption of 128-bit data with AES-128 key.

**Keywords:** Cryptography, Cipher, Reconfiguration, Encryption, Decryption.

## 1. INTRODUCTION

Cryptography is the science of information and communication security. Cryptography is the science of secret codes, enabling the confidentiality of communication through an insecure channel. It protects against unauthorized parties by preventing unauthorized alteration of use. Generally speaking, it uses a cryptographic system to transform a plaintext into a cipher text, using most of the time a key. One has to notice that there exist certain ciphers that don't need a key at all. A famous example is ROT13 (abbreviation from Rotation 13), a simple Caesar-cipher that obscures text by replacing each letter with the letter thirteen places down in the alphabet. Since our alphabet has 26 characters, it is enough to encrypt the cipher text again to retrieve the original message. Let me just mention briefly that there are secure public-key ciphers, like the famous and very secure Rivest-Shamir-Adleman (commonly called RSA) that uses a public key to encrypt a message and a secret key to decrypt it. Cryptography is a very important domain in computer science with many applications. The most famous example of cryptography is certainly the Enigma machine, the legendary cipher machine used by the German Third Reich to encrypt their messages, whose security breach ultimately led to the defeat of their submarine force. Before continuing, please read carefully the legal issues involving cryptography as in several countries even the domestic use of cryptography is prohibited: Cryptography has long been of interest to intelligence gathering agencies and law enforcement agencies. Because of its facilitation of privacy, and the diminution of privacy attendant on its prohibition, cryptography is also of considerable interest to civil rights supporters. Accordingly, there has been a history of controversial legal issues surrounding cryptography, especially since the advent of inexpensive computers has made possible widespread access to high quality cryptography. In some countries, even the domestic use of

cryptography is, or has been, restricted. Until 1999, France significantly restricted the use of cryptography domestically. In China, a license is still required to use cryptography. Many countries have tight restrictions on the use of cryptography. Among the more restrictive are laws in Belarus, Kazakhstan, Mongolia, Pakistan, Russia, Singapore, Tunisia, Venezuela, and Vietnam.[31] In the United States, cryptography is legal for domestic use, but there has been much conflict over legal issues related to cryptography. One particularly important issue has been the export of cryptography and cryptographic software and hardware. Because of the importance of cryptanalysis in World War II and an expectation that cryptography would continue to be important for national security, many western governments have, at some point, strictly regulated export of cryptography. After World War II, it was illegal in the US to sell or distribute encryption technology overseas; in fact, encryption was classified as a munitions, like tanks and nuclear weapons.[32] Until the advent of the personal computer and the Internet, this was not especially problematic. Good cryptography is indistinguishable from bad cryptography for nearly all users, and in any case, most of the cryptographic techniques generally available were slow and error prone whether good or bad. However, as the Internet grew and computers became more widely available, high quality encryption techniques became well-known around the globe. As a result, export controls came to be seen to be an impediment to commerce and to research.

### 1.1 Description of AES Algorithm:

Advanced Encryption Standard is the successor of Data Encryption Standard which was in use during the early 1977 to 1990. In DES encryption is based on a symmetric key algorithm that uses a 56-bit key. However by the mid 1990's, it was clear that the DES with 56-bit is insecure for many applications since the key is very small. Then it

was upgraded to Triple DES which was believed to be practically secure although there are theoretical attacks. Thus in Nov-26-2001 the FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 197(FIPS 197) specifies an algorithm called Advanced Encryption Standard (AES). AES is based on the principle known as Substitution Permutation network (SP-network) which means there will be a series of linked mathematical operations in the block cipher algorithm. AES encrypts a data block of 128-bits which is fixed with three different key sizes 128,192,256 bits.

The operations are based on Rijndael algorithm. The input of AES algorithm is 128-bit or 16 byte data which can be specified as a block. The basic unit of processing in the AES algorithm is a byte. All byte values in the AES algorithm will be presented as the concatenation of its individual bit values (0 or 1) between the braces in the order (b7, b6, b5, b4, b3, b2, b1, b0). These bytes are interpreted as finite field elements using a polynomial representation as follows

Algorithm	Key length (Nk words)	Block Size (Nb words)	Number of rounds (Nr)
AES-128	4	4	10
AES-192	6	4	12
AES-256	8	4	14

$$b_7X^7 + b_6X^6 + b_5X^5 + b_4X^4 + b_3X^3 + b_2X^2 + b_1X + b_0 = \sum_{i=0}^7 b_i X^i$$

Internally in AES algorithm operations are performed on a two-dimensional array of bytes called the state. The state consists of four rows of bytes, each containing Nb bytes, where Nb is the block length divided by 32 (4 for 128-bit key, 6 for 192-bit key, 8 for 256-bit key). Likewise the key length and number of rounds (iterations) differ from key to key as shown in table 1.

Figure1: Different keys and its attributes

### AES Encryption:

Encryption is the process of converting the plain text into a format which is not easily readable and is called as cipher. The cipher is got by doing a series of mathematical operations iteratively.

#### a) Sub Bytes:

In this sub bytes step the data in the plain text is substituted by some pre-defined values from a substitution box. The substitution box is invertible.

#### b) Shift Rows:

In shift rows operation the rows in the 4×4 matrix is shifted to left r bits and r varies with the rows of the matrix (r=0 for row1, r=1 for row2, r=2 for row3, r=3 for row 4). This process is illustrated in fig 2. This has the effect of moving positions of lower positions in the row, while the lowest bytes wrap around to the top of the row.

#### c) Mix Columns:

Mix column is calculated using the below formula.

Here a0, a1, a2, a3 is calculated using the polynomials as below

$$a(x) = \{2\}x^3 + \{3\}x^2 + \{1\}x + \{1\}.$$

The mix column transformation operates on the state column by column, treating each column as a four term polynomial. The columns are considered as polynomials over GF (2<sup>8</sup>) and multiplied modulo x<sup>4</sup> + 1 with a fixed polynomial a(x) which is got from the above formula. This can also written as a matrix multiplication

$$s'(x) = a(x) s(x)$$

#### d) Add Round Key:

In the add round key step the 128 bit data is xored with the sub key of the current round using the key expansion operation. The add round key is used in two different places one during the start that is when round r=0 and then during the other rounds that is when 1 ≤ round ≤ Nr, where Nr is the maximum number of rounds. The formula to perform the add round key is S'(x) = S(x) R(x)

where S'(x) – state after adding round key ,S(x) – state before adding round key and R(x) – round key

#### e) Key Expansion:

The key expansion has three steps: Byte Substitution subword(), Rotation rotword() and Xor with RCON (round constant). The input to key schedule is the cipher key K. Key expansion generates a total of Nb(Nr + 1) words. The algorithm requires an initial set of Nb words, and each of the Nr rounds requires Nb words of key data. The resulting key schedule consists of a linear array of 4-byte words, denoted [wi], with i in the range 0 ≤ i < Nb(Nr + 1). The subword() function takes a four byte input and applies the byte substitution operation and produces an output word. The rotword() takes a word [a0, a1, a2, a3] as input and performs a cyclic permutation to produce [a1, a2, a3, a0] as output word. The round constant word array rcon[i] is calculated using the below formula in rijndale finite field.

$$rcon[i] = x^{(2^i - 1)} \text{ mod } x^4 + x + 1$$

The first Nk words of the expanded key are filled with the cipher key. Every following word w[i] is equal to the xor of previous word w[i-1] and the word Nk positions earlier w[i-Nk]. For words in positions that are a multiple of Nk, a transformation is applied to w[i-1] prior to the XOR, followed by an XOR with a round constant Rcon[i]. This transformation consists of a cyclic shift of the bytes in a word rotword() and byte substitution subword().

But in key expansion of 256 -bit cipher if Nk=8 and i-4 is a multiple of Nk then subword() function is applied to w[i-1] prior to the xor. The algorithm for the key expansion routine is given in table 2. Thus with all the above operations the algorithm for the encryption of the data is as follows. Since it begins and ends with the add round key operation there is no wasted unkeyed step in the beginning or the end. The table 3 shows the algorithm for implementation of all three AES encryption.

### Algorithm for key expansion

```

KeyExpansion(byte key[4*Nk], word w[Nb*(Nr+1)], Nk)
begin
word temp
i=0
while(i<Nk)
{
w[i] = word(key[4*i], key[4*i+1], key[4*i+2], key[4*i+3])
i = i+1
}
end while
I = Nk
while (i < Nb * (Nr+1))
{
temp = w[i-1]
if (i mod Nk = 0)
temp = SubWord(RotWord(temp)) xor Rcon[i/Nk]
else if (Nk > 6 and i mod Nk = 4)
temp = SubWord(temp)
end if

w[i] = w[i-Nk] xor temp
i = i + 1
}
end while
end

```

### Algorithm for encryption

```

byte state[4,Nb]
state = in
AddRoundKey(state, key Schedule[0, Nb-1])
for round = 1 step 1 to Nr-1
{
SubBytes(state)
ShiftRows(state)
MixColumns(state)
AddRoundKey(state, key Schedule[round*Nb, (round+1)*Nb-1])
}
SubBytes(state)

ShiftRows(state)
AddRoundKey(state, key Schedule[Nr*Nb, (Nr+1)*Nb-1])
out = state

```

## 2. HARDWARE IMPLEMENTATION OF ENCRYPTION ALGORITHM

Many hardware implementation of encryption algorithm using VHDL is available. In most of the case hardware implementations of AES uses only the AES-128 candidate. Some software implementation of AES192 and AES-256 are available. In the proposed architecture all candidates of AES i.e. AES-128, AES192 and AES-256 are implemented in the same device. The proposed design is implemented using VHDL coding in Xilinx ISE 9.2. Iterative looping techniques is followed to implement the entire design modules of AES encryption algorithm to the

minimize hardware utilization. The key controller unit, key expansion unit, and round function unit and mix column unit everything are implemented in hardware.

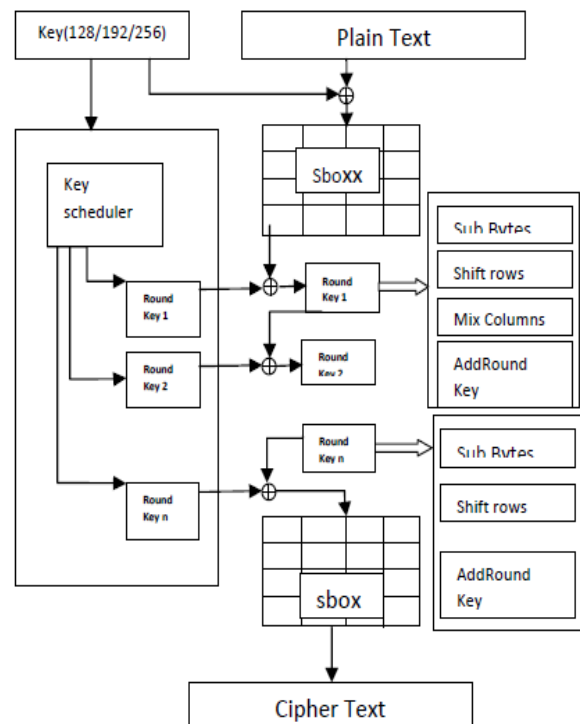


Figure 2. Flow Chart of AES Encryption Implementation

The table 2 Show s that proposed design outperforms all designs based on iterative looping in terms of area and throughput. The performance of AES-192 and AES 256 is also verified and simulation results are given. The novel architecture to implement all AES candidates in same hardware proposed is shown in figure 2. the simulation results of various AES key lengths is shown in figure 3,4 and 5.



Figure 3. Schematic Diagram AES-128/192/256 Architecture

## 3. CONCLUSION

The AES algorithm is an iterative private key symmetric block cipher that can process data block of 128-bits through the use of cipher keys with key length 128,192 and 256 bits. An efficient FPGA implementation of 128 bit block and keys 128, 192 and 256 bits of AES –

Rijndael algorithm has been presented in this paper. Optimized and synthesizable VHDL code is developed for implementation of all AES-128/192/256 bit key encryption and is verified using xilinx ISE 9.2 simulation tool. All the transformations of algorithm are simulated using an iterative design approach in order to minimize the hardware utilization. Thus it can reduce the space by enclosing three different encryption standards in a single architecture and the power consumption can also be reduced which makes it usable in battery operated network devices having Bluetooth and wireless communication devices like software radio. Throughput reaches the value of 666.7Mbps for AES-128 encryption with FPGA device XC2V6000BF957-6.

## REFERENCES

- [1] J. Daemen, V.Rijmen: The Rijndael Block Cipher: AES Proposal : First AES Candidate Conference (AES1) : August 20-22, 1998
- [2] A. Dandalis, V.K. Prasanna, J.D.P. Rolim: A Comparative Study of Performance of AES Candidates Using FPGAs: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [3] T. Ichikawa, T. Kasuya, M. Matsui: Hardware Evaluation of the AES Finalists: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [4] K. Gaj, P. Chodowiec: Comparison of the Hardware Performance of the AES Candidates using Reconfigurable Hardware: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [5] Xilinx VirtexTM-E 1.8V Field Programmable Gate Arrays: URL: <http://www.xilinx.com>: November 2000.
- [6] M.McLoone, J.V. McCanny: Apparatus for Selectably Encrypting and Decrypting Data: UK Patent Application No. 0107592.8: Filed March 2001.
- [7] B. Weeks, M. Bean, T. Rozylowicz, C. Ficke: Hardware Performance Simulations of Round 2 Advanced Encryption Standard Algorithms: The Third Advanced Encryption Standard (AES3) Candidate Conference, 13-14 April 2000, New York, USA.
- [8] Announcing the ADVANCED ENCRYPTION STANDARD (AES)" Federal Information Processing Standards Publication 197 November 26, 2001

## BIOGRAPHY



**B. Venkata Ramudu** completed B. Tech (CSE) M. Tech (CSE) from JNTU Hyderabad. He is working as associate professor in Malla Reddy Engineering College (Autonomous) Hyderabad. His research includes data mining.



**Sanjeeva Polepaka** completed B. Tech (CSE) from Andhra University and M. Tech (CSE) from Acharya Nagarjuna University. He is pursuing Ph D from JNTU Hyderabad; He is working as associate professor in Malla Reddy Engineering College (Autonomous) Hyderabad. His research includes digital image processing.

# Distributed Databases Rules of Mining in the field of Association

**K.Rajeshwar Rao<sup>1</sup>, A.Prakash<sup>2</sup>, B.Venkat Ramudu<sup>3</sup>**

Asst.Professor, Department of CSE, MREC(A), Hyderabad, India<sup>1,3</sup>

Asso.Professor, Department of CSE, St. Martyn's Engineering College, Hyderabad, India<sup>2</sup>

**Abstract:** Association Rule mining is one among the data mining technique employed in distributed databases. This technique disclose some attention grabbing relationship between locally massive associate degreed globally massive itemsets and proposed an algorithm, Fast distributed mining of association(FDM), unsecured distributed version of Apriori algorithm and generates a low range of candidate sets and well reduces the amount of message to be passed at mining association rules. Two secure multi party algorithms are main ingredients in our protocol. One that computes the union of private subsets that every of the interacting players hold, and another that tests the inclusion of element by one player in a subset held by another. This protocol offers increased privacy with relation to the protocol and additionally, it is less complicated and considerably a lot of efficient in terms of Communication rounds, communication cost and computation cost.

**Index terms:** Privacy Preserving Data Mining, Association Rules, Distributed Databases.

## I. INTRODUCTION

Aim of data mining is to extract vital information from massive datasets, however typically these datasets are split among varied parties. Data mining is defined as the technique for extracting hidden, predictive and knowledge data from large distributed databases. The technology that has emerged as method of identifying patterns and trends from large quantities of knowledge. This paper studies the matter of association rule mining in horizontally distributed databases. In the distributed databases, there are many players that hold same databases that share same schema however hold data totally on different entities. The goal is to search out all association rules with support  $s$  and confidence  $c$  to attenuate the data disclosed regarding the personal databases command by those players [1].

Kantarcioğlu and Clifton studied the problem whenever a lot of suitable security definitions that permit parties to settle on their desired level of security are required, effective solutions that maintain the required security [2]. So they devised a protocol for its solution. The main part of that protocol is the sub protocol for secure computation of union of private subsets that are controlled by the various players. It makes the protocol expensive and its implementation depends upon cryptographic primitive's strategies, oblivious transfer and hash function also the leakage of data renders the protocol not secured perfectly.

This paper proposed an algorithm privacy preserving distributed data mining algorithm for horizontally distributed data sets and to get most interesting association or correlation relationships among an outsized or large set of data items and to incorporate encryption security techniques to reduce the information that goes to shared with others, while adding little overhead to the mining task [1]. In the proposed association rule that hold within the unified database with support and confidence no smaller than the given threshold  $s$  and  $c$ , accordingly.

The data that would like to secure during this paper is not only individual transaction within the completely different databases, however more public or global information like what association rules are supported locally in every of these databases. The proposed protocol improves upon that in Kantarcioğlu and Clifton [2] in terms of simplicity efficiency such as privacy and security.

## II. LITERATURE REVIEW

Data mining in centralized model assumes that required data is either available at or can sent to a central site and data mining algorithm applied at central site. A simple approach not to share data is, perform data mining tool at each site independently and combine the results. However, it fails in globally valid results because of data skewness and disparity between local and global results. Individual sites are unable to detect cross-site correlations in data mining. Because the same item may be duplicated at different sites and there will be an overweight in the results.

FP tree is a compact data structure and without candidate set generation it finds the frequent itemset by traversing itemsets through FP tree. Paper [10] uses the FP tree algorithm for association rule mining in distributed environment to find the frequent itemsets. Data encryption standards are used to generate the privacy to database by using two keys. With the key1, first party encrypts the dataset. This encrypted dataset again encrypted with key2. At the receiver side at first, it decrypt the dataset with key2 and then after with key1. It provides high security to the databases compared with other cryptographic techniques and it is also known as Double encryption. With data leakage of zero percentage it provides higher privacy to the database; however it applicable to homogeneous databases only.

The paper [11] uses apriori algorithm with extended Distributed RK sum secure protocol for privacy preserving data mining. At first, on given dataset Apriori algorithm is applied then to find the frequent itemsets of individual parties and then to get global results Extended distributed RK sum secure algorithm is applied. Distributed RK sum secure protocol used for secure multi party protocol but drawback is more than two parties join together, they can know the data of some party. The drawback is reduced in extended distributed RK sum protocol with more privacy to the data, the communication and computation complexity  $O(n)$ .

The paper [13] studied five different algorithms such as Apriori algorithm, MCISI algorithm, MS algorithm, Apriori with systematic rules and HMT (Hash Mapping Table). Apriori algorithms combined with hashing techniques to reduce the time and space complexity. MSApriori (minimum support) method provides different minimum support itemset values for different items. Apriori systematic rules used to save the time by scanning the database only once rather than multiple scans.

### Privacy Preserving Data Mining

In [4] studied the problem in privacy preserving association rule mining in horizontally distributed databases among  $n$  ( $n > 2$ ) sites, with no sites can be treated as trusted. The proposed protocol uses hash based cryptographic techniques to find the global frequent itemsets. Privacy preserving rule mining categorized into three types based on privacy preserving/protection techniques.

1. Heuristic based techniques
2. Reconstruction based techniques
3. Cryptographic based techniques

Heuristic based techniques uses adaptive modification and it only modifies the selected values, utility loss is minimized. Cryptographic based techniques are used for both horizontal distributed databases and vertically distributed databases and it is based on the secure multi party computation. Here no database knows anything except his input database and the result, so that computation is secured. Reconstruction based techniques are used for different data such as numerical, binary and categorical data, its work on the problem of privacy preservation by perturbing the data. Original distribution of data is constructed from randomized data by Reconstruction based techniques.

### Association Rule:

Association rule mining is well known and popular method finding interesting relationships between from large databases. Today association rules are used in many applications including intrusion detection, biometrics, web usage mining and continuous production. The concept was popularized by Agrawal et al. article in 1993. It is a two step process i.e. Minimum support and confidence, minimum support used to find all frequent itemsets from given database and the constraint confidence used to form

the rules. The association rule mining problem is stated/defined as Let  $I = \{i_1, i_2, i_3, \dots, i_n\}$  are the set of items, database  $D$  consists set of transaction  $T$ . Each transaction  $T$  is a set of items from  $I$  such that  $T \subseteq I$ , with a unique transaction id TID. Each transaction  $T$  contains  $X$ , and a set of some items in  $I$ , if  $X \subseteq T$ . An association rule is an implication of the form  $X \Rightarrow Y$ , where  $X \subset I$ ,  $Y \subset I$  and  $X \cap Y = \emptyset$ . The association rule  $X \Rightarrow Y$  holds in the transaction set  $D$ , with confidence  $c$  if  $c\%$  of transaction in  $D$  that contain  $X$  also contain  $Y$ . The rule  $X \Rightarrow Y$  has support  $s$  in the transaction set  $D$  if  $s\%$  of transactions in  $D$  contain  $X \cup Y$ . In shopping centers association rules are used to place the items side by side, so that more items will be sold. In Amazon, association rules mining to recommended you to buy the products based on present item buying or browsing. Google uses association rules for mining, when you type the words for search it will give frequent associated words.

### Apriori Algorithm

In transactional databases, apriori algorithm is used for mining of frequent itemset and association rule learning. Apriori algorithm proceeds by identifying the frequent itemsets and extending them to larger and larger itemsets in database. It identifies the itemset at least  $C$  transaction in the database. Apriori algorithm uses the bottom up approach and proposed by Agarwal and Srikanth in 1994. Its usage is more in market basket analysis. To count candidate itemset efficiently Apriori algorithm uses breadth-first search and hash tree structure. Apriori algorithm works based on Apriori principle and to generate candidate  $k$ -item sets from  $(k-1)$  frequent itemsets, pruning technique used to avoid the measuring certain itemsets. Hash based methods with Apriori algorithms are used to reduce the time and space complexity.

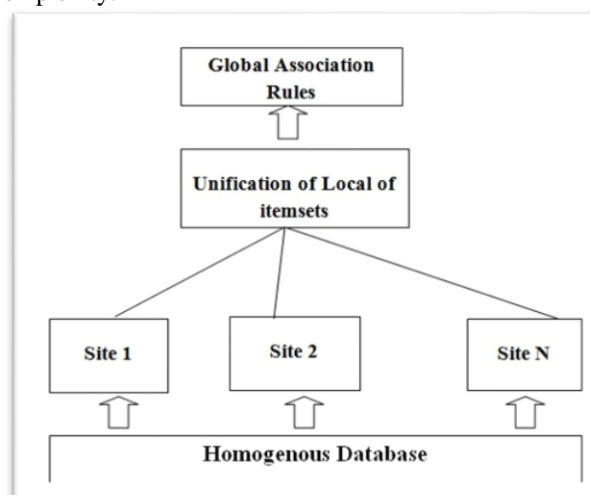


Fig1 Distributed Databases Architecture of the system

### Synthetic Database Generation

To calculate the performance of algorithms over a very large range of data characteristics the synthetic transactions are used and the process involved here is data anonymization. It is used in data mining process. Synthetic data are generated to get certain conditions and

to meet specific needs that may not be found in real, original data. Following are the parameters used in synthetic data generation and these values are similar to used in [2], [5], [12].

### FDM Algorithm

Following are the steps for FDM algorithm

1. Initialization
2. Candidate set generation
3. Local Pruning
4. Unifying the candidate itemset
5. Computing local support
6. Broadcast the mining results

Cheung [4] proposed Fast distributed mining for association rules, for distributed databases, Apriori algorithm for association rule mining for individual databases. In [2] proposed an insecure version of FDM algorithm but algorithm violates privacy in two stages, in step four and step six. In step 4 whenever the players broadcast the itemsets that are locally frequent in their private databases. In step 6 where they broadcast the sizes of the local supports of candidate itemsets. Kantarcioglu & Clifton [2] proposed secure implementations of these steps and also in [1] describe the various implementations and proceed to analyze implementations in terms of privacy, efficiency and compare them. Finally it show that our protocol offers better/enhanced privacy which is less complicated and is considerably additional efficient in terms of communication rounds, communication cost and computation cost.

### III. DISCUSSIONS

Following are the steps to secure computation itemsets using FDM K&C protocol.

1. Selection of Cryptographic primitive cipher.
  - Players choose the required commutative cipher and its corresponding private key and selected hash function applied on all itemsets for encryption.
  - A lookup table with hash values is generated to compute pre image of hash values.
  - Encrypt all the itemsets.
2. Merging of itemsets
  - Odd players send its encrypted set to player1.
  - Even players send its encrypted set to player2.
  - Player1 unifies all itemsets were sent by odd players and duplicates are removed.
  - Players 2 unifies all itemsets were sent by odd players and duplicates are removed.
  - Player 2 sends his permuted list of itemsets to player 1.
  - Player 1 unifies his list of itemsets and the list received from player 2 and from the unified list duplicates are removed.
  - Denote the final list as  $EC_S^k$ .
3. Decryption.

Parameter	Interception
N	No. of transactions in database
L	No. of items
$A_t$	Average size of transaction
$A_f$	Maximal potentially large itemsets average size
$N_f$	No. of maximal potentially large itemsets
CS	Clustering size
PS	Pool size
C or	Correlation level
M F	Multiplying factor

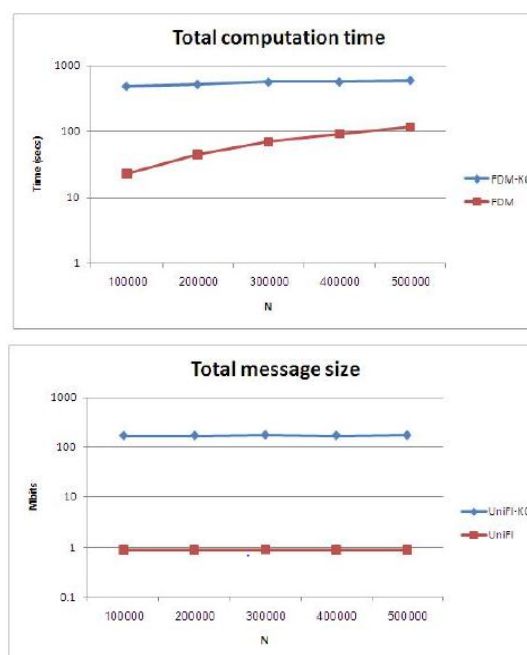


Fig2 Computation and communication cost versus number of transactions N

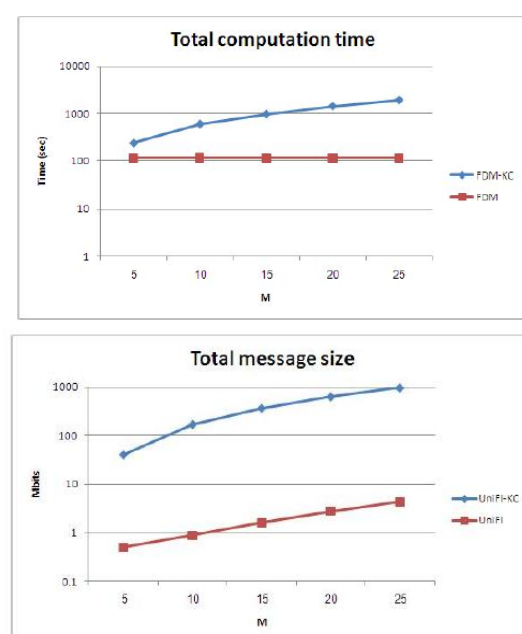


Fig3 Computation and communication costs versus the number of players M

The fig. 3 illustrates the leading protocol K and C. In this protocol main ingredients are two novel secure multiparty algorithms in which union and intersection operations. This protocol works if the number of players is more than two.

#### IV. CONCLUSION

The proposed protocol secure mining of association rules in horizontally distributed databases offers enhanced privacy and security than the leading protocol [2]. Two secure multi party algorithms are main ingredients in our protocol.

One that computes the union of private subsets that every of the interacting players hold, and another that tests the inclusion of element by one player in a subset held by another. Proposed protocol is more significant in communication cost, communication rounds and computation cost

#### REFERENCES

- [1] Tamir Tassa, "secure Mining of Association rules in Horizontally Distributed Databases" IEEE trans. Knowledge and Data Engg. Vol. 26, no. 2, April 2014.
- [2] M. Kantarcioglu & C. Clifton," Privacy preserving distributed mining of association rules on horizontally partitioned data. IEEE Transactions on Knowledge and Data Engineering, 16:1026-1037, 2004.
- [3] Agarwal and R. Srikant "Fast algorithms for mining association rules in large databases. In VLDB, pages 487-499, 1994.
- [4] D.W.L Cheung, V.T.Y. Ng, A.W.C. Fu, and Y. Fu. Efficient mining of association rules in distributed databases. *IEEE Trans. Knowl. Data Eng.*, 8(6):911-922, 1996Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.
- [5] D.W.L. Cheung, J. Han, V.T.Y. Ng, A.W.C. Fu, and Y. Fu. A fast distributed algorithm for mining association rules. In *PDIS*, pages 31-42, 1996.
- [6] A.V. Evfimievski, R. Srikant, R. Agrawal, and J. Gehrke. Privacy preserving mining of association rules. In *KDD*, pages 217-228, 2002.
- [7] Larry A. Dunning, Member, IEEE, and Ray Kresman,"Privacy Preserving Data Sharing With Anonymous ID Assignment", IEEE Transaction On Information Forensics and security, VOL. 8, NO. 2, FEBRUARY 2013.
- [8] J. Vaidya and C. Clifton, "Privacy preserving association rule mining in vertically partitioned data," in *The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Edmonton, Alberta, Canada, July 23-26 2002, pp. 639-644.
- [9] H. Grosskreutz, B. Lemmen, and S. R'uping. Secure distributed subgroup discovery in horizontally partitioned data. *Transactions on Data Privacy*, 4:147-165, 2011.
- [10] J Jyotirmayee Rautaray, Raghavendra Kumar, "Privacy Preserving Databases Using Data Encryption (DES) " International Journal of Innovative Research in Science , Engineering and Technology Vol. 2 , Issue 3, March 2013.
- [11] Meers Treesa Mathews, Manju E.V," Extended Distributed RK Secure Sum Protocol in Apriori Algorithm for Privacy Preserving" International Journal of Engineering and Advanced Technology(IJEAT), Volume-3, Issue-4, April 2014.
- [12] J.S. Park, M.S. Chen, and P.S. Yu. An effective hash based algorithm for mining association rules. In *SIGMOD Conference*, pages 175-186, 1995.
- [13] Priyanka Asthana, Anju Singh , Diwakar Singh," A Survey on Association Rule Mining Using Apriori Based Algorithm and Hash Based Methods ", International Journal of Advanced Research in Computer Science and Software Engineering. Volume 3, Issue 7, July 2013.

#### BIOGRAPHIES

**K .RAJESHWAR RAO** ,Working as Asst.professor, in Department of CSE in Malla Reddy Engineering College , Hyderabad . He has vast experience in teaching. His research includes data mining , Web technologies, Artificial intelligence , Software Engineering and Computer networks.

**A.PRAKASH**, currently working as Head of the Department, CSE, St.Martin's Engineering College. He completed his M.Tech in Software Engineering from JNTUH. He had an experience of more than 12 years in teaching and his areas of interest are Education Technologies, Computer Networks, Cryptography and security.

**B VENKATA RAMUDU**, is working as Asst.professor Malla Reddy Engineering College (Auto nous) Hyderabad Education Technologies, Computer Networks, Cryptography and security.



# A Query Based Approach to Find Proportionate Entities

Mr. M. Ajith<sup>1</sup>, Mr. Ch. Ramesh Babu<sup>2</sup>

<sup>1</sup> M.Tech Student, Dept. of cse, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> Associate Professor, Dept. of cse, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>1</sup>ajithreddymatta5@gmail.com

<sup>2</sup>rameshbabu.ch@mrec.ac.in

---

**Abstract**— Comparing entities by user is difficult to decide whether a query is comparable or not. To help higher intellectual process and to check entities share common functionality however has characteristics. One potential approach is a query based approach to find proportionate entities. The method is used bootstrapping approach which identifies the proportional queries and extracts the proportionate entities. This can be done by decision makers to check whether the given query is comparative or not. A successive pattern introduced that is Indicative Extraction Pattern (IEP) it may be accustomed and identify proportional queries and extract comparator pairs. These techniques identify relative question identification and equivalent entity mining. Ranking technique is used to rank the equivalent entities for user's input entity and also the results show extremely connection to uses evaluation intention.

**Keywords**— Bootstrapping algorithm, Proportionate entity finding, Information extraction, IEP, Sequential pattern mining.

---

## I. INTRODUCTION

In decision-making process, comparing various things with other one is one among the required steps that feature out. But this requires high data experience. For example throughout on-line searching a portable electronic items should have elaborated data and its specifications like Printers, Scanners, digital cameras, Graphic cards etc. In such case, it becomes troublesome for an individual with negligible data to create an honest call on that portable computer to shop for and further comparison the choice for constant. Magazine like shopper Reports, laptop Magazine and online media like CNet.com create efforts in providing editorial comparison content and survey to satisfy this would like. A judgment activity, in World Wide net era, typically involves and looks at for applicable websites insertion data relating to the targeted product, discovering difficult product, and acknowledge execs constraints. In this paper focus is on finding a collection of comparable entities provided a user's input entity. As an example, provided associate degree entity, Nokia (cell phone), wish to seek out comparable entities like Micromax, iPhone, blackberry etc. To extract comparable entities from relative matter, should always first recognize whether they are relative or not.

### Terms and concepts

#### A. Information Extraction

The process of identifying the structured information from unstructured or a semi- structured readable document is that term called as information Extraction. Methods used for information extraction

- 1) Rule-based
- 2) Pattern based
- 3) Supervised learning



#### **B. Ordered Pattern mining**

Sequential Pattern mining is primarily involved with finding statistical relevant patterns between knowledge examples whenever the values are delivered in an exceptionally sequence.

#### **C. Comparable Entity mining**

Proportionate entity mining achieves through the extract comparable entities from the text, questions or web corpus.

#### **D. POS Tags (Part-of-speech)**

Part-of-Speech of a word could be a linguistic class outlined by its syntactical or morphological behavior. Common POS classes are: noun, verb, adverb, adjective, pronoun, preposition and conjunction. Then there are several classes that arise from totally different styles of these classes. POS tags in this paper are NN: Noun, NP: noun phrases, NNP: suitable name, NNS: Noun plural, PRP: closed-class word, VBZ: Verb, present, person particular. JJR: relative Adjective. JJS: exceptional Adjective, CC: correlative combination. Our effort on comparable entity mining is expound to the study on entity and relation removal in information taking out. In step of our definition, a comparative question has got to be a question with intention to have distinction of minimum of 2 entities. Have a tendency to exploit this approach and develop a weakly supervised bootstrapping algorithm. The main aim of this algorithm is to compare queries and extract comparable entities at a time.

#### **E. Comparative Questions**

A matter whose purpose is matches two or more entities and these entities are expressly mentioned within the question.

#### **F. Comparator**

An entity is a particularly comparative question that is to be compared. In step of the definitions, Q1 & Q2 below aren't comparative queries where as Q3 is. "Chennai" and "Hyderabad" are comparable key words.

Q1. "Which one is better?"

Q2. "Is Hyderabad the most effective city?"

Q3. "Which city is healthier Chennai or Hyderabad?"

The results are very helpful serving to user's exploration alternative of different decisions by suggesting the comparable entities supported other previous user's requests

## **II. INFORMATION EXTRACTION**

In terms of discovering connected things to associate a tending entity, the work is similar to the analysis on recommended systems that suggest things to a user. Recommended systems primarily consider similarities between things and their applied math correlations in user log information. Take an example of Amazon recommend artifact to its customers supported their own history. However recommending associate in tending item isn't similar to find equivalent item. In case of Amazon, the aim of this paper is to get their customers to feature more things to their looking carts by suggesting similar or connected things. Bootstrapping process has been shown to be a very effective in previous data mining analysis [9]. Work is related to them in terms of method discrimination bootstrapping technique is to extract entities with a selected relation. However, our task is totally different from their needs. Only extracted entities (comparator extraction) guarantee that the entities are extracted from proportionate queries (proportional query identification), which is usually not needed in the task.

## **III. WEAKLY SUPERVISED METHOD FOR COMPARATOR MINING**

Weakly supervised algorithm is pattern based approach similar to J&L method, but it is different in many aspects as a replacement for using separate CSR (Class Sequential Rule) and LSR (Label Sequential Rule) [10] This method aims to learn sequential patterns are used to identify proportional question and mine comparators at the same time.

### **Mining Indicative Extraction Patterns**

The indicative extraction pattern mining approach is based on two assumptions

- 1) Many reliable pairs are extracted by make use of sequential pattern comparator pairs, it is very likely to be an IEP.
- 2) An IEP is used for extract the comparator pairs.

The reliable pairs are mainly based on above two assumptions. Bootstrapping algorithm starts with a single IEP. It extracts a set of initial comparator Pairs. For each comparison pair, all questions containing the keywords are retrieved from a question collection and regard as relative entities from the proportional questions and comparator pairs, based on the reliability scores sequential pattern is generated and evaluated. Patterns evaluated through reliable Indicative extraction pattern and added into an IEP storage area. Then new evaluation pairs are extracting from the related group of questions using the new IEP. The new comparators are added to a consistent comparator repository and used as new seeds for pattern learning in the next iteration. The overview of bootstrapping algorithm is shown below, where the database store seed of pairs and questions archive from relevant data. Entities are extracted from reliable Questions. Comparators are also extracted from query collection to allow finding new patterns efficiently in further iterations. The process iterates until no more new patterns can be found from the related queries collection. There are two key steps in this method

- 1) Pattern Generation
- 2) Pattern Evaluation

#### **1) Pattern Generation**

To generate sequential pattern, we used surface text pattern mining method introduced in [9]. For any given proportionate query and its comparator pairs, equivalent key words in the query Replaced with symbol \$Cs. #start and #end, special words with symbols are attached to the starting and the ending of a sentence in the question. Then the following three kinds of sequential patterns are generated from sequences of queries

#### **Lexical Pattern**

Lexical pattern indicate sequential patterns consisting of only words and special symbols like (\$C, #start, and #end). They are generated with make use of suffix tree algorithm [5] with two constraints: every pattern must contain at least two special symbols like \$C, and its frequency collection should be more than an analytically determined by number. These patterns mainly concentrate on general related information and extract the related information.

#### **Generalized Pattern**

Generalized pattern can be formed from lexical pattern this pattern contain N words consequently, generalize lexical patterns by replacing the words with their POS tags and one or more including special symbols \$Cs. POS consist of what bases to compare the entities, these rules are specified.

#### **Specialized Pattern**

In a few cases, a pattern can also be too general. For example, although a query have two entities like “Nokia or Samsung?” equivalent question, pattern “<\$C or \$C>” is general. And there can be many non equivalent queries matching. Pattern for instance. “True or false?” For this reason, perform pattern specialization by adding POS tags to all comparator slots. For example, from the previous pattern “<\$C or \$C>” and the question “Nokia or Samsung?”, “<\$C/NN or \$C/NN?>” will be produced as a specialized pattern.

#### **2) Pattern Evaluation**

According to first assumption, a reliability score  $R^k(p_i)$  for a candidate pattern follows:  $\forall$  at iteration k can be defined as

$$R^k(p_i) = \frac{N_Q(p_i \rightarrow c p_j) \forall c p_j \in c p^{k-1}}{N_Q(p_i \rightarrow *)} \quad (1)$$

Where  $p_i$  is extract known reliable comparator pair's.  $cp^{k-1}$  indicates the reliable comparator pair repositories accumulate until the  $(k-1)$  iteration ( $x$ ) means the number of questions satisfying a condition  $x$  the  $p_i \rightarrow cp_j$  denotes that  $cp_j$  can be extract from a appropriate question by applying pattern  $p_i$ . The condition  $p_i \rightarrow *$  denotes every question contain pattern. However, Equation (1) can suffer from imperfect information about reliable comparator pairs. Example Very few predictable pairs are generally discovered in early stage of bootstrapping. In this case, the value of Equation (1) might be underestimated which could affect the performance of equation (2) an individual Indicative extraction pattern from non-reliable patterns. Reduce this problem by applying above procedure. Let us denote the set of candidate patterns at the iteration  $k$  by  $R^k$ . define the support for comparator pair  $cp_i$  which can be extracted by  $R^k$  and does not exist in the current reliable set

$$S(cp_i) = N(p_k \rightarrow cp_i) \quad (2)$$

Where  $p_k \rightarrow cp_i$  means that one of the patterns can extract  $cp_i$  in certain questions. Naturally, if  $cp_i$  can be extracted by several candidate patterns in, it is likely to be extracted as a reliable one in the next iteration. Based on this perception a pair  $cp_i$  whose support is more than a threshold  $\alpha$  is regarded as a likely reliable pair. Look ahead reliability score  $Rp_i$  is defined:

$$R^k(p_i) = \frac{N_q(p_i \rightarrow cp_i) \vee Rel^k}{N_q(p_i \rightarrow *)} \quad (3)$$

Where  $Rel^k$  indicates a set of acceptable strong pairs based on  $RK$ . Include in the Equation (1) and (3), the final reliability score  $R(p_i)$  final  $k$  for a pattern is defined as follows:

$$(p_i) \text{ final } k = \lambda \cdot R^k + (1-\lambda) \cdot R(p_i) \quad (4)$$

Evaluate all candidate patterns and select pattern whose score is more than the threshold has IEPs. All necessary parameter values are empirically determined and are determine based on values.

#### IV. MINING INDICATIVE EXTRACTION PATTERN

Proportional query analysis is a grammatical expression used to make a request for information extraction or the request made using finding proportionate entities. Questions assess proportional judgment are often phrased as directed comparison, that is, one entity is compared with other one. The equivalent queries have keywords. Admin analyze the weather queries are comparative or not. IEP implementation this defines the sequential pattern identifies the starting and ending of the sentences.

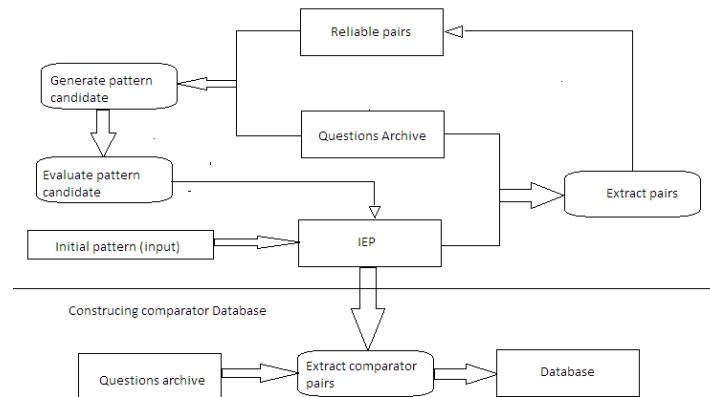


Figure: Overview of the bootstrapping Algorithm (curtsey [1])

A sequent pattern is named indicative extraction pattern if it will won't to determine proportional queries and extract comparators with high responsibilities. Then formally outline the unreliability score of a pattern. Once a matter matches an IEP, It's classified as a proportional query and therefore the token sequences admire the comparator slots within the IEP square measure extract as comparators.



### **1) Comparator Extraction**

The bootstrapping approach is prototype based mostly this approach to analyze the seed pairs. It's would like to extract the options supported the seed pairs. Review the patterns with many options. New comparator pair's mine from the queries collection exploitation the most recent IEP. The new comparators square measure value-added to a consistent comparator storage area and used as new seed for pattern learning within the next iteration. The method iterates till no additional new patterns will be found from the question assortment.

### **2) Decision Making**

Given a matter, choose the longest one in all patterns which may be applied to the question. Offer the choice creating supported pattern analysis. If a comparator is compared to several alternative significant comparators which may be again compared to the input entity, thought a valuable comparator in ranking.

### **3) Performance Evaluation**

Performance evaluation assess smart and the great proportionate query identification pattern and take out the great comparators and a decent comparator combine ought to occur in good comparative investigation to bootstrap the extraction and identification of methods and Then calculate Frequency pattern rate to see whether the choice creating is correct or not.

## **V. CONCLUSION**

This paper planned for analysis and enables several patterns to be enclosed and compared to other one. The equivalent entities are extracted from the proportionate queries that are obtained from the user. The simplest comparator is detected and comparators ranking methodologies is enforced for analyzing all the products and seek out their competitors. It's won't to dissolve the anomaly in entities. The results of comparator mining are often used for exchange search or product suggestion method.

Advantages of this paper

Improve recall and accuracy values Remove inconsistency in entities by exploitation comparator analysis with patterns. In future may be chance to improve pattern applications and mine the exceptional proportionate entities. And identify the comparator pair's aliases.

## **REFERENCES**

- [1] Shasha li, Chin-yew lin and Young-in Song 2013. "Comparable Entity Mining from Comparative Questions".
- [2] Raymond J. Mooney and Razvan Bunescu. 2005. Mining knowledge from text using information extraction. ACM SIGKDD Exploration Newsletter, 7(1):3-10.
- [3] Dragomir Radev, Weiguo Fan, Hong Qi, and Harris Wu and Amardeep Grewal. 2002. Probabilistic question answering on the web. Journal of the American Society for Information Science and Technology, pages 408-419.
- [4] Deepak Ravichandran and Eduard Hovy. 2002. Learning surface text patterns for a question answering system.
- [5] Gusfield .D, Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology. Cambridge Univ. Press, 1997.
- [6] Haveliwala.T.H, "Topic-Sensitive Page rank," Proc. 11th Int'l Conf. 2002.
- [7] Jeh.G and Widom.J, "Scaling Personalized Web Search," Proc. 12th Int'l Conf. World Wide Web (WWW '02), pp. 271-279, 2003.
- [8] Stephen Sunderland. 1999. Learning information extraction rules for semi-structured and free text. Machine Learning, 34(1-3):233-272.
- [9] D. Ravichandran and E. Hovy, "Learning Surface Text Patterns for a Question Answering System," Proc. 40th Ann. Meeting on Assoc. For Computational Linguistics (ACL '02), pp. 41-47, 2002.
- [10] N. Jindal and B. Liu, "Mining Comparative Sentences and Relations," Proc. 21st Nat'l Conf. Artificial Intelligence (AAAI '06), 2006.



## International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

[www.iasir.net](http://www.iasir.net)

### An Automatic File Replication Mechanism in Computational Grid

Ch. Ramesh Babu<sup>a</sup>, Dr. Ch. D. V. Subba Rao<sup>b</sup>

<sup>a</sup> Research Scholar, JNTU Kakinada, AP, India.

<sup>b</sup> Professor, Dept of CSE, SV University Tirupati, AP, India.

**Abstract:** The massive dynamic virtual computing systems often generate large number of files as checkpoints for fault tolerance. In computational grid the vulnerability of node is high, so nodes more prone to frequent failures due to heterogeneous and autonomic nature of grid computational resource. To facilitate this, grid resources takes checkpoints to recover loss several hours/days of computation using replication. Replication is one of the widely used methods to provide fault tolerance in computational grid environment. In this work we have developed an automation of checkpoint replication by using Two Level Scheduling Strategy model, so as to select optimal failure resources to reduce checkpoint replication in computational grid. The failure can be systematically restarted with minimal interval of time for each grid resource using automatic replica (context file). The experimental results has shown that the proposed Automatic File Replication- AFR methodology considerably reduces the replication overhead, delay, storage space and turn-around time when compared to adaptive replication approaches.

**Keywords:** automatic replication, fault tolerance, computational grid.

#### I. Introduction

Computational grid is an independent multiple-owner heterogeneous architecture to accomplish huge computationally intensive tasks. The computational grids are the current trend advanced technology to execute the large scale application in assorted heterogeneous resources. Advanced computing technologies are changing to virtual resources sharing, resource pooling in dynamic environment [4, 6]. Most frequently computational grid faces problems with delay and job failures due to heterogeneity and autonomic nature of grid's resources. Many times computational grids are prone to failure due to lack of proper resource scheduling and fault tolerance mechanism. So the failure rate of the nodes is relative to the number of processors. The computational grid workload-logs study says that system failure occurred 2 to 5 times per day on a large computing system which has about on average of 5000 processors [18]. Even though the number of failures or failure rate of each processor is somewhat low, it might affect the node's reliability. As a result it is need to make a fault tolerant system. Most of the computational problems, the complexity is habitually exacted using the number of failure nodes with respect to the number of calculations necessary to solve them and to find the complexity. So the large computational problems take much time to solve it. In many algorithms computational problems which have more calculation and computational capability such as climatologic forecast, analysis of weather forecasting, scientific simulations, seismic analysis and the genome sequencing like many computational applications needs to investigate more number of parameters which requires large scale parallel processing systems. High performance computing (HPC) can be used effectively to solve and address these types of issues. In HPC the jobs are allocated to complex heterogeneous group of dissimilar nodes, in order that jobs might be executed concurrently in independent nodes or processors [1, 27].

Typically these sophisticated computational resources require fault tolerance mechanism. According to the large scale, dynamic virtual computational resources such as NASA iPSC, LANL CM5 replica-logs traces shows that job delay or job failure is the major challenging task to provide fault tolerance [2, 3, 7]. Since very robust fault tolerant checkpointing, replication and scheduling algorithms are used to handle various resource allocations in computational grids. However there are some performance issues needs to enhance the parameters such as number of replication overhead, replication process context storage space and low throughput. Consequently the major issue in fault-tolerance is to effectively tolerate failures using job replication and fault-tolerant job scheduling to make efficient failure handling mechanism in presence of faults. Currently using techniques for fault tolerance in the widely held computational applications are adaptive checkpointing and replication. Since most of the job replication techniques are not merely based on scheduling algorithm [3, 14]. Since very robust scheduling algorithms are used to handle various resource allocation in computational grids. However, it is necessary in earlier studies to remedy the failures and delay of executing jobs with respect to resource availability, which can handle scheduling and failures in any large scale high performance computational applications [15, 16]. The major issue is fault-tolerance with regard to job scheduling and failure handling mechanism. Currently using techniques for fault tolerance in the

widely held computational applications are adaptive replication or without checkpoint automation and replication [21].

The main objective of computational grids is to execute large computational tasks more effectively to improve throughput and to save processing time. Therefore the user submits jobs to the Grid Scheduler (GS) along with their Quality of Service (QoS) requirements. These requirements may include the deadline in which users want jobs to be executed, the type of the resources required to execute the job and the type of the platform needed. The GS of the present scheduling systems allocates each job to the most suitable resource. In case of fault free resource, the results of executing the job are returned to the user after completion of the job. If a grid resource fails during the execution of a job, then that job is rescheduled on another resource which starts executing the job from scratch. This leads to more time consumption for the job than expected time so that the user's QoS requirements may not satisfy [4, 24]. In this paper we have presented heuristics that resolves number of replica overhead issue so as to provide high job throughput in presence of failure [25].

## II Automatic File Replication Service

The major drawbacks in most of grid simulators such as SimJava, Bricks, MicroGrid, SimGrid, MicroGrid, SimGrid, GangSim, MONARC, OptorSim, EcoGrid, Opportunistic Grid, SchMng are based on scheduling of jobs to grid resources and lack of synchronization among the simulation components and the grid middleware. These are all based on to simulate scheduling of jobs, job migration, replication analysis etc. Most of these simulators don't analyses overhead of grids completely and not support for fault tolerant analysis such as checkpointing and replication, whenever a node fails it has to recover from last saved checkpoint to overcome job delay. But many times there is an overhead due to many replicas files in without appropriate file replication automation approaches. The simulators GridSim for scheduling of jobs and OptorSim data replication analysis toolkit are java based simulation toolkits. In order to tackle this we have proposed an optimal checkpoint replication automation strategy which effectively changes the dynamic behavior of the grid middleware working mechanism according to the grid resource failure information [9, 20, and 25].

The problem of transferring a checkpoint from one site to another in real time takes bandwidth consumption and access delay. Data replication technique is a frequently reduce bandwidth consumption and access latency by taking one or more checkpoints to other sites. Two kinds of replication methods are possible. The static replication creates and manages the replicas manually. The dynamic replication changes the location of replicas and creates new replicas to other sites automatically. Judging from the fact that resources or data checkpoints are always changing, it is clear that dynamic replication is more appropriate for the Data Grid [5, 8]. In this we have observed several dynamic replication strategies proposed in earlier [20, 26]. The multi-tier hierarchical Data Grid architecture supports an efficient method for sharing of data, computational and other resources. In [11, 12], two dynamic replication algorithms had been proposed for the multi-tier Data Grid, which are Simple Bottom-Up (SBU) and Aggregate Bottom-Up (ABU). The basic concept of SBU is to create the replicas as close as possible to the clients that request the data checkpoints with a rate exceeding a predefined threshold. If the number of requests for checkpoint  $f$  exceeds the threshold and  $f$  has existed in the parent node of the client which has the highest request rate, then there is might not to replicate [13, 17 and 23]. The SBU replication algorithm has the disadvantage that it does not consider the relations among these historical records but processes the records individually. After aggregation, the checkpoint with the highest request rate will be replicated if the rate is above a threshold. With the exception of aggregation steps, all concepts in ABU are similar to the SBU replication algorithm [1, 3, and 14].

## III Automatic Replication File model for computational grid

The aim of this work is to optimize the performance of the grid in the presence of faults and to improve throughput value. When a fault occur a grid resource may not complete its job within the given QoS. The main strategy of the proposed AFR mechanism is to minimize the effect of grid faults and to reduce the fault recovery time using optimal automation of replication of context so as to minimize the amount of checkpoint overheads. To evaluate the above prototype we have considered faults history so that each resource scheduled based on the next sequence of pattern failures. The failure patterns can be scheduled by Two Level Scheduling Strategy to assign checkpoint interval and also to provide automatic failure replica (context file of checkpoint) to the grid resource. The interaction among different components of the AFR is shown in the following Figure.1. The AFR optimizes the checkpoints to improve the efficiency over the execution of the failed job from the last saved checkpoint [25]. Thus it reduces the response time of the job by reducing the time wasted in additional checkpoints storage. A grid contains multiple grid resources that provide computing services to users. The main component of the AFR is the cluster information to reduce the data transfer time by reducing number of intercommunications [19, 22]. It receives the jobs with their information from users. Job information includes job number, job type, and job size. Also the user submits QoS requirements for grid application, such as the deadline to complete its execution, the number of required resources and the type of these resources.

The main function of Work\_Manager is to find and sort the most suitable resources that can execute the job and satisfy user QoS requirements. In order to perform this function, the Work\_Manager connects to the External\_Schedulers to get information of all available grid resources to execute the job. The Work\_Manager uses response time, resource failure rate and resource failure time to construct the list of suitable resources that can execute the job.

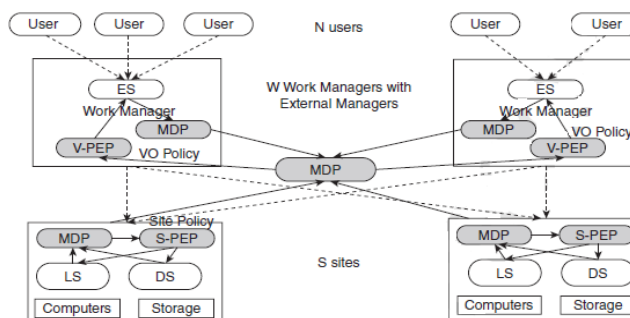


Figure 1: GridSim Components.

In Figure.1 GridSim architecture contains the components such as Work\_Manager, External\_Schedulers, Virtual\_Organization Points, and autonomous computational resources as portable batch system message oriented middleware (Local\_Schedulers and Data\_Schedulers). Each component accomplishes set of tasks, explained in 3.2 and 3.3.

To solve grid application, the grid user submits the jobs along with QoS which requires large computational power for computation. Virtual\_Organization Points gets the list of failure resources and checkpoint file information using computational resource replica-logs. The Work\_Manager gets the failure patterns, checkpoint intervals and available resources from Virtual\_Organization Points, optimal checkpoint automation library and External\_Schedulers respectively. The External\_Schedulers gets the all the information about the resources like response time, failure rate and sorts the list of resources according to resource response time and make new list of available resources with QoS requirements to find checkpoint interval [10]. The Work\_Manager finds the optimal checkpoints for each resource and sends the list to External\_Schedulers to allocate the jobs to resources. The External\_Schedulers allocates the jobs to the selected resources. Based on the checkpoint interval the Virtual\_Organization Points takes the snapshots (context file) from computing nodes. If any failure occurs in a node then latest context job file can be retransmitted automatically. If more number of failures occurs in the same grid site then the computational resource can be rescheduled to another site. Optimal checkpoint automation library reschedules the jobs to new grid resource through External\_Schedulers. External\_Schedulers collects aggregated results and ships it to grid user.

This component takes grid application from the user which typically requires large scale heterogeneous resources, processing nodes with QoS of an application. External\_Schedulers contains information of all available resources in the grid required by the Work\_Manager. The information includes resource speed, current load, resource failure rate and total failure time of each resource. For each job  $j$  dispatched by the Virtual\_Organization Points, if Virtual\_Organization Points receives a job completion message then it sends a message to the External\_Schedulers to increment  $S$  (Success) or if Virtual\_Organization Points receives a response as failure message then send it a message to External\_Schedulers to increment  $F$  (Failure), if there is a checkpoint stored then Virtual\_Organization Points dispatches the not completed part of the job along the checkpoint status to the second resource in the resources list, else Virtual\_Organization Points dispatches the whole job along to the second resource in the resource list (see Figure 2).

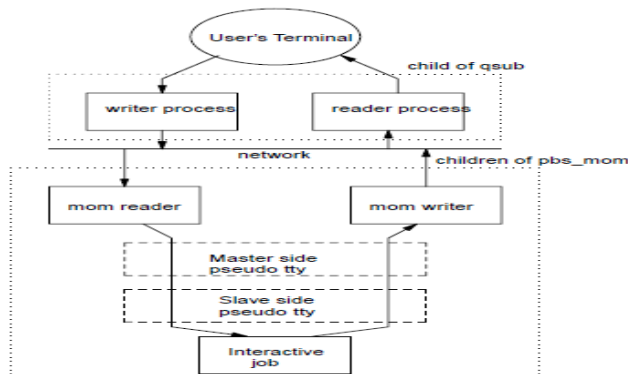


Figure 2: Interactive Job Communication Flow.

The External\_Schedulers receives and stores partially executed results of a job from children of Local\_Schedulers and Data\_Schedulers. These intermediate results are called checkpoint status. For each job there is only one record of checkpoint status. When Virtual\_Organization Points receives a new checkpoint status it overwrites the old one and if it receives a job completion message from the resource it removes the record of such job.

Virtual\_Organization Points is an important component of AFR. The main functions of Virtual\_Organization Points are determining the number of checkpoints for each Local\_Schedulers and Data\_Schedulers and the checkpoint interval for each job. Virtual\_Organization Points receives a job with its assigned list of resources from Work\_Manager. It connects to Virtual\_Organization Points to get information about the failure history of grid resources assigned to the job. Based on failure rate of the resource, the Virtual\_Organization Points determines the number of checkpoints and the checkpoint intervals for each job. Then, it submits the job to the first grid resource in the resources list. The Virtual\_Organization Points calculates the number of checkpoints and the checkpoint interval for each resource using replica log files which are collected.

The Work\_Manager is the major component in grid architecture. The Work\_Manager initiates the application with n-number of jobs to allocate grid resource. Now for each set of job the Work\_Manager gets the list of available resources from External\_Schedulers. The Work\_Manager sorts the suitable list of resources according to resource response time -RRT, failure rate -FR and average failure time -AFT. Afterwards Work\_Manager dispatches the list to checkpoint header. The Work\_Manager schedules the resources using optimal checkpointing approach-AFR with respect to job list acquired from Virtual\_Organization Points.

The goal optimizing checkpoint intervals is to minimize the number of checkpoints for each resource. This work presents the design and analysis of AFR using GridSim and implementation of the automatic replication checkpoint process context using OpenMPI toolkit, including high level complex applications context file recording, failure detection and application recovery, using BLCR framework in TORQUE.

$$\text{MIN } TT_{(X)} = \sum_{i=0}^N (T_{ij}^1 + T_{ij}^1 * F_i^1) \quad (1)$$

$$\text{Where } T_{ij}^1 = t_{js} + t_{je} + t_{jr} \quad (2)$$

$$F_i = t_{jc} + t_{jr} \quad (3)$$

Subject To:

$$(i). T_{ij}^1 < RT_{ij}^i$$

$$(ii). 0 < F_i < 1$$

Where,  $TT_{(X)}$  is the estimated minimum turnaround time for j resources to carry out application X of i jobs to checkpoint in optimal manner.  $T_{ij}^1$  is the estimated minimum turnaround time for j resources to carry out job i,  $F_i^1$  is the failure resource,  $t_{js}$  scheduling time for  $i^{th}$  job on resource j,  $t_{je}^{ij}$  execution time for  $i^{th}$  job on resource j,  $t_{jr}$  restart time for  $i^{th}$  job on resource j.

Find the failure probability density function of resource using exponential distribution as

$$P(k, j | ) = \begin{cases} kje^{-kix}, & x \geq 1 \\ 0, & \text{Otherwise} \end{cases} \quad (4)$$

Each resource x checkpoint interval I can be generated as using Bayes formula

$$CI_i = P(x_i | k, j) = P(k, j | x_i) * P(x_i) \quad (5)$$

Differentiate interval I(x) with respect to mean and last failure of resources.

$$l1 = \frac{d^*CI_i}{dj} \quad (6)$$

$$l2 = \frac{d^*CI_i}{dk} \quad (7)$$

Optimal checkpoint interval can be calculated as

$$CI_{i\_new} = (l1+l2)/2 \quad (8)$$

Where,  $CI_{i\_new}$  is the new optimal checkpoint interval. The above equations 4-8 are used to simulate grid failure resource with optimal number of checkpoints by using the parameters LastFailure (k) and MeanFailure (j). Here  $CI_{i\_new}$  is the optimal checkpoint interval which reduces the number of checkpoints, thus the checkpoint overheads are reduced.

#### IV Replication Automation

In our mechanism, fabrication of snapshot points are generated automatically (checkpoints) in a frequency predicted by the job scheduling strategies, based on OpenMPI collective function calls. The implementation of the automatic checkpoint operations is based on threads created in runtime, which execute the new functionalities inserted into OpenMPI using TORQUE. We have implemented different threads for different applications. The following are the scheduling strategy and thread workflow mechanism. The cycle of steps among the components of planned tactic of checkpoint plotting and replication are explained in equations 9 to 14. The Hidden Markov Model (HMM) is a state machine. Here the failure states of the model are represented as nodes and the failure transition are represented as edges. The HMM have become the well-known and widely used statistical approach to characterizing the spectral properties failure prediction approaches. HMM is a stochastic modeling tool having an advantage of providing the high reliable and natural way of failure analysis for resources. HMM integrates into the systems involving information related to last and mean failure approaches, currently it is predominant approach for the optimal checkpointing in computational grid. HMM provides a method which directly estimates the conditional probability of index of failures in resource given a hypothesized identity for index of failures in resource. HMMs consists of two processes namely Hidden and the Observed process. The Hidden process consists of a collection of failure states connected by the resources. And each of these transactions is described by two sets of probabilities:

Steps involved in making HMM's work:

Estimating conditional probabilities to last and mean failure sequence to given a model AFR methodology.

Finds the best checkpoint interval which is closely matches the input sequence. This enables to assign optimal checkpoint interval from last failures pattern. Prepare a model to using the mean and last failure parameters of checkpoints and its corresponding transition probabilities to best account for a checkpoint library. The failure patterns of the resources can find using the formula

$$w_j = P(x_i | r_j) = P(r_j | x_i) * P(x_i) / P(r_j). \quad (9)$$

The grid Resource are sorted by using

$$(r_j, w_j) \quad r_j = r_j * w_j \quad (10)$$

Finding the checkpoint interval

$$CI = \sum_{i,j=1}^{n,m} (TT_{(X)}) / N_r^i; \quad (11)$$

Finding number of checkpoints

$$CN = T_j^i / CI_i; \quad (12)$$

According to number of checkpoints the resources can be rescheduled using

$$CT_r^i \% CI_i = 0 \ \&\& \ RE_r^i > CI_i \quad (13)$$

Finding the turnaround time of the application  $X_{i=1}^n$ .

$$TT_j = T_r^i * (1 + CN_r^i) \quad (14)$$

Throughput of the application is

$$n = N / \sum TT_j^{tp} \quad (15)$$

Where,  $X_{i=1}^n$  Computational jobs  $(x_1, x_2, \dots, x_n)$ ,  $CI_j$  -  $j^{th}$  resource failure weight,  $n$ - Throughput of an application  $X$ ,  $CN_j$  - Number of Checkpoints for resource  $j$ ,  $TT_{(X)}$  - Total turnaround time of application  $X$ .

Computational grid User can submit jobs through the grid user interface. The application interface receives user jobs and transforms to scheduler. Typically the job information consists of job number, job type, and job size

also receives Quality of Service requirements of each job such as the deadline to complete its execution, the number of required resources and the type of these resources. The scheduler assigns each job to the most reliable, suitable, and available resource to execute the job. The most reliable resource is the resource that has a lower fault rate.

### V. Performance Evaluation

In this experiment, applications with 10 to 200 jobs with 10-200 faults are modelled. The size of each job replica context is approximately selected from 1 KB up to 250 MB. The number of resources in the grid can reach up to 4 clusters with cluster have 30 nodes. Different simulation experiments have been conducted with variation in the total number of jobs submitted to the grid and measuring the throughput, turnaround time and the tendency of resources to fail. The proposed AFR approach is compared with the combination of Two level replication strategy (TLRS) with Two level scheduling strategy (TLSS), Structure based on internet hierarchy (BHR) with TLSS, Least Recently Used (LRU) with TLSS the details of that experiment can be found in Table-1. Based on the experiments the following graphs Figure 3 and Figure 4 are plotted. In this experiment, applications with 100 to 2000 jobs with 10-200 faults are modeled. The size of each job is randomly selected from 1 KB up to 10 MB. The number of resources in the grid can reach up to 10. Different simulation experiments have been conducted with variation in the total number of jobs submitted to the grid and measuring the throughput, turnaround time and the tendency of resources to fail. The proposed AFR approach is compared with the other replications methods, the details of the experiments are be found in Table-1&s. Based on the experiments the following graphs Figure 3 and Figure 4 are plotted.

Replication Algorithm	Number of Jobs				
	10	50	100	150	200
AFR	350	1750	3600	4700	7200
TLRS+TLSS	360	1800	3800	4800	7500
BHR+TLSS	365	1850	3900	4940	7650
LRU+TLSS	370	1880	4100	5100	7900

**Table 1: Average replication time (sec).**

Replication Algorithm	Number of Jobs				
	10	50	100	150	200
AFR	14	64.81481	124.1379	156.6667	218.1818
TLRS+TLSS	14.4	66.66667	131.0345	160	227.2727
BHR+TLSS	14.6	68.51852	134.4828	164.6667	231.8182
LRU+TLSS	14.8	69.62963	141.3793	170	239.3939

**Table 2: Average storage usage (mbs).**

The Table-1& 2 shows the comparison between AFR approach with TLRS, TLSS, BHR and LRU with TLSS strategy [7, 28]. Most of the existing grid simulators like GridSim, SimGrid do not support fault-tolerant scheduling and checkpointing. To assess the optimal replica file automation operation, we have arranged heterogeneous nodes called NUMA Non-uniform memory architecture environment in TORQUE resource manager which composed of 50 nodes, running Open SUSE Linux 12.2. Each node has 1 GB RAM, 1 GB swap and a Gigabit network card. We have taken the first five nodes to operate Open SUSE Linux with TORQUE resource manager and the remaining nodes to execute Windows-XP to access and operate the jobs through Putty software. The TORQUE resource manager assigns the jobs to the idle resources. We have used the Networking File System (NFS) to store replica context into the checkpoint server so that if any one node fails then chronologically the last saved context file can be accessed automatically. At present our claim methodology ropes optimal replica automation for large scale grid application small statistical samples. We have installed TORQUE resource manager service on all nodes and applied auto replica generation based on optimal replication interval. Since it improves the performance of the grid application in presence of a failure when compare to TLRS&TLSS approaches.

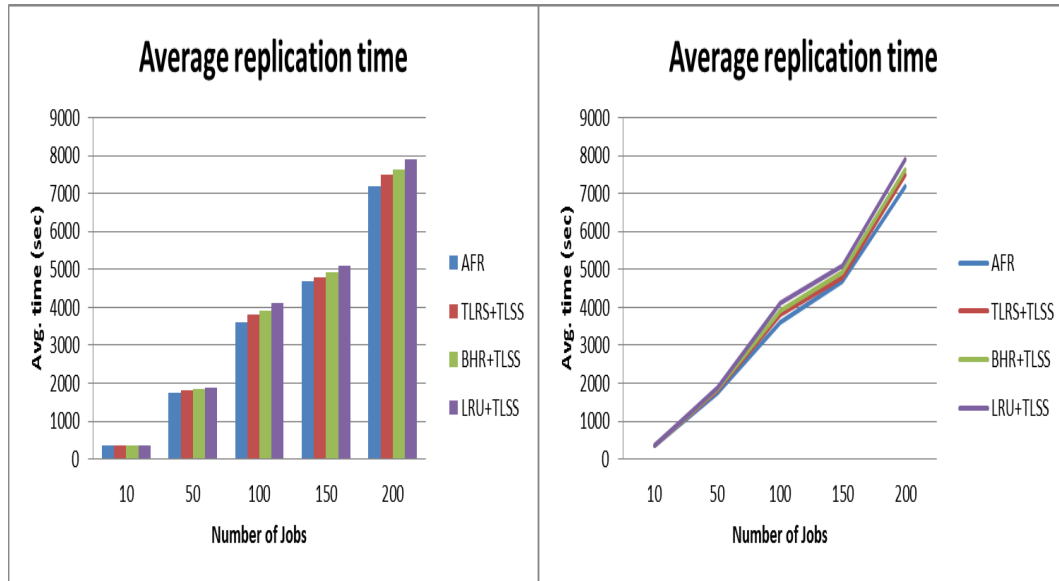


Figure 3: Processing time comparison by varying number of jobs.

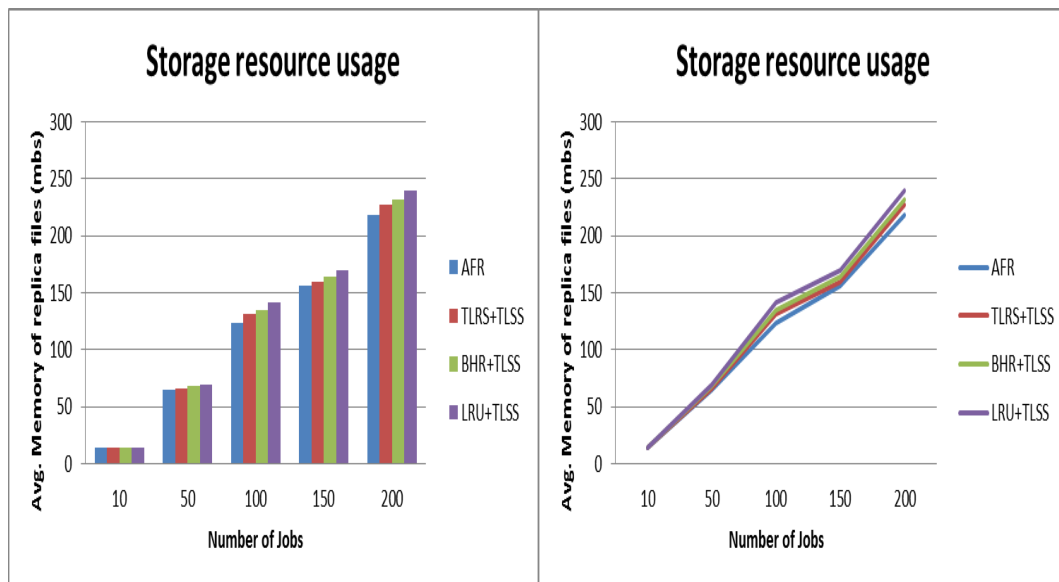


Figure 4: Space comparison by varying number of jobs

The Figure 3, Figure. 4 depict processing time and space comparison of checkpoint algorithms respectively. The proposed system depends on average failure time and last failure rate to make optimal checkpointing decisions using HMM model. The checkpoint interval is calculated merely using resource failure rate prior information. The performance of AFR algorithm is compared with LRU, BHR and TLRS in which these algorithms depend on TLSS scheduling to make replication. But the AFR method uses resource fault index to make replication using Hidden Markov Model states (HMM) to make replica scheduling for calculating effective prediction based replication files. Experimental results show that AFR approach effectively replicates jobs in the presence of failures. It improves the turnaround time and throughput when compared with the adaptive checkpointing algorithms. Moreover the failure tendency for the proposed AFR methodology is far better than the TLRS&TLSS algorithm. Thus, it can be concluded that the proposed scheduling system provides better performance when compared to adaptive Algorithm. In this approach the failure history of resources must be known in prior to make the optimal replication approach using fault tolerant scheduling approach. If the computational resources are new then checkpointing interval might be unpredictable.

## VI. Conclusions

In this work we have proposed an Automatic File Replication mechanism for computational grids. Here the optimal replica files are calculated using the predicted failure patterns of resources using HMM. The unique optimal replica context files are generated for each resource, which yields reduction in number of context replicas. Hence this methodology considerably reduces the replication overheads by reducing number of context files. The developed AFR model which shows better turnaround time, storage space and throughput of different

jobs compare to LFU and TLSS algorithms. Experimental results show that AFR methodology effectively tolerates the resource failures. Therefore we conclude that the developed checkpointing system provides better performance when compare to adaptive replication approaches.

## References

- [1]. Jianhua Jiang, Huifang Ji, "scheduling algorithm with potential behaviors", Journal of Computers, VOL. 3, NO. 12 , December 2008.
- [2]. Ming Tang, Bu-Sung Lee, Xueyan Tang, Chai-Kiat Yeo. "The Impact of Data Replication on Job Scheduling Performance in the Data Grid", Future Generation Computer Systems, Volume 22, Issue 3, February 2006, Pages 254-268
- [3]. Ali Elghirani, Riky Subrata, Albert Y. Zomaya, and Ali Al Mazari., "Performance Enhancement through Hybrid Replication and Genetic Algorithm Co-Scheduling in Data Grids", Advanced Networks Research Group, School of Information Technologies, University of Sydney, NSW 2006 Australia.
- [4]. Sang-Min Park, Jai-Hoon Kim, Young-Bae Ko: "Dynamic Grid Replication Strategy based on Internet Hierarchy", Book Series Lecture Notes in Computer Science, Grid and Cooperative computing book, Publisher Springer, August 2005, Volume 3033/2004, Pages 838-846
- [5]. Ruay-Shiung Chang, Jih-Sheng Chang, Shin-Yi Lin, "Job scheduling and data replication on data grids", Future Generation Computer Systems, Volume 23, Issue 7, August 2007, Pages 846
- [6]. Ashish Kumar Singh, Shashank Srivastava and Udai Shanker, "A Survey on Dynamic Replication Strategies for Improving Response Time in Data Grids", JBSTR, JULY 2013.
- [7]. Ruay-Shiung Chang ; Nat. Dong Hwa Univ., Hualien ; Hui-Ping Chang ; Yun-Ting Wang, "A dynamic weighted data replication strategy in data grids ", Computer Systems and Applications, AICCSA 2008, IEEE/ACS 2008.
- [8]. A. Chervenak, I. Foster, C. Kesselman, C. Salisburly, and S. Tuecke, "The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets," Journal of Network and Computer Application, vol. 23, pages 187-200, 2000.
- [9]. D. G. Cameron, R. C. Schiaffino, P. Millar, C. Nicholson, K. Stockinger, and F. Zini, "OptorSim: A Grid Simulator for Replica Optimization," UK e-Science All Hands Conference 31 August - 3 September 2004.
- [10]. I. Foster, "Globus Toolkit Version 4: Software for Service-Oriented Systems," IFIP International Conference on Network and Parallel Computing, Springer-Verlag LNCS 3779, pp 2-13, 2005.
- [11]. W. Hoschek, F. J. Jaen-Martinez, A. Samar, H. Stockinger, K. Stockinger, "Data management in an International data grid project", Proceedings of the First IEEE/ACM International Workshop on Grid Computing (GRID '00), Lecture Notes in Computer Science, vol. 1971, pages 77-90, Bangalore, India, December 2000.
- [12]. The European Data Grid Project, <http://eudatagrid.web.cern.ch/eu-datagrid/Winter>.
- [13]. Nhan Nguyen Dang, Sang Boem Lim: "Combination of Replication and Scheduling in Data Grids", IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.3, March 2007.
- [14]. T. Phan, K. Ranganathan, and R. Sion, "Evolving toward the perfect schedule: Co-scheduling job assignments and data replication in wide-area systems using a genetic algorithm", Job scheduling strategies for parallel processing (11th international workshop), JSSPP 2005, Cambridge MA, 2005.
- [15]. Klaus Krauter, Rajkuar Buyya and Muthucumaru Maheswaran, "A taxonomy and survey of grid resource management systems for distributed computing", 2002; 32:135-164 (DOI: 10.1002/spe.432).
- [16]. Klaus Krauter and Manzur Murshed, GridSim a toolkit for the modelling and simulation of distributed resource management and scheduling for Grid computing, CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE Concurrency Computat.: Pract. Exper. 2002; 14:1175-1220 (DOI: 10.1002/cpe.710).
- [17]. The Data Grid Project. <http://www.eu-datagrid.org>
- [18]. Parallel workload Project. <http://www.parallelworkload.org>.
- [19]. The European data grid project. <http://eudatagrid.web.cern.ch/eu-datagrid/Winter>.
- [20]. W.H. Bell, D.G. Cameron, L. Capozza, P. Millar, K. Stockinger, F. Zini, Simulation of dynamic grid replication strategies in OptorSim, in: Proceedings of the Third ACM/IEEE International Workshop on Grid Computing, Grid2002, Baltimore, USA, in: Lecture Notes in Computer Science, vol. 2536, 2002, pp. 46-57.
- [21]. E. Deelman, H. Lamahmedi, B. Szymanski, S. Zujun, Data replication strategies in grid environments, in: Proceedings of 5th International Conference on Algorithms and Architecture for Parallel Processing, ICA3PP'2002, IEEE Computer Science Press, Beijing, China, 2002, pp. 378-383.
- [22]. H.H. Mohamed, D.H.J. Epema, An evaluation of the close-to-files processor and data coallocation policy in multiclustes, in: 2004 IEEE International Conference on Cluster Computing, IEEE Society Press, San Diego, California, USA, 2004, pp. 287-298.
- [23]. M. Carman, F. Zini, L. Serafini, K. Stockinger, Towards an economy based optimisation of file access and replication on a data grid, in: Proceedings of 2nd IEEE/ACM International Symposium on Cluster Computing and the Grid, CCGrid 2002, IEEE-CS Press, Berlin, Germany, 2002, pp. 340-345.
- [24]. P. Kunszt, E. Laure, H. Stockinger, K. Stockinger, Advanced replica management with rector, in: Proceedings of 5th International Conference on Parallel Processing and Applied Mathematics, PPAM 2003, Czestochowa, Poland, September 2003, pp. 848-855.
- [25]. D.G. Cameron, A.P. Millar, C. Nicholson, OptorSim: A simulation tool for scheduling and replica optimisation in data grids, in: Proceedings of Computing in High Energy Physics, CHEP 2004, Interlaken, Switzerland, September 2004.
- [26]. K. Ranganathan and I. Foster, "Identifying Dynamic Replication Strategies for a High Performance Data Grid". In Proceedings of the International Grid Computing Workshop, Denver, Colorado, USA, 2001. International Journal of Grid Computing & Applications (IJGCA) Vol.1, No.1, September 2010
- [27]. I. Foster, K. Ranganathan, Decoupling computation and data scheduling in distributed data intensive applications, in: Proceedings of the 11th IEEE International Symposium on High Performance Distributed Computing, HPDC-11, IEEE, CS Press, Edinburgh, UK, 2002, pp. 352-3
- [28]. Somayeh Abdi, Somayeh Mohamadi, "Two Level Job Scheduling and Data Replication in Data Grid", IJGCA-Sep-2010.

## **An Analytical Model of Consistent Check Pointing In Computational Grid**

**Ch. Ramesh Babu<sup>a1</sup>, Dr. Ch. D. V. Subba Rao<sup>b</sup>**

<sup>a</sup> *Research Scholar, JNTU Kakinada, AP, India. E-mail:crb.challagundla@gmail.com*

<sup>b</sup> *Professor, Dept of CSE, SV University Tirupati, AP, India.*

### **Abstract**

The TORQUE resource manager tool provides an environment for open-source batch processing jobs and distributed computing nodes, which simulates grid based operating system based on Open SUSE Linux. To generate consistent checkpointing for grid applications we used a distributed grid check pointing service called Torque GCP. This service is considered to support various checkpointing protocols and different checkpointing packages such as BLCR, Open MPI in an efficient manner. In this paper, we combined both the features of last failure and mean failure checkpointing protocol with a FTCS scheduling strategy to generate consistent check pointing using Torque GCP service. To assess this prototype we run it within a heterogeneous environment composed of autonomous computing nodes in different LAN cluster. The experimental results prove that the capability of the Torque GCP service to integrate different optimal checkpointing protocol in a distributed application within a heterogeneous computational grid. Furthermore, the performance valuations also prove that our solution better performs than the existing checkpointing approaches in terms of space and throughput.

### **Introduction**

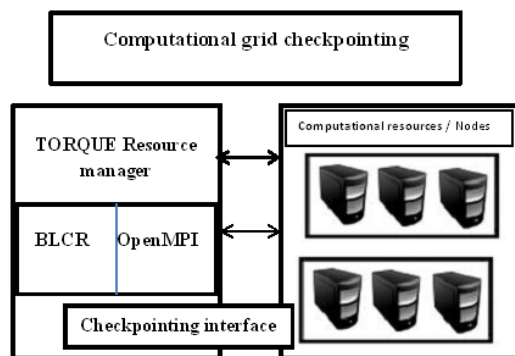
A Grid is a type of distributed computational environment which supports resource sharing, coordinated use of geographically dispersed nodes and autonomous management of physical and dynamic virtual organizations that share the common goal of solving large-scale applications. The increasing demands of computational power of large scale applications, grids become popular managing and running large computing resources during the last fifteen years. Computational Grids provide a large-scale distributed environment for the execution of various forms of computational and data relevant applications such as particle system simulation studies, weather prediction etc., across many varied resources [5]. These large applications executing on Grid or cluster architectures consisting of computational nodes often creates problems with reliability. The source of the problems is node

failures and job delay which leads to dynamic configuration over extensive runtime [2].

Computational Grids and Cluster Computing architectures have gained popularity for computationally intensive parallel applications. However, the complexity of the organization, consisting of autonomous computational nodes, large storage, and interconnection networks, poses great difficulties with respect to overall system reliability. Simple tools of reliability analysis show that as the difficulty of the system increases, its reliability, and thus, Mean Time to Failure (MTF), decreases [8]. The reliability of the computational grid systems is computed as the product of the reliabilities of all system components or nodes. For applications executing on large clusters or a Grid, the long-time execution tasks may exceed the MTF of the infrastructure and, thus, reduce the execution infeasible. The high failure probabilities are due to the fact that, in the absence of appropriate fault tolerance mechanisms, the failure of a single node will cause the entire execution to fail or delay. Hence this simple example does not even reflect network failures, which are typically more likely than computer failure. Fault tolerance is thus a necessity factor to consider to avoid failure in large applications, such as found in scientific computing, executing on a Grid, or large cluster [6]. Efficient technique of fault tolerance for distributed applications is periodic checkpointing. But this checkpointing approach further leads to more number of checkpoints overhead due to appropriate checkpointing algorithm. The following sections discuss the various checkpointing studies, limitations and proposed solution with results [10].

### Torque Grid Check pointing Service

In this section we briefly describe the architecture of the Torque Grid Check pointing service (Torque GCP) and its components related to optimal checkpointing [23]. Fault tolerance technique based on checkpointing can be used to save the process state of the application on secondary storages so as to recover checkpoint in the time of failure or job delay. Moreover, checkpointing and replications are crucial in the process of virtual grid resource in order to distribute the jobs across multiple resources the Open MPI is used [11,12, 15, 17].



**Figure 1:** Torque Grid Check pointing architecture

In figure 1 it shows that the computational jobs large application chunks can be distributed across many resources, most of the times it is not sufficient to exclusively rely on the discrete checkpoints of the processes in order to compute consistent global state. The messages needed to be handled otherwise it can lead to inconsistent checkpoints [5]. The checkpoint context files are process states whose receive events are part of the destination process checkpoint but the corresponding send events are lost. In case of a recovery the destination process would receive that checkpoint context twice, which could result in unpredictable application behaviour. On the other hand, checkpoint context occur when the send events are part of the sender-side checkpoint, however the receiving events are lost. As a consequence, checkpointing protocols and adaptive checkpointing systems are needed in order to address the problem of inappropriate checkpointing. In the two common checkpointing protocols, the processes uncoordinated their checkpoints or take them independently. Some examples for adaptive cluster-level checkpointing include BLCR and Open MPI [21, 22], which supports both protocols within MPI which integrates the checkpointing protocol [18, 19 ].

Computational Grids are large-scale distributed and heterogeneous resources which can be composed out of hundreds or thousands of nodes. It is clear that accuracy of each grid node to use one checkpointing approach. Therefore, in order to achieve fault tolerance for large computational applications in grids, TORQUE deploys a heterogeneous grid check pointing service called Torque GCP. This service is designed to support different checkpointing protocols and address the fundamental checkpointing packages in a obvious manner through a uniform interface. Even with the general design of this service, until recently it only implemented the coordinated checkpointing protocol [9, 10]. But with computational grid environment where data and computing resources of one application can be dispersed across hundreds or thousands of grid nodes, coordinated checkpointing can be expensive in terms of scalability [10]. Also, the entire grid application needs to be replicated in the event of a single node failure. Thus, in additional more number of checkpoints overhead and consistency of checkpoints is added to the replication and storage infrastructure [20].

In this paper, we present the design and implementation of a solution for Optimal check pointing within Torque GCP. Therefore, we describe all the necessary steps towards this goal. This involves consistent dependency tracking among processes, transparent selection of the underlying checkpoint interval, monitoring the process failures, computing a consistent global state of the information and retrieval of an application. Our solution is not bound to a specific checkpointing approach and can be obviously used on any existing checkpointing package.

The result of this work is an extended Torque GCP service with the support of two checkpointing protocols, which is able to checkpoint an application distributed across different checkpointing packages. This is the first work dealing with integration aspects of Optimal checkpointing in grids [7]. We have confidence in our results can help future works to identify the issues related to implementing consistent checkpointing protocols in large heterogeneous environments such as computational grids or cloud computing etc.

The next sections of paper are organized as follows. Section 3 outlines the integration of Optimal check pointing within Torque GCP. Section 4 describes the results from the performance evaluation of our prototype. Finally, Section 5 reports the paper with conclusions and future work.

### Implementation of Optimal Check pointing Within Torque GCP

Check pointing is the commonly used technique for fault tolerance and improving system availability and reliability. It stores the status of current application state, and then it can be used for restarting the execution in case of failure avoiding the job to start from beginning.

The main drawback of unconditional classical checkpointing approach called periodic checkpointing, which performs checkpoint operation periodically whether the job is executed on a stable resource. This approach generates lot of checkpoints, sometimes the checkpoint operation time may exceed the normal processing times. According to the adaptive checkpointing scheme, it uses the information about the remaining job execution time, time left before the deadline and the expected remaining needed, to decide whether checkpointing is to be done or not [13, 14]. It also gives information on when the next checkpoint request need to be given. There are two methods in adaptive checkpointing scheme Last Failure Dependent Check pointing (Last Failure CP) and Mean Failure Dependent Check pointing (Mean Failure CP). In Last Failure CP algorithm it omits unnecessary checkpoint placement with reference to the total execution time and failure frequency of the resource. This algorithm keeps a time stamp LFr that gives the time when the last failure had occurred. Initially checkpointing request will be given at time interval I and then request will be executed by Grid Scheduler by comparing whether  $t_c - LF \leq E_r^j$ , where  $t_c$  is the current time and  $E_r^j$  is the execution time of job j on resource, if the condition is true then checkpointing is allowed otherwise checkpointing is omitted [1]. In fault tolerant checkpoint scheduling system (FTCS) [24], this effectively schedules the jobs in the presence of failures [3]. This approach depends on mean failure time and the failure amount of resources combined with resource response time to make checkpoint scheduling decisions. In this approach, if any one of mean failure rate or resource response time or response time of resource fails then this approach fails to get better checkpointing scheduling decision [15, 16].

Now our approach integrates the adaptive checkpointing schemes [1, 3 and 24] to get better consistent checkpointing approach. To make this the last failure rate and mean failure rate are considered from the resources, and then finds the appropriate checkpoint interval for each node. Let resource failure rate RF [24], last failure rate LF, mean failure rate MF, and fix k and j values derived from the grid workload logs [4], then find probability density function(pdf) of an exponential distribution as.

$$f(k; j) = \begin{cases} kje^{-kx}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad \dots (1)$$

Now, differentiate the probability density functions with respect to last failure rate and mean failure rate combine to get predictive checkpoints with optimum checkpoint intervals.

$$LFR = \frac{d(f(k,j))}{dk} = 0 \quad \dots (2)$$

$$MFR = \frac{d(f(k,j))}{dj} = 0 \quad \dots (3)$$

Then we have found consistent checkpoint interval using the following equation.

$$CFR = \frac{(RF)}{(LF + MF)} \quad \dots (4)$$

By using the above formulae [24], we have found consistent checkpoint interval for job failures in computational grid. In such a case, the Torque GCP service computes consistent global states based on the last failure and mean failure with integration of overall resource failure to integrate each job unit and restores the job units afterwards. The consistent checkpointing scheme identifies the job units belonging to the job based on the workload log meta information. Then, the last and mean failure information is loaded and the consistent checkpoint of the application is generated. We use the Torque GCP mechanism to stores a restart replication context file and restart them within this replication before starting the job checkpoint execution.

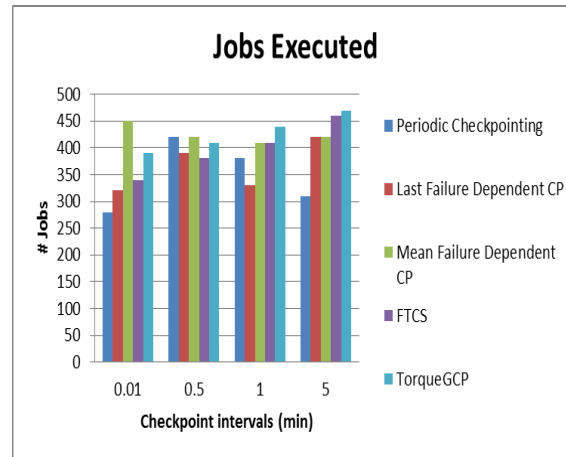
## Performance Evaluation of Torque GCP

According to the presumed settings we have inspected different applications with adaptive method such as periodic, Last Failure CP, Mean Failure CP, and Fault tolerant checkpoint scheduling. Here we have experimented mathematical, computers and mechanical engineering applications in 50 nodes with varying number of processes (i.e., threads or subtasks using forks) and checkpoints.

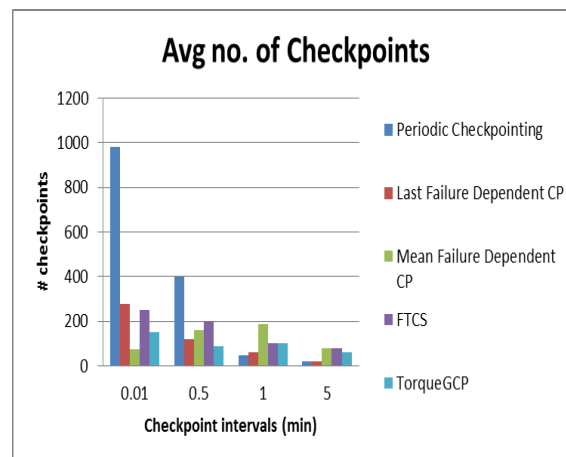
**Table 1:** Check pointing approaches comparisons

Checkpointing Approaches	Jobs Submitted	No. of Jobs Executed in different intervals				Avg no. of Checkpoints in different intervals				Avg. Job Execution time (min) in different intervals				Avg CKPT Space (kb.) in different intervals			
		0.01	0.5	1	5	0.01	0.5	1	5	0.01	0.5	1	5	0.01	0.5	1	5
Periodic Checkpointing	500	280	420	380	310	980	400	50	20	9	7.8	8	8.2	122.5	50	6.25	2.5
Last Failure Dependent CP	500	320	390	330	420	280	120	60	20	7.8	7.6	7.5	7.8	35	15	7.5	2.5
Mean Failure Dependent CP	500	450	420	410	420	75	160	190	80	6.5	7	7	7.2	9.3	21	24	9.8
FTCS	500	340	380	410	460	250	200	100	80	7.4	7.4	7.2	7.3	32	25	12.4	9.6
Torque GCP	500	390	410	440	470	150	90	100	60	7.3	7.3	7.2	7.1	18	12	11	7.8

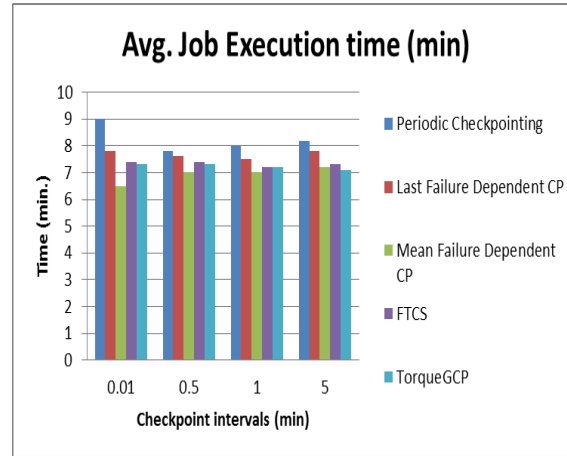
We have evaluated the Torque GCP check pointing protocol using the different approaches such as shown in the Table I with some application within a heterogeneous environment consisting of single PC-nodes and Linux cluster. Therefore, we have measured the time, it takes to checkpoint this application when it opens up to 50 communication channels and sends jobs takes the checkpoints at intervals of 0.01,0.5,1 and 5 minutes. The result from this evaluation is that it takes approximately 4.25 hours to synchronize the processes. However, the results show that the adaptive checkpointing approaches have inconsistent over number of jobs executed, average number of checkpoints generated, the average time execution of application and the average storage space of a checkpoint context file etc as shown in the Figure 2,3,4and 5. Our approach Torque GCP got consistency over all checkpoint duration.



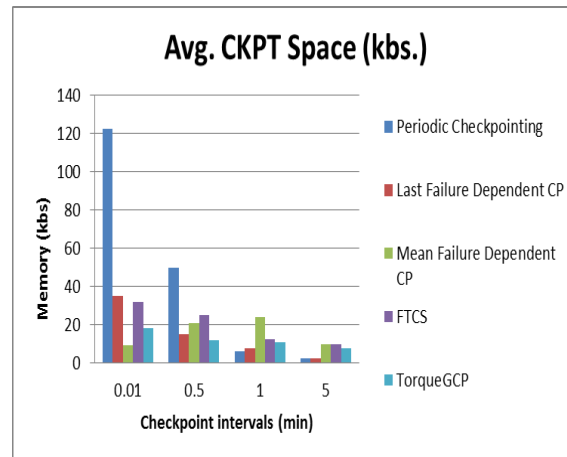
**Figure 2:** Jobs successfully executed



**Figure 3:** Average number of Checkpoints



**Figure 4:** Average Job Execution time (min)



**Figure 5:** Average CKPT Space (kb.)

## Conclusions

In this paper, a Torque GCP check pointing system for computational grids is proposed and evaluated. The performance evaluation of consistent check pointing approach Torque GCP got better consistency when compared with the periodic CP, Last FCP, Mean FCP and FTCS schemes. The metrics used for evaluation are processing time, throughput and checkpoint storage space. The experimental results show that Torque GCP have consistency over adaptive approaches. It is observed that the throughput and over all processing time for the proposed system got better results over Mean FCP. Thus, it is concluded that the proposed Torque GCP provides performance consistency over the adaptive approaches.

## References

- [1] Maria Chtepen, Filip H.A. Claeys, Bart Dhoedt, "Adaptive Task Checkpointing and Replication: Toward Efficient Fault-Tolerant Grids", *IEEE Transactions on Parallel and Distributed Systems*, vol.20, no.2, february 2009.
- [2] Brent Rood, John Paul Walters, Vipin Chaudhary, and Michael J. Lewis, "Failure Prediction and Scalable Checkpointing for Reliable Large-Scale Grid Computing", *The 16<sup>th</sup> IEEE transactions*, June 2007.
- [3] M. Chtepen, F. Claeys, B. Dhoedt, F. De Turck, P. Vanrolleghem, and P. Demeester, "Dynamic Scheduling of Computationally Intensive Applications on Unreliable Infrastructures," *Proc. Second European Modeling and Simulation Symp. (EMSS '06)*, Oct. 2006.
- [4] D. Feitelson, *Parallel Workloads Archive*, <http://www.cs.huji.ac.il/labs/parallel/workload/>, 2008.
- [5] Rajkumar Buyya and Srikumar venugopal, "A Gentle Introduction to Grid Computing and Technologies", <http://www.buyya.com/papers/gridintro-CSI2005.pdf>.
- [6] Gokuldev S, Valarmathi M, "A New Checkpoint Approach for Fault Checkpoint Approach for Fault Checkpoint Approach for Fault Tolerance in Computing" *IJCNS Volume 2, Issue 3*, June 2013
- [7] J. Young, "A First Order Approximation to the Optimum Checkpoint Interval", *Comm. ACM*, vol. 17, no.9, pp.530-531, Sept 1974.
- [8] T. Ozaki, T. Dohi, H. Okamura and N. Kaio, "Min-max checkpoint Placement under Incomplete Failure Information," *Proc. Int'l Conf. Dependable Systems and Networks (DSN '04)*, June-July 2004.
- [9] Suri P.K., Singh, M. "An Efficient Decentralized Load Balancing Algorithm for Grid," *IEEE 2nd International, Advance Computing Conference*, 201, pp. 10-130. Feb 2010.
- [10] Oliner, A.; Rudolph, L.; Sahoo, R.; "Cooperative checkpointing theory", *Parallel and Distributed Processing Symposium*, 2006. IPDPS 2006.
- [11] Menno Dobber, Rob van der Mei, and Ger Koole, "Dynamic Load Balancing and Job Replication in a Global-Scale Grid Environment: A Comparison" *IEEE transaction on parallel and distributed systems*, vol. 20, no. 2, February 2009.
- [12] Chang, R.-S., National Dong Hwa University, Hualien, Chang, H.-P., Wang, and Y. T., "A dynamic weighted data replication strategy in data grids", In: *Computer Systems and Applications, AICCSA 2008, IEEE/ACS 2008*.
- [13] Saurabh Agarwal, Rahul Garg, Joes Moreira, "Adaptive Incremental checkpointing for Massively Parallel Systems", *ACM*, 2004.
- [14] T. Ozaki, T. Dohi, H. Okamura and N. Kaio, "Min-Max Checkpoint Placement under In-complete Failure Information", *Proc. Int'l Conf. Dependable Systems and Networks (DSN '04)*, June-July 2004.

- [15] Dilli babu, S., Ramesh Babu, Ch., Subba Rao, Ch.D.V., “An efficient fault tolerance technique using checkpointing and replication in grids using data logs.”, In: publications of problems and application in engineering research—ijpaper.com, vol 04. Special issue 01, 2013.
- [16] Ch. Ramesh Babu, Ch. D. V. Subba Rao, “Automatic checkpoint based fault tolerance in computational grid”, IEEE Conference COMMANTEL-2014.
- [17] G. Molto, V. Hernández, J.M. Alonso, “Automatic replication of WSRF-based Grid services via operation providers” J. of FGCS-2009.
- [18] Greg Bronevetsky, Daniel Marques, Keshav Pingali, Paul Stodghill, "C3: A System for Automating Application-level Checkpointing of MPI Programs", Department of Computer Science, Cornell University.
- [19] Gengbin Zheng, Chao Huang and Laxmikant V. Kal'e, "Performance Evaluation of Automatic Checkpointbased Fault Tolerance for AMPI and Charm++", Department of Computer Science, University of Illinois at Urban a Champaign.
- [20] Najme Mansouri, "An Effective Weighted Data Replication Strategy for Data Grid", Australian Journal of Basic and Applied Sciences- 2012.
- [21] [Http://www.open-mpi.org/](http://www.open-mpi.org/)
- [22] [Http://crd.lbl.gov/groups-depts/ftg/projects/current-projects/BLCR](http://crd.lbl.gov/groups-depts/ftg/projects/current-projects/BLCR)
- [23] [Http://www.adaptivecomputing.com/](http://www.adaptivecomputing.com/)
- [24] Mohammed Amoon, “A job checkpointing system for computational grids”, CEJES Springer, 2013.



# CLOUD SERVICE NEGOTIATION IN INTERNET OF THINGS ENVIRONMENT: A MIXED APPROACH

Satya Veni Matta<sup>1</sup>, Dr.Y.Venkateswarulu<sup>2</sup>, Dr U Ravi Babu<sup>3</sup>

<sup>1</sup>Mtech Student, CSE, Giet Engineering College, Rajahmundry, A,P, India  
Professor and HOD, Dept of CSE, Giet Engineering College, Rajahmundry, A,P, India  
Professor, Dept of CSE, MREC(A)

## ABSTRACT

Internet of Things (IoT) is allowed to communicate the connected objects via the Internet. IoT can benefit from the boundless capabilities and resources of cloud computing. Due to the cloud market becomes more competitive and open, Quality of Service (QoS) will be more important. However, cloud providers and cloud consumers have different, and sometimes opposite, preferences. To balance effectiveness and accomplishment rate, the present paper propose a mixed approach for cloud service negotiation, which is based on the “game of chicken”. To evaluate the effectiveness of this approach, experiments are conducted on extensive simulations. Results show that a mixed negotiation approach can achieve a higher utility than a concession approach.

## INTRODUCTION

Internet of Things (IoT) allows objects like computers, sensors, mobile phones, etc. to communicate via the Internet. IoT is expected to be a worldwide network of interconnected objects [1]. It is characterized by limited capacities and constrained devices, and its development depends on new technologies including cloud computing. IoT can benefit from the unlimited

capabilities and resources of cloud computing. Also, when coupled with IoT, cloud computing can in turn deal with real world things in a more distributed and dynamic manner.

Cloud services are easier to access and use, cost-efficient, and environmentally sustainable. As they eliminate large upfront expenses in hardware and expensive labor costs for maintenance, cloud services are beneficial to small- and medium-sized enterprises. Cloud services are Internet-based IT services. Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS) are three representative examples [2], [3]. However, cloud providers and cloud consumers have different and sometimes opposite preferences. For example, a cloud consumer usually prefers a high reliability, whereas a cloud provider may only guarantee a less than maximum reliability in order to reduce costs and maximize profits. If such a conflict occurs, a Service Level Agreement (SLA) cannot be reached without negotiation.

To create a proposal, a negotiation agent can adopt two strategies i.e. concession and tradeoff. However, if information is incomplete, it may cause miscalculations, and so underperform the concession one in

terms of success rate. To balance utility and success rate, in this paper propose a mixed approach for cloud service negotiation, which is based on the “game of chicken.” In other words, if a party’s counterpart uses a concession strategy, it is best to adopt a tradeoff one; if a party’s counterpart uses a tradeoff strategy, it is best to adopt a concession one; and if a party is uncertain about the strategy of its counterpart, it is best to mix concession and tradeoff. In fact, those are the three Nash equilibrium of a negotiation game with two pure strategies. A mixed negotiation approach based on the “game of chicken,” which can balance utility and success rate. In particular, if a party has no knowledge of which strategy that its counterpart will play, it is best to mix concession and tradeoff in negotiation to evaluate the effectiveness of the mixed negotiation approach. We first test the impact of different parameters on negotiation results and then conduct Monte Carlo simulations. Results show that the mixed negotiation approach can achieve a higher utility than a concession approach, while incurring fewer failures than a tradeoff approach, which demonstrates its effectiveness.

The rest of the paper is structured as follows. Section II describes multi-attribute bilateral negotiations where concession and tradeoff strategies are detailed and proposes a mixed approach for cloud service negotiation, which is based on the “game of chicken.” Section III results and discussions analyzes results. Conclusion are given in section IV

## Section II: Proposed Approach

To balance usefulness and accomplishment rate, the present propose a mixed negotiation approach for cloud service negotiation, which is based on the “game of chicken.”

## Two-Player Negotiation Game

In a negotiation game, a selfish agent’s utility remains the same with a tradeoff strategy, whereas its utility is decreased with a concession one. As the agent attempts to maximize its utility, it seems that it should stick to the tradeoff strategy instead of the concession one. If the agent and its counterpart both adopt the tradeoff strategy, unfortunately, it is very likely that a failure happens, whereupon both receive the worst utility. It thus becomes a dilemma. This indicates that how to play concession and tradeoff strategies is of utmost importance. However, to the best of our knowledge, no previous work deals with this problem. In fact, we first identify the problem and model it with the “game of chicken,” which goes as follows [2]. Two boys, say Alan and Bob, want to prove their manhood. They drive toward each other at breakneck speed. The one who swerves loses face and becomes a “chicken,” whereas the other who stays, of course, proves his manhood and becomes a hero to his friends. If both swerve, nothing is proved. If neither swerves, they crash into each other with potentially disastrous results.

A possible payoff matrix of the game of chicken is shown in Table II, where a number only has a relative significance, namely, the greater the number, the higher the payoff. A Nash equilibrium is “a situation in which each player in a game chooses the strategy that yields the highest payoff, given the strategies chosen by the other players” [4]. The “game of chicken” has two pure strategy Nash equilibrium. One is for Alan to swerve and for Bob to stay, whereas the other is for Alan to stay and for Bob to swerve. In fact, if Alan swerves, Bob is better off staying (payoff 1) than swerving (payoff 0). Conversely, if Alan stays, Bob is better off

swerving (payoff  $-1$ ) than staying payoff  $-10$ ). So, those are the two pure strategies Nash equilibrium. Below, we give a formal description for Nash equilibrium [5].

**Definition 1 (Nash Equilibrium):** A Nash equilibrium is a strategy profile

$s^* = (s^*_1, s^*_{-1})$  such that each player,  $i (i=1, 2, 3, \dots, n)$ , has no incentive to deviate from its current strategy,  $s^*_i$ , given the strategy profile,  $s^*_{-i}$ , of the other players.

A general payoff matrix of a two-player negotiation game with concession and tradeoff strategies is shown in Table III, where  $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2$  belongs to  $\mathbb{R}$  and  $a_1 \geq a_2 > b_1 \geq b_2 > c_1 \geq c_2 > d_1 \geq d_2$ . It should be noted that, here, the game is asymmetric, in that the two players are distinguishable from each other, and is more applicable, in that it generalizes the “game of chicken.”

**Proposed Algorithmic Description:**

Algorithm 1 implements a mixed negotiation approach. It works as follows. First, in line 1, agent  $i$  sends  $V$ —its initial proposal—to agent  $j$ , and waits for a response. If  $j$  does not accept  $V$  and  $j$ 's counter proposal is not acceptable to  $i$ , then  $i$  adopts a mixed approach in the while loop of lines 2–15 to create a new proposal; otherwise, true is returned in line 16. Here, a party's acceptance criterion is that the utility received from a proposal is no less than that of its reserved proposal, and the values received from the proposal do not go beyond its reserved values.

Next, in line 4, uses function random to generate a random number between 0 and 1 for variable  $r$ . In lines 5–10, if  $r < 1-p$ , which implies that a concession strategy is triggered,  $i$  uses function concession to create a new proposal, where  $P_r \{ r < 1-p \} = 1-p$ . In

line 6, is increased by one, each time the condition is triggered. It should be mentioned that concession is a function that implements a concession strategy of a multi-attribute negotiation. Refer to [30] for its algorithmic description. If  $r \geq 1-p$ , which implies that a tradeoff strategy is triggered, uses function tradeoff to create a new proposal, where  $P_r \{ r \geq 1-p \} = p$ . In line 9,  $k_2$  is increased by 1, each time the condition is triggered. The Mixed approach algorithm is shown below

**Algorithm:** Mixed Approach ( $V, W, F, \lambda_1, \lambda_2, p$ )

**Input:** array  $V$  with raw values of  $n$  attributes

array  $W$  with weights of  $n$  attributes

array  $F$  with flags of  $n$  attributes

A flag indicates whether an attribute is higher-is-better

parameters  $\lambda_1$  and  $\lambda_2$  ( $0 < \lambda_1, \lambda_2 < 1$ ) which indicate the rate of concession and the rate of tradeoff at a time, respectively;

parameter  $p$  ( $0 < p < 1$ ) which indicates the probability of playing tradeoff, or  $p$ -for short

**Output:** true if succeed and false otherwise

Step 1: agent sends  $V$  to agent and waits for a response

Step 2: while agent  $j$  does not accept  $V$  and  $j$ 's counter proposal is not

Step 3: acceptable to agent  $i$

Step 4:  $r = \text{random}(0,1)$

Step 5: if  $r$  is less than  $1 - p$  then

Step 6:  $k_1 = k_1 + 1$

Step 7:  $V = \text{concession}(V, W, F, k_1, \lambda_1)$

Step 8: otherwise

Step 9:  $k_2 = k_2 + 1$

Step 10:  $V = \text{tradeoff}(V, W, F, k_2, \lambda_2)$

Step 11:  $k = k_2 + k_1$

Step 12: if  $V$  is out of bounds then

Step 13: return FALSE

Step 14: otherwise

Step 15: agent  $i$  sends  $V$  to agent  $j$  and waits for a response

Step 16: return TRUE

It should also be mentioned that tradeoff is a function that implements a tradeoff strategy of a multi-attribute negotiation. Refer to [6] for its algorithmic description. In line 11,  $k$  counts the total number of negotiation rounds.

Finally, in lines 12–15, if  $V$  is out of bounds, false is returned; otherwise, agent sends, whose values are adjusted, to agent  $j$  as a new proposal, and waits for a response again. The process repeats until either success or failure occurs. In this process, utility of the current proposal can remain the same (moves along its current indifference curve) or be reduced (moves down to its next indifference curve). It can be proved that Algorithm 1 converges and terminates in a

finite number of rounds. Refer to [6] for the proof.

It should be noted that the mixed approach we adopt in negotiation exhibits a certain degree of intelligence. Just as Turing [7] pointed out in Computing Machinery and Intelligence, “Intelligent behavior presumably consists in a departure from the completely disciplined behavior involved in computation, but a rather slight one, which does not give rise to random behavior, or to pointless repetitive loops.”

## EVALUATION AND ANALYSIS

We conduct extensive simulations to evaluate the mixed approach for cloud service negotiation. First, we describe the experimental setup. Next, we describe the parameter setup. Finally, we report and analyze simulation results.

### Experimental Setup

All simulations are conducted on a Lenovo Think Centre desktop with a 2.80-GHz Intel Pentium Dual-Core CPU and a 2.96-GB RAM, running Microsoft Windows 7 Professional Operating System. The simulations are implemented with Java under NetBeans IDE 7.2.1 with JDK 7u13. An alternating-offers protocol is adopted as the negotiation protocol, and a mixed negotiation strategy is compared with concession and tradeoff strategies. The negotiation process works as follows. First, without loss of generality, a SP sends its initial proposal to a SC. Next, if the proposal is accepted by the SC, negotiation ends successfully; otherwise, the SC uses either mixed, tradeoff, or concession negotiation approach to create a counter proposal. After that, the SC sends back the counter proposal to the SP, and the negotiation process repeats. The

process ends once a proposal or a counter proposal is accepted, and it fails if no proposal is acceptable to both parties. Java multithreading, which allows multiple tasks in a program to be executed concurrently, is the ideal technique to simulate the negotiation process. A thread is the flow of execution, from beginning to end, of a task. We model the behaviors of the SP and the SC as two threads. In particular, we use thread synchronization techniques to coordinate their behaviors, and a shared object to exchange their proposals and counter proposals. In our software prototype, there is a QoS matrix to be negotiated, where the SP and the SC can specify their QoS requirements, i.e., their preferred values, reserved values, and weights over quality dimensions AVAL, REL, RESP, SECY, and ELAS. In a real negotiation, those values would be kept private.

**TABLE III**

**TWO-PLAYER NEGOTIATION GAME**

		Player 2	
		Concession	Tradeoff
Player 1	Concession	$b_1, b_2$	$c_1, a_1$
	Tradeoff	$a_2, c_2$	$d_1, d_2$

In our simulations, we attempt to resolve QoS conflicts in the motivating example. In other words, we use Table I as the QoS matrix to be negotiated. Also, there is a parameter variance  $\delta (0 < \delta < 1)$  that can be used to generate a random number within a certain interval of a value, such that the impact of a specific data set on negotiation results can be reduced, if not completely removed.

As to negotiation strategies, the SP and the SC can choose a concession, a tradeoff, or a mixed approach. So, in total, there exist nine combinations, i.e., CC, CT, CM, TC, TT, TM, MC, MT, MM, respectively, where

stands for a concession approach, a tradeoff one, and a mixed one. Also, there are parameters the rate of concession  $\ll$ , the rate of tradeoff  $\ll$ , and the probability of playing tradeoff. As to negotiation results, success or failure can happen. In the case that success occurs, QoS conflicts are resolved, and the new values agreed to by the SP and the SC are output. In the case that failure happens, relevant information about the failure is output. Also, there is parameter tolerance  $\ll$ , within which a solution whose values go beyond a party's reserved values, but its utility is no less than its reserved utility is still acceptable. So, a rigid cutoff value is avoided, and the chance of success increased. In fact, it is the acceptance criterion that we adopt in our simulations. To fully understand the impact of different parameters on negotiation results, we conduct a series of simulations where the SP and the SC both adopt a concession, a tradeoff, and a mixed approach, respectively. Refer to [30] for those parts. We keep, without further details, in our simulations. Unless specified otherwise, we keep the gap in preferred values as 0.11, 0.20, 0.60, 0.05, and 0.06, and the gap in reserved values as 0.19, 0.10, 0.20, 0.15, and 0.14 for AVAL, REL, RESP, SECY, and ELAS, respectively, in our simulations. Also, we set the maximum negotiation round, as most negotiations can finish in no more than 20 rounds. In our simulations, agents' preferences are kept private so we cannot apply Theorem 2 here. However, it does give us some hint on how to choose. As a general rule, if competition is high, a small value is preferred; otherwise, a large value is preferred.

**CONCLUSIONS:**

A tradeoff approach can outperform a concession one in terms of utility, but may incur more failures if information is

incomplete. To balance utility and success rate, we propose a mixed approach for cloud service negotiation, which is based on the “game of chicken.” In particular, if a party is uncertain about the strategy of its counterpart, it is best to mix concession and tradeoff strategies. In fact, it is a mixed strategy Nash equilibrium of a negotiation game with two pure strategies, which provides the theoretical basis for our approach. To demonstrate the effectiveness of the mixed approach, we conduct extensive simulations. Results show that when a party has no knowledge of the strategy of its counterpart, a mixed approach outperforms a concession one in terms of utility, and it outperforms a tradeoff one in terms of success rate. It should be noted that the mixed approach works under incomplete information, and so is applicable for real negotiations, where information is generally not complete. In conclusion, when one is uncertain about the strategy of its counterpart, a mixed negotiation approach, which exhibits a certain degree of intelligence, can achieve a higher utility than a concession one, while incurring fewer failures than a tradeoff one. It thus becomes a promising approach for cloud servicenegotiation.

## REFERENCES:

- [1] Q. Li et al., “Applications integration in a hybrid cloud computing environment: Modelling and platform,” *Enterpr. Inf. Syst.*, vol. 7, no. 3, pp. 237–271, 2013
- [2] M. Armbrust et al., “A view of cloud computing,” *Commun. ACM*, vol. 40, no. 4, pp. 50–58, 2010.
- [3] L. M. Vaquero et al., “A break in the clouds: Towards a cloud definition,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 39, no. 1, pp. 50–55, 2009
- [4] D. Besanko and R. R. Braeutigam, *Microeconomics*, 3rd ed. Hoboken, NJ, USA: Wiley, 2008.
- [5] K. Leyton-Brown and Y. Shoham, *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. San Rafael, CA, USA: Morgan & Claypool, 2008.
- [6] X. Zheng, “QoS Representation, Negotiation and Assurance in Cloud Services,” Ph.D dissertation, School of Computing Queen’s Univ., Kingston, ON, Canada, Feb. 2014
- [7] A. M. Turing, “Computing machinery and intelligence,” *Mind*, vol. 59, no. 236, pp. 203–233, 1950.

# **CARTOON-TEXTURE IMAGE DECOMPOSITION USING BLOCK WISE LOW-RANK TEXTURE CHARACTERIZATION**

**Meenakshi Tummala<sup>1</sup>, Dr.Y.Venkateswarulu<sup>2</sup>, Dr. U Ravi Babu<sup>3</sup>**

<sup>1</sup>Mtech Student ,CSE, Giet Engineering College, Rajahmundry, A,P, India

<sup>2</sup>Professor and HOD, Dept of CSE, Giet Engineering College, Rajahmundry, A,P, India  
 Professor, Dept of CSE, MREC(A)

## **ABSTRACT**

We propose an Image decomposition method that can viably deteriorates a picture into its cartoon and composition segments by utilizing a portrayal of surface. The portrayal rests on our perception that the surface segment appreciates a blockwise low-rank nature with conceivable cover and shear, on the grounds that composition, all in all, is universally different however by regional standards decently designed. We set up a cartoon composition disintegration demonstrate as a raised improvement issue, where the synchronous estimation of the toon and surface parts from a given picture or debased perception is executed by minimizing the aggregate variety and BNN. Moreover, the model can deal with different sorts of corruption happening in picture handling, including smear + missing pixels with a few sorts of clamor. By revising the issue through variable part, the supposed rotating heading strategy for multipliers gets to be relevant, bringing about an effective algorithmic answer for the issue. Numerical illustrations outline that the proposed model is exceptionally particular to examples of composition, which improves it deliver results than cutting edge disintegration models.

Index Terms-Cartoon-texturedecompo -

sition, convexoptimization, image restoration, low-rank interpretation, texturecharacterization.

## **INTRODUCTION:**

One of the most important and longstanding problems in image processing is image decomposition, which plays a central role in a wide range of applications such as image restoration, biomedical engineering, astronomical imaging, remote sensing, pattern recognition, and computer vision. In image decomposition, an image is often modeled as the superposition of two meaningful components, namely, the cartoon component and the texture component. The cartoon component is the piecewise-smooth part having the global structural information of the image, and the texture component is the locally-patterned oscillating part. A popular and effective strategy to achieve such decomposition is to formulate it as a convex optimization problem, in which the components are characterized by appropriate convex priors. Indeed, under a variety of scenarios, image decomposition based on the said strategy has been extensively studied [1-10].

Very recently, Schaeffer and Osher proposed to interpret texture using what is called the Low Patch-Rank [11]. Their approach is mainly motivated by Robust PCA [12]–[15] and models the texture

component as an alignment of patches which are almost linearly dependent. For its special capability of capturing patterned structure, their cartoon-texture decomposition model, which we shall refer to as the LPR model, is shown to be superior to other existing ones in terms of texture characterization and at the same time outperforms the non-local TV regularization [16]–[18] in the restoration of well-textured images. However, it is also true that, because of its fully-global (not locally-adaptive) nature, the LPR model is not much suitable for characterizing

texture having various different patterns, which is typical in many images, and thereby sometimes producing undesirable patterned-artifacts

The very nature of BNN in the proposed model not only realizes a reasonable texture characterization by its low-rank interpretation but also overcomes the limitation of the LPR model by its local adaptivity. In addition, the proposed model is designed to accept various degradation scenarios, including blur + missing pixels with noise. Such a scenario was considered in a recent study under a Gaussian noise assumption. Notably the proposed model can also handle several non-Gaussian noise cases in a unified way with the associated convex optimization problem solvable using proximal splitting techniques.

The rest of the paper is organized into 3 sections section 2 describe the proposed method and results are discussed in section 4 and conclusions are given in section 4

## PROPOSED METHOD: IMAGE DECOMPOSITION MODEL

This section is devoted to the proposal of an image decomposition model as a convex optimization problem, where we

characterize the texture component using multiple BNN with different shear angles and present a guarantee of the existence of a minimizer. An optimization scheme based on ADMM for the proposed model is also provided with its convergence property and efficient implementation.

**Formulation:** Consider to the estimation of ideal cartoon and texture components  $C, T \in \mathbb{R}^N$  ( $N = n_v n_h$ ,  $n_v \times n_h$  corresponds to the image size) from an observation  $\mathbf{v} := \mathbf{h}(\mathbf{u}_{\text{org}}) \in \mathbb{R}^M$  ( $M \in \mathbb{N}$ ), where  $\mathbf{u}_{\text{org}} = C + T \in \mathbb{R}^N$  is an original image,  $\mathbf{h} \in \mathbb{R}^{M \times N}$  a linear observation operator representing some deterioration process (e.g., blur and/or missing pixels), and  $\mathbf{N} : \mathbb{R}^M \rightarrow \mathbb{R}^M$  a noise contamination being not necessarily additive. The proposed cartoon-texture decomposition (possibly with degradation) is then formulated as follows.

## (Cartoon-texture decomposition using BNN)

$$\begin{aligned} & \text{Minimize} \\ & \mathbf{c} \in \mathbb{R}^N, \mathbf{t}_k \in \mathbb{Z}, \mathbf{c} + \sum_{k=1}^K \mathbf{t}_k \in \mathcal{D} \quad \mathcal{J}_{\text{obj}}, \\ & \mathcal{J}_{\text{obj}} := \lambda \|\mathbf{c}\|_{\text{TV}} + \sum_{k=1}^K \|\mathbf{t}_k\|_{\text{mBNN}}^{\mathbf{w}, \delta, \theta_k} + \mathcal{F}_{\mathbf{v}}(\Phi(\mathbf{c} + \sum_{k=1}^K \mathbf{t}_k)), \end{aligned}$$

where  $\mathbf{t}_k$  ( $k = 1, \dots, K$ ) are the sub-texture components,  $\mathcal{Z} \subset \mathbb{R}^N$

is the set of all zero-average vectors defined by  $\mathcal{Z} := \{\mathbf{x} \in \mathbb{R}^N \mid \sum_{i=1}^N x_i = 0\}$ ,  $\mathcal{D} \subset \mathbb{R}^N$  a

(normalized) dynamic range constraint given by  $\mathcal{D} := \{\mathbf{x} \in \mathbb{R}^N \mid x_i \in [0, 1] \ (i=1, \dots, N)\}$ ,  $\lambda \in \mathbb{R}_{++}$ , and  $\mathcal{F}_{\mathbf{v}} \in \mathbb{R}^M$

is a certain data fidelity function regarding the observation  $\mathbf{v}$ , the proximity operator of which is available. Note that the texture component  $\mathbf{T}$  is modeled as

$\mathbf{T} := \sum_{k=1}^K \mathbf{t}_k$ , where each  $\mathbf{t}_k$  is characterized by BNN with a different shear angle, i.e., consisting of patterns extending in a specific direction.

- (a) Noiseless case: Since  $\mathbf{v} = \mathbf{u}_{\text{org}}$ , a suitable fidelity is  $F_{\mathbf{v}}(\mathbf{x}) = \iota_{C_{\mathbf{v}}}(\mathbf{x})$ , where  $\iota_{C_{\mathbf{v}}}$  is the indicator function of the closed convex set  $C_{\mathbf{v}} := \{\mathbf{x} \in \mathbb{R}^M \mid \mathbf{x} = \mathbf{v}\}$ . The associated proximity operator is given by  $\text{prox}_{\gamma F_{\mathbf{v}}}(\mathbf{x}) = \text{PC}_{\mathbf{v}}(\mathbf{x}) = \mathbf{v}$ , where PC stands for the metric projection onto a nonempty closed convex set CI
- (b) Gaussian noise case: A standard choice is to use the additive  $\ell_2$  fidelity:  $F_{\mathbf{v}} := \mu/2 \|\cdot - \mathbf{v}\|_2^2$  ( $\mu \in \mathbb{R}_{++}$ , and  $\|\cdot\|_2$  denotes the  $\ell_2$  norm), with the associated proximity operator given by  $\text{prox}_{\gamma F_{\mathbf{v}}}(\mathbf{x}) = \mu\gamma\mathbf{v} + \mathbf{x}/(\mu\gamma + 1)$ . An alternative choice is the constrained type fidelity:  $F_{\mathbf{v}}(\mathbf{x}) := \iota_{B_{\mathbf{v},\varepsilon}^2}(\mathbf{x})$ , where  $B_{\mathbf{v},\varepsilon}^2 := \{\mathbf{x} \in \mathbb{R}^M \mid \|\mathbf{x} - \mathbf{v}\|_2 \leq \varepsilon\}$  ( $\varepsilon \in \mathbb{R}_{++}$ ). The associated proximity operator is the metric projection onto  $B_{\mathbf{v},\varepsilon}^2$ , i.e.,  $\text{prox}_{\gamma F_{\mathbf{v}}}(\mathbf{x}) = P_{B_{\mathbf{v},\varepsilon}^2}(\mathbf{x}) = \mathbf{x}$ , if  $\mathbf{x} \in B_{\mathbf{v},\varepsilon}^2$ ;  $\mathbf{v} + \varepsilon(\mathbf{x} - \mathbf{v})/\|\mathbf{x} - \mathbf{v}\|_2$ , otherwise.
- (c) Impulsive noise case: It is well known that using the  $\ell_1$  norm as fidelity measure is robust to impulsive noise contamination. The additive  $\ell_1$  fidelity is given by  $F_{\mathbf{v}} := \mu \|\cdot - \mathbf{v}\|_1$  ( $\mu \in \mathbb{R}_{++}$ , and  $\|\cdot\|_1$  denotes the  $\ell_1$  norm), the proximity operator of which can be computed by the soft-thresholding: for  $i = 1, \dots, M$ , by

$$[\text{prox}_{\gamma F_{\mathbf{v}}}(\mathbf{x})]_i = v_i + \text{sgn}(x_i - v_i) \max\{|x_i - v_i| - \gamma\mu, 0\}$$

where  $\text{sgn}$  denotes the signum function. As in (a), the constrained type alternative is  $F_{\mathbf{v}}(\mathbf{x}) := \iota_{B_{\mathbf{v},\varepsilon}^1}(\mathbf{x})$ , where  $B_{\mathbf{v},\varepsilon}^1 := \{\mathbf{x} \in \mathbb{R}^M \mid \|\mathbf{x} - \mathbf{v}\|_1 \leq \varepsilon\}$ . The associated proximity operator can be computed efficiently by a fast  $\ell_1$  ball projection algorithm [19].

- (d) Poisson noise case: It has been shown that under Poisson noise contamination, the so-called generalized KulbackLeibler divergence is suitable for  $F_{\mathbf{v}}$ . The definition and the computation of the associated proximity operator can be found in [20].

## RESULTS AND DISCUSSIONS

We examine the effectiveness of the proposed cartoon texture decomposition model in several scenarios. All experiments were performed using MATLAB (R2013a), on a Windows 7 (64bit) desktop computer with an Intel Core i7 2.8 GHz processor and 8.0 GB of RAM. The dynamic range of test images ( $256 \times 256$ ) are normalized to  $[0, 1]$ . In (5), we use  $\gamma = 0.1$  and  $\mathbf{s}(0) = \mathbf{d}(0) = \mathbf{0}$ , and for the stopping criterion, adopt  $\|\mathbf{r}(n+1) - \mathbf{r}(n)\|_2 < 0.1$ . The tolerance value and max iteration number of the preconditioned conjugate gradient method are set to  $1.0 \times 10^{-6}$  and 20, respectively.

### Decomposition

Here we consider pure decomposition, i.e., no degradation, to analyze the characteristics of the proposed model. In this case, the observation  $\mathbf{v}$  is the original image  $\mathbf{u}_{\text{org}}$ , so that a suitable data-fidelity function is  $\iota_{C_{\mathbf{v}}}$ . We first use the synthesized image Sakura, shown in the top left of Fig. 1, to give a clear insight into the behavior of the proposed decomposition, i.e., the sub-texture separation capability. For reference, we also apply the nuclear norm model, i.e., simply using the nuclear norm (i.e., BNN with  $k = 1$ ,  $\mathbf{0}\mathbf{1} = (*, 0)$ , and  $m\mathbf{v} = \delta\mathbf{v} = n\mathbf{v}$ ,  $m\mathbf{h} = \delta\mathbf{h} = n\mathbf{h}$ ) as a prior for texture component, and the LPR model to the image, where we optimize them by using the proposed algorithm modified for these models. We choose the blocksize  $\mathbf{m}$  and the shift step number  $\delta$  of

BNN as (16, 16) and (8, 8), and the number of sub-texture components is set to  $K = 3$ . In each model, the parameter controlling the balance between TV and texture prior (the nuclear norm, LPR or BNN) is selected in such a way that the decomposed cartoon and texture components are closest to the oracle ones in the sense of the Euclidean distance. Note that such parameter optimization is only available when we use synthesized images.

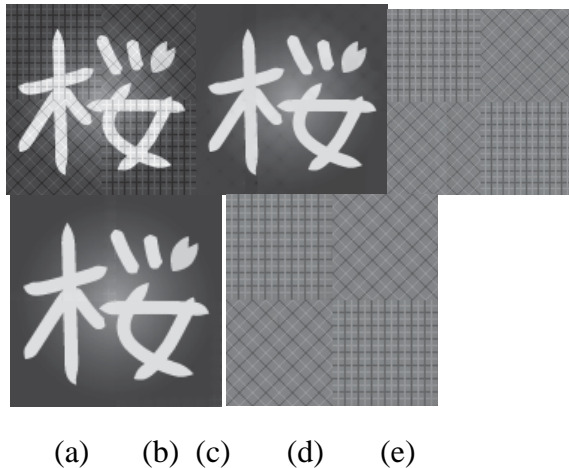


Fig1: Pure decomposition results using the synthesized image a) original Image b) cartoon Image c) texture by cartoon d) cartoon LPR e) texture by LPR

As can be seen in Fig. 1, all the models produce reasonable cartoon-texture decompositions. This is because the Sakura image consists of both simple cartoon and texture, and thus can be easily decomposed. We observe that patterns extending in different directions are extracted as sub-texture components in an almost completely-separated manner by using the proposed model with the exact shear angle of  $45^\circ$ , implying that the proposed texture characterization based on BNN with different shear angles appropriately works. We also show decomposition

results of using shear angles which do not exactly match the directions in the third and fourth

rows of Fig. 1. As expected, the larger difference of shear angle from the exact one leads to the more different sub-texture separation. One may think that, in the case of simple texture such as in the Sakura image, similar sub-texture components can be generated by applying some directional analysis transform (e.g., directional wavelet) to the texture component obtained by the nuclear or LPR model. Actually, this kind of things may be possible by using a combination of some suitable methods, but, compared with such sequential methods, the use of the proposed model has the following advantages: i) To begin with, extracting a reasonable texture component from a given image is difficult and important, and indeed, the succeeding experimental results demonstrate that the proposed model is more suitable to this task than existing models; ii) The proposed model provides cartoon-texture decomposition and sub-texture separation in a unified manner, i.e., by just solving one convex optimization problem, so that we can guarantee the optimality in the sense of the model. Next we compare the proposed model with the LPR one via non-synthesized images Barbara and House [the top left of Fig. 2(a) and (b)]. Note that, in this case, the oracle decomposition of each image is unavailable, so that we cannot optimize the parameters of the models as in the case of Sakura. To conduct a fair comparison, we produce decomposition results by the use of the LPR model with different

parametersettings, in order to illustrate the superiority of the proposedmodel in terms of extracting patterned texture.The results are depicted in Fig. 3 (from top to bottom, cartoon component, texture component, and their close-ups).In Fig. 3(b), we see that the texture component extracted bythe proposed model only contains patterned fabric and nocontours, implying that our model is very selective to patterns.Such decomposition cannot be achieved by the LPR modelas shown in Fig. 3(a), where if we set parameters to extractsufficient texture then the resulting texture component containscontours (left), and if set to not contain contours then patternedtexture remains in the cartoon component (right). In the caseof House, our texture component [Fig. 3(d)] exhibits patterned

Structure like roof tiles and windows but there are very fewnon-patterned objects, which also indicates the selectivity ofthe proposed model. By contrast, the LPR model with anyparameter setting does not lead to a decomposition similar to the proposed one [Fig. 3(c)]. We also present the correspondingsub-texture components obtained by the proposedmodel in Fig. 5, in which we observe that patterns extending in different directions are extracted separately. These results verify a special capability of BNN, namely, that of capturingglobally dissimilar but locally well-patterned nature of texture.

The required iteration number and CPU time for the proposedmodel on Barbara are 66 and 37[sec], and those for the LPRmodel 151 and 32[sec], being almost the same on House.

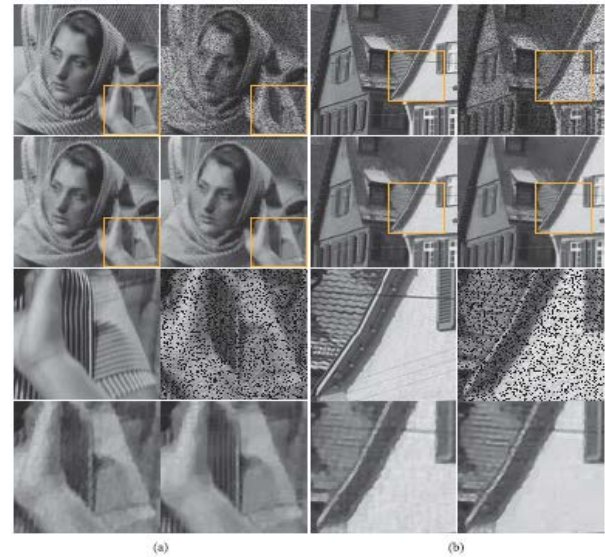


Fig. 2: Comparison of the results of deblurring ( $3 \times 3$  Gaussian blur) with 20% missing pixels for each group, from left to right, top to bottom: original, observation, LPR's result, our result, and their close-ups): The proposed model is locally-adaptive, and hence it can restore fine texture without generating patterned artifacts observed in the LPR's results. (a) Results on Barbara. (b) Results on House.

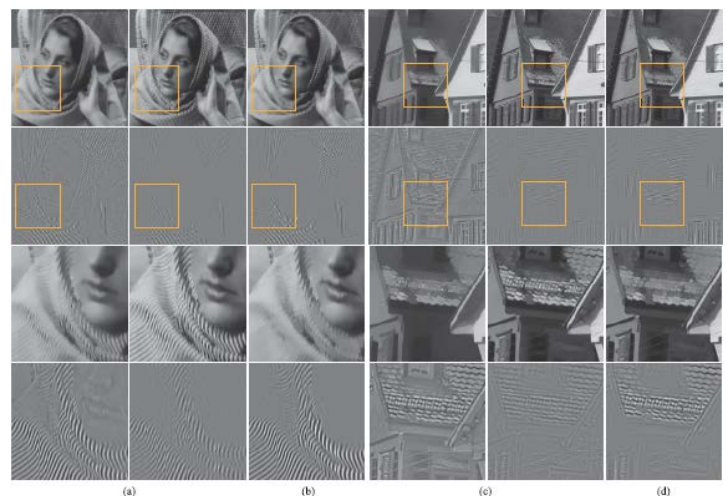


Fig. 4. Comparison of decomposition results (for each column, from top to bottom, cartoon; texture; close-up of cartoon; close-up of texture): The proposed model extracts

sufficient texture without containing edges and contours, which cannot be achieved by the LPR model using any parameter setting. (a) LPR with two different parameter settings. (b) Proposed. (c) LPR with two different parameter settings. (d) Proposed.

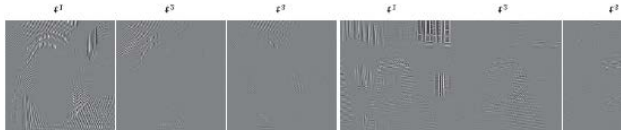


Fig. 5: Sub-texture components of the results in Fig. 3(b) and (d): Patterns extending in different directions are extracted separately. Note that they are magnified by a factor of three for visibility.

We plot in Fig. 6 the improvement of PSNR and SSIM by the proposed model and the LPR model from the best average performance of the G-norm model [12], where we examine the proposed method using different block sizes and shift step numbers with respect to varying  $\lambda$  in (3). From the results in Fig. 7, the best average performance of the proposed model with all the block sizes and shift step numbers exceed those of the G-norm and LPR models, which implies that the proposed model agrees well with a variety of images compared to the other models.

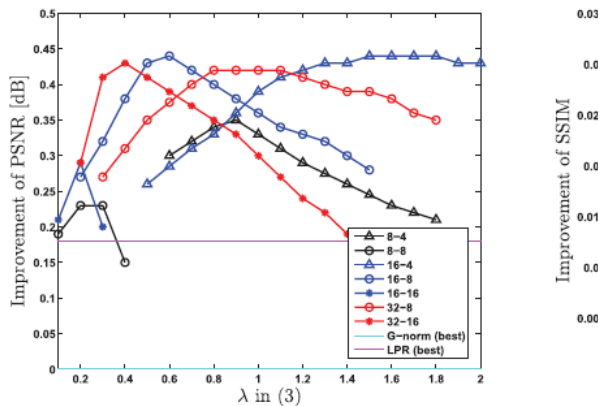


Fig. 6. Improvement of PSNR [dB] and SSIM by the proposed model and the LPR model from the best average performance of the G-norm (averaged over 200 images) with respect to varying  $\lambda$  in (3): The legend of black, blue and red lines indicate the block size and shift step number of the proposed method, e.g., 16 – 8 means  $mv = mh = 16$  and  $\delta v = \delta h = 8$ .

Observations from the proposed method:

- The value of lambda giving the best average performance depends on the block size and shift step number, as indicated in Fig. 6. Specifically, first, the more redundant setting, i.e., the smaller shift step number, requires the larger  $\lambda$  to achieve the best average performance, implying that the value of BNN is proportional to the redundancy rate  $(mvmh)/(\delta v\delta h)$ . Second, based on one redundancy rate, the smaller block size needs a larger  $\lambda$  for the best average performance, which means the increase of the number of blocks leads to the increase of the value of BNN.
- Among all the block sizes, the smallest one ( $mv = mh = 8$ ) results in the worst performance, as observed in Fig. 6. This would be because an excessively small block size cannot capture and reconstruct patterns well.
- The non-overlapped settings (the lines of '8-8' and '16-16' in Fig. 7) perform poorly, which agrees with the intuition that it tends to produce undesirable blocky effects. Basically, a higher overlap level leads to a better performance but with a more expensive computational cost.
- It is natural to suppose that a large K leads to a more accurate

characterization of patterns extending in various directions, and thus results in a better restoration. The experimental results in Table II match this expectation. At the same time, the computational complexity grows because of the increase in the number of the separated components.

## CONCLUSIONS:

We have proposed a cartoon-texture decomposition model with a novel texture prior named the Block Nuclear Norm (BNN). Using BNN, our model interprets the texture component as the combination of blockwise low-rank matrices with possible overlap and shear, which leads to a suitable characterization of globally dissimilar but locally well-patterned nature of texture. The convex optimization problem associated with the proposed model is efficiently solved by ADMM. Numerical examples demonstrate its effectiveness both in pure decomposition and restoration. Future works include an efficient implementation based on the blockwise nature of BNN, extension to color image decomposition with the incorporation of the combination with more involved data-fidelity constraints under various scenarios of noise contamination, and the incorporation into hierarchical convex optimization for selecting a better pair of cartoon and texture components among all the solutions to the proposed model.

## REFERENCES:

1. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, nos. 1–4, pp. 259–268, 1992.
2. Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, vol. 22. Providence, RI, USA: AMS, 2001.
3. L. Vese and S. Osher, "Modeling textures with total variation minimization and oscillating patterns in image processing," *J. Sci. Comput.*, vol. 19, no. 1, pp. 553–572, 2003.
4. S. Osher, A. Solè, and L. Vese, "Image decomposition and restoration using total variation minimization and the  $H^{-1}$  norm," *Multiscale Model. Simul.*, vol. 1, no. 3, pp. 349–370, 2003.
5. M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, "Simultaneous structure and texture image inpainting," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 882–889, Aug. 2003.
6. J.-F. Aujol, G. Aubert, L. Blanc-Fèrau, and A. Chambolle, "Image decomposition into a bounded variation component and an oscillating component," *J. Math. Imag. Vis.*, vol. 22, no. 1, pp. 71–88, 2005.
7. J.-F. Aujol and A. Chambolle, "Dual norms and image decomposition models," *Int. J. Comput. Vis.*, vol. 63, no. 1, pp. 85–104, 2005.
8. J. L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.

9. I. Daubechies and G. Teschke, "Variational image restoration by means of wavelets: Simultaneous decomposition, deblurring, and denoising," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 1, pp. 1–16, Jul. 2005.
10. J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher, "Structure-texture image decomposition—Modeling, algorithms, and parameter selection," *Int. J. Comput. Vis.*, vol. 67, no. 1, pp. 111–136, 2006.
11. H. Schaeffer and S. Osher, "A low patch-rank interpretation of texture," *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 226–262, 2013.
12. Y. Peng, A. Ganesh, J. Wright, and Y. Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," in *Proc. IEEE Conf. CVPR*, Jun. 2010, pp. 7633–770.
13. S. Gandy and I. Yamada, "Convex optimization techniques for the efficient recovery of a sparsely corrupted low-rank matrix," *J. Math. Ind.*, vol. 2, pp. 147–156, Aug. 2010.
14. E. J. Candès, X. Li, . Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–22, 2011.
15. Z. Zhang, X. Liang, A. Ganesh, and Y. Ma, "TILT: Transform invariant low-rank textures," *Int. J. Compt. Vis.*, vol. 99, no. 1, pp. 1–24, 2012.
16. G. Gilboa and S. Osher, "Nonlocal linear image regularization and supervised segmentation," *Multiscale Model. Simul.*, vol. 6, no. 2, pp. 595–630, 2007.
17. G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing," *Multiscale Model. Simul.*, vol. 7, no. 3, pp. 1005–1028, 2009.
18. X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction," *SIAM J. Imag. Sci.*, vol. 3, no. 3, pp. 253–276, 2010.
19. J. Duchi, S. S. Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the  $\ell_1$ -ball for learning in high dimensions," in *Proc. 25th ICML*, 2008, pp. 272–279.
20. P. L. Combettes and J.-C. Pesquet, "A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 564–574, Dec. 2007.

# TEXTURE CLASSIFICATION BASED ON OVERLAPPED TEXTON CO-OCCURRENCE MATRIX (OTCoM) FEATURES

Patnala S. R. Chandra Murthy<sup>1</sup>, U Ravi Babu<sup>2</sup>, R Venkatalakshmi<sup>3</sup>

<sup>1</sup>Assistant Professor, Department of CSE, University College of Engineering and Technology, ANU-Guntur

<sup>2</sup>Professors, Department of CSE, Malla Reddy Engg. College (Autonomous) TS, INODA

<sup>3</sup>Research Scholar, JNTUK. Kakinada, India

## Abstract

**Abstract:** The pattern identification problems such as stone, rock categorization and wood recognition are used texture classification technique due to its valuable usage in it. Generally, texture analysis can be done one of the two ways i.e. statistical and structural approaches. More problems are occurred when working with statistical approaches in texture analysis for texture categorization. One of the most popular statistical approaches is Gray Level Co-occurrence Matrices (GLCM) approach. This approach is used to discriminating different textures in images. This approach gives better accuracy results but this takes high computational cost. Usually, texture analysis method depends upon how the texture features are extracted from the image to characterize image. Whenever a new texture feature is derived it is tested whether it is precisely classifies the textures or not. Texture features are most important for precise and accurate texture classification and also important that the way in which they are extracted and applied. The present paper derived a new co-occurrence matrix based on overlapped textons patterns. The present paper generates overlapped texton patterns and generates co-occurrence matrices derived a new matrix called Overlapped Texton Co-occurrence Matrices (OTCoM) for stone texture classification. The present paper integrates the advantages of co-occurrence matrix and texton image by representing the attribute of co-occurrence. The co-occurrence features extracted from the OTCoM provides complete texture information about a texture image. The proposed method is experimented on Vistex, Brodatz textures, CURET, Mayang, Paul Brooke, and Google color texture images. The experimental results indicate the proposed method classification performance is superior to that of many existing methods.

**Keywords:** co-occurrence matrix, texton, Texture Classification

\*\*\*

## 1. INTRODUCTION

Texture classification and segmentation is an important research area from industrial to bio-medical images. The classification problem is basically the problem of identifying an observed textured sample as one of several possible texture classes by a reliable but computationally attractive texture classifier. This implies that the choice of the textural features should be as compact as possible and yet as discriminating as possible. In other words, the extraction of texture features should efficiently embody information about the textural characteristics of the image. The ultimate goal of texture characterization systems is to recognize different textures. To design an effective algorithm for texture classification, it is essential to find a set of texture features with good discriminating power. Previously a number of different texture analysis methods have been introduced namely statistical, structural, transform based and model based methods [1, 2, 3] Normally textures are studied through statistical and syntactical methods. The statistical method measures the coarseness and the directionality of textures in terms of averages on a window of the image [4, 5, 6]. On the other hand syntactical method describes the shape and distribution of the entities. The statistical method has the main features which are to be extracted that includes the autocorrelation function, Fourier transform domain, Markov random field models, local linear transforms, power spectra, difference gray level statistics, co-occurrence matrices and from sum and different statistics [7, 8, 9, 10, 11, 12, 13].

Initially, texture analysis was based on the first order or second order statistics of textures. The co-occurrence matrix features were first proposed by Haralick [6]. Weszka [14] compared texture feature extraction schemes based on the Fourier power spectrum, second order gray level statistics, the co-occurrence statistics and gray level run length statistics. The co occurrence features were found to be the best of these features. This fact is demonstrated in a study by Connors and Harlow [15]. In [16], Haralick features are obtained from wavelet decomposed image yielding improved classification rates.

S.S Sreeja mole [17] in this method classifies the textures on a pixel basis, where each pixel is associated with textural features extracted from co-occurrence matrices that differs the pixel itself. Here the windows related with the adjacent pixels are mostly overlapping resulting the pixels can be obtained by updating values already found. The classification rate in this method is 90%. Jing Yi Tou [18] proposed a method. In this method two popular texture analysis methods i.e. Gabor filters and the Grey Level Co-occurrence Matrices (GLCM). By using this method achieved a recognition rate of 88.52%. Guang-Hai Liu [19] proposed another method. In this method uses the Textons concept and the Grey-level Co-occurrence Matrices (GLCM) techniques used for texture categorization. The preset method uses the combination of the Grey-level Co-

occurrence Matrices (GLCM) and textons are used for stone texture classification. This method can achieve higher classification rate compare to existing methods. The present paper derived a new co-occurrence matrix based on overlapped textons for texture classification. The new co-occurrence matrix is called as Overlapped Texton Co-occurrence Matrix (OTCoM)

This paper is organized as follows. In Section 2, OTCoM and texture features are proposed. Section 3 discusses results and discussions. Conclusions are given in Section 4.

## 2. GENERATION OF OVERLAPPED TEXTON CO-OCCURRENCE MATRIX (OTCoM) AND EXTRACTION OF FEATURES

Various algorithms are proposed by many researchers to extract color, texture and other features. Color is the most distinguishing important and dominant visual feature. That's why color histogram techniques remain popular in the literature. The main drawback of this is, it lacks spatial information. The proposed method consists of three steps which are listed below. In the first step of the proposed method is the color image is converted in to grey level image by using any HSV color model. The following section describes the RGB to HSV conversion procedure

### 2.1 RGB to HSV Color Model Conversion

Recent literature reveal various color models in color image processing. In order to extract facial image features from color image information, the proposed method utilized the HSV color space. In the RGB model, images are represented by three components, one for each primary color – red, green and blue. Hue is a color attribute and represents a dominant color. Saturation is an expression of the relative purity or the degree to which a pure color is diluted by white light. HSV color space describes more accurately the perceptual color relationship than RGB color space because it is adopted with a non-linear transform. The present paper has used HSV color space model conversion, because the present study is aimed to classify the human age in to four groups with a gap of 15 years.

HSV color space is created by Hue (H), saturation (S) and value (V). Hue is the property of color such as red, green and blue. Saturation is the intensity of a specific color. Value is brightness of a specific color. However, HSV color space separates the color into three categories i.e. hue, saturation, and value. Separation means variations of color observed individually.

The transformation equations for RGB to HSV color model conversion is given below.

$$V = \max(R, G, B) \quad (1)$$

$$S = \frac{V - \min(R, G, B)}{V} \quad (2)$$

$$H = \frac{G-B}{6S} \quad \text{if } V = R \quad (3)$$

$$H = \frac{1}{3} + \frac{B-R}{6S} \quad \text{if } V = G \quad (4)$$

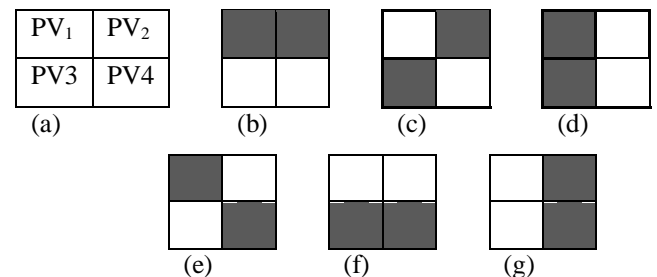
$$H = \frac{1}{3} + \frac{R-G}{6S} \quad \text{if } V = B \quad (5)$$

The range of color component Hue (H) is [0,255], the component saturation (S) range is [0,1] and the Value (V) range is [0,255]. In this work, the color component Hue (H) is considered as color information for the classification of facial images. Color is an important attribute for image processing applications.

### 2.2 Overlapped Texton Matrix Detection

The texton patterns are defined as a set of blobs or growing patterns sharing a common property on the image [21, 22]. Based on the texton theory, texture can be decomposed into elementary units. Julesz's texton theory mainly focuses on analyzing regular textures, while the overlapped textons can be considered as the extension of Julesz's textons. Since overlapped texton involve texture and shape (edge) information, they can better present features for texture classification.

The present paper utilized a 2×2 sub window texton pattern as shown in Fig 1(a). In figure 1(a), the four pixel values of a 2×2 sub window are denoted as PV<sub>1</sub>, PV<sub>2</sub>, PV<sub>3</sub> and PV<sub>4</sub>. If two pixels are highlighted in gray color of same value then the grid will form a texton. The six texton types denoted as TP<sub>1</sub>, TP<sub>2</sub>, TP<sub>3</sub>, TP<sub>4</sub>, TP<sub>5</sub> and TP<sub>6</sub> are shown in figure 1(b) to 1(g).



**Fig 1** Six special types of Textons: a) 2×2 grid b) TP<sub>1</sub> c) TP<sub>2</sub> d) TP<sub>3</sub> e) TP<sub>4</sub> f) TP<sub>5</sub> and g) TP<sub>6</sub>.

Julesz's texton theory mainly focuses on analyzing regular textures, but in all those texture analysis use non overlapped texton features for texture analysis. The disadvantage of non overlapped texton pattern, texton matrix consists of more number of zeros so that more information about the image lost. Another major disadvantage is that non overlapped textons does not consider the neighboring texton elements, even though they form the texton pattern with neighboring texton elements. Form Fig. 2 we observe that non overlapped texton image consists more number of zeros. To overcome the above disadvantages overlapped texton concept applies that causes the texton image consists less number of zeros and more information about the image is available for precise texture classification.

151	143	143	143	152	153	146	146
151	143	143	143	152	153	146	146
155	142	142	138	147	153	148	148
157	143	143	135	142	151	149	149
157	143	143	135	142	151	149	149
154	146	146	140	143	148	146	146
154	146	146	140	143	148	146	146
145	149	149	150	147	144	143	143

(a)

151	143	143	143	152	153	146	146
151	143	143	143	152	153	146	146
0	0	0	0	0	0	148	148
0	0	0	0	0	0	149	149
0	0	0	0	0	0	149	149
0	0	0	0	0	0	146	146
0	0	0	0	0	0	146	146
0	0	0	0	0	0	143	143

(b)

**Fig 2.** Non overlapped texton matrix example (a) original matrix (b) non overlapped matrix

There are many types of texton patterns in texture images. In this paper, we define six types of texton patterns and co-occurrence matrix derived from the overlapped texton pattern image for texture analysis.

V <sub>1</sub>	V <sub>2</sub>
V <sub>3</sub>	V <sub>4</sub>

151	143	143	143	152	153	146	146
151	143	143	143	152	153	146	146
155	142	142	138	147	153	148	148
157	143	143	135	142	151	149	149
157	143	143	135	142	151	149	149
154	146	146	140	143	148	146	146
154	146	146	140	143	148	146	146
145	149	149	150	147	144	143	143

151	143	143	143	152	153	146	146
151	143	143	143	152	153	146	146
0	142	142	0	0	153	148	148
157	143	143	135	142	151	149	149
157	143	143	135	142	151	149	149
154	146	146	140	143	148	146	146
154	146	146	140	143	148	146	146
0	149	149	0	0	0	143	143

**Fig.3:** Illustration of the OTM process: (a) 2×2 grid (b) Original image (c) Overlapped Texton Matrix (OTM)

### 2.3 Co-occurrence Matrix and Features

Recently Texton Co-occurrence matrix (TCM) is proposed in the literature [27] for higher retrieval rate. TCM is defined to be the distribution of co-occurring texton with a given offset over the texton index image. More over TCM is a computationally expensive procedure. To overcome this, the present paper considers Overlapped Texton Matrix (OTM), which is directly obtained from the original image. To extract precise texture features, the present study computes co occurrence matrix for OTM. Due to **co-occurrence matrices are typically large and sparse** they are used to measure the texture image. GLCM is proposed by Haralick et al back in 1973 [6]. It is widely used for various texture analysis applications, such as texture Analysis [24], rock texture classification, wood classification and etc. GLCM is a popular statistical technique for extracting textural features from different types of images. In order to find the spatial relationships effectively, the classification method is used and Grey-level co-occurrence matrix (GLCM) is one of the most widely used statistical texture measures. The idea of the method is to consider the relative frequencies for which two neighboring pixels are separated by a distance on the image.

Since the GLCM collects information about pixel pairs instead of single pixels and which is called by a name as second-order statistics.

The GLCM is generated by cumulating the total numbers of grey pixel pairs from the images. Each GLCM will be generated by defining a spatial distance  $d$  and an orientation, which can be 0 degree, 45 degree, 90 degree or 135 degree at a selected grey level  $G$ . The GLCM produced will be of size  $G \times G$ . When the GLCM is constructed,  $C_d(r,n)$  represents the total pixel pair value where  $r$  represents the reference pixel value and  $n$  represents the neighboring pixel value according to the spatial distance and orientation defined. Co occurrence matrix is generated from the OTM is called Overlapped Texton Co-occurrence Matrix (OTCoM). Based on this, OTCoM with different orientations  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  are formed as shown in Fig.4(a)-(e) respectively. Textural features are extracted from the OTCoM for classification process. There are a total of fourteen features for GLCM [25]. The textural features used in this method are energy, entropy, contrast, local homogeneity, correlation, and inertia are shown in Eq (1) to Eq (6) [5].

3	0	0	0	0	0	0	4		0	0	0	0	0	0	0	7
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	47		3	0	0	0	0	0	0	46
////////(a)									(b)							
0	0	0	0	0	0	0	7		0	0	0	0	0	0	0	5
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0		0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	40		3	0	0	0	0	0	0	41
(c)									(d)							

**Fig. 4:** a) Overlapped Texton matrix (b) (c), (d) and (e) represents the co occurrences on OTCOM of  $0^0$ ,  $45^0$ ,  $90^0$  and  $135^0$ .

$$\text{Entropy} = \sum_{i,j=0}^{N-1} -\ln(P_{ij})P_{ij} \quad (6)$$

$$\text{Energy} = \sum_{i,j=0}^{N-1} -\ln(P_{ij})^2 \quad (7)$$

$$\text{Contrast} = \sum_{i,j=0}^{N-1} P_{ij} (i - j)^2 \quad (8)$$

$$\text{Local Homogeneity} = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1 + (i - j)^2} \quad (9)$$

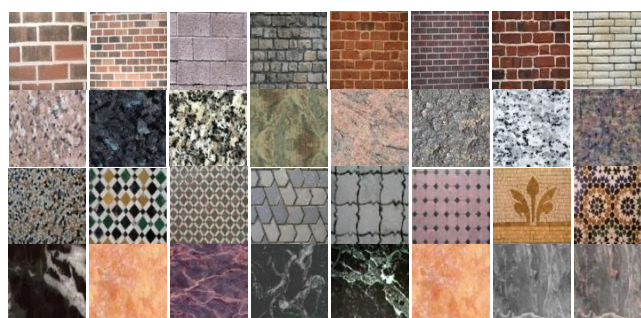
$$\text{Correlation} = \sum_{i,j=0}^{N-1} P_{ij} \frac{(i - \mu)(j - \mu)}{\sigma^2} \quad (10)$$

$$\text{Inertia} = \sum_{i,j=0}^{N-1} (i - j)^2 P(i, j) \quad (11)$$

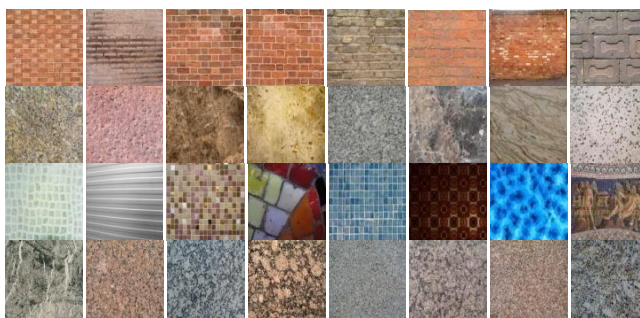
### 3. RESULTS AND DISCUSSIONS

The present paper carried out the experiments on two Datasets. The Dataset-1 consists of various Mosaic, Granite, Marble and Brick stone textures with resolution of  $256 \times 256$  collected Vistex, Brodatz textures, Mayang, Google color texture images and also from natural images taken by using digital camera. Some of stone texture images in Dataset-1 are shown in the Fig. 5. The Dataset-2 consists of various Mosaic, Granite, Marble and Brick stone textures with resolution of  $256 \times 256$  collected from CURET, Paul Bourke, and also from natural images taken by using digital camera. Some of images in Dataset2 are shown in the Fig. 6. Dataset-1 and Dataset-2 contains 80 and 96 original color texture images respectively.

Every texture image is subdivided into 16 sub images of non-overlapped image regions of size  $(64 \times 64)$ . This results into a total of 2816 sub image regions. The classification is done for all sub texture regions derived from each texture image in Dataset-1 and Dataset-2. Feature set leads to representation of the training and testing textures features. The absolute difference of the feature vector values of the query texture and database textures are also calculated.



**Fig.5:** Input texture group of 8 samples of Brick, Granite, Mosaic, Marble with size of  $256 \times 256$



**Fig 6:** Input texture group of 8 samples of Brick, Granite, Mosaic, Marble with size of 200×150

To classify the relevant textures, fixed threshold,  $K$ -NN classifier is used to measure the similarity between query texture and the database textures. In case of fixed threshold, the threshold values are computed for different query textures. The best threshold value is chosen as the threshold of that particular texture feature. The Euclidean distance between these FVs helps in classifying the texture into correct group.

The results from two datasets are obtained in Table 1 & 2 which shows the average classification rates of the proposed OTCoM method.

**Table 1a:** Database-1: (%) mean classification rate of brick and marble stone textures

Sno	Texture Name	Classification Rate	Texture Name	Classification Rate
1	Brick1	96.75	marble1	95.45
2	Brick2	92.81	marble2	97.47
3	Brick3	90.34	marble3	95.12
4	Brick4	96.28	marble4	96.58
5	Brick5	97.47	marble5	91.78
6	Brick6	96.9	marble6	87.57
7	Brick7	90.92	marble7	93.65
8	Brick8	92.71	marble8	93.78
9	Brick9	91.29	marble9	97.42
10	Brick10	96.62	marble10	87.53
11	Brick11	94.74	marble11	95.86
12	Brick12	93.17	marble12	91.79
13	Brick13	91.71	marble13	95.89
14	Brick14	92.76	marble14	95.67
15	Brick15	91.76	marble15	91.17
16	Brick16	93.37	marble16	96.37
17	Brick17	91.76	marble17	90.22
18	Brick18	91.79	marble18	96.23
19	Brick19	91.71	marble19	95.53
20	Brick20	91.76	marble20	91.75
<b>Average</b>		<b>93.331</b>	<b>Average</b>	<b>93.8415</b>

**Table 1b:** Database-1: (%) mean classification rate of mosaic and granite stone textures

Sno	Texture Name	Classification Rate	Texture Name	Classification Rate
1	granite1	91.51	mosiac1	93.87
2	granite2	91.72	mosiac2	90.38
3	granite3	99.68	mosiac3	97.33
4	granite4	87.56	mosiac4	91.76
5	granite5	90.81	mosiac5	89.73
6	granite6	95.56	mosiac6	97.19
7	granite7	79.29	mosiac7	83.37
8	granite8	83.34	mosiac8	91.77
9	granite9	90.53	mosiac9	95.98
10	granite10	96.27	mosiac10	94.71
11	granite11	91.72	mosiac11	98.24

12	granite12	99.74	mosiac12	96.91
13	granite13	97.62	mosiac13	97.52
14	granite14	91.75	mosiac14	92.74
15	granite15	96.93	mosiac15	97.63
16	granite16	97.19	mosiac16	97.48
17	granite17	98.23	mosiac17	91.91
18	granite18	96.83	mosiac18	97.24
19	granite19	91.85	mosiac19	91.56
20	granite20	91.78	mosiac20	97.64
<b>Average</b>		<b>94.248</b>	<b>Average</b>	<b>94.248</b>

**Table 2a:** Database-2: (%) mean classification rate of brick and marble stone textures

Sno	Texture Name	Classification Rate	Texture Name	Classification Rate
1	Brick1	92.90	marble1	93.00
2	Brick2	98.13	marble2	91.7
3	Brick3	93.63	marble3	91.65
4	Brick4	95.37	marble4	94.17
5	Brick5	93	marble5	94.13
6	Brick6	93.57	marble6	89.53
7	Brick7	96.13	marble7	97.53
8	Brick8	94.77	marble8	88.37
9	Brick9	93.07	marble9	95.37
10	Brick10	93.97	marble10	94.6
11	Brick11	93.75	marble11	87.5
12	Brick12	97.13	marble12	98.03
13	Brick13	91.25	marble13	88.37
14	Brick14	88.37	marble14	99.83
15	Brick15	88.37	marble15	96.43
16	Brick16	88.37	marble16	97.43
17	Brick17	88.37	marble17	88.37
18	Brick18	94.37	marble18	92.70
19	Brick19	97.67	marble19	88.37
20	Brick20	89.5	marble20	89.6
21	Brick21	93.50	marble21	91.75
22	Brick22	91.90	marble22	95.85
23	Brick23	95.20	marble23	92.03
24	Brick24	91.13	marble24	95.70
<b>Average</b>		<b>93.06</b>	<b>Average</b>	<b>93.00</b>

**Table 2a:** Database-2: (%) mean classification rate of granite and marble stone textures

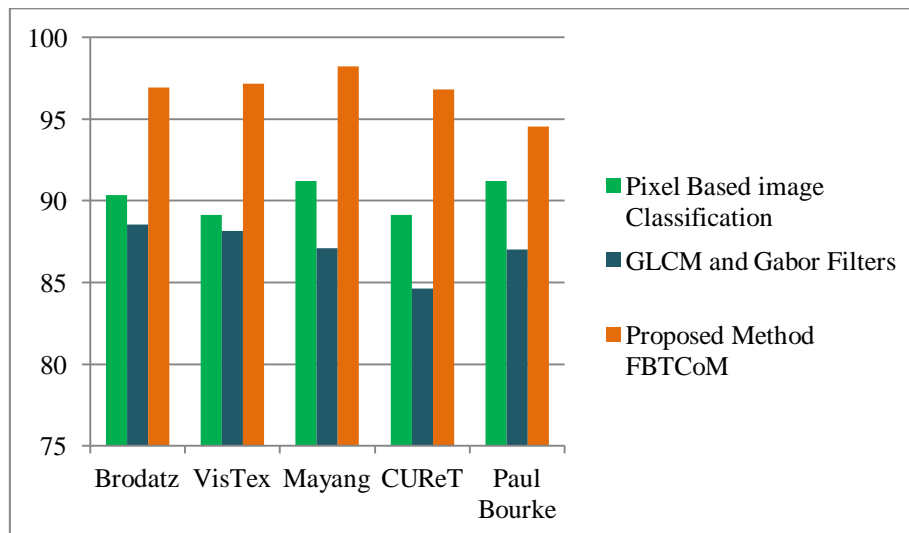
Sno	Texture Name	Classification Rate	Texture Name	Classification Rate
1	granite1		mosiac1	94.3
2	granite2	93.75	mosiac2	93.53
3	granite3	91.6	mosiac3	88.37
4	granite4	94.97	mosiac4	93.50
5	granite5	88.37	mosiac5	91.45
6	granite6	88.37	mosiac6	90.80
7	granite7	93.43	mosiac7	96.83
8	granite8	88.37	mosiac8	90.17
9	granite9	96.67	mosiac9	95.53
10	granite10	93.07	mosiac10	90.20
11	granite11	94.55	mosiac11	88.37
12	granite12	93.07	mosiac12	85.45
13	granite13	95.33	mosiac13	95.23
14	granite14	94.83	mosiac14	93.37
15	granite15	90.73	mosiac15	95.23
16	granite16	93.57	mosiac16	91.7
17	granite17	93.40	mosiac17	91.7

18	granite18	90.60	mosiac18	94.43
19	granite19	89.6	mosiac19	91.7
20	granite20	92.17	mosiac20	96.65
21	granite21	98.83	mosiac21	97.33
22	granite22	91.70	mosiac22	98.13
23	granite23	92.43	mosiac23	97.43
24	granite24	96.20	mosiac24	97.33
<b>Average</b>		<b>92.85</b>	<b>Average</b>	<b>93.28</b>

Mean classification rates for the proposed OTCoM and the other existing methods for classification stone textures using K-NN classifier is shown in Table 3 which clearly indicates that the proposed OTCoM outperforms the other existing methods. Fig.7 shows the comparison chart of the existing methods which are specified in table 3 and proposed OTCoM method.

**Table 3:** Mean classification rates for the two different texture image datasets using k-NN classifier

Image Dataset	Pixel Based image Classification	GLCM and Gabor Filters	Proposed Method FBTCoM
Brodatz	90.34	88.53	96.93
VisTex	89.12	88.15	97.19
Mayang	91.23	87.08	98.23
CUReT	89.12	84.61	96.83
Paul Bourke	91.23	87.01	94.55



**Fig 7:** Classification accuracy comparison of *K-NN* classifier obtained in Brodatz Vistex, Paul Bourke, CURET, and Mayang dataset for Pixel Based image Classification, GLCM and Gabor Filters and proposed method.

#### 4. CONCLUSION

Paragraph The present paper derived a new co-occurrence matrix called as Overlapped Texton Co-occurrence Matrix (OTCoM) for rotation invariant texture classification. Julesz [21] proposed texton which represents the patterns of texture which is useful in texture analysis. The disadvantage of TCM is that, the computationally expensive. To overcome this problem, the present paper considered overlapped Texton Matrix (OTM), which is directly obtained from an original image and to extract precise texture features. The features on color textures are extracted by means of GLCM [17] statistical method though the concept of overlapping window for neighboring pixels. So that, it is the computationally expensive. The present paper statistical method though the concept of non overlapping window for neighboring pixels. The experimental results clearly indicate the efficacy of the proposed OTCoM over the various existing methods.

#### REFERENCES

- [ 1 ] "Measuring Texture Classification Algorithms" by G. Smith and I. Burns, , Pattern Recognition Letters, 1997, vol. 18, pp. 1495-1501
- [ 2 ] "Statistical and Structural Approaches to Texture," by Haralick, R.M., Proceedings of the IEEE, 67, pp. 786-804, 1979
- [ 3 ] 'A theoretical comparison of texture algorithms' by Conners, r.w., and harlow, C.A., IEEE Transactions., May 1980, PAMI-2, pp. 20.5222
- [ 4 ] Texture Classification By Using Advanced Local Binary Patterns And Spatial Distribution Of Dominant Patterns", by Shu Liao and Albert C. S. Chung "in 2007.pp.1221-1224
- [ 5 ] "A comparative study of texture measures for terrain classification" by Weszka, j., dyer, c., and rosenfeld, A in IEEE Transactions. April 1976, SMC- 6, (4), pp. 269-285

- [ 6 ] Haralick, r.m., shanmugam, k., and dinstein, "Textural features for image Classification", IEEE Transactions, November 1973, SMCJ, (6), pp. 610-621
- [ 7 ] Unser, M.: 'Sum and difference histograms for texture classification', IEEE Trans., January 1986, PAMI-8, (1), pp. 118-125
- [ 8 ] Davis, l.s., johns, s.a., and aggarwal, J.K.: "Texture analysis using generalized co-occurrence matrices", IEEE Trans., July 1979, PAMI-1, (3), pp. 251-259
- [ 9 ] Alparone, l., argenti, f., and benelli, g.: 'Fast calculation of co-occurrence matrix parameters for image segmentation', Electron. Lett., January 1990, 26, (1), pp. 23-24
- [ 10 ] A. Laine and J. Fan, "Texture classification by wavelet packet signatures", IEEE Transactions. on PAMI, 15(11), 1993, pp. 1186-1190.
- [ 11 ] "Multichannel Texture Analysis Using Localized Spatial Filters" by A. Bovik , M. Clark , W. S. Geisler, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12 (1), 1990, pp. 55-73.
- [ 12 ] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters", in Pattern Recognition, 24(12), 1991, pp. 1167-1186.
- [ 13 ] M. Unser, M. Eden, "Multiresolution feature extraction and selection for texture segmentation", IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, 1989, pp. 717-728
- [ 14 ] J. S. Weszka, C. R. Dyer, A. Rosenfeld, "A comparative study of texture measures for terrain classification", IEEE Transactions in System Man Cybernat, 6(4), 1976, pp. 269-285.
- [ 15 ] R. W. Connors and C. A. Harlow. "A Theoretical Comparison Texture Algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2, 1980, pp. 204-222.
- [ 16 ] "Texture classification using wavelet transform" by S.Arivazhagan, L. Ganesan, , Pattern Recognition Letters, 24, 2003, pp. 1513-1521.
- [ 17 ] "Texture classification using Wavelet Packet Decomposition", by P. S.Hiremath, and S. Shivashankar, ICGSTs GVIP Journal, 6(2), 2006, pp. 77-80.
- [ 18 ] "Pixel Based Classification On Color Images In Statistical Texture Analysis" by S.S Sreeja Mole, Dr.L.Ganesan, IJCSET -2010, vol 1 no 2 pages 41-46.
- [ 19 ] "Gabor Filters and Grey-level Co-occurrence Matrices in Texture Classification" by Jing Yi Tou, Yong Haur Tay and Phooi Yee Lau Pattern Recognition 34, 2001, pp. 727-739.
- [ 20 ] Julesz B., —Textons, The Elements of Texture Perception, and their Interac-tions," Nature, vol.290 (5802): pp.91-97, 1981.
- [ 21 ] Julesz B., —Texton gradients: the texton theory revisited," Biological Cybernet-ics, vol.54 pp.245–251, 1986.
- [ 22 ] Guang-Hai Liu, Zuo-Yong Li, Lei Zhang, Yong Xu, "Image retrieval based on micro-structure descriptor," Pattern Recognition, vol. 44, pp:2123-2133, 2011.
- [ 23 ] M. Tuceryan, and A. K. Jain, "Texture Analysis, The Handbook of Pattern Recognition and Computer Vision, Ed. 2", World Scientific Publishing Co., 1998.
- [ 24 ] "Design of an Intelligent Wood Recognition System for the Classification of Tropical Wood Species", by Y. L. Lew, Master of Engineering (Electrical) thesis, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Malaysia, 2005
- [ 25 ] M. Petrou, and P. G. Sevilla, "Image Processing: Dealing with Texture", Wiley, 2006
- [ 26 ] N. Jhanwar, S. Chaudhuri, G. Seetharaman, B. Zavidovique "Content based image retrieval using motif cooccurrence matrix", Image and Vision Computing 22 (2004) 1211–1220
- [ 27 ] Guang-Hai Liu\*, Jing-Yu Yang "Image retrieval based on the texton co-occurrence matrix", Pattern Recognition 41 (2008) 3521 -- 3527

## BIOGRAPHIES



**Dr. Patnala S. R. Chandra Murty** completed B.Tech in the year of 2005 from JNTU Hyderabad and M.Tech from ANU, Guntur in the year of 2008. He received Ph D degree in Computer Science and Engineering from JNTUK, Kakinada in year 2013. He published 8 papers in National and International Journals and 7 papers in National and International Conferences. Mail Id chandra\_psr@rediffmail.com



**Dr. U Ravi Babu** completed M.Tech from RVD university in the year of 2005. He received Ph D degree in Computer Science and Engineering from ANU, Guntur in year 2013. He published 10 papers in National and International Journals and 9 papers in National and International Conferences. He working as professor in the depart of CSE at Malla Reddy Engendering College (Autonomous). He is life member of ISCA, IRS, IAENG and CSTA. Mail id: uppu.ravibabu@gmail.com



**R. Venkata Lakshmi** received B.Tech degree in Computer Science & Engineering from Acharya Nagarjuna University. She received her M.Tech degree in Computer Science & Engineering from Andhra University. Presently she is doing her Ph.D in JNTU Kakinada. Her Area of interest is Image Processing.

# Boundary and Region based Moments Analysis for Image Pattern Recognition

Dr. U Ravi Babu <sup>1</sup>, Dr Md Mastan <sup>2</sup> and Dr Y Venkateswarlu <sup>3</sup>

<sup>1</sup> Professor, Department of CSE. MREC, Secendrabad, TS, India

<sup>2</sup> Professor and HOD, Department of CSE. MREC, Secendrabad, TS, India

<sup>3</sup> Professor and HOD, Department of CSE. GIET Engg College, RJY, AP, India

*(Received October 09, 2014, accepted December 28, 2014)*

**Abstract.** In a number of pattern recognition application, moments have been used to recognize image patterns. The recognition process involves effective shape representation method. So, the present paper analyzes the recognition rate in two different shape representation methods called external and internal representation. With the proper shape representation of the given image pattern, the present paper has computed 7 boundary (external) based and 10 region (internal) based Hu moments. The experimental results on five different pattern image groups (Brick, Circle, Curve, Line and Zigzag) are precisely recognized by both boundary based and region based moments of order 1 and 10 respectively.

**Keywords:** Skeleton, Hu Moment, recognition

## 1. Introduction

Image retrieval is becoming a more important problem with the rapid increase of media information. Users want to provide query images and obtain a set of similar images. In content-based image retrieval systems, several low-level image features, such as color, texture, shape or the combination of these features, describe images. Shape is an important low-level image feature.

There are generally two types of shape descriptors: external representation based shape descriptors and internal representation based shape descriptors. External representation based shape descriptors use only the boundary of the objects shape, while the internal representation based shape descriptors use the internal region details in addition to the boundary [1].

Moments due to its ability to represent global features have found extensive applications in the field of image processing [2]–[10]. In 1961, Hu [2] introduced moment invariants. Based on the theory of algebraic invariants he derived a set of moment invariants, which are position, size and orientation independent. Dudani et al. [4] used Hu's moment invariants up to the third order in the recognition of images of aircraft. The same invariants were also used for recognition of ships [5]. Markandey et al. [10] developed techniques for robot sensing based on high dimensional moment invariants and tensors. Gang Xu et al. [11] has proposed a new image recognition algorithm by using region based shape representation. They proposed new region based moments based on skeletons. Zhihu Huang et al. [12] has performed an analysis of boundary based Hu's Moment invariants on image scaling and rotation. Hongbo Mu [13] used boundary and region based Hu moments for recognizing different types of defects in wood pattern images. Cecilia Di Ruberto et al. [14] has combined morphological image features with the moment invariants for classification.

The organization of the paper is as follows. Section 2 deals with the methodology of boundary and region based moment computation, the results and discussions are presented in section 3 and last section deals with conclusions.

## 2. Methodology

Shape representation is an important issue in image processing and computer vision, because it provides the foundation for developing algorithms for shape-related processing such as image coding, shape matching and object/pattern recognition, content-based video processing and image data retrieval.

The present paper uses two types of shape representation namely external representation and internal representation. In external representation, the shape of the given object is represented by the boundary while in the internal representation, the entire region is represented by skeleton.

From the second- and third-order normalized central moments, a set of seven invariant moments, which are invariant to translation, scale change and rotation, has been derived by Hu as given in Equations (1)-(7). The present paper has computed the 7 Hu moments on the boundary of the given image pattern.

$$BM_1 = \eta_{20} + \eta_{02} \quad (1)$$

$$BM_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (2)$$

$$BM_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (3)$$

$$BM_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (4)$$

$$BM_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + 3(\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (5)$$

$$BM_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{11})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{11})(\eta_{21} + \eta_{03}) \quad (6)$$

$$BM_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (7)$$

Where the normalized central moment of order (p+q) is given in the Equation (8)

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q+2}{2}}} \quad (8)$$

The central moment of order (p+q) is given by the Equations (9)-(11).

$$\mu_{pq} = \int \int_{x,y \in c} (x - \bar{x})^p (y - \bar{y})^q f(x,y) dx dy \quad (9)$$

$$\bar{x} = \frac{m_{10}}{m_{00}} = \frac{\int \int_{x,y \in c} x f(x,y) dx dy}{\int \int_{x,y \in c} f(x,y) dx dy} \quad (10)$$

$$\bar{y} = \frac{m_{01}}{m_{00}} = \frac{\int \int_{x,y \in c} y f(x,y) dx dy}{\int \int_{x,y \in c} f(x,y) dx dy} \quad (11)$$

The 10 extended Hu moments given in the Equations (12)-(21) are computed on the skeleton of the given image.

$$RM1 = \frac{\sqrt{BM2}}{BM1} \quad (12)$$

$$RM2 = \sqrt{\frac{BM1 + \sqrt{BM2}}{BM1 - \sqrt{BM2}}} \quad (13)$$

$$RM3 = \frac{\sqrt{BM3}}{\sqrt{BM4}} \quad (14)$$

$$RM4 = \sqrt{\frac{BM3}{\sqrt{|BM5|}}} \quad (15)$$

$$RM5 = \sqrt{\frac{BM4}{\sqrt{|BM5|}}} \quad (16)$$

$$RM6 = \sqrt{\frac{|BM6|}{BM1 \times BM3}} \quad (17)$$

$$RM7 = \sqrt{\frac{|BM6|}{BM1 \times \sqrt{BM5}}} \quad (18)$$

$$RM8 = \sqrt{\frac{|BM6|}{BM1 \times BM4}} \quad (19)$$

$$RM9 = \sqrt{\frac{|BM6|}{\sqrt{BM2 \times |BM5|}}} \quad (20)$$

$$RM10 = \sqrt{\frac{|BM5|}{BM2 \times BM3}} \quad (21)$$

### 3. Results and Discussions

The experiments are conducted with 5 different image pattern groups namely Brick, Circle, Curve, Line and Zigzag patterns, 10 images per group of each size  $256 \times 256$ , collected as shown in Figure (1)-(5). In the first method, the 7 boundary based Hu moments are calculated and within each group, the average of all 10 images is computed and are represented in Table 1. In the second method, the 10 region based extended Hu moments are calculated and within each group, the average of all 10 images is computed and are represented in Table 2.

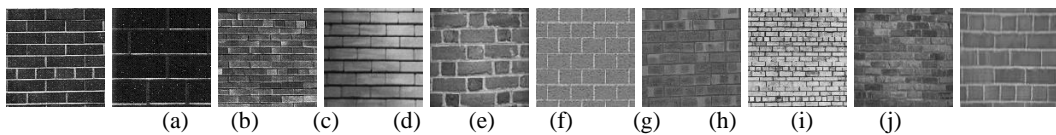


Figure 1. Input Images of Brick Pattern Images (a) Brick1 (b) Brick2 (c) Brick3 (d) Brick4 (e) Brick5 (f) Brick6 (g) Brick7 (h) Brick8 (i) Brick9 (j) Brick10.

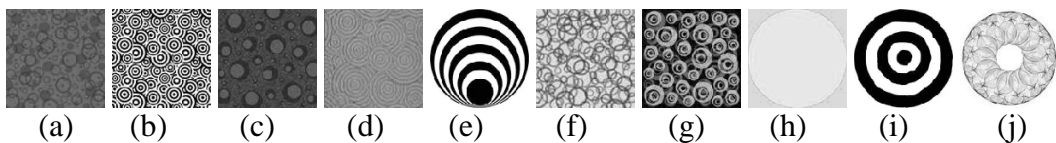


Figure 2. Input Images of Circle Pattern Images (a) Circle1 (b) Circle2 (c) Circle3 (d) Circle4 (e) Circle5 (f) Circle6 (g) Circle7 (h) Circle8 (i) Circle9 (j) Circle10.

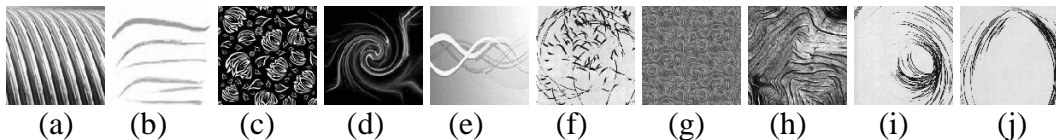


Figure 3. Input Images of Curve Pattern Images (a) Curve1 (b) Curve2 (c) Curve3 (d) Curve4 (e) Curve5 (f) Curve6 (g) Curve7 (h) Curve8 (i) Curve9 (j) Curve10.

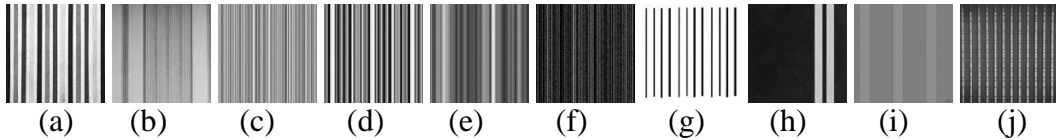


Figure 4. Input Images of Line Pattern Images (a) Line1 (b) Line2 (c) Line3 (d) Line4 (e) Line5 (f) Line6 (g) Line7 (h) Line8 (i) Line9 (j) Line10.

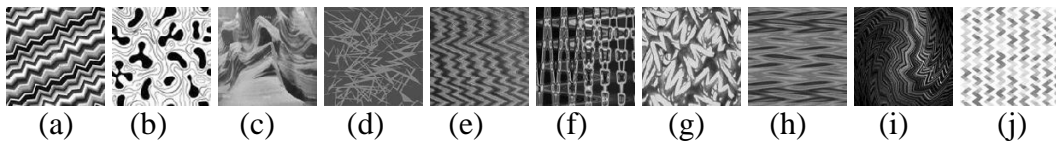


Figure 5. Input Images of Zigzag Pattern Images (a) Zigzag1 (b) Zigzag2 (c) Zigzag3 (d) Zigzag4 (e) Zigzag5 (f) Zigzag6 (g) Zigzag7 (h) Zigzag8 (i) Zigzag9 (j) Zigzag10.

Table 1. Average Boundary based Hu moment values of Image Pattern Groups.

Image Name	BM1	BM2	BM3	BM4	BM5	BM6	BM7
Brick	0.8873	0.0397	0.0393	0.0035	0.0000	-0.0007	0.0000
Circle	1.4598	0.0005	0.0324	0.1274	0.0639	0.0002	-0.0057
Curve	1.2604	0.0787	0.2987	0.2461	-0.0557	-0.0940	0.3726
Line	2.1213	2.4965	0.1688	0.1627	0.2216	0.0214	0.0099
Zigzag	0.9981	0.0206	0.0203	0.0057	0.0000	-0.0003	-0.0004

Table 2. Average Region based Extended Hu moment values of Image Pattern Groups.

Image Name	RM1	RM2	RM3	RM4	RM5	RM6	RM7	RM8	RM9	RM10
Brick	0.0739	1.0873	2.2204	2.0890	1.1205	0.1411	0.1657	0.1613	0.9113	191.3938
Circle	0.0240	1.0244	2.2818	4.1238	1.6135	0.1508	0.1746	0.1079	1.4050	316.7487
Curve	0.1688	1.1959	1.2629	1.6468	1.3142	0.2768	0.3598	0.2828	1.0213	55.8664
Line	0.1457	1.4402	1.2398	1.3584	1.2829	0.3139	0.3283	0.2385	1.1754	535.7405
Zigzag	0.0821	1.0888	3.6058	2.7394	0.9756	0.2596	0.2132	0.2035	0.8462	60.8954

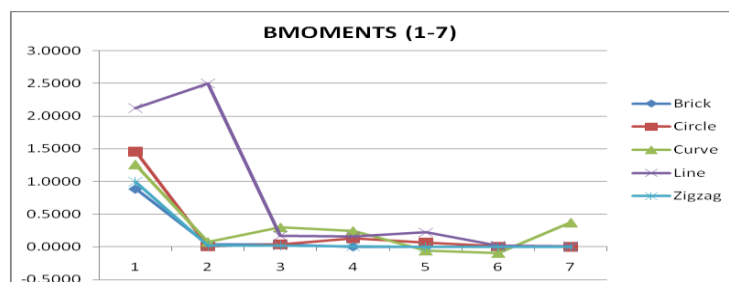


Figure 6. Recognition graph for Image Pattern Groups by Boundary based Moments.

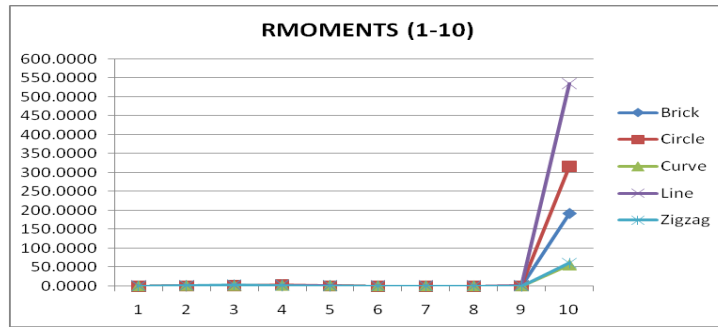


Figure 7. Recognition graph for Image Pattern Groups by Region based Moments.

The graphs of Figure 6 and 7 show the recognition rate of all five different image pattern groups. All five groups are precisely recognized by using boundary based moments of order 1 and region based moments of order 10. The Table 3 shows the average computation time (in second) used by the boundary based and region based moments per each group of image pattern.

Table 3. Computation Time (in seconds) used by Boundary based and Region based Moments.

Input Image Pattern Group	Boundary based Moment	Region based Moment
Brick	2.21466	1.49968
Circle	2.20549	2.16217
Curve	2.22533	2.38623
Line	2.22407	1.63937
Zigzag	2.20469	1.66281

## 4. Conclusions

All the five types of images are clearly classified by using boundary and region based Moments. The images are precisely classified by using Boundary Moment of order 1 and by using Region Moment of order 10. In boundary and region moments, Line shape images are having maximum values. Though the images are precisely classified by boundary and region moments, the region moments are efficient because the average computation time is less compared to boundary moments.

## 5. References

- [1] Irina Mocanu, "Image Retrieval by Shape Based on Contour Techniques - A Comparative Study," IEEE Conference, pp. 219-223, 2007.
- [2] M. K. Hu, "Visual pattern recognition by moment invariants," IRE Trans. Inform. Theory, vol. IT-8, pp. 179-187, Feb. 1962.
- [3] M. R. Teague, "Image analysis via the general theory of moments," J. Opt. Soc. Amer., vol. 70, pp. 920-930, Aug. 1962.
- [4] S. Dudani, K. Breeding, and R. McGhee, "Aircraft identification by moment invariants," IEEE Trans. Comput., vol. 26, pp. 39-45, Feb. 1977.
- [5] D. Casasent and R. Cheatham, "Image segmentation and real image tests for an optical moment-based feature extractor," Opt. Commun., vol. 51, pp. 227-230, Sept. 1984.
- [6] A. Khotanzad and J. J. H. Liou, "Recognition and pose estimation of unoccluded three-dimensional objects from a two-dimensional perspective view by banks of neural networks," IEEE Trans. Neural Networks, vol. 7, pp. 897-906, Sept. 1996.
- [7] S. O. Belkasim, M. Shridhar, and M. Ahmadi, "Pattern recognition with moment invariants: A comparative study and new results," Pattern Recognit., vol. 24, no. 12, pp. 1117-1138, 1991.
- [8] Shape recognition using Zernike moment invariants," in Proc. 23rd Annu. Asilomar Conf. Signals Systems and Computers, Oct.-Nov. 1989, pp. 167-171.

- [9] S. Ghosal and R. Mehrotra, "Orthogonal moment operators for subpixel edge detection," in Proc. 23rd Annu. Asilomar Conf. Signals Systems and Computers, vol. 26, 1993, pp. 295–306.
- [10] V. Markandey and R. J. P. Figueiredo, "Robot sensing techniques based on high dimensional moment invariants and tensors," IEEE Trans. Robot Automat., vol. 8, pp. 186–195, Feb. 1992.
- [11] Gang Xu, Yuqing Lei, "A New Image Recognition Algorithm based on Skeleton," IEEE 2008, pp. 777-782.
- [12] Zhihu Huang, Jinsong Leng, "Analysis of Hu's Moment Invariants on Image Scaling and Rotation," IEEE 2<sup>nd</sup> International Conference on Computer Engineering and Technology, Vol. 7, pp. 476-480, 2010.
- [13] Hongbo Mu, Dawei Qi, "Pattern Recognition of Wood Defects Types based on Hu Invariant Moments," IEEE 2009.
- [14] Cecila Di Ruberto et al., "Moment based Techniques for Image Retrieval," IEEE 19<sup>th</sup> International Conference on Database and Expert Systems Application, pp. 155-159, 2008.

# A Novel Method for Rotational and Pose Invariant Child or Adult Age Classification based on Morphological Pattern Representation Schemes

<sup>1</sup>Dr. Pullela S V V S R Kumar, <sup>2</sup> Ms. A V Lakshmi Dhanisetty, <sup>3</sup>Dr, U Ravi Babu

<sup>1</sup> Professor of CSE, Sri Aditya Engineering College, Surampalem, pullelark@yahoo.com

<sup>2</sup> Computer Science & Engineering, Sri Aditya Engineering College,  
Email ID : srilakshmimca4@gmail.com

<sup>3</sup> Professor, Dept. of CSE, MREC(A), India, uppu.ravibabu@mrec.ac.in

**Abstract--** The present paper proposes an approach for classify the facial images either child or adult based on local texture features extracted on facial images. The local texture features used in this paper are Morphological-prehistoric Patterns with grain components (MP-g). these patterns are extracted on a Local Directional Pattern (LDP) values. Edge response values in all eight directions from a 3×3 local window are used to calculate the LDP values. The local descriptor LDP is more constant in the incidence of noise and lighting changes, since edge response level is more stable than pixel intensity. The proposed method is rotationally & poses invariant when compared to pattern trends that represents a shape. The proposed method is tested on large set of images from different databases like FgNet, Morph, Google and scanned images. The present paper proves the efficiency of the proposed method.

**Indexed Term—** Morphological Prehistoric Patterns, Local Directional Pattern, invariant, pose invariant.

## 1. INTRODUCTION

Computer vision and psychophysics researchers faced so many challenges by face recognition of human and representation the typical features of human faces. Human faces be in the right place to a 3-D objects. Therefore, to develop an accurate representation for description that consider for lighting; face variations, facial expressions, etc.; of facial images. In face recognition system, large database maintained and search an image in that database. Day to day increase the size of the database and also increase the computational cost for the classification of human face into two categories i.e. either child or adult.

To address this problem some procedures that highlights the important of facial development over a period of instance and considered a representation to minimize the dissimilarity between test databases and images which are used at the time of method development. In appearance-based method, a face image is usually considered as a point in the high-dimensional space. Many linear subspace learning methods, such as LDA/FKT (linear discriminant analysis/Fukunaga–Koontz transform) [6], Eigenface [1, 2], C-LDA (complete LDA) [4], Fisherface [3], MMSD (multiple maximum scatter difference) [5], and Laplacian face [7] are typical dimensionality reduction methods to find a low-dimensional feature space. First, Local Binary Pattern LBP operator was introduced by Ojala et al. [8] for texture categorization and texture analysis.

The LBP given good results especially in texture analysis so that LBP mostly suitable in texture examination and its applications. The LBP operator has huge acceptance against lighting changes. The main advantage of LBP is, It was obtained with sample calculations. From the above properties, it is well suitable for real-world applications like image texture analysis. The concept of LBP was also extended in applications such as face recognition and age classification [9, 10, 11].

In above proposed methods are applied on small set images and get comparative results but those methods are not suitable for all kinds of databases. The proposed method is applied on large set of images and different kinds of databases. No such method is available to classify the facial images into two categories i.e. either child or adult in effective and efficient manner of different database like FgNet, Morph Google and scanned image databases.

The paper is organized as follows: Section 2 explains the proposed method and result analysis, experimental details and discussions are explained in Section 3. Conclusion part are given in the section 4.

## 2. THE PROPOSED MORPHOLOGY PREHISTORIC PATTERNS WITH GRAIN COMPONENTS ON LDP FOR CHILD AND ADULT AGE CLASSIFICATION

The present paper identified MP-g based on changes in the facial skin. The skin of a human face tends to more alters with growing age is observed by the present method. These changes in the skin are subjugated by MP-g. The other important feature of the present method is to find which MP-g causes, the quick changes in the skin that leads to a higher age classification rate. The proposed novel scheme contains five major steps. In the first step, the facial color image is cropped to cover only skin area of the face. Convert the RGB into gray scale image in 2<sup>nd</sup> step. Gray level facial image is converted into a two valued image by using LDP approach in the 3<sup>rd</sup> step. In the fourth step, identify the Morphological prehistoric patterns with one to eight-grain components (MP-1g to MP-8g). In the final step, derive an algorithm for classifying the facial image into either child or adult. The schematic diagram of the entire process is illustrated in figure 1.

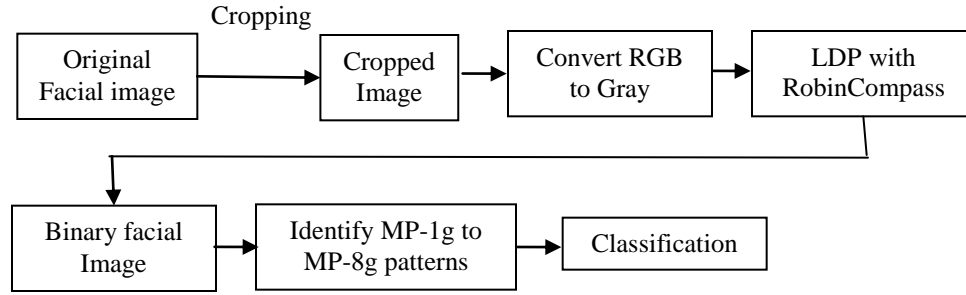


Fig. 1. Block diagram of the child and adult classification system

**Step 1: Cropping:** In this proposed method cropping is necessary because of identifying the facial skin edges. The proposed method based on the skin area of the human face. The given database, each facial image consist not only skin area but also hair, neck, and other areas also included. So to remove the unwanted area from the facial image cropping is necessary for effective results. The input facial image is cropped to cover the entire skin area of the face based on the location of two eyes in the first step as shown in figure 2.

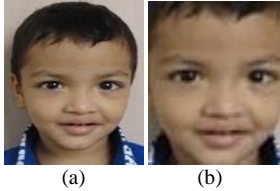


Fig. 2. Crop operation a) Input image b) Output image after cropping

### Step 2: RGB to Gray scale conversion:

So many color models are available in color image processing. The facial image is converted to gray scale image for identifying the morphological patterns on the facial color image. The present paper uses HSV color model for converting the facial color image into gray scale, because the present study is aimed to classify the human age into four groups with a gap of 15 years based on changes on facial skin are identified on gray scale image. HSV color space separates the color into three categories i.e. hue, saturation, and value. Separation means variations of color are observed individually. The converting equations for RGB to grey level conversion are given below.

$$V = \max(R, G, B) \quad (1)$$

$$S = \frac{V - \min(R, G, B)}{V} \quad (2)$$

$$H = \frac{G-B}{6S} \text{ if } V = R \quad (3)$$

$$H = \frac{1}{3} + \frac{B-R}{6S} \text{ if } V = G \quad (4)$$

$$H = \frac{1}{3} + \frac{R-G}{6S} \text{ if } V = B \quad (5)$$

In this work, the color component Hue (H) is considered as grey information for the classification of facial images. And this value lies between 0 and 255.

### Step 3: Local Directional Pattern

Local Directional Pattern (LDP) [12] concept is used in this present study because it has more advantages compare to LBP approach. The LDP approach is more suitable for age group classification because this approach considers the edge response values in all different directions instead of surrounding neighboring pixel intensities like LBP. This provides more consistency in the presence of noise, and illumination changes since edge response magnitude is more stable than pixel intensity. The LDP is based on LBP. In the LBP operator, a gray-scale invariant texture prehistoric, has gained significant popularity for describing the texture [13]. It labels each pixel of an image by thresholding its P-neighboring values with the center value by converting the result into a binary number by using Equation 6.

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(ESP - CP) 2^p,$$

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

Where CP denotes the gray value of the center pixel  $(x_c, y_c)$  and ESP corresponds to the gray values of P equally spaced pixels on the circumference of a circle with radius R.

### Local Directional Pattern with Robin Compass Masks Response

The LDP generates eight-bit binary code assigned to each  $3 \times 3$  sub window of an input image. These patterns are calculated by comparing the relative edge response value of a pixel in different directions by using Robin compass masks. The Robin compass masks in eight distinct orientations  $(r_0 \sim r_7)$  centered on its own position. The Robin compass masks are shown in the Fig.3.

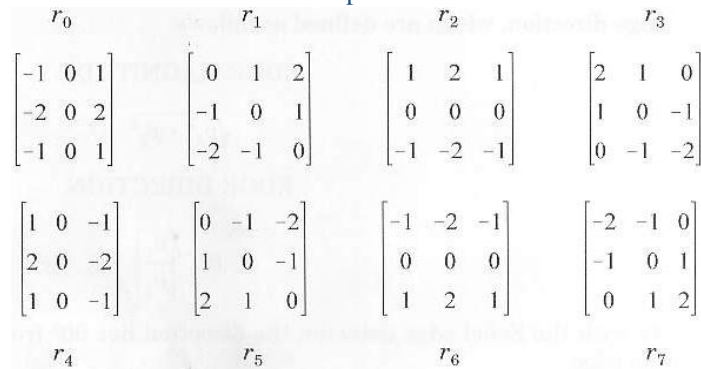


Fig. 3. Robin Compass masks in eight directions.

Applying Robin masks on  $3 \times 3$  masks, eight mask values  $V_0, V_1, \dots, V_7$  are obtained, each representing the edge significance in its respective direction. The mask values are not equally important in all directions. In robin mask values corner or edge pixels show higher values because those pixels are more important in particular direction.

The LDP code produces the more firm pattern in the presence of noise, illumination changes and various

conversion schemes of color facial images into gray images. For example, Fig. 4 shows an original image and the corresponding image with illumination changes. After illumination change, 5<sup>th</sup> bit of LBP changed from 1 to 0, Thus LBP pattern changed from uniform to a non-uniform code. Since gradients are more stable than gray value, LDP pattern provides the same-pattern value even in the presence of noise and non-monotonic illumination changes.

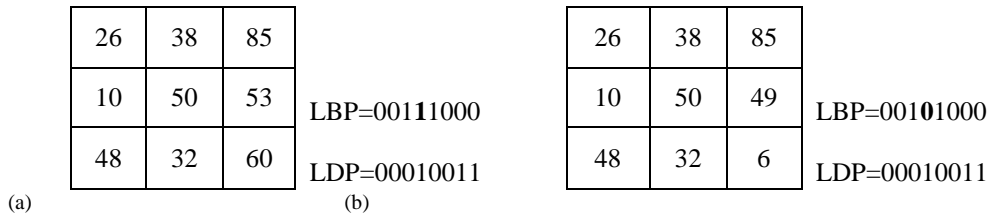
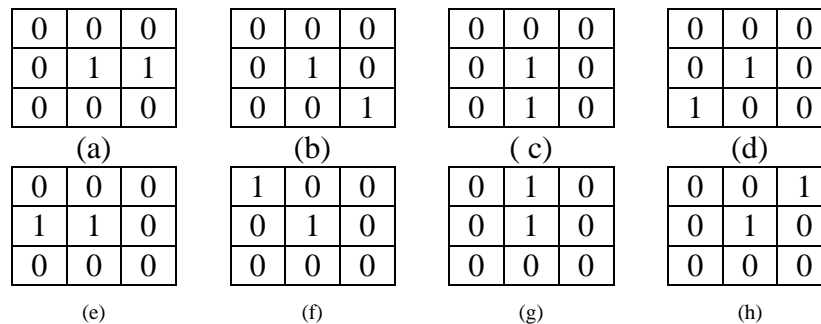


Fig.4.Calculating the LDP and LBP values in two different situations (a) Original Image (b) Image with Noise

#### Step 4: Evaluation of Morphological-prehistoric Patterns with Grain Components (MP-g) on LDP

On the binary LDP facial texture images of the previous step, the present study evaluated the incidence of MP-g on a  $3 \times 3$  mask. The present study classify the facial images into either Child or adult based on the number of grains incidence on a  $3 \times 3$  sub-window LDP facial image in any orientation. This makes the present method as rotationally and poses invariant. The present method counts frequency incidence of MP-g if and only if the central pixel of the  $3 \times 3$  window is 1 and it is treated as a grain. If the central pixel is zero (0) then  $3 \times 3$  window is treated as not-a-grain. In the following figures '0' indicates no grain and '1' indicates a grain. There can be eight combinations of MP-1g, which are shown in the Fig. 5. By any rotation, the MP-1g may change its position in 8 ways

on a  $3 \times 3$  mask as shown in Fig.5. The present method counts the frequency incidence of MP-1g on a  $3 \times 3$  mask irrespective of its position. Therefore, the present method is rotationally invariant. There will be seven different formations of MP's with 2pixel-grain components (MP-2g) by fixing one of the grains at pixel location (0,0) on a  $3 \times 3$  mask as shown in Fig. 6. In a similar way, there will be six formations of MP-2g by locating one of the grains at the pixel location (0,1) as shown in Fig. 7. Thus, there will be 7! ways of forming MP-2g for a  $3 \times 3$  window. In the same way, there will be 6!, 5!, 4!, 3!, 2! and 1! ways of forming MP-g of 3, 4, 5, 6, 7 and 8 respectively, on a  $3 \times 3$  mask irrespective of their rotational positions.

Fig. 5. Representation of  $3 \times 3$  window with MP-1g.

1	0	0
0	1	1
0	0	0

(a)

1	0	0
0	1	0
0	0	1

(b)

1	0	0
0	1	0
0	1	0

(c)

1	0	0
0	1	0
1	0	0

(d)

1	0	0
1	1	0
0	0	0

(e)

1	1	0
0	1	0
0	0	0

(f)

1	0	1
0	1	0
0	0	0

(g)

Fig. 6. Representation of MP-2g by fixing grain at (0,1)

0	1	0
0	1	1
0	0	0

(a)

0	1	0
0	1	0
0	0	1

(b)

0	1	0
0	1	0
0	1	0

(c)

0	1	0
0	1	0
1	0	0

(d)

0	1	0
1	1	0
0	0	0

(e)

0	1	1
0	1	0
0	0	0

(f)

Fig. 7. Representation of MP-2g by fixing one of the grain component at (0,1).

The frequency incidence of MP-1g to MP-8g are computed by the proposed Algorithm 1 on a 3×3 non-overlapped mask of LDP facial texture images, and stored in the facial database.

The pseudo code entire method is illustrated below

Step 1: Read input image

Step 2: Crop the facial image

Step 3: if the input image is RGB then convert the image into gray level image as described by the step2 in section2 with the size of N×N

Step 4: Define the RobinMasks( $r_0, r_1, r_2, r_3, r_4, r_5, r_6, r_7$ ) based on masks given in the figure 3.

Step 5: Convert the Gray image into Binary by using LDP Robin Masks

- Apply 2-D convolution of image  $G(i,j)$  and mask  $M_N(i,j)$  with size 3×3  $F(i,j) = G(i,j) * M_N(i,j)$
- Calculate the threshold,  $th = F(i+1,j+1)$
- Initialize the variables  $p=0$  and  $q=0$
- if  $F(p,q) \geq th$  then  $BW(p,q) = 1$  otherwise  $BW(p,q) = 0$
- Initialize the variables  $l=0$  and  $m=0$
- for  $a = l$  to  $l+2$  and  $b = m$  to  $m+2$  do the following procedure
- if  $BW(l+1,m+1) == 1$  then Count the frequencies of incidence of Grain Features from 1,2,...,8 otherwise update  $m = m+3$

- if  $m < N$  then go to step (f) otherwise update  $l = l+3$  and  $m = 0$

- if  $l < N$  then go to step (f) otherwise Store the Grain features in the face recognition database and stop the procedure

### 3. RESULTS AND DISCUSSION

To find the consequence of the present method, the present study has evaluated MP-1g to MP-8g on facial LDP images of child and adults from different poses of 1002 FgNet ageing database images, 12000 Morph data base images, 1500 scanned images and 750 Google imagestotally, its leads of 15252 images. Some of the child and adult images in different databases are shown in figure8 and Figure9. The present study considered that the childhood agesbetween 0 and 15 years and adulthood is over30 years.

Table 1 and Table 2 indicate the frequency incidence of MP-g for sample database of FgNet. The facial image is recognized as child or adult based on Algorithm 1. From the computed frequency of incidence of MP-1g to MP-8g by the Algorithm 1 the present study observed that only twoMP-gS are exhibiting successful child-adulthood classification rates. The MP-1g and MP-4g have proved to have significant, precise and accurate classification rates than others MP-g's. The present study suggests that it is not necessary to evaluate frequency incidence of MP-2g, MP-3g, MP-5g, MP-6g, MP-7g and MP-8g on facial LDP images of the child and adult for classification. Since the facial images are of different poses, the proposed method is posed invariant. To prove the proposed method is rotationally invariant the MP-g are evaluated with different rotations 30, 45 and 135 degrees

and the obtained results are scheduled in Tables from 3 to 8. Even by rotation with different angles, the Algorithm 1 based on MP-1g and MP-4g classifies the child and adult. This proves that the present method is rotationally invariant. Thus,

the present method has overcome the disadvantage of pattern based and also previous methods that are rotational and pose variant.



Fig. 8. Sample facial images of adults with different poses of FGNet aging database.

Table I  
Frequency incidence of child facial images using MP-g.

Child images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
015A01	14689	765	679	964	1047	2015	1371	2076
002A03	15365	864	837	952	964	2135	1129	3168
009A03	13649	752	913	994	899	2049	1078	7045
069A03	13451	1468	1188	1228	2209	4066	2155	895
053A06	14478	1090	861	1052	1233	1763	1186	548
066A06a	19661	1170	929	1138	1454	2221	1379	828
019A07	14369	1092	982	1163	1676	3527	2956	3015
016A08	17206	788	658	1009	833	1216	363	138
023A09	21243	1459	990	1183	1129	1521	774	581
073A09	10871	1186	1164	1122	2258	4396	3901	2482
065A09	11821	806	661	856	930	1774	1501	10531
011A11	18020	958	687	919	733	1153	611	230
022A11	22196	789	799	910	1092	1325	863	706
012A12	16829	923	690	1023	854	9080	617	195
008A13	14843	1421	1046	1162	1305	1389	746	299
Sci-img-01	13568	963	914	984	865	1423	865	203
Sci-img-02	14768	768	863	913	1025	1564	1365	276
Sci-img-03	15649	854	786	927	973	1579	1278	536
Sci-img-04	18637	913	597	1013	918	1679	1135	895
Sci-img-05	19037	963	618	1019	879	2123	1496	9563
mor-img-01	17433	876	723	1101	1123	2094	1526	8768
mor-img-02	15684	961	831	978	1079	1998	1428	1963
mor-img-03	13081	1293	1176	1164	2372	4605	2996	1493
mor-img-04	16566	620	640	917	1065	2490	2168	4414
mor-img-05	14595	926	726	935	1153	2255	2626	5664

Table II  
Frequency incidence of adult facial images using MP-g.

Adult images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
027A41	13658	1075	865	1364	1123	1450	1211	384
062A41	15638	1349	913	1268	1235	3623	1466	295
071A42	17698	1745	1035	1495	1465	2359	1355	406
033A44	18486	2146	1583	2037	2140	2855	923	406
047A45	22360	1140	924	1273	1088	1400	514	281
028A46	20731	861	760	1182	1368	2120	1011	647
045A48	18649	1948	964	2018	2013	1365	1217	594
003A49	20134	1956	813	2134	2348	1433	1175	712
048A50	23498	896	854	1204	1065	1637	885	337
039A52	16538	2313	2169	2406	2635	3013	1143	359
004A53	18649	2013	979	1769	1645	1956	1259	535
005A61	16498	2146	1237	1649	2156	1765	1567	478
006A61	15253	1594	1477	2057	2424	3758	1534	783
Sci-img-101	22308	1374	981	1437	1105	1233	251	191
Sci-img-102	20125	2237	1165	1548	1652	2016	1468	776
Sci-img-103	21237	2379	1179	1788	1847	2347	1346	649
Sci-img-104	19858	1290	1067	1621	1569	2188	852	435
Sci-img-105	14628	1643	1629	1936	2414	4082	2934	1310
mor-img-501	19556	1065	910	1278	1395	2373	1587	716
mor-img-502	14390	2159	1912	2513	3067	4187	1770	578
mor-img-503	19679	2140	1204	1499	1769	2246	1646	662
mor-img-504	20860	1351	952	1771	1637	2585	998	522
mor-img-505	23198	1298	1076	1254	1286	1654	586	224

Algorithm 1: Rotational and pose invariant child and adult age classification using MP-g on LDP.

```

if (MP-1g<=13700)
    print (facial image is of Child)
elseif (((MP-1g >14000) && (MP-1g <22300)) && (MP-4g<=1200))
    print (facial image is of Child)
elseif (((MP-1g >14000) && (MP-1g <23500)) && ((MP-4g >1300) && (MP-4g < 2350)))
    print (facial image is of the Adult)
else
    print (facial image is not of the child and the adult)
end

```

The Algorithm 1 classifies child from adult, based only on the frequency incidence of MP-1g and MP-4g values. If a MP-1g value is less than 13700 then the facial image is treated as child, else they form group 2 entries. By considering both MP-1g and MP-4g values if a MP-1g count is inbetween

14000 to 23500 and a MP-4g count is less than 1300 then it classifies as a child otherwise adult as specified in the Algorithm 1. The same thing is also true for all rotational facial texture images.

Table III  
Frequency incidence of a MP-g with 30° rotation on LDP for child facial images.

Child images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
002A03	10346	613	634	1064	1348	2216	2349	16494
008A13	9768	718	728	978	1346	2575	2659	13462
009A03	10769	834	531	867	927	1567	1946	12346
011A11	15722	794	562	637	774	799	520	18924
012A12	16014	1268	888	938	889	914	574	20005
015A01	7568	513	1365	846	1035	936	694	19864
016A08	8316	648	1230	813	778	1365	1346	23156
019A07	9769	716	1065	679	684	1364	1864	23458
022A11	10325	839	943	681	1235	2346	1964	22467
023A09	12356	746	1032	1035	1365	2867	2034	19467
073A09	6106	1029	1094	1138	2230	2255	3884	22269
069A03	14389	911	785	932	1199	1224	2624	31622
066A06a	6139	685	734	1010	1349	1374	2439	17932
065A09	7024	539	517	617	756	781	1371	30261
053A06	11737	480	471	743	982	1007	2062	24097
Sci-img-01	13645	943	978	1094	1649	3121	2264	18631
Sci-img-02	16819	852	662	842	994	1019	527	20343
Sci-img-03	13971	1431	1085	1169	1196	1221	717	19158
Sci-img-04	11261	984	788	999	1179	1204	1038	17023
Sci-img-05	11455	585	876	1113	1869	2348	2019	17345
mor-img-01	8301	1149	1227	1004	2370	2395	2722	21617
mor-img-02	14290	950	751	992	1268	1293	1123	20056
mor-img-03	9658	579	528	663	673	698	364	12336
mor-img-04	8923	732	567	703	765	790	295	11841
mor-img-05	13257	1595	1287	1145	2333	2358	2002	25244

Table IV  
Frequency incidence of a MP-g with 30° rotation on LDP for adult facial images.

Adult images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
003A49	11235	1358	1265	1547	1649	2065	1425	23456
005A61	13465	1649	1349	1463	1563	2164	1268	22658
006A61	12346	1346	1765	1429	1429	2318	1361	31456
027A41	17960	812	1326	1823	1034	652	670	32992
028A46	18893	1196	1280	1399	1255	980	580	30986
033A44	13495	574	1270	2214	1099	1649	1345	20349
039A52	15144	2391	2726	2318	2624	2114	1227	27725
045A48	10395	974	1794	1358	1037	2495	1278	25648
047A45	10970	1526	3704	1850	2327	1540	3038	31314
048A50	12586	618	1873	2199	1267	1765	1994	21546
071A42	18556	1167	910	1296	1155	888	295	24098
062A41	18032	1081	1110	1246	1008	835	419	24834
004A53	14167	833	1972	1967	1274	694	1070	27388
Sci-img-101	13644	529	1979	2035	1348	1649	1236	20658
Sci-img-102	15264	1137	1905	1213	1471	995	943	33341
Sci-img-103	9034	2093	3372	2285	2930	1832	1600	33123
Sci-img-104	11630	1679	1973	1607	1783	1317	957	35323
Sci-img-105	10349	936	1864	1647	945	2020	1649	26690
mor-img-501	12453	817	1945	1950	865	1367	1721	27988
mor-img-502	11358	928	2034	1713	765	1428	1833	26485
mor-img-503	11173	1693	3195	1942	2371	1517	1629	24845
mor-img-504	14159	916	1957	1999	1343	803	1267	26921
mor-img-505	14129	1234	1716	1339	1530	997	603	26817

Table V  
Frequency incidence of a MP-g with 45° rotation on LDP for child facial images.

Child images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
002A03	12649	815	1344	611	1214	1458	1127	18188
008A13	7510	1243	1294	1035	2210	2766	2771	22845
009A03	13432	946	826	1065	1115	1169	1174	21530
011A11	9468	798	1037	864	1679	1864	1095	18465
012A12	12466	1638	1397	1183	2153	2008	2013	27594
015A01	15160	1270	873	914	864	731	736	21532
016A08	9846	575	613	865	894	1649	1465	24683
019A07	11235	1035	1175	599	2133	2213	845	20316
022A11	8697	661	535	579	601	274	279	13086
023A09	10356	897	1239	798	2041	1946	799	17645
073A09	5228	1067	1183	1194	2169	3836	3841	24038
069A03	8764	596	915	1403	1035	1764	1564	23446
066A06a	13568	929	837	981	1176	2537	2542	34117
065A09	14665	1123	1094	764	1449	2103	911	19465
053A06	12356	946	1211	849	1365	2317	1094	23157
Sci-img-01	13091	1475	1121	1181	1126	881	886	20495
Sci-img-02	8091	714	630	735	642	403	408	12142
Sci-img-03	7264	613	846	943	1863	2031	1346	26491
Sci-img-04	6497	678	769	1097	1486	1946	1294	19485
Sci-img-05	14763	784	609	665	742	451	456	20094
mor-img-01	15992	865	691	813	933	673	678	21366
mor-img-02	10398	1030	829	1013	1062	1039	1044	17735
mor-img-03	5298	660	740	1035	1347	2394	2399	18580
mor-img-04	6109	592	554	618	736	1344	1349	32259
mor-img-05	10904	485	526	757	901	2153	2158	25795

Table VI  
Frequency incidence of a MP-g with 45° rotation on LDP for adult facial images.

Adult images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
003A49	10349	864	649	946	563	1786	812	15546
005A61	16542	1165	1021	2156	2564	2035	945	19468
006A61	15612	914	794	1949	1032	1739	849	23095
027A41	12654	1326	867	876	946	2349	723	16866
028A46	9332	1765	1637	2043	2516	3233	1477	23837
033A44	11751	1283	1415	1714	2157	3618	2918	29048
039A52	14687	1037	913	2300	2013	1649	846	23456
045A48	17491	840	703	1811	776	1094	553	24003
047A45	13456	1034	918	785	1945	3015	698	16489
048A50	18249	483	490	1624	690	1064	716	23524
071A42	12145	1348	765	648	1762	1034	714	23456
062A41	12295	1660	1637	1913	2284	3218	1652	27397
004A53	14627	775	741	1921	1112	2005	1268	23110
Sci-img-101	10346	584	596	663	703	1325	613	21354
Sci-img-102	19090	1059	908	1315	952	1138	424	26558
Sci-img-103	13464	1256	864	1550	1648	1766	765	20164
Sci-img-104	19684	1334	1059	1212	1080	958	271	27478
Sci-img-105	12062	581	571	1672	783	1562	697	18408
mor-img-501	13459	1360	972	1948	1246	2649	1034	18465
mor-img-502	14654	867	877	975	1645	2789	465	17563
mor-img-503	11687	1359	1068	1356	1182	1652	582	19738
mor-img-504	16286	2554	2217	2278	2400	2728	1418	31735
mor-img-505	14703	1186	1012	1296	1256	1668	774	23089

Table VII  
Frequency incidence of aMP-g with 135° rotation on LDP for child facial images.

Child images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
002A03	16908	857	758	940	889	1108	574	22910
008A13	14038	1642	1104	1144	1129	1426	595	22186
009A03	11282	1149	856	956	1061	1651	1056	19238
011A11	9605	665	550	613	582	962	327	14585
012A12	9022	757	651	693	644	788	369	13645
015A01	13230	1650	1475	700	2120	3777	2491	29253
016A08	14458	1012	832	948	1177	2483	2532	35607
019A07	6239	744	744	1014	1435	2865	2362	19941
022A11	15688	1292	894	873	900	1173	652	23040
023A09	16120	1098	1201	1175	2164	4652	3839	25542
073A09	6154	1259	1315	1058	2254	4164	2819	24337
069A03	8419	1259	1315	1058	2254	4164	2819	24337
066A06a	14332	1020	787	1017	1146	2058	1140	23021
065A09	7009	639	570	613	723	1543	1367	33761
053A06	11791	554	529	721	917	2219	2208	27286
Sci-img-01	10395	1037	594	894	1013	966	1345	23456
Sci-img-02	14656	1123	913	931	1216	2034	1649	22151
Sci-img-03	12349	978	795	1027	1324	2246	1864	19457
Sci-img-04	11256	981	981	1346	1525	3156	1447	16485
Sci-img-05	10356	1035	596	1294	1646	4023	1552	14863
mor-img-01	9468	1210	637	1246	2112	3947	1336	13498
mor-img-02	8864	971	779	1495	2001	3129	1649	32467
mor-img-03	9862	968	891	1221	1940	798	975	23622
mor-img-04	10354	811	730	1323	899	1039	869	19468
mor-img-05	11349	722	616	1424	905	1164	713	17644

Table VIII  
Frequency incidence of a MP-g with 135° rotation on LDP for adult facial images.

Adult images	MP-1	MP-2	MP-3	MP-4	MP-5	MP-6	MP-7	MP-8
003A49	13954	1178	849	846	916	2167	775	22156
005A61	10580	1353	1035	1273	1124	1672	599	18462
006A61	9467	1532	741	943	761	2011	1038	23141
027A41	8879	711	951	694	696	1036	1465	19468
028A46	14446	942	812	1933	982	1752	626	22165
033A44	9846	917	756	862	851	853	1094	20946
039A52	15052	2492	2354	2335	2356	2800	1036	30865
045A48	1465	1265	1094	946	813	2034	669	20134
047A45	16654	886	692	1839	759	1082	504	22529
048A50	10902	635	551	1649	763	1474	1013	17023
071A42	12649	937	963	864	945	1943	913	27895
062A41	11128	1824	1655	1946	2186	3201	1711	26079
004A53	10317	943	753	953	762	1864	1311	23498
Sci-img-101	10686	1405	1425	1771	2017	3750	2917	27607
Sci-img-102	17147	552	488	1639	633	1070	745	22240
Sci-img-103	13495	1399	794	1067	1034	1846	844	23145
Sci-img-104	18619	1361	1052	1282	990	1005	240	26201
Sci-img-105	10346	1434	852	824	848	1765	966	26499
mor-img-501	17969	1154	873	1227	935	1182	399	25279
mor-img-502	8649	1034	355	761	515	1946	1132	21346
mor-img-503	13542	1273	1012	1317	1223	1595	600	22196
mor-img-504	13500	808	711	1991	1031	2035	948	22209
mor-img-505	8233	1795	1686	2114	2305	3420	1414	22547

### 3.1 Comparison of Results with othersProposedmethods results:

The efficiency of the present method is compared with others proposed methods like“geometric properties” approach [14] and SPBPLME[15] methods. Geometric properties approach

extracts the geometric facture of Wavelet facial image. The overall percentage of this method is 91.88%.The proposed SPBPLME method estimated incidence rates for zero transitions on prominent binary patterns on SBPLME. The overall proportion of classification rate of the SPBPLME is

about 96.13. The proportion of classification in each group of the proposed method and others methods are listed out in Table 9. The graphical representation of the classification results is shown in figure9. The Table 9 clearly indicates that the proposed method yields better performance rate when compared with the existing methods.

Table IX

Overall % of classification rates of the proposed method and others methods.

Image Dataset	Geometric properties Approach	SPBPLME method	Proposed Method
FgNet	91.34%	96.37%	97.73%
Morph	90.79%	95.46%	96.96%
Scanned	89.15%	96.13%	98.15%
Google	91.28%	96.67%	97.97%

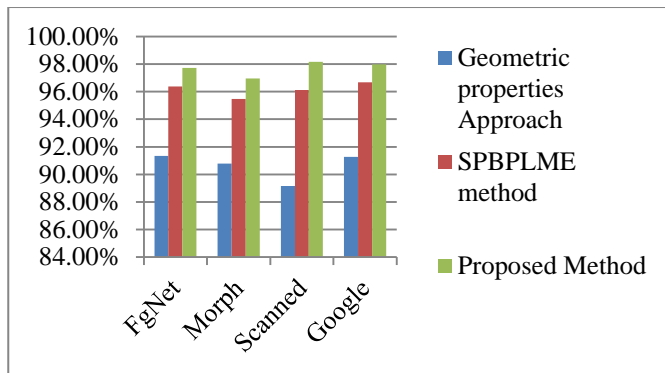


Fig. 9. Comparison graph of proposed method with others methods

#### 4. CONCLUSIONS

The present paper proposed a novel method for age classification of child and adults based on the MP-gSon LDP facial skin. The novelty of the present method is, it is rotationally, pose, noise, illumination invariant due to basic principles of LDP and the proposed MP-g. The present approach outlines that one need not necessarily evaluate the frequency of incidence of MP-2g, MP-3g and MP-5g to MP-8g for the age classification. The MP-1g and MP-4g contains more textural and topological information of the facial skin, that is the reason these two texture features are classifying the child and adult.

#### REFERENCE

- [1] Turk M. and Pentland A. "Eigenfaces for recognition," *Int. J. Cognitive Neurosci.* 3(1) (1991) 71–86.
- [2] Turk M. and Pentland A. "Pattern Recognition Learning and Thought," pp.486, Prentice Hall, New Jersey, 1973.
- [3] Belhumer P. N., Hespanha J.P. and Kriegman D.J. "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.* 19 (7) (1997) 711–720.
- [4] Yang J. and Yang J.Y. "Why can LDA be performed in PCA transformed space," *Pattern Recognition* 36 (2) (2003) 563–566.
- [5] Song F., Zhang D., Mei D. and Guo Z., "A multiple maximum scatter difference discriminant criterion for facial feature extraction," *IEEE Trans. Syst. Man Cybern. Part B Cybern.* 37 (6) (2007) 1599–1606.
- [6] Zhang S. and Sim T., "Discriminant subspace analysis: a Fukunaga-Koontz approach," *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (10) (2007) 1732–1745
- [7] He X., Yan S., Hu Y., Niyogi P. and Zhang H., "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (3) (2005) 328–340.
- [8] T Ojala, M Pietikainen, and D Harwood published a paper entitled "A Comparative Study of Texture Measures with Classification Based on Feature Distributions" in the *Journal of Pattern Recognition*, Year: 1996 volume: no : 29 Pages 51 to 59.
- [9] M Chandra Mohan, V Vijaya Kumar and B Sujatha published a paper entitled "Classification of child and adult based on geometric features of face using linear wavelets," in the *Journal of Signal and Image Processing*, volume no: 1 and issue no:3, year : 2010 Pages:211 to 220.
- [10] "Texton Based Shape Features on Local Binary Pattern for Age Classification", Chandra Sekhar Reddy et al.; *IJIGSP* 2012, 7, 54-60
- [11] Novel method of adult age classification using linear wavelet transforms", Chandra Mohan, VijayaKumar V., Venkata Krishna V., *IJCSNS*, pp:1-8, 2010.
- [12] Abdel-Mottaleb M. and Elgammal A., "Face Detection in Complex Environments from Color Images," *IEEE International Conference Image Processing*, pp. 622-626, 1999.
- [13] Pantic M. and Rothkrantz L.J.M. "Automatic analysis of facial expressions: the state of the art," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12) (2000) 1424-1445
- [14] "Classification of child and adult based on geometric features of face using linear wavelets", Chandra Mohan M., Vijaya Kumar V., Sujatha B., *IJSIP*, vol.1, Iss.3, pp:211-220, 2010.
- [15] "Novel Approach for Child and Adulthood Classification based on Significant Prominent Binary Patterns of Local Maximum Edge (SPBPLME)" Rajendra Babu .Ch, DrSreenivasa Reddy. E and DrPrabhakara Rao. B, *IJ. Information Technology and Computer Science*, 2015, 06, 30-37

# Novel Technique for the Handwritten Digit Image Features Extraction for Recognition

<sup>1</sup>Dr. U Ravi Babu and <sup>2</sup>Dr. G Charles babu

<sup>1</sup> Professor, Dept. of CSE, MREC(A), TS, India, uppu.ravibabu@mrec.ac.in

<sup>2</sup> Professor, Dept. of CSE, MREC(A), TS, India, charlesbabu26@gmail.com

**Abstract**—This paper proposes a novel approach for handwritten digit recognition system. The present paper extracts digit image features based on distance measure and derives an algorithm to classify the digit images. The distance measure can be performed on the thinned image. Thinning is the one of the preprocessing technique in image processing. The present paper mainly concentrated on an extraction of features from digit image for effective recognition of the numeral. The main objective of the paper is exact features are extracted for good classification results. To find the effectiveness of the proposed method, tested on MNIST database, CENPARMI, CEDAR, and newly collected data. The proposed method is implemented on more than one lakh digit images and got good comparative recognition results. The percentage of the recognition is achieved about 97.32%.

**Index Term**—Handwritten digit recognition, Distance measure, MNIST Database, image features.

## I. INTRODUCTION

In the field of image processing and pattern recognition [1] handwriting recognition has one of the hottest and demanding directions in the recent years. Day to day new technologies and innovative methods have been proposed continuously. With the development of the smart phone operation system, the application in handwritten recognition has aroused more and more attention from researchers. In general, handwritten character recognition is classified into two types of offline and online recognition methods [2]. An optical character recognition (OCR) system with a good recognition performance needs to maintain a very high recognition rate, and at the same time, to obtain a very high reliability, or a very low error rate [3, 4]. Recent developments on classifiers and feature extraction have significantly increased the recognition accuracy of handwritten digit recognition systems.

Gorgevik et al. [5] proposed handwritten digits recognition system by using Support Vector Machines (SVM). They extracted four types of features from each digit image 1) projection histograms, 2) contour profiles, 3) ring-zones and 4) Kirsch features. They reported 97.27% recognition accuracy on National Institute of Standards and Technology (NIST) [6] handwritten digits database [7] when four types of features were used collectively. In [8] Chen et al. proposed maximum posterior pseudo-probabilities framework for handwritten digits recognition. They extracted 256-D directional features from the input image. Finally, 256-D features were transformed into 128-D feature using Principal Component

Analysis (PCA). They reported recognition accuracy of 98.76% on NIST database [7].

A few works have also been reported in the literature on handwritten digit recognition of Indian scripts. Most of the available works on handwriting recognition of Indian scripts are based on small databases with a few basic characters and collected in laboratory environments. At present, a few large databases of handwritten digits of major Indic scripts are available from Computer Vision and Pattern Recognition Unit (CVPR), Indian Statistical Institute, Kolkata. Offline handwritten numerals database is available for Devanagari, Bangla, and Oriya scripts. Basic characters, vowel modifiers and compound characters about Devanagari, Bangla, and Oriya scripts are also available. Both online and offline forms of the dataset are available for research purpose. An offline version of the data is also available in the form of bi-level TIFF images, generated from the online data using simple piecewise linear interpolation with a constant thickening factor applied. The paper considers only the Basic English numerals for recognition purpose. The dataset for the experiment was collected from different individuals of various professions in the states of Andhra Pradesh and Telangana. The proposed method overcome the disadvantages of various techniques and find the effective feature extraction process for the recognition handwritten digits of different databases in an effective manner.

The paper is organized as follows: Section 2 contains database creation and the preprocessing. Section 3 describes feature extraction methods and the proposed algorithm is presented in Section 4. The experimental details and result analysis is presented in Section 6. Section 6 contains the conclusion part.

## II. PROPOSED METHOD

The proposed method is mainly consists of 5 steps. In the first step, collecting the numerals data from various data bases and gathering images from various people in AP and Telangana state. After collecting the numeral data preprocess data i.e. elimination of noise and conversion of gray scale images into binary images and also the normalization of the binary images by using the normalization techniques. In the third step, binary images are converted into single pixel-width images i.e. thinning of the digit images. In this step, various thinning operations are performed and found out the best thinning techniques suitable for the digit images. In the fourth step, extracted the features of the thinned image by using distance measure. In the last step, derive an algorithm for recognition of handwritten numerals system. The block

diagram of the proposed method is shown below figure 1.

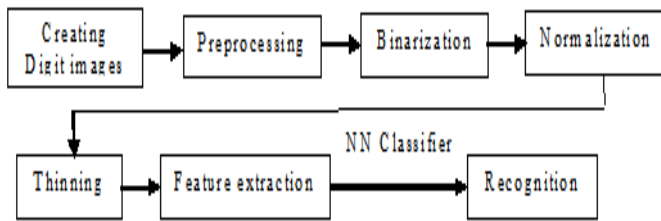


Fig. 1. Block diagram of the Handwritten Digit Recognition system

### Step 1: Collection of numeral image database

Several standard datasets of digits are found in English. Some of them are CENPARMI, CEDAR, and MNIST datasets. The CENPARMI (Centre for Pattern Recognition and Machine Intelligence) digit dataset [9] is available from CENPARMI, Concordia University. In this dataset 4000 images (400 samples per class) are specified for training purpose and 2000 images are used for testing purpose. These digit datasets were collected from United States Postal Service (USPS). The Center of Excellence for Document Analysis and Recognition (CEDAR) digit dataset is available from CEDAR, The State University of New York, Buffalo. The training and test sets contain 18468 and 2711 digits, respectively. The number of samples in both training and test sets differ for each class. The Modified National Institute of Standards and Technology (MNIST) dataset [10] was extracted from the NIST datasets SD3 and SD7. The training and test sets are composed of both SD3 and SD7. Samples are normalized into 20×20 grayscale images with aspect ratio reserved, and the normalized images are located in a 28×28 frame. The number of training and test samples is 60,000 and 10,000 respectively. The sample images of the MNIST dataset is shown in figure 2.

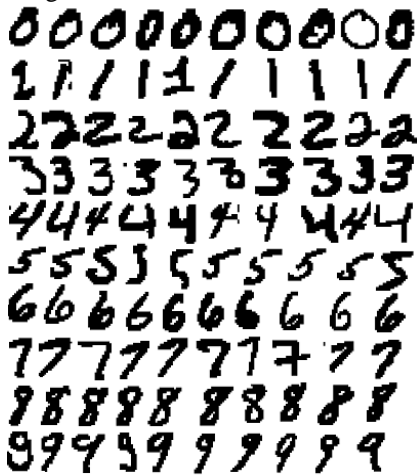


Fig. 2. Sample digit images of MNIST database

The plain paper was used for data collection. Each person was instructed to write the digits (fully unconstrained) along the vertical direction. The dataset contains about 100 isolated samples each of 10 numerals written by 1000 native writers including university graduates, high school children, and adults. Around fifty percent data is from high school children. A flatbed scanner was used for digitization, with images in gray tone at 300 dpi. These were stored as Bit Map File (BMP) format using a standard technique for converting them

into monochrome images. Data was manually extracted from scanned images and normalized into 50×50 size using a standard bicubic approach. After processing scanned images about digits and a total of 100000 (100×1000) images of numerals are obtained. Dataset developed planned to be made available publicly for research purpose. Some of the sample images after extracting from the scanned image are shown in figure 3.



Fig. 3. Sample Scanned document of digit images

### Step 2: Digit Image Preprocessing

Data capture of documents by optical scanning or by digital video yields a file of picture elements, which is the raw input to document analysis process. The first step in document analysis is to perform a preprocessing on this image to prepare it for further analysis. Such processing includes thresholding to convert a gray scale (or color image) to a binary image, reduction of noise to reduce extraneous data, skew estimation of a document image if document suffers from tilt (skewed), thinning, enable subsequent detection of pertinent features of the object of interest and then segmentation of text line to individual digit character. It is assumed that characters are already isolated and preprocessing steps such as Binarization, noise removal, normalization and thinning are to be done. Initially, the document is captured using gray level mode. Character extraction from the scanned document is done manually. Then Otsu's histogram-based global threshold approach is applied for digit image Binarization. Figure 4 shows the gray level picture of the character and resultant of global thresholding approach.



Fig. 4. Result of Binarization Technique

The next step is to normalize the given character image into a standard size. For that, a standard nearest neighbor interpolation method is used. Figure 5 shows some sample images after applying normalization technique.



Fig. 5. Samples of Digit Images after Normalization

### Step 3: Thinning of the digit image

Skeletonization has been a part of image processing for a wide variety of applications. Thinning is the process of reducing the thickness of each line of pattern to just a single pixel width. Thinning is usually used as the first step in applications such as optical character recognition to improve the recognition rate. It has been applied to many fields [11], such as inspection of printed circuit boards, counting of asbestos fibers on air filters, analysis of chromosomes, classification of fingerprints, recognition of characters, application in intelligent copying and facsimile transmission systems, data reduction of map storage, etc. The major advantages of thinning in image processing and pattern recognition are: reduction of an amount of data as input binary image and preservation of fundamental skeleton, which is topologically equivalent to the original object. Many thinning algorithms have been introduced for decades. These algorithms have both advantages and disadvantages. Some algorithms can obtain good quality skeletons, but they run slowly [12]. None of the methods address rotation invariant is thinning. Many are specific to digits, characters, or letters, written in English, Chinese, Arabic, or any other scripts. However, to solve the aforementioned problems, the rotation invariant is rule-based thinning algorithm for character recognition is proposed by [13]. This is a generalized algorithm, which is used to thin the symbols irrespective of the scripts. The advantage of this method is it is invariant to rotation. The algorithm fails on two-pixel wide lines. To overcome this drawback, an improved rotation-invariant thinning algorithm was proposed by [14]. Recently, improved two-step thinning algorithm is described by Aradhya et.al in the year 2005, which is an improved version of standard two-step thinning algorithm [15]. The proposed algorithm thins the symbols to their central lines and has the rotation invariant property. Compared to two-step and rotation rule-based methods, the improved two-step method is better with respect to time, computations, and preserving topology. Hence in this work, improved two-step thinning algorithm is used for better representation of character images and to extract features from the character images. The resultant of thinning algorithms is shown in figure 6.

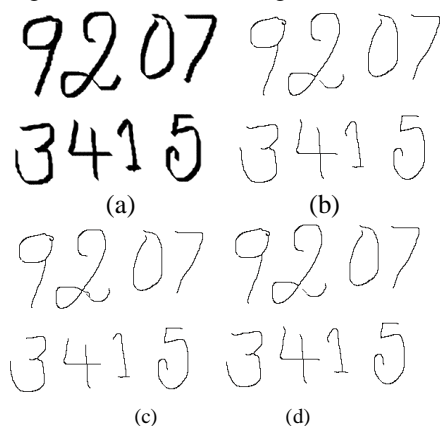


Fig. 6. Result of different thinning algorithms

### III. A NOVEL FEATURE EXTRACTION TECHNIQUE BASED ON DISTANCE MEASURE FOR HANDWRITTEN DIGITS

Digit image preprocessing has also been done for effective representation of characters. In this paper, a novel feature extraction technique based on distance measure is presented. In this proposed method features are extracted based on distance measure technique. Feature extraction is one of the most important steps in any recognition system. The aim of feature extraction is to represent the image in terms of some quantifiable measurements that may be easily utilized in the classification stage. To a large extent, the accuracy of a recognition system depends on the discriminative capability of features and the generalization performance of a designed classifier. Document image representation methods are categorized in three major groups [16].

#### A. Global Transformation and Series Expansion:

Geometrical and topological features can represent various global and local properties of numerals images with high tolerance to distortions and style variations. The methods on topological and geometrical representations can be grouped into four categories i.e. Coding, Graphs & Trees, Counting & Extracting Topological Structures and Measuring & Approximating the Geometrical Properties.

#### B. Statistical Representation:

Representation by the use of statistical distribution of points will take care of the style variations to some extent. There are three important statistical features used for character representation: Zoning, Crossing and Distances and Projections

#### C. Geometrical and Topological Representation:

A continuous signal generally contains more information that needed to be represented for the purpose of classification. One way to represent a signal is by a linear combination of a series of simpler well-defined functions. The coefficients of the linear combination provide a compact encoding known as transformation or series expansion. Some common transform methods used in the character recognition field are Fourier transform, Gabor transform, Moments and Karhunen-Loeve transform.

In this paper, novel feature extractions schemes are developed i.e. Feature Extraction Method (FEM) based on distance measure are presented.

FEM mainly consists of X major steps

Step 1: Thinning is applied to the given input digit image. In this method two-step & improved two-step thinning algorithms are used for thinning purpose. The resultant thinned images for a given character are shown in Figure 7.

Step 2: The thinned digit image is scanned from top-left corner to top-right corner (row-wise). While scanning the image, the black pixel positions are identified in each row by using the following procedure.

(a) If any black pixel is encountered in this process, the algorithm computes the distance between the

points of the region where the black pixel is encountered.

(b) If the algorithm fails to encounter any black pixel in the row, the distance is considered as zero value.

This procedure of finding black pixel in the specified region is continued till the end of the row. Figure 8 shows the working procedure of the proposed method. If the size of the image is  $50 \times 50$ , then 50 feature values along each row is encountered.

Step 3: Similarly distances are calculated column wise i.e. from the topmost pixel to the bottommost pixel of every column. Figure 9 shows the working procedure of the proposed method. If the size of the image is  $50 \times 50$ , then 50 feature values along each row is encountered.

Totally 100 (i.e.,  $50 + 50$ ) features are extracted from each input character.

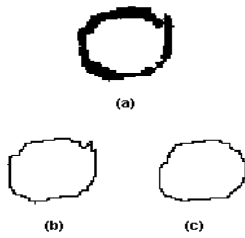


Fig. 7. Thinning Results (a) Original Image (b) 2 Step result (c) Result Using Improved 2 Step

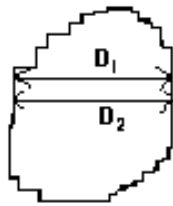


Fig. 8. Distance Computing along Row-wise

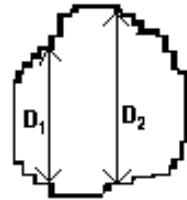


Fig. 9. Distance Computing along Column-wise

The major goal of representation is to extract and select a set of features, which maximizes the recognition rate with the least amount of elements. Feature extraction and selection is defined as extracting the most representative information from the raw data, which minimizes the intra class pattern variability while enhancing the inter-class pattern variability.

#### IV. RESULTS AND DISCUSSIONS

The proposed feature extraction methods are experimentally evaluated with the dataset containing various handwritten numerals collecting from MNIST data base, CENPARMI data base, CEDAR data base and most of the images from scanned images. Collectively, the data base contains 176000 digit

images. No method has tested using such type of large database. In this paper, Nearest Neighbor Classifier (NNC) is used for classification purpose. All experiments are carried out on a PC machine with i3 processor 2.7GHz CPU and 2 GB RAM memory under MatLab 10.0 platform. 20 percentage of the each data base is used for training and remaining 80 percentage images are used for testing purpose i.e. 35200 images are used for training purpose and 140800 images are used for testing purpose.

The proposed method is evaluated when two different thinning algorithms are used for thinning a digit image i.e. two-step and improved two step thinning algorithms. The percentage of recognition of the proposed method when two-step method is applied is listed out in tables 1, 2, 3 and 4. The percentage of recognition of the proposed method when improved two-step method is applied is listed out in table 5, 6, 7 and 8. The overall percentage of the proposed recognition system when two thinning technique are applied is shown in a table and the comparison graph of the two thinning algorithms is shown in figure 10. From the figure 10, it is clear that, using two-steps thinning algorithm the overall recognition rate is 94.25 achieved whereas in case of improved thinning algorithm the performance of the system is 96.28%.

TABLE I  
MNIST DATA BASE RECOGNITION PERCENTAGE WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	6573	6323	250	96.20
1	6715	6464	251	96.26
2	6580	6308	272	95.87
3	6600	6302	298	95.48
4	6442	6172	270	95.81
5	6575	6365	210	96.81
6	6705	6287	418	93.77
7	6715	6465	250	96.28
8	6605	6262	343	94.81
9	6490	6304	186	97.13

TABLE II  
CEPARMI DATA BASE RECOGNITION PERCENTAGE WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	320	308	12	96.25
1	320	307	13	95.94
2	320	308	12	96.25
3	320	303	17	94.69
4	320	306	14	95.63
5	320	305	15	95.31
6	320	309	11	96.56
7	320	303	17	94.69
8	320	307	13	95.94
9	320	310	10	96.88

TABLE III

CEDAR DATABASE PERCENTAGE OF RECOGNITION WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	160	150	10	93.75
1	160	149	11	93.13
2	160	150	10	93.75
3	160	151	9	94.38
4	160	152	8	95.00
5	160	155	5	96.88
6	160	151	9	94.38
7	160	153	7	95.63
8	160	154	6	96.25
9	160	156	4	97.50

TABLE IV

SCANNED IMAGES DATABASE PERCENTAGE OF RECOGNITION WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	7985	7803	182	97.72
1	8025	7733	292	96.36
2	8123	7800	323	96.02
3	7816	7579	237	96.97
4	7923	7738	185	97.67
5	8050	7711	339	95.79
6	8023	7711	312	96.11
7	8075	7703	372	95.39
8	8115	7731	384	95.27
9	7865	7441	424	94.61

TABLE V

MNIST DATA BASE RECOGNITION PERCENTAGE WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	6573	6503	70	98.94
1	6715	6490	225	96.65
2	6580	6509	71	98.92
3	6600	6423	177	97.32
4	6442	6299	143	97.78
5	6575	6406	169	97.43
6	6705	6513	192	97.14
7	6715	6560	155	97.69
8	6605	6491	114	98.27
9	6490	6324	166	97.44

TABLE VI

CEPARMI DATA BASE RECOGNITION PERCENTAGE WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	320	312	8	97.50
1	320	315	5	98.44
2	320	305	15	95.31
3	320	307	13	95.94
4	320	312	8	97.50
5	320	314	6	98.13
6	320	312	8	97.50
7	320	313	7	97.81
8	320	316	4	98.75
9	320	315	5	98.44

TABLE VII

CEDAR DATABASE PERCENTAGE OF RECOGNITION WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	160	156	4	97.50
1	160	157	3	98.13
2	160	155	5	96.88
3	160	155	5	96.88
4	160	156	4	97.50
5	160	153	7	95.63
6	160	152	8	95.00
7	160	155	5	96.88
8	160	153	7	95.63
9	160	157	3	98.13

TABLE VIII

SCANNED IMAGES DATABASE PERCENTAGE OF RECOGNITION WHEN IMPROVED 2-STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	7985	7833	152	98.10
1	8025	7811	214	97.33
2	8123	7837	286	96.48
3	7816	7632	184	97.65
4	7923	7803	120	98.49
5	8050	7737	313	96.11
6	8023	7820	203	97.47
7	8075	7819	256	96.83
8	8115	7810	305	96.24
9	7865	7641	224	97.15

TABLE IX

%GE OF RECOGNITION OF THE PROPOSED METHOD WHEN TWO THINNING APPROACHES ARE APPLIED

Database	%ge of Recognition	
	When Two-step thinning	When improved two-step thinning
CEPARMI	95.84	97.76
CEDAR	95.81	97.53
MNIST	95.06	96.81
Scanned Images	96.19	97.18

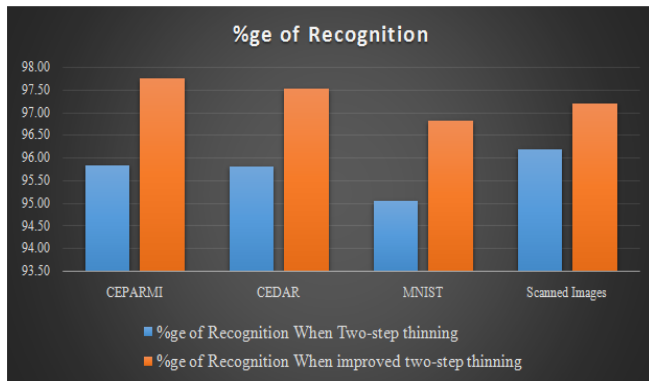


Fig. 10. Comparing Recognition Accuracy with two Different Thinning Algorithms

From the figure 10 and table 9, it is clear that improved two step thinning algorithm has performed better compared to standard two step thinning algorithm. The proposed method seeks more structural features and the size of the feature vector is also small. The overall recognition rate of the considered digit database is shown in table 10 and the overall recognition percentage is 97.44.

TABLE X  
CONSIDERED DATABASE PERCENTAGE OF RECOGNITION WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	15038	14804	234	98.44
1	15220	14773	447	97.06
2	15183	14806	377	97.52
3	14896	14517	379	97.46
4	14845	14570	275	98.15
5	15105	14610	495	96.72
6	15208	14797	411	97.30
7	15270	14847	423	97.23
8	15200	14770	430	97.17
9	14835	14437	398	97.32

#### 4.1 Analysis of the proposed method:

To analyze the proficiency of the proposed system, the outcomes of the proposed method are analyzed in gotten Ten Cycle Cross Validation (TCCV) approach.

**Ten Cycle Cross Validation (TCCV)** approach: In TCCV approach results analysis strategy, the entire digit data base i.e. 150800 are divided into 10 sets. Each set consists of 15080 digit images mix up with the images of CEPRM, MNIST, CEDAR images, and Scanned Images. Every set must contain ten classes (0-9) of digit images. In TCCV approach results are analyzed in 10 Cycles. In cycle 1, first set is treated as a sample set and remaining 9 sets are taken as a test dataset. Compute the % of image grouping for test set. In cycle 2, second set is dealt with as a sample set and staying nine sets are taken as a test set. Compute the % of grouping for test set. The same methodology is applied remaining cycles also. This approach Strengthens the proposed method. The % of digit grouping of the proposed strategy in ten cycles are listed in tables XI to XX individually.

TABLE XI  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-1 OF TCCV APPROACH

Digit	Cycle-1			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13389	145	98.93
1	13698	13467	231	98.31
2	13665	13486	179	98.69
3	13406	13167	239	98.22
4	13361	13205	156	98.83
5	13595	13382	213	98.43
6	13687	13413	274	98.00
7	13743	13444	299	97.82
8	13680	13386	294	97.85
9	13352	13137	215	98.39

TABLE XII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-2 OF TCCV APPROACH

Digit	Cycle-2			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13391	143	98.94
1	13698	13485	213	98.45
2	13665	13468	197	98.56
3	13406	13187	219	98.37
4	13361	13121	240	98.20
5	13595	13419	176	98.71
6	13687	13559	128	99.06
7	13743	13544	199	98.55
8	13680	13467	213	98.44
9	13352	13177	175	98.69

TABLE XIII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-3 OF TCCV APPROACH

Digit	Cycle-3			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13370	164	98.79
1	13698	13501	197	98.56
2	13665	13468	197	98.56
3	13406	13237	169	98.74
4	13361	13215	146	98.91
5	13595	13412	183	98.65
6	13687	13473	214	98.44
7	13743	13584	159	98.84
8	13680	13462	218	98.41
9	13352	13176	176	98.68

TABLE XIV  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-4 OF TCCV  
APPROACH

Digit	Cycle-4			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13421	113	99.17
1	13698	13520	178	98.70
2	13665	13501	164	98.80
3	13406	13191	215	98.40
4	13361	13085	276	97.93
5	13595	13449	146	98.93
6	13687	13472	215	98.43
7	13743	13567	176	98.72
8	13680	13416	264	98.07
9	13352	13206	146	98.91

TABLE XVII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-7 OF TCCV  
APPROACH

Digit	Cycle-7			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13415	119	99.12
1	13698	13497	201	98.53
2	13665	13501	164	98.80
3	13406	13283	123	99.08
4	13361	13223	138	98.97
5	13595	13399	196	98.56
6	13687	13469	218	98.41
7	13743	13550	193	98.60
8	13680	13493	187	98.63
9	13352	13218	134	99.00

TABLE XV  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-5 OF TCCV  
APPROACH

Digit	Cycle-5			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13320	214	98.42
1	13698	13504	194	98.58
2	13665	13501	164	98.80
3	13406	13209	197	98.53
4	13361	13215	146	98.91
5	13595	13456	139	98.98
6	13687	13473	214	98.44
7	13743	13490	253	98.16
8	13680	13531	149	98.91
9	13352	13215	137	98.97

TABLE XVIII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-8 OF TCCV  
APPROACH

Digit	Cycle-8			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13420	114	99.16
1	13698	13485	213	98.45
2	13665	13497	168	98.77
3	13406	13190	216	98.39
4	13361	13224	137	98.97
5	13595	13431	164	98.79
6	13687	13552	135	99.01
7	13743	13532	211	98.46
8	13680	13537	143	98.95
9	13352	13198	154	98.85

TABLE XVI  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-6 OF TCCV  
APPROACH

Digit	Cycle-6			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13389	145	98.93
1	13698	13485	213	98.45
2	13665	13489	176	98.71
3	13406	13272	134	99.00
4	13361	13223	138	98.97
5	13595	13436	159	98.83
6	13687	13530	157	98.85
7	13743	13484	259	98.12
8	13680	13472	208	98.48
9	13352	13186	166	98.76

TABLE XIX  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-9 OF TCCV  
APPROACH

Digit	Cycle-9			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13321	213	98.43
1	13698	13522	176	98.72
2	13665	13473	192	98.59
3	13406	13239	167	98.75
4	13361	13222	139	98.96
5	13595	13408	187	98.62
6	13687	13538	149	98.91
7	13743	13606	137	99.00
8	13680	13541	139	98.98
9	13352	13155	197	98.52

TABLE XX  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-1 OF TCCV  
APPROACH

Digit	Cycle-10			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13343	191	98.59
1	13698	13468	230	98.32
2	13665	13489	176	98.71
3	13406	13171	235	98.25
4	13361	13212	149	98.88
5	13595	13391	204	98.50
6	13687	13471	216	98.42
7	13743	13474	269	98.04
8	13680	13431	249	98.18
9	13352	13101	251	98.12

### Comparison of the proposed method with other existing methods:

The efficiency of the proposed method is compared to other existing methods like twin minimax probability machine (TWMPM) proposed by Zhijie et.al [17], transformation based features proposed by Syed et.al [18], Back Propagation with Neural Network approach[19] and selection, reproduction, mutation and crossover methods with Genetic Algorithm (GA) proposed by Devikanniga et.al[20]. The TWMPM method generates two hyper-planes of the digit images and it also avoids making distributional assumptions about the class-conditional densities of the digit. The performance of the TWMPM method is evaluated on two data sets only i.e UTC and MNIST and the overall percentage of the proposed method 88.07%. The method proposed by the Syed utilizes the Discrete Cosine Transform (2D-DCT) for feature extraction and Hidden Markov models (HMMs) used for classification. The syed proposed method is applied on only MNIST database and got 95.95% if feature vector size is more. More feature vector causes more computational cost. Sakshica proposed to classify the handwritten digits by using the features and their spatial relationship in the pattern with Hopfield Neural Network. A small number of images are tested by using this method and got 90.23% only. Devikanniga proposed a method to classify the handwritten digit using GA and got the overall performance is only 87%. The performance evolution of the proposed method with other existing methods is listed out in table 21 and the classification graph is represented in figure 11. From table 10 and figure 11, it is clearly evident that, the proposed method exhibits a high recognition rate than the existing methods.

TABLE XXI  
OVERALL %GE OF THE DIFFERENT RECOGNITION SYSTEMS

Data Base	TWMP M method [17]	2d-DCT with HMM Approac h [18]	NN with Back propagati on [19]	different NN Approach es [20]	propose d method
CEPARMI	87.18	95.28	85	89.23	95.11
CEDAR	91.23	96.45	89	90.24	95.43
MNIST	88.73	95.92	91	91.27	97.63
Scanned Images	85.15	96.18	84	90.87	96.95

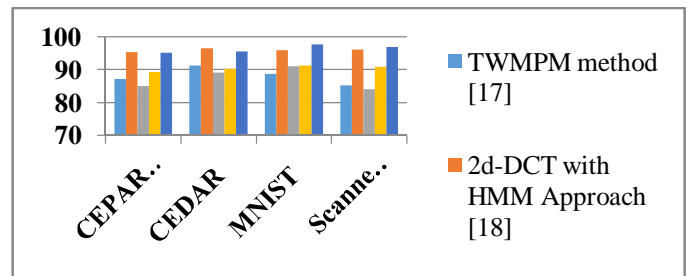


Fig. 11. Graphical representation of the % recognition of the proposed method and other existing methods

### V.CONCLUSION

In this paper a very large database of unconstrained handwritten digit images for experiment purpose was introduced. No author tested with such a large data base. It is proposed to make this database available for research purpose. 100000 handwritten characters collected from 1000 different individuals of different age groups and different areas. No author has attempted to such collection. The proposed method utilizes a small algorithms for preprocessing such as Binarization, normalization and thinning algorithms. The proposed method takes very less amount processor time for recognition of digit image. The main object of the proposed method is an efficient feature extraction method is derived for handwritten digit recognition. For thinning digit image, a well-known technique is used i.e. improved two-step thinning method. The overall percentage of the proposed method achieves 97.36%. The proposed method is evaluated by Ten Cycle Cross Validation (TCCV) approach to find the effectiveness of the proposed method No such method is available up to now utilizing the large set data base for testing.

### REFERENCES

- [1] J. pradeep, E. Strinivasan, and S. Himavathi, Neural network based handwritten character recognition system with feature extraction. International conference on computer, communication and electrical technology- ICCET 2011, 18th & 19th Mar. 2011.
- [2] W. Wu and Y. Bao, Online handwritten magnolia words recognition based on multiple classifiers, 2009.
- [3] C. L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, Pattern Recognition, vol. 37, iss. 2, pp. 265-279, 2004
- [4] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai, and L. Lam, Computer recognition of unconstrained handwritten numerals, Proc. IEEE, vol. 80, iss. 7, pp. 1162-1180, 1992.
- [5] D. Gorgevik and D. Cakmakov, "Handwritten Digit Recognition by Combining SVM Classifiers," in The International Conference on Computer as a Tool (EUROCON), 2005.
- [6] M. D. Garris, J. L. Blue and G. T. Candela, "NIST form-based handprint recognition system," NIST, 1997.
- [7] P. J. Grother, "NIST special database 19 hand printed forms and characters database," National Institute of Standards and Technology, 1995.
- [8] X. Chen, X. Liu and Y. Jia, "Learning Handwritten Digit Recognition by the Max-Min Posterior Pseudo-Probabilities Method," in Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), 2007.
- [9] Xu, L., A. Krzyzak and C.Y. Suen (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. IEEE Trans. Syst. Man Cybernet. 22 (3), 418-435.

- [10] Y. LeCun and Y. Bengio. Convolutional networks for images, speech, and time-series. In M. A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1995
- [11] B T Chen, Y S Chen and WH Hsu, "A Parameterized Fuzzy Processor and its Applications," *Fuzzy sets and systems*, Vol 59, pp. 149-172, 1993
- [12] Sehultze-Lam S, Harauz G, Beveridge TJ, "Participation of a cyanobacterial S-layer in fine grain mineral formation", 1992
- [13] Maher Ahmed, Rabab Ward, "A Rotation Invariant Rule-Based Thinning Algorithm for Character Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 24 Issue 12, December 2002, Page 1672-1678
- [14] Peter I. Rockett, 2005, An Improved Rotation- Invariant Thinning Algorithm, *IEEE, Trans. Patt.Anal.MachineIntll.*, Oct, Vol. 27.No.10. pp. 1671-1674.
- [15] Gonzalez R.C and Woods R. E, 2002 "Digital Image Processing text book" 2<sup>nd</sup> edition 2002
- [16] Nafiz Arica and Fatos T. Yarman-Vural, "An Overview of Character Recognition Focused on Off-Line Handwriting", *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, Vol. 31, No. 2, May 2001
- [17] Zhijie Xu, Jianqin Zhang and Hengyou Wang, "Twin Minimax Probability Machine for Handwritten Digit Recognition", *International Journal of Hybrid Information Technology* Vol.8, No.2 (2015), pp.31-40
- [18] Syed Salman Ali, Muhammad Usman Ghani "Handwritten Digit Recognition using DCT and HMMs", *CONFERENCE PAPER* · JANUARY 2014
- [19] Sakshica, Dr. Kusum Gupta, "Handwritten Digit Recognition using various Neural Network Approaches", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 2, February 2015
- [20] Ms. D. Devikanniga, Dr. M. Thangamani, "Handwritten Digit Recognition using Genetic Algorithm", *International Journal of Innovations & Advancement in Computer Science*, Volume 4, Special Issue May 2015

## New approaches for the Features Extraction on Handwritten Digit Image for Recognition

<sup>1</sup>Dr. Pullela S V V S R Kumar, <sup>2</sup> Ms. Appana Naga Lakshmi, <sup>3</sup>Dr, U ravi Babu

<sup>1</sup> Professor of CSE, Sri Aditya Engineering College, Surampalem, pullelark@yahoo.com

<sup>2</sup> Computer Science & Engineering, Sri Aditya Engineering College,  
nagalakshmi08506@gmail.com

<sup>3</sup>. Professor, Dept. of CSE, MREC(A), India, uppu.ravibabu@mrec.ac.in

### Abstract

*The present paper proposes a novel approach for handwritten digit recognition system. The present paper extract digit image features based on distance measure and derive an algorithm to classify the digit images. The distance measure can be performed on the thinned image. Thinning is the one of the preprocessing technique in image processing. The present paper mainly concentrated on extraction of features from digit image for effective recognition of the numeral. The main goal of the proposed method is that for effective classification correct measure of feature extraction is needed. To find the effectiveness of the proposed method, tested on MNIST database, CENPARMI, CEDAR and newly collected data. The proposed method is implemented on more than one lakh digit images and get good comparative recognition results. The percentage of the recognition is achieved about 96.3%.*

**Keywords:** Handwritten Digit Recognition, Distance Measure, MNIST Database, Image Features

### 1. Introduction

In the field of image processing and pattern recognition [1] handwriting recognition has one of the hottest and demanding directions in the recent years. Day to day new technologies and innovative methods have been proposed continuously. With the development of the smart phone operation system, the application in handwritten recognition has aroused more and more attention from researchers. In general, handwritten character recognition is classified into two types of offline and online recognition methods [2]. An optical character recognition (OCR) system with a good recognition performance needs to maintain a very high recognition rate, and at the same time, to obtain a very high reliability, or a very low error rate [3, 4]. Recent developments on classifiers and feature extraction have significantly increased the recognition accuracy of handwritten digit recognition systems.

Support Vector Machines (SVM) based handwritten digits recognition system proposed by Gorgevik et al. [5]. They extracted four types of features from each digit image 1) projection histograms, 2) contour profiles, 3) ring-zones and 4) Kirsch features. They reported 97.27% recognition accuracy on National Institute of Standards and Technology (NIST) [6] handwritten digits database [7], when four types of features were used collectively. In [8] Chen et al. proposed maximum posterior pseudo-probabilities framework for handwritten digits recognition. They extracted 256-D directional features from the input image. Finally 256-D features were transformed into 128-D feature using Principal Component Analysis (PCA). They reported recognition accuracy of 98.76% on NIST database [7].

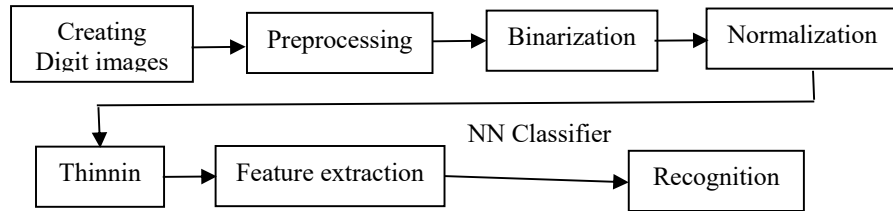
A few works have also been reported in the literature on handwritten digit recognition of Indian scripts. Most of the available works on handwriting recognition of Indian scripts are based on small databases with a few basic characters and collected in laboratory environments. At present, a few large databases of handwritten digits of major Indic scripts are available from Computer Vision and Pattern Recognition Unit (CVPR), Indian Statistical Institute, Kolkata. Offline handwritten numerals database is available for Devanagari, Bangla, and Oriya scripts. Basic characters, vowel modifiers and compound characters pertaining to Devanagari, Bangla, and Oriya scripts are also available. Both online and offline forms of dataset are available for research purpose. An offline version of the data is also available in the form of bi-level TIFF images, generated from the online data-using simple piecewise linear interpolation with a constant thickening factor applied. The paper considers only the Basic English numerals for recognition purpose. The dataset for the experiment was collected from different individuals of various professions in the states of Andhra Pradesh and Telangana.

The main objective of the proposed system is that when using the standard classification algorithms for classifying the digit images suitable features are extracts with low computational cost. Previous proposed approaches used the standard classification system for classifying the numerical images so much of time took for both extraction and also for classification system. In this proposed approach reduce the time consuming for classification system also reduces the feature extraction time. The main of the proposed method is that for developing a system for classification is only depends on the type of features extracted not on the classification system. If suitable features are extracted then less time should take classification algorithm for classifying digits into classes for large datasets. The proposed method overcome the disadvantages of the various techniques and find the effective feature extraction process for the recognition handwritten digits of different databases in effective manner.

The paper is organized as follows: Section 2 contains database creation and the preprocessing. Section 3 describes feature extraction methods and the proposed algorithm is presented in Section 4. The experimental details and result analysis are presented in Section 6. Section 6 contains the conclusion part.

## 2. Proposed Method

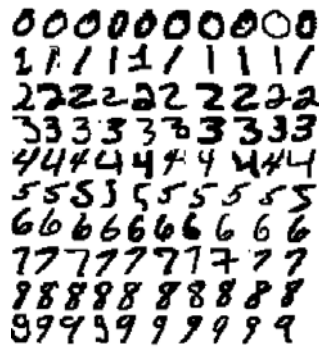
The propose method is mainly consists of 5 steps. In the first step is collecting the numerals data from various data bases and also gather the images from various people in AP and Telangana state. After collecting the numeral data preprocess data i.e. elimination of noise and conversion of grey scale images into binary images and also do the normalization of the binary images by using the normalization techniques. In the third step, binary images are converted into single pixel width images i.e. thinning of the digit images. In this step various thinning operations are performed and find out the best thinning techniques suitable for the digit images. In the fourth step, extract the features of the thinned image by using distance measure. In the last step, derive an algorithm for recognition of handwritten numerals system. The block diagram of the proposed method is shown below figure 1.



**Figure 1.** Block diagram of the Handwritten Digit Recognition system

### Step 1: Collection of numeral image database

Several standard datasets of digits are found in English. Some of them are CENPARMI, CEDAR, and MNIST datasets. The CENPARMI (Centre for Pattern Recognition and Machine Intelligence) digit dataset [9] is available from CENPARMI, Concordia University. In this dataset 4000 images (400 samples per class) are specified for training purpose and 2000 images are used for testing purpose. These digit data sets are collected from United States Postal Service (USPS). The Center of Excellence for Document Analysis and Recognition (CEDAR) digit dataset is available from CEDAR, The State university of New York, Buffalo. The training and test sets contain 18468 and 2711 digits, respectively. The number of samples in both training and test sets differ for each class. The Modified National Institute of Standards and Technology (MNIST) dataset [10] was extracted from the NIST datasets SD3 and SD7. The training and test sets are composed from both SD3 and SD7. Samples are normalized into 20×20 grayscale images with aspect ratio reserved, and the normalized images are located in a 28×28 frame. The number of training and test samples is 60,000 and 10,000 respectively. The sample images of the MNIST dataset is shown in figure 2.



**Figure 2.** Sample digit images of MNIST database

Plain paper was used for data collection. Each individual was instructed to write the digits (fully unconstrained) along vertical direction. The dataset contains approximately 100 isolated samples each of 10 numerals written by 1000 native writers including university graduates, high school children, and adults. Around fifty percent data are from high school children. A flatbed scanner was used for digitization, with images in gray tone at 300 dpi. They were stored as Bit Map File (BMP) format using standard technique for converting them into monochrome images. Data was manually extracted from scanned images and normalized into 50×50 size using standard bicubic approach. After processing scanned images pertaining to digits and a total of 100000 (100×1000) images of numerals are obtained. Dataset developed planned to be made available publicly for research purpose. Some of the sample images after extracting from the scanned image are shown in figure 3.



**Figure 3.** Sample Scanned document of digit images

## Step 2: Digit Image Preprocessing

Data capture of documents by optical scanning or by digital video yields a file of picture elements, which is the raw input to document analysis process. The first step in document analysis is to perform a preprocessing on this image to prepare it for further analysis. Such processing includes thresholding to convert a gray scale (or color image) to a binary image, reduction of noise to reduce extraneous data, skew estimation of a document image if document suffers from tilt (skewed), thinning, enable subsequent detection of pertinent features of the object of interest and then segmentation of text line to individual digit character. It is assumed that characters are already isolated and preprocessing steps such as Binarization, noise removal, normalization and thinning are to be done. Initially, the document is captured using gray level mode. Character extraction from the scanned document is done manually. Then Otsu's histogram-based global threshold approach is applied for digit image Binarization. Figure 4 shows the gray level picture of the character and resultant of global thresholding approach.



**Figure 4. Result of Binarization Technique**

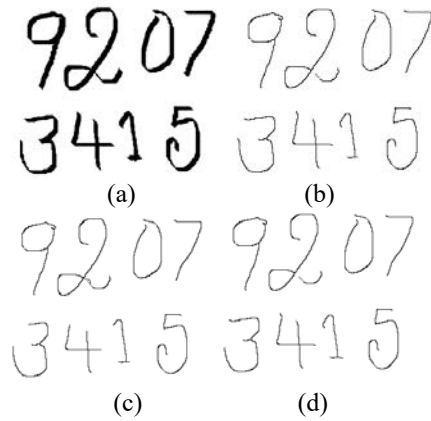
The next step is to normalize the given character image into standard size. For that, standard nearest neighbor interpolation method is used. Figure 5 shows some the sample images after applying normalization technique.



**Figure 5. Samples of Digit Images after Normalization**

### **Step 3: Thinning of the digit image**

Skeletonization has been a part of image processing for a wide variety of applications. Thinning is the process of reducing thickness of each line of pattern to just a single pixel width. Thinning is usually used as the first step in applications such as optical character recognition to improve the recognition rate. It has been applied to many fields [11], such as inspection of printed circuit boards, counting of asbestos fibers on air filters, analysis of chromosomes, classification of fingerprints, recognition of characters, application in intelligent copying and facsimile transmission systems, data reduction of map storage, etc. The major advantages of thinning in image processing and pattern recognition are: reduction of amount of data as input binary image and preservation of fundamental skeleton, which is topologically equivalent to the original object. Many thinning algorithms have been introduced for decades. These algorithms have both advantages and disadvantages. Some algorithms can obtain good quality skeletons, but they run slowly [12]. None of the methods address rotation invariant thinning. Many are specific to digits, characters, or letters, written in English, Chinese, Arabic, or any other scripts. However, to solve the aforementioned problems, the rotation invariant rule-based thinning algorithm for character recognition is proposed by [13]. This is a generalized algorithm, which is used to thin the symbols irrespective of the scripts. The advantage of this method is it is invariant to rotation. The algorithm fails on two-pixel wide lines. To overcome this drawback, an improved rotation-invariant thinning algorithm was proposed by [14]. Recently, improved two-step thinning algorithm is described by Aradhya et al in the year 2005, which is an improved version of standard two-step thinning algorithm [15]. The proposed algorithm thins the symbols to their central lines and has the rotation invariant property. Compared to two step and rotation rule based methods, the improved two-step method is better with respect to time, computations, and preserving topology. Hence in this work, improved two-step thinning algorithm is used for better representation of character images and to extract features from the character images. The resultant of thinning algorithms is shown in figure 6.



**Figure 6.** Result of different thinning algorithms

### 3. A Novel Feature Extraction Technique based on Distance Measure for Handwritten Digits

Digit image preprocessing has also been done for effective representation of characters. In this paper, a novel feature extraction technique based on distance measure will be presented. In this proposed method features are extracted based on distance measure technique. Feature extraction is one of the most important steps in any recognition system. The aim of feature extraction is to represent the image in terms of some quantifiable measurements that may be easily utilized in the classification stage. To a large extent, the accuracy of a recognition system depends on the discriminative capability of features and the generalization performance of a designed classifier. Document image representation methods are categorized in three major groups [16]:

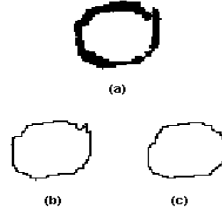
- a) Global Transformation and Series Expansion: Geometrical and topological features can represent various global and local properties of numerals images with high tolerance to distortions and style variations. The methods on topological and geometrical representations can be grouped into four categories i.e. Coding, Graphs & Trees, Counting & Extracting Topological Structures and Measuring & Approximating the Geometrical Properties.
- b) Statistical Representation: Representation by the use of statistical distribution of points will take care of the style variations to some extent. There are three important statistical features used for character representation: Zoning, Crossing and Distances and Projections.
- c) Geometrical and Topological Representation: A continuous signal generally contains more information that needed to be represented for the purpose of classification. One way to represent a signal is by a linear combination of a series of simpler well-defined functions. The coefficients of the linear combination provide a compact encoding known as transformation or series expansion. Some common transform methods used in the character recognition field are Fourier transform, Gabor transform, Moments and Karhunen-Loeve transform.

In this paper, novel feature extractions schemes are developed i.e. Feature Extraction Method1 (FEM) based on distance measure are presented.

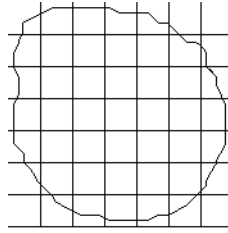
FEM mainly consists of 3 major steps

- Step 1: Thinning is applied to the given input digit image. In this method two step & improved two step thinning algorithms are used for thinning purpose. The resultant thinned images for a given character are shown in Figure 7.
- Step 2: Digit thinned image is divided into non-overlapping zones. In this paper each digit image is divided into  $7 \times 7$  blocks. The zoning of the digit 'zero' is shown in figure 8.
- Step 3: Calculate the "digit centroid" using the Equation given in algorithm 1. The distance is calculated between the character centroid point to the pixel present in each block of the

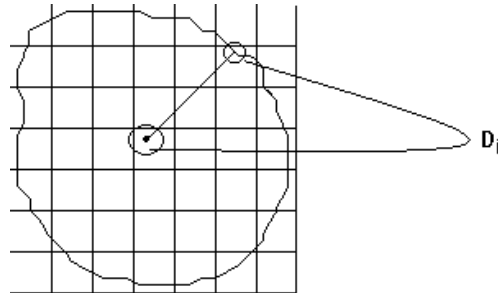
thinned image and it is graphically shown in figure 9. In this paper, city block distance is used for finding the distance. The sum of the distances of each pixel present in the block to the centroid pixel is calculated and this value is added to the feature vector. This procedure is repeated for all the blocks in the character image. Finally, 49 features are extracted from each given thinned image. These features are stored in a database for our further classification purpose. The pseudo code of the third step is shown algorithm 1.



**Figure 7.** Thinning Results (a) Original Image (b) 2 Step result (c) Result Using Improved 2 Step



**Figure 8.** Result of the digit image into 7x7 Zones



**Figure 9.** Distance Calculations from Centroid to Pixel in each Block

From the above feature extraction methods in this paper utilize the distance measure. The major goal of representation is to extract and select a set of features, which maximizes the recognition rate with the least amount of elements. Feature extraction and selection is defined as extracting the most representative information from the raw data, which minimizes the intra class pattern variability while enhancing the inter class pattern variability.

#### 4. Proposed Algorithm for Feature extraction process

**Algorithm 1:** Feature Extraction based on Distance Measure

**Input:** Thinned Image

**Output:** Feature Vector

**Method Begins**

Step 3.1: Find the character centroid (C) of the thinned image.

Step 3.2: Sum ( $S_x$ ) all X coordinate of image using the equation,  $S_x = \sum_{i=1}^n X_i$ , where n is the number of black pixels.

Step 3.3: Sum ( $S_y$ ) all Y coordinates of image using the equation,  $S_y = \sum_{i=1}^n Y_i$

Step 3.4: X coordinate of C =  $C_x = \frac{S_x}{N}$ , where N is the total number of black pixels

Step 1.4: Y coordinate of C =  $C_y = \frac{S_y}{N}$ .

Step 3.5: Find out the distance between character centroid point to the black pixels in each zone using city block distance.

Step 3.6: Sum all the distance of each block and represent as a feature value.

Step 3.7: Repeat this procedure for all the blocks.

**Method ends**

## 5. Results and Discussions

The proposed feature extraction methods are experimentally evaluated with the dataset containing various handwritten numerals collecting from MNIST data base, CENPARMI data base, CEDAR data base and most of the images from scanned images. Collectively, the data base contains 176000 digit images. No method has tested using such type of large database. In this paper, for testing the proposed method, Nearest Neighbor Classifier (NNC) is used for classification purpose. All experiments are carried out on a PC machine with i3 processor 2.7GHz CPU and 2 GB RAM memory under MatLab 10.0 platform. 20 percentage of the each database is used for training and reaming 80 percentage images are used for testing purpose i.e. 35200 images are used for training purpose and 140800 images are used for testing purpose.

The proposed method is evaluated when two different thinning algorithms are used for thinning a digit image i.e. two step and improved two step thinning algorithms. The percentage of recognition of the proposed method when two-step method is applied is listed out in tables 1, 2, 3 and 4. The percentage of recognition of the proposed method when improved two step method is applied is listed out in table 5, 6, 7 and 8. The overall percentage of the proposed recognition system when two thinning technique are applied is shown in table and the comparison graph of the two thinning algorithms is shown in figure 10. From the figure 10, it is clear that, using two steps thinning algorithm the overall recognition rate is 94.25 achieved whereas in case of improved thinning algorithm the performance of the system is 96.28%. From the above discussion strongly say that correct classification percentage not only depends on the type of feature extraction it also depends on type previous steps (Preprocessing i.e. Thinning) in the proposed method. So that the proposed method evaluate the results in two different preprocessing techniques for proving the above the statement.

**Table 1.** MNIST data base recognition percentage when two step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	6573	6315	258	96.07
1	6715	6451	264	96.07
2	6580	6297	283	95.70
3	6600	6294	306	95.36
4	6442	6159	283	95.61
5	6575	6354	221	96.64
6	6705	6279	426	93.65
7	6715	6452	263	96.08
8	6605	6251	354	94.64
9	6490	6296	194	97.01

**Table 2.** CEPARMI data base recognition percentage when two step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	320	305	15	95.31
1	320	297	23	92.81
2	320	298	22	93.13
3	320	291	29	90.94
4	320	301	19	94.06
5	320	297	23	92.81
6	320	299	21	93.44
7	320	301	19	94.06
8	320	296	24	92.50
9	320	294	26	91.88

**Table 3.** CEDAR database percentage of recognition when two step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	160	148	12	92.50
1	160	147	13	91.88
2	160	149	11	93.13
3	160	148	12	92.50
4	160	151	9	94.38
5	160	147	13	91.88
6	160	145	15	90.63
7	160	150	10	93.75
8	160	146	14	91.25
9	160	145	15	90.63

**Table 4.** SCANNED Images Database percentage of Recognition when two step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	7985	7785	200	97.50
1	8025	7717	308	96.16
2	8123	7783	340	95.81
3	7816	7561	255	96.74
4	7923	7722	201	97.46
5	8050	7694	356	95.58
6	8023	7693	330	95.89
7	8075	7687	388	95.20
8	8115	7714	401	95.06
9	7865	7423	442	94.38

**Table 5.** MNIST data base recognition percentage when improved two step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	6573	6495	78	98.81
1	6715	6482	233	96.53
2	6580	6501	79	98.80
3	6600	6415	185	97.20
4	6442	6291	151	97.66
5	6575	6398	177	97.31
6	6705	6505	200	97.02
7	6715	6552	163	97.57
8	6605	6483	122	98.15
9	6490	6316	174	97.32

**Table 6.** CEPARMI data base recognition percentage when improved two step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	320	310	10	96.88
1	320	305	15	95.31
2	320	297	23	92.81
3	320	298	22	93.13
4	320	305	15	95.31
5	320	303	17	94.69
6	320	307	13	95.94
7	320	309	11	96.56
8	320	305	15	95.31
9	320	304	16	95.00

**Table 7.** CEDAR database percentage of recognition when improved two step thinning applied

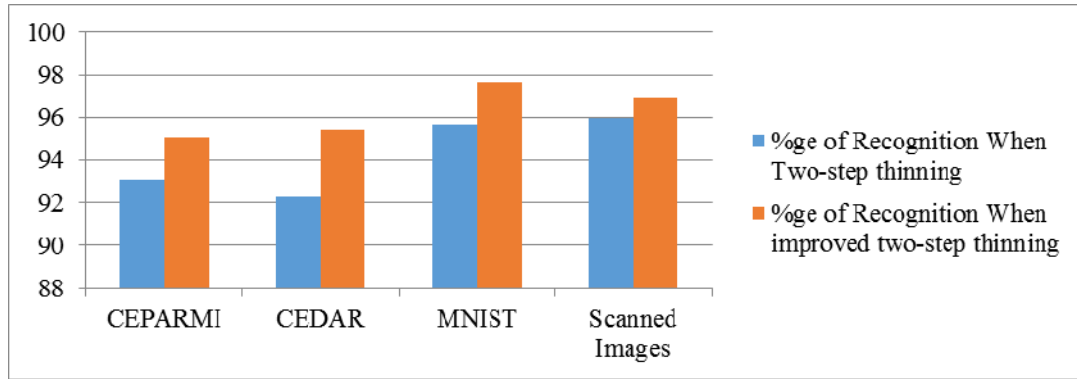
Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	160	154	6	96.25
1	160	156	4	97.50
2	160	153	7	95.63
3	160	154	6	96.25
4	160	154	6	96.25
5	160	149	11	93.13
6	160	151	9	94.38
7	160	154	6	96.25
8	160	152	8	95.00
9	160	150	10	93.75

**Table 8.** Scanned Images Database percentage of Recognition when improved 2-step thinning applied

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	7985	7815	170	97.87
1	8025	7793	232	97.11
2	8123	7819	304	96.26
3	7816	7614	202	97.42
4	7923	7785	138	98.26
5	8050	7719	331	95.89
6	8023	7802	221	97.25
7	8075	7801	274	96.61
8	8115	7792	323	96.02
9	7865	7623	242	96.92

**Table 9.** %ge of recognition of the proposed method when two thinning approaches are applied

Database	%ge of Recognition	
	When Two-step thinning	When improved two-step thinning
CEPARMI	93.09	95.11
CEDAR	92.25	95.43
MNIST	95.68	97.63
Scanned Images	95.98	96.95



**Figure 10.** Comparing Recognition Accuracy with two Different Thinning Algorithms

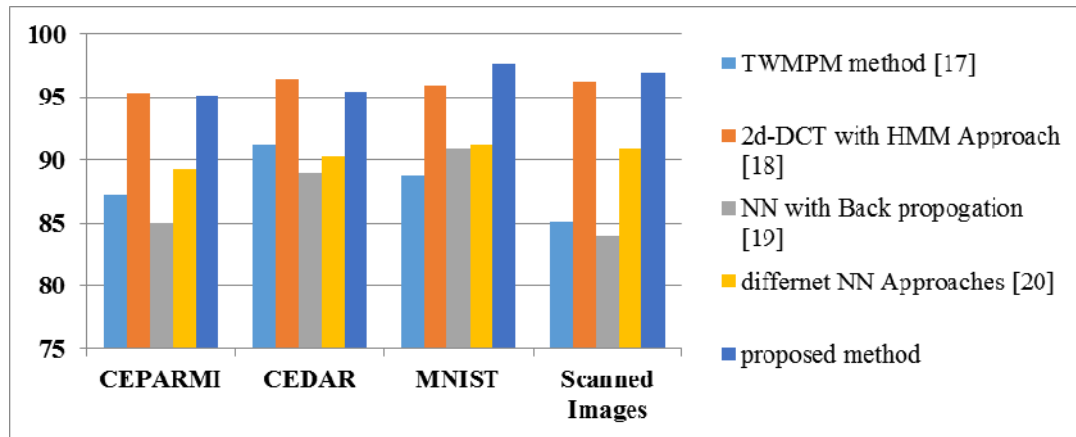
From the figure 10 and table 9, it is clear that improved two step thinning algorithm has performed better compared to standard two step thinning algorithm. The proposed method seeks more structural features and the size of the feature vector is also small.

#### Comparison of the proposed method with other existing methods:

The efficiency of the proposed method is compared with other existing methods like twin minimax probability machine (TWMPM) proposed by Zhijie et.al [17], transformation based features proposed by Syed et.al [18], Back Propagation with Neural Network approach[19] and selection, reproduction, mutation and crossover methods with Genetic Algorithm (GA) proposed by Devikanniga et.al[20]. The TWMPM method generates two hyper-planes of the digit images and it also avoids making distributional assumptions about the class-conditional densities of the digit. The performance of the TWNP method is evaluated on two data sets only i.e UTC and MNSIT and the overall percentage of the proposed method 88.07%. The method proposed by the Syed utilizes the Discrete Cosine Transform (2D-DCT) for feature extraction and Hidden Markov models (HMMs) used for classification. The syed proposed method is applied on only MNIST database and got 95.95% if feature vector size is more. More feature vector causes more computational cost. Sakshica proposed to classify the handwritten digits by using the features and their spatial relationship in the pattern with Hopfield Neural Network. Small number of images are tested by using this method and got 90.23% only. Devikanniga proposed a method to classify the handwritten digit using GA and got the overall performance is only 87%. The performance evolution of the proposed method with other existing methods is listed out in table 10 and the classification graph is represented in figure 11. From table 10 and figure 11, it is clearly evident that, the proposed method exhibits a high recognition rate than the existing methods.

**Table 10:** Overall %ge of the Different Recognition systems

Data Base	TWMPM method [17]	2d-DCT with HMM Approach [18]	NN with Back propagation [19]	different NN Approaches [20]	proposed method
CEPARMI	87.18	95.28	85	89.23	95.11
CEDAR	91.23	96.45	89	90.24	95.43
MNIST	88.73	95.92	91	91.27	97.63
Scanned Images	85.15	96.18	84	90.87	96.95



**Figure 11.** Graphical representation of the % recognition of the proposed method and other existing methods

## 6. Conclusions

In this paper a very large database of unconstrained handwritten digit images for experiment purpose was introduced. No author tested with such a large data base. It is proposed to make this database available for research purpose. 100000 handwritten characters collected from 1000 different individuals of different age groups and different areas. No author has attempted to such collection. The proposed method utilizes small algorithms for preprocessing such as Binarization, normalization and thinning algorithms. The proposed method takes very less amount processor time for recognition of digit image. The main object of the proposed method is an efficient feature extraction method is derived for handwritten digit recognition. For thinning digit image, a well-known technique is used i.e. improved two-step thinning method. The overall percentage of the proposed method achieves 96.36%. No such method is available up to now utilize the large set data base for testing. The proposed method proved that good classification not depend on the number of features. It's depends on only suitable feature are extract from an image. This paper shows the percentage of correct classification is not only depend on the feature extraction but also on the preprocessing stage (before the feature extraction) also important

## 7. References

- [1] J. pradeep, E. Strinivasan, and S. Himavathi, Neural network based handwritten character recognition system with feature extraction. International conference on computer, communication and electrical technology- ICCET 2011, 18th & 19th Mar. 2011.
- [2] W. Wu and Y. Bao, Online handwritten magnolia words recognition based on multiple classifiers, 2009.
- [3] C. L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, Pattern Recognition, vol. 37, iss. 2, pp. 265-279, 2004
- [4] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai, and L. Lam, Computer recognition of unconstrained handwritten numerals, Proc. IEEE, vol. 80, iss. 7, pp. 1162-1180, 1992.
- [5] D. Gorgevik and D. Cakmakov, "Handwritten Digit Recognition by Combining SVM Classifiers," in The International Conference on Computer as a Tool (EUROCON), 2005.
- [6] M. D. Garriss, J. L. Blue and G. T. Candela, "NIST form-based handprint recognition system," NIST, 1997.
- [7] P. J. Grother, "NIST special database 19 hand printed forms and characters database," National Institute of Standards and Technology, 1995.

- [8] X. Chen, X. Liu and Y. Jia, "Learning Handwritten Digit Recognition by the Max-Min Posterior Pseudo-Probabilities Method," in Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), 2007.
- [9] Xu, L., A. Krzyzak and C.Y. Suen (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. IEEE Trans. Syst. Man Cybernet. 22 (3), 418-435.
- [10] Y. LeCun and Y. Bengio. Convolutional networks for images, speech, and time-series. In M. A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1995
- [11] B T Chen, Y S Chen and WH Hsu, "A Parameterized Fuzzy Processor and its Applications," Fuzzy sets and systems, Vol 59, pp. 149-172, 1993
- [12] Sehultze-Lam S, Harauz G, Beveridge TJ, "Participation of a cyanobacterial S-layer in fine grain mineral formation", 1992
- [13] Maher Ahmed, Rabab Ward, "A Rotation Invariant Rule-Based Thinning Algorithm for Character Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence , Volume 24 Issue 12, December 2002 , Page 1672-1678
- [14] Peter I. Rockett, 2005, An Improved Rotation- Invariant Thinning Algorithm, IEEE, Trans. Patt.Anal.Machine Intl., Oct, Vol. 27.No.10. pp. 1671-1674.
- [15] Gonzalez R.C and Woods R. E, 2002 "Digital Image Processing text book" 2<sup>nd</sup> edition 2002
- [16] Nafiz Arica and Fatos T. Yarman-Vural, "An Overview of Character Recognition Focused on Off-Line Handwriting", IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 31, No. 2, May 2001
- [17] Zhijie Xu, Jianqin Zhang and Hengyou Wang, "Twin Minimax Probability Machine for Handwritten Digit Recognition", International Journal of Hybrid Information Technology Vol.8, No.2 (2015), pp.31-40
- [18] Syed Salman Ali, Muhammad Usman Ghani "Handwritten Digit Recognition using DCT and HMMs", CONFERENCE PAPER · JANUARY 2014
- [19] Sakshica, Dr. Kusum Gupta, "Handwritten Digit Recognition using various Neural Network Approaches", International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 2, February 2015
- [20] Ms. D. Devikanniga, Dr. M. Thangamani, "Handwritten Digit Recognition using Genetic Algorithm", International Journal of Innovations & Advancement in Computer Science, Volume 4, Special Issue May 2015

## UNCOVERING OF ANONYMOUS ATTACKS BY DISCOVERING VALID PATTERNS OF NETWORK

Dr G.Charles Babu

Professor  
Dept.of CSE  
MRE College  
Secunderabad, India.  
[charlesbabu26@gmail.com](mailto:charlesbabu26@gmail.com)

N.Chennakesavulu

Assoc.Professor  
Dept.of CSE  
Wesley PG College  
Secunderabad, India  
[keshav-reddy@yahoo.com](mailto:keshav-reddy@yahoo.com)

R.Venkateswara Gandhi

Lecturer  
Dept.of CSE  
Jigjiga University  
Jigjiga, Ethiopia.  
[ravegag@gmail.com](mailto:ravegag@gmail.com)

### ABSTRACT:

Efficient recognition of denial of service attacks is compulsory towards protecting online services. Quite a lot of efforts were made to attain novelty-tolerant recognition systems and build up an additional superior concept, specifically, anomaly based detection. The traffic attacks of denial of service acts in a different method from practical network traffic, and network traffic performance is revealed by its statistical assets. We put forward a system of denial of service attacks recognition scheme that utilizes multivariate correlation analysis for exact network traffic description by means of mining geometrical correlations among features of network traffic. Proposed approach provider obtains benefits in the direction of data analysis and does not require

information of historic traffic in performing analysis. Proposed system of multivariate correlation analysis based denial of service attacks detection utilizes anomaly based recognition in attack recognition thus capable of identifying recognized and unidentified attacks efficiently by means of learning patterns of reasonable network traffic. Triangle-area-based approach of multivariate correlation analysis was implemented to analyze reasonable network traffic, and generated triangle area maps are then used to provide quality features in support of common profile generation.

**Keywords:** *Denial of service attacks, Multivariate correlation analysis, Triangle-area, Network traffic, online services.*

### 1. INTRODUCTION:

Denial of service attacks strictly degrades accessibility of a victim and imposes demanding computation tasks towards

victim by means of exploiting its system susceptibility. Hence, successful detection of denial of service attacks is necessary in the direction of protecting online services [1]. Efforts on denial of service attacks recognition generally spotlight on improvement of network-based recognition methods. Detection systems of network-based are classified into two most important categories, specifically, misuse-based as well as anomaly based detection systems. Misuse-based recognition systems identify attacks by means of monitoring network performance and search for matches with existing attack signatures. Misuse-based recognition systems are simply evaded by any new attacks regardless of having high detection rates to recognized attacks. Several research community, searched to attain novelty-tolerant recognition systems and build up an additional superior concept, specifically, anomaly based detection. Modern studies have focused on analysis of feature correlation. Detection systems of network-based are loosely fixed with operating systems working on protecting host machines thus configurations of network basis recognition systems are less difficult than that of host-based recognition systems. Tan et al. projected complicated non-pay load-based denial of service attacks recognition system by

means of multivariate correlation analysis (MCA) [2][3]. In our work we suggest denial of service attacks recognition system that utilizes multivariate correlation analysis for exact network traffic description by means of mining geometrical correlations among features of network traffic. Projected system of triangle-area-based multivariate correlation examination withstands difficulty and provides description for particular network traffic records to a certain extent than model system traffic behaviour of network traffic records. Our multivariate correlation analysis approach providers benefits in the direction of data analysis and do not require information of historic traffic in performing analysis.

## **2. DETECTION SYSTEM**

### **ARCHITECTURE OF DOS ATTACK:**

A triangle area method is developed to improve and to accelerate the procedure of multivariate correlation analysis. Our multivariate correlation analysis based denial of service attacks detection system as shown in fig1 utilizes opinion of anomaly based recognition in attack recognition hence capable of identifying recognized and unidentified DoS attacks efficiently by means of learning patterns of reasonable network traffic. Projected detection system can make available

effectual protection to all of systems by means of considering their commonality. The general idea of projected denial of service attacks recognition system is given that consists of three most important steps. In the initial step, fundamental features are produced from ingress network traffic towards internal network where protected servers exist in and forms traffic records for a definite time period. Monitoring of destination network decreases transparency of detecting malevolent actions by concentrating on applicable inbound traffic. It enables our detector to make available protection which is best fit for targeted internal network. Second step is analysis of multivariate correlation, in which module of triangle area map generation is functional to mine the correlations among two distinct features within every traffic record coming from initial step. The entire extracted correlations, specifically, triangle areas stored within triangle area maps, are subsequently used to restore the original fundamental features to symbolize traffic records [4]. This differentiates among legitimate as well as illicit traffic records. In the third step, mechanism of anomaly based detection is assumed in decision making which facilitates detection of any denial of service attacks without necessitating any attack applicable

information. Particularly, two phases such as training phase as well as test phase are concerned in decision making. The normal profile generation component is functional in training phase to produce profiles for a variety of types of legitimate traffic records, and produced regular profiles are stored within a database. The tested profile generation component is used in test phase to construct profiles meant for individual observed traffic records. Tested profiles are surrendered to attack detection module, which evaluates individual tested profiles by particular stored common profiles.

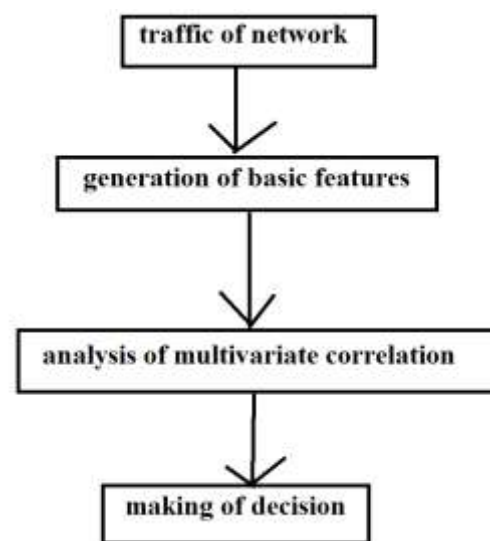


Fig1: An overview of proposed system.

### 3. AN OVERVIEW OF ANALYSIS OF MULTIVARIATE CORRELATION:

Denial of service attack traffic behaves in a different way from reasonable network traffic performance is revealed by its statistical assets. To explain statistical

properties, we suggest a novel multivariate correlation analysis approach that utilizes triangle area designed for mining of correlative information among features within an observed data object. Technique of triangle area is developed to improve and to accelerate the procedure of multivariate correlation analysis. When triangle area maps are compared, they are imagined as two images symmetric all along their major diagonals [5]. Any differentiations, identified on upper triangles of images, are found on lower triangles hence, to carry out a speedy comparison of two triangle area maps, we prefer to examine moreover upper triangles or else lower triangles of the triangle area maps. We apply projected triangle-area-based multivariate correlation analysis approach to analyze justifiable network traffic, and generated triangle area maps are then used to provide quality features in support of common profile generation. Our multivariate correlation analysis approach provides benefits in the direction of data analysis. It does not necessitate information of historic traffic in performing analysis. Contrasting from approaches of covariance matrix which is susceptible to linear change of all features, projected triangle-area-based multivariate correlation analysis withstands difficulty. It provides description for particular

network traffic records to a certain extent than model system traffic behaviour of network traffic records [6]. This consequence in lower latency in decision making and facilitate sample-by-sample detection. The correlations among separate pairs of features are exposed all the way through geometrical structure analysis. Changes of these structures might happen when anomaly behaviours come into view in network that provides a momentous signal to set off an alert.

#### **4. CONCLUSION:**

Attempts which are made on denial of service attacks recognition generally spotlight on improvement of network-based recognition methods. Our work recommends denial of service attacks recognition system that utilizes multivariate correlation analysis for exact network traffic description by means of mining geometrical correlations between features of network traffic. The system can make available effectual protection to all of systems by means of considering their commonality. The novel approach utilizes triangle area designed for mining of correlative information among features within an observed data object. Contrasting from the existing approaches projected triangle-area-based multivariate correlation analysis withstands difficulty

and makes available description for particular network traffic records to a certain extent than model system traffic behaviour of network traffic records. A triangle area system improves and accelerates the procedure of multivariate correlation analysis. Our multivariate correlation analysis approach provides benefits in the direction of data analysis and do not call for information of historic traffic in performing analysis.

## REFERENCES

- [1] C. Yu, H. Kai, and K. Wei-Shinn, "Collaborative Detection of DDoS Attacks over Multiple Network Domains," IEEE Trans. Parallel and Distributed Systems, vol. 18, no. 12, pp. 1649-1662, Dec. 2007.
- [2] G. Thatte, U. Mitra, and J. Heidemann, "Parametric Methods for Anomaly Detection in Aggregate Traffic," IEEE/ACM Trans. Networking, vol. 19, no. 2, pp. 512-525, Apr. 2011.
- [3] S.T. Sarasamma, Q.A. Zhu, and J. Huff, "Hierarchical Kohonen Net for Anomaly Detection in Network Security," IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 35, no. 2, pp. 302-312, Apr. 2005.
- [4] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R.P. Liu, "Triangle- Area-Based Multivariate Correlation Analysis for Effective Denial-of-Service Attack Detection," Proc. IEEE 11th Int'l Conf. Trust, Security and Privacy in Computing and Comm., pp. 33-40, 2012.
- [5] S.J. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P.K. Chan, "Cost- Based Modeling for Fraud and Intrusion Detection: Results from the JAM Project," Proc. DARPA Information Survivability Conf. And Exposition (DISCEX '00), vol. 2, pp. 130-144, 2000.
- [6] G.V. Moustakides, "Quickest Detection of Abrupt Changes for a Class of Random Processes," IEEE Trans. Information Theory, vol. 44, no. 5, pp. 1965-1968, Sept. 1998.

## BIOGRAPHY



**Dr.G.Charles Babu**, Ph.D in CSE from ANU in 2015 and Presently working as a Professor in CSE department in Malla Reddy Engineering College (Autonomous) and has 16 years of Academic experience. His Research interests include Data Mining, Software Engineering, Networks and Cloud Computing.



**N.Chennakesavulu**, Post Graduated in Computer Science & Engineering (**M.Tech**) From JNT University, Hyderabad in 2013. He is working as a Associate Professor in Department of Computer Science & Engineering in **Wesley PG College, Secunderabad** and he has 10 years of experience in Teaching in India. His Research Interests Include Image Processing, Network Security, Software Engineering.



**R.VENKATESWARA GANDHI**, Post Graduated in Computer Science & Engineering (**M.Tech**) From JNT University, Hyderabad in 2009 and Graduated in Computer Science & engineering (B.Tech) form JNTU, Hyderabad, 2007. He has worked as an Lecturer in Department of Computer Science & Engineering in **JigjigaUnivesity, Jigjiga, Ethiopia**. He has 5+ years of Teaching Experience in India and 1 academic year experience abroad. His Research Interests Include Network Security, Cloud Computing & Data Warehousing and Data Mining.

# Novel Technique for the Handwritten Digit Image Features Extraction for Recognition

<sup>1</sup>Dr. U Ravi Babu and <sup>2</sup>Dr. G Charles babu

<sup>1</sup> Professor, Dept. of CSE, MREC(A), TS, India, uppu.ravibabu@mrec.ac.in

<sup>2</sup> Professor, Dept. of CSE, MREC(A), TS, India, charlesbabu26@gmail.com

**Abstract**— This paper proposes a novel approach for handwritten digit recognition system. The present paper extracts digit image features based on distance measure and derives an algorithm to classify the digit images. The distance measure can be performed on the thinned image. Thinning is the one of the preprocessing technique in image processing. The present paper mainly concentrated on an extraction of features from digit image for effective recognition of the numeral. The main objective of the paper is exact features are extracted for good classification results. To find the effectiveness of the proposed method, tested on MNIST database, CENPARMI, CEDAR, and newly collected data. The proposed method is implemented on more than one lakh digit images and got good comparative recognition results. The percentage of the recognition is achieved about 97.32%.

**Index Term**— Handwritten digit recognition, Distance measure, MNIST Database, image features.

## I. INTRODUCTION

In the field of image processing and pattern recognition [1] handwriting recognition has one of the hottest and demanding directions in the recent years. Day to day new technologies and innovative methods have been proposed continuously. With the development of the smart phone operation system, the application in handwritten recognition has aroused more and more attention from researchers. In general, handwritten character recognition is classified into two types of offline and online recognition methods [2]. An optical character recognition (OCR) system with a good recognition performance needs to maintain a very high recognition rate, and at the same time, to obtain a very high reliability, or a very low error rate [3, 4]. Recent developments on classifiers and feature extraction have significantly increased the recognition accuracy of handwritten digit recognition systems.

Gorgevik et al. [5] proposed handwritten digits recognition system by using Support Vector Machines (SVM). They extracted four types of features from each digit image 1) projection histograms, 2) contour profiles, 3) ring-zones and 4) Kirsch features. They reported 97.27% recognition accuracy on National Institute of Standards and Technology (NIST) [6] handwritten digits database [7] when four types of features were used collectively. In [8] Chen et al. proposed maximum posterior pseudo-probabilities framework for handwritten digits recognition. They extracted 256-D directional features from the input image. Finally, 256-D features were transformed into 128-D feature using Principal Component

Analysis (PCA). They reported recognition accuracy of 98.76% on NIST database [7].

A few works have also been reported in the literature on handwritten digit recognition of Indian scripts. Most of the available works on handwriting recognition of Indian scripts are based on small databases with a few basic characters and collected in laboratory environments. At present, a few large databases of handwritten digits of major Indic scripts are available from Computer Vision and Pattern Recognition Unit (CVPR), Indian Statistical Institute, Kolkata. Offline handwritten numerals database is available for Devanagari, Bangla, and Oriya scripts. Basic characters, vowel modifiers and compound characters about Devanagari, Bangla, and Oriya scripts are also available. Both online and offline forms of the dataset are available for research purpose. An offline version of the data is also available in the form of bi-level TIFF images, generated from the online data using simple piecewise linear interpolation with a constant thickening factor applied. The paper considers only the Basic English numerals for recognition purpose. The dataset for the experiment was collected from different individuals of various professions in the states of Andhra Pradesh and Telangana. The proposed method overcome the disadvantages of various techniques and find the effective feature extraction process for the recognition handwritten digits of different databases in an effective manner.

The paper is organized as follows: Section 2 contains database creation and the preprocessing. Section 3 describes feature extraction methods and the proposed algorithm is presented in Section 4. The experimental details and result analysis is presented in Section 6. Section 6 contains the conclusion part.

## II. PROPOSED METHOD

The proposed method is mainly consists of 5 steps. In the first step, collecting the numerals data from various data bases and gathering images from various people in AP and Telangana state. After collecting the numeral data preprocess data i.e. elimination of noise and conversion of gray scale images into binary images and also the normalization of the binary images by using the normalization techniques. In the third step, binary images are converted into single pixel-width images i.e. thinning of the digit images. In this step, various thinning operations are performed and found out the best thinning techniques suitable for the digit images. In the fourth step, extracted the features of the thinned image by using distance measure. In the last step, derive an algorithm for recognition of handwritten numerals system. The block

diagram of the proposed method is shown below figure 1.

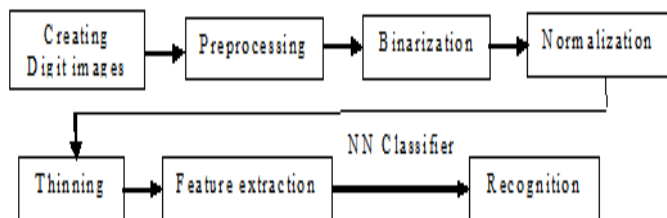


Fig. 1. Block diagram of the Handwritten Digit Recognition system

### Step 1: Collection of numeral image database

Several standard datasets of digits are found in English. Some of them are CENPARMI, CEDAR, and MNIST datasets. The CENPARMI (Centre for Pattern Recognition and Machine Intelligence) digit dataset [9] is available from CENPARMI, Concordia University. In this dataset 4000 images (400 samples per class) are specified for training purpose and 2000 images are used for testing purpose. These digit datasets were collected from United States Postal Service (USPS). The Center of Excellence for Document Analysis and Recognition (CEDAR) digit dataset is available from CEDAR, The State University of New York, Buffalo. The training and test sets contain 18468 and 2711 digits, respectively. The number of samples in both training and test sets differ for each class. The Modified National Institute of Standards and Technology (MNIST) dataset [10] was extracted from the NIST datasets SD3 and SD7. The training and test sets are composed of both SD3 and SD7. Samples are normalized into 20×20 grayscale images with aspect ratio reserved, and the normalized images are located in a 28×28 frame. The number of training and test samples is 60,000 and 10,000 respectively. The sample images of the MNIST dataset is shown in figure 2.

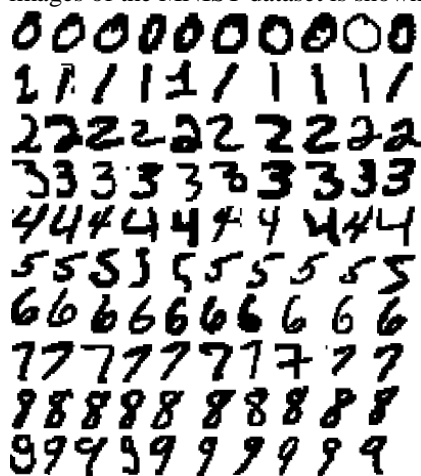


Fig. 2. Sample digit images of MNIST database

The plain paper was used for data collection. Each person was instructed to write the digits (fully unconstrained) along the vertical direction. The dataset contains about 100 isolated samples each of 10 numerals written by 1000 native writers including university graduates, high school children, and adults. Around fifty percent data is from high school children. A flatbed scanner was used for digitization, with images in gray tone at 300 dpi. These were stored as Bit Map File (BMP) format using a standard technique for converting them

into monochrome images. Data was manually extracted from scanned images and normalized into 50×50 size using a standard bicubic approach. After processing scanned images about digits and a total of 100000 (100×1000) images of numerals are obtained. Dataset developed planned to be made available publicly for research purpose. Some of the sample images after extracting from the scanned image are shown in figure 3.



Fig. 3. Sample Scanned document of digit images

### Step 2: Digit Image Preprocessing

Data capture of documents by optical scanning or by digital video yields a file of picture elements, which is the raw input to document analysis process. The first step in document analysis is to perform a preprocessing on this image to prepare it for further analysis. Such processing includes thresholding to convert a gray scale (or color image) to a binary image, reduction of noise to reduce extraneous data, skew estimation of a document image if document suffers from tilt (skewed), thinning, enable subsequent detection of pertinent features of the object of interest and then segmentation of text line to individual digit character. It is assumed that characters are already isolated and preprocessing steps such as Binarization, noise removal, normalization and thinning are to be done. Initially, the document is captured using gray level mode. Character extraction from the scanned document is done manually. Then Otsu's histogram-based global threshold approach is applied for digit image Binarization. Figure 4 shows the gray level picture of the character and resultant of global thresholding approach.



Fig. 4. Result of Binarization Technique

The next step is to normalize the given character image into a standard size. For that, a standard nearest neighbor interpolation method is used. Figure 5 shows some sample images after applying normalization technique.



Fig. 5. Samples of Digit Images after Normalization

### Step 3: Thinning of the digit image

Skeletonization has been a part of image processing for a wide variety of applications. Thinning is the process of reducing the thickness of each line of pattern to just a single pixel width. Thinning is usually used as the first step in applications such as optical character recognition to improve the recognition rate. It has been applied to many fields [11], such as inspection of printed circuit boards, counting of asbestos fibers on air filters, analysis of chromosomes, classification of fingerprints, recognition of characters, application in intelligent copying and facsimile transmission systems, data reduction of map storage, etc. The major advantages of thinning in image processing and pattern recognition are: reduction of an amount of data as input binary image and preservation of fundamental skeleton, which is topologically equivalent to the original object. Many thinning algorithms have been introduced for decades. These algorithms have both advantages and disadvantages. Some algorithms can obtain good quality skeletons, but they run slowly [12]. None of the methods address rotation invariant is thinning. Many are specific to digits, characters, or letters, written in English, Chinese, Arabic, or any other scripts. However, to solve the aforementioned problems, the rotation invariant is rule-based thinning algorithm for character recognition is proposed by [13]. This is a generalized algorithm, which is used to thin the symbols irrespective of the scripts. The advantage of this method is it is invariant to rotation. The algorithm fails on two-pixel wide lines. To overcome this drawback, an improved rotation-invariant thinning algorithm was proposed by [14]. Recently, improved two-step thinning algorithm is described by Aradhya et.al in the year 2005, which is an improved version of standard two-step thinning algorithm [15]. The proposed algorithm thins the symbols to their central lines and has the rotation invariant property. Compared to two-step and rotation rule-based methods, the improved two-step method is better with respect to time, computations, and preserving topology. Hence in this work, improved two-step thinning algorithm is used for better representation of character images and to extract features from the character images. The resultant of thinning algorithms is shown in figure 6.

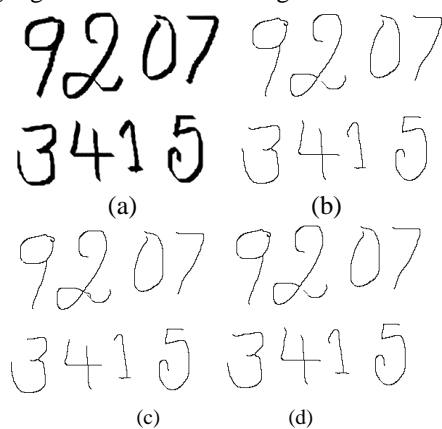


Fig. 6. Result of different thinning algorithms

### III. A NOVEL FEATURE EXTRACTION TECHNIQUE BASED ON DISTANCE MEASURE FOR HANDWRITTEN DIGITS

Digit image preprocessing has also been done for effective representation of characters. In this paper, a novel feature extraction technique based on distance measure is presented. In this proposed method features are extracted based on distance measure technique. Feature extraction is one of the most important steps in any recognition system. The aim of feature extraction is to represent the image in terms of some quantifiable measurements that may be easily utilized in the classification stage. To a large extent, the accuracy of a recognition system depends on the discriminative capability of features and the generalization performance of a designed classifier. Document image representation methods are categorized in three major groups [16].

#### A. Global Transformation and Series Expansion:

Geometrical and topological features can represent various global and local properties of numerals images with high tolerance to distortions and style variations. The methods on topological and geometrical representations can be grouped into four categories i.e. Coding, Graphs & Trees, Counting & Extracting Topological Structures and Measuring & Approximating the Geometrical Properties.

#### B. Statistical Representation:

Representation by the use of statistical distribution of points will take care of the style variations to some extent. There are three important statistical features used for character representation: Zoning, Crossing and Distances and Projections

#### C. Geometrical and Topological Representation:

A continuous signal generally contains more information that needed to be represented for the purpose of classification. One way to represent a signal is by a linear combination of a series of simpler well-defined functions. The coefficients of the linear combination provide a compact encoding known as transformation or series expansion. Some common transform methods used in the character recognition field are Fourier transform, Gabor transform, Moments and Karhunen-Loeve transform.

In this paper, novel feature extractions schemes are developed i.e. Feature Extraction Method (FEM) based on distance measure are presented.

FEM mainly consists of X major steps

Step 1: Thinning is applied to the given input digit image. In this method two-step & improved two-step thinning algorithms are used for thinning purpose. The resultant thinned images for a given character are shown in Figure 7.

Step 2: The thinned digit image is scanned from top-left corner to top-right corner (row-wise). While scanning the image, the black pixel positions are identified in each row by using the following procedure.

(a) If any black pixel is encountered in this process, the algorithm computes the distance between the

points of the region where the black pixel is encountered.

(b) If the algorithm fails to encounter any black pixel in the row, the distance is considered as zero value.

This procedure of finding black pixel in the specified region is continued till the end of the row. Figure 8 shows the working procedure of the proposed method. If the size of the image is  $50 \times 50$ , then 50 feature values along each row is encountered.

Step 3: Similarly distances are calculated column wise i.e. from the topmost pixel to the bottommost pixel of every column. Figure 9 shows the working procedure of the proposed method. If the size of the image is  $50 \times 50$ , then 50 feature values along each row is encountered.

Totally 100 (i.e.,  $50 + 50$ ) features are extracted from each input character.

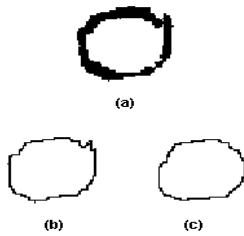


Fig. 7. Thinning Results (a) Original Image (b) 2 Step result (c) Result Using Improved 2 Step

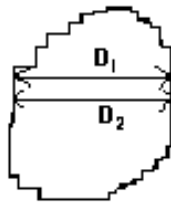


Fig. 8. Distance Computing along Row-wise

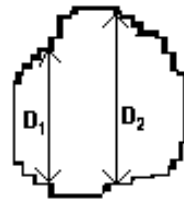


Fig. 9. Distance Computing along Column-wise

The major goal of representation is to extract and select a set of features, which maximizes the recognition rate with the least amount of elements. Feature extraction and selection is defined as extracting the most representative information from the raw data, which minimizes the intra class pattern variability while enhancing the inter-class pattern variability.

#### IV. RESULTS AND DISCUSSIONS

The proposed feature extraction methods are experimentally evaluated with the dataset containing various handwritten numerals collecting from MNIST data base, CENPARMI data base, CEDAR data base and most of the images from scanned images. Collectively, the data base contains 176000 digit

images. No method has tested using such type of large database. In this paper, Nearest Neighbor Classifier (NNC) is used for classification purpose. All experiments are carried out on a PC machine with i3 processor 2.7GHz CPU and 2 GB RAM memory under MatLab 10.0 platform. 20 percentage of the each data base is used for training and remaining 80 percentage images are used for testing purpose i.e. 35200 images are used for training purpose and 140800 images are used for testing purpose.

The proposed method is evaluated when two different thinning algorithms are used for thinning a digit image i.e. two-step and improved two step thinning algorithms. The percentage of recognition of the proposed method when two-step method is applied is listed out in tables 1, 2, 3 and 4. The percentage of recognition of the proposed method when improved two-step method is applied is listed out in table 5, 6, 7 and 8. The overall percentage of the proposed recognition system when two thinning technique are applied is shown in a table and the comparison graph of the two thinning algorithms is shown in figure 10. From the figure 10, it is clear that, using two-steps thinning algorithm the overall recognition rate is 94.25 achieved whereas in case of improved thinning algorithm the performance of the system is 96.28%.

TABLE I  
MNIST DATA BASE RECOGNITION PERCENTAGE WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	6573	6323	250	96.20
1	6715	6464	251	96.26
2	6580	6308	272	95.87
3	6600	6302	298	95.48
4	6442	6172	270	95.81
5	6575	6365	210	96.81
6	6705	6287	418	93.77
7	6715	6465	250	96.28
8	6605	6262	343	94.81
9	6490	6304	186	97.13

TABLE II  
CEPARMI DATA BASE RECOGNITION PERCENTAGE WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	320	308	12	96.25
1	320	307	13	95.94
2	320	308	12	96.25
3	320	303	17	94.69
4	320	306	14	95.63
5	320	305	15	95.31
6	320	309	11	96.56
7	320	303	17	94.69
8	320	307	13	95.94
9	320	310	10	96.88

TABLE III

CEDAR DATABASE PERCENTAGE OF RECOGNITION WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	160	150	10	93.75
1	160	149	11	93.13
2	160	150	10	93.75
3	160	151	9	94.38
4	160	152	8	95.00
5	160	155	5	96.88
6	160	151	9	94.38
7	160	153	7	95.63
8	160	154	6	96.25
9	160	156	4	97.50

TABLE IV

SCANNED IMAGES DATABASE PERCENTAGE OF RECOGNITION WHEN TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	7985	7803	182	97.72
1	8025	7733	292	96.36
2	8123	7800	323	96.02
3	7816	7579	237	96.97
4	7923	7738	185	97.67
5	8050	7711	339	95.79
6	8023	7711	312	96.11
7	8075	7703	372	95.39
8	8115	7731	384	95.27
9	7865	7441	424	94.61

TABLE V

MNIST DATA BASE RECOGNITION PERCENTAGE WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	6573	6503	70	98.94
1	6715	6490	225	96.65
2	6580	6509	71	98.92
3	6600	6423	177	97.32
4	6442	6299	143	97.78
5	6575	6406	169	97.43
6	6705	6513	192	97.14
7	6715	6560	155	97.69
8	6605	6491	114	98.27
9	6490	6324	166	97.44

TABLE VI

CEPARMI DATA BASE RECOGNITION PERCENTAGE WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	320	312	8	97.50
1	320	315	5	98.44
2	320	305	15	95.31
3	320	307	13	95.94
4	320	312	8	97.50
5	320	314	6	98.13
6	320	312	8	97.50
7	320	313	7	97.81
8	320	316	4	98.75
9	320	315	5	98.44

TABLE VII

CEDAR DATABASE PERCENTAGE OF RECOGNITION WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	160	156	4	97.50
1	160	157	3	98.13
2	160	155	5	96.88
3	160	155	5	96.88
4	160	156	4	97.50
5	160	153	7	95.63
6	160	152	8	95.00
7	160	155	5	96.88
8	160	153	7	95.63
9	160	157	3	98.13

TABLE VIII

SCANNED IMAGES DATABASE PERCENTAGE OF RECOGNITION WHEN IMPROVED 2-STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	7985	7833	152	98.10
1	8025	7811	214	97.33
2	8123	7837	286	96.48
3	7816	7632	184	97.65
4	7923	7803	120	98.49
5	8050	7737	313	96.11
6	8023	7820	203	97.47
7	8075	7819	256	96.83
8	8115	7810	305	96.24
9	7865	7641	224	97.15

TABLE IX

%GE OF RECOGNITION OF THE PROPOSED METHOD WHEN TWO THINNING APPROACHES ARE APPLIED

Database	%ge of Recognition	
	When Two-step thinning	When improved two-step thinning
CEPARMI	95.84	97.76
CEDAR	95.81	97.53
MNIST	95.06	96.81
Scanned Images	96.19	97.18

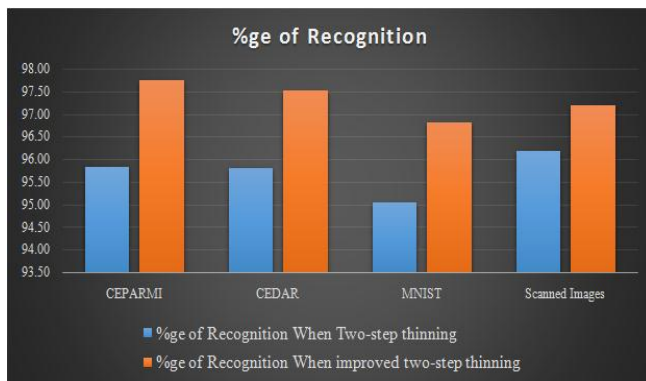


Fig. 10. Comparing Recognition Accuracy with two Different Thinning Algorithms

From the figure 10 and table 9, it is clear that improved two step thinning algorithm has performed better compared to standard two step thinning algorithm. The proposed method seeks more structural features and the size of the feature vector is also small. The overall recognition rate of the considered digit database is shown in table 10 and the overall recognition percentage is 97.44.

TABLE X  
CONSIDERED DATABASE PERCENTAGE OF RECOGNITION WHEN IMPROVED TWO STEP THINNING APPLIED

Digit	Total no of Digits	correctly classified	Not Correctly Classified	% Recognition
0	15038	14804	234	98.44
1	15220	14773	447	97.06
2	15183	14806	377	97.52
3	14896	14517	379	97.46
4	14845	14570	275	98.15
5	15105	14610	495	96.72
6	15208	14797	411	97.30
7	15270	14847	423	97.23
8	15200	14770	430	97.17
9	14835	14437	398	97.32

#### 4.1 Analysis of the proposed method:

To analyze the proficiency of the proposed system, the outcomes of the proposed method are analyzed in gotten Ten Cycle Cross Validation (TCCV) approach.

**Ten Cycle Cross Validation (TCCV)** approach: In TCCV approach results analysis strategy, the entire digit data base i.e. 150800 are divided into 10 sets. Each set consists of 15080 digit images mix up with the images of CEPRM, MNIST, CEDAR images, and Scanned Images. Every set must contain ten classes (0-9) of digit images. In TCCV approach results are analyzed in 10 Cycles. In cycle 1, first set is treated as a sample set and remaining 9 sets are taken as a test dataset. Compute the % of image grouping for test set. In cycle 2, second set is dealt with as a sample set and staying nine sets are taken as a test set. Compute the % of grouping for test set. The same methodology is applied remaining cycles also. This approach Strengthens the proposed method. The % of digit grouping of the proposed strategy in ten cycles are listed in tables XI to XX individually.

TABLE XI  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-1 OF TCCV APPROACH

Digit	Cycle-1			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13389	145	98.93
1	13698	13467	231	98.31
2	13665	13486	179	98.69
3	13406	13167	239	98.22
4	13361	13205	156	98.83
5	13595	13382	213	98.43
6	13687	13413	274	98.00
7	13743	13444	299	97.82
8	13680	13386	294	97.85
9	13352	13137	215	98.39

TABLE XII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-2 OF TCCV APPROACH

Digit	Cycle-2			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13391	143	98.94
1	13698	13485	213	98.45
2	13665	13468	197	98.56
3	13406	13187	219	98.37
4	13361	13121	240	98.20
5	13595	13419	176	98.71
6	13687	13559	128	99.06
7	13743	13544	199	98.55
8	13680	13467	213	98.44
9	13352	13177	175	98.69

TABLE XIII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-3 OF TCCV APPROACH

Digit	Cycle-3			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13370	164	98.79
1	13698	13501	197	98.56
2	13665	13468	197	98.56
3	13406	13237	169	98.74
4	13361	13215	146	98.91
5	13595	13412	183	98.65
6	13687	13473	214	98.44
7	13743	13584	159	98.84
8	13680	13462	218	98.41
9	13352	13176	176	98.68

TABLE XIV  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-4 OF TCCV  
APPROACH

Digit	Cycle-4			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13421	113	99.17
1	13698	13520	178	98.70
2	13665	13501	164	98.80
3	13406	13191	215	98.40
4	13361	13085	276	97.93
5	13595	13449	146	98.93
6	13687	13472	215	98.43
7	13743	13567	176	98.72
8	13680	13416	264	98.07
9	13352	13206	146	98.91

TABLE XVII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-7 OF TCCV  
APPROACH

Digit	Cycle-7			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13415	119	99.12
1	13698	13497	201	98.53
2	13665	13501	164	98.80
3	13406	13283	123	99.08
4	13361	13223	138	98.97
5	13595	13399	196	98.56
6	13687	13469	218	98.41
7	13743	13550	193	98.60
8	13680	13493	187	98.63
9	13352	13218	134	99.00

TABLE XV  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-5 OF TCCV  
APPROACH

Digit	Cycle-5			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13320	214	98.42
1	13698	13504	194	98.58
2	13665	13501	164	98.80
3	13406	13209	197	98.53
4	13361	13215	146	98.91
5	13595	13456	139	98.98
6	13687	13473	214	98.44
7	13743	13490	253	98.16
8	13680	13531	149	98.91
9	13352	13215	137	98.97

TABLE XVI  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-6 OF TCCV  
APPROACH

Digit	Cycle-6			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13389	145	98.93
1	13698	13485	213	98.45
2	13665	13489	176	98.71
3	13406	13272	134	99.00
4	13361	13223	138	98.97
5	13595	13436	159	98.83
6	13687	13530	157	98.85
7	13743	13484	259	98.12
8	13680	13472	208	98.48
9	13352	13186	166	98.76

TABLE XVIII  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-8 OF TCCV  
APPROACH

Digit	Cycle-8			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13420	114	99.16
1	13698	13485	213	98.45
2	13665	13497	168	98.77
3	13406	13190	216	98.39
4	13361	13224	137	98.97
5	13595	13431	164	98.79
6	13687	13552	135	99.01
7	13743	13532	211	98.46
8	13680	13537	143	98.95
9	13352	13198	154	98.85

TABLE XIX  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-9 OF TCCV  
APPROACH

Digit	Cycle-9			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13321	213	98.43
1	13698	13522	176	98.72
2	13665	13473	192	98.59
3	13406	13239	167	98.75
4	13361	13222	139	98.96
5	13595	13408	187	98.62
6	13687	13538	149	98.91
7	13743	13606	137	99.00
8	13680	13541	139	98.98
9	13352	13155	197	98.52

TABLE XX  
% OF RECOGNITION OF THE PROPOSED METHOD IN CYCLE-1 OF TCCV  
APPROACH

Digit	Cycle-10			
	Total	Properly classified	Not correctly classified	% of Classification
0	13534	13343	191	98.59
1	13698	13468	230	98.32
2	13665	13489	176	98.71
3	13406	13171	235	98.25
4	13361	13212	149	98.88
5	13595	13391	204	98.50
6	13687	13471	216	98.42
7	13743	13474	269	98.04
8	13680	13431	249	98.18
9	13352	13101	251	98.12

### Comparison of the proposed method with other existing methods:

The efficiency of the proposed method is compared to other existing methods like twin minimax probability machine (TWMPM) proposed by Zhijie et.al [17], transformation based features proposed by Syed et.al [18], Back Propagation with Neural Network approach[19] and selection, reproduction, mutation and crossover methods with Genetic Algorithm (GA) proposed by Devikanniga et.al[20]. The TWMPM method generates two hyper-planes of the digit images and it also avoids making distributional assumptions about the class-conditional densities of the digit. The performance of the TWMPM method is evaluated on two data sets only i.e UTC and MNIST and the overall percentage of the proposed method 88.07%. The method proposed by the Syed utilizes the Discrete Cosine Transform (2D-DCT) for feature extraction and Hidden Markov models (HMMs) used for classification. The syed proposed method is applied on only MNIST database and got 95.95% if feature vector size is more. More feature vector causes more computational cost. Sakshica proposed to classify the handwritten digits by using the features and their spatial relationship in the pattern with Hopfield Neural Network. A small number of images are tested by using this method and got 90.23% only. Devikanniga proposed a method to classify the handwritten digit using GA and got the overall performance is only 87%. The performance evolution of the proposed method with other existing methods is listed out in table 21 and the classification graph is represented in figure 11. From table 10 and figure 11, it is clearly evident that, the proposed method exhibits a high recognition rate than the existing methods.

TABLE XXI  
OVERALL %GE OF THE DIFFERENT RECOGNITION SYSTEMS

Data Base	TWMP M method [17]	2d-DCT with HMM Approac h [18]	NN with Back propagati on [19]	different NN Approach es [20]	propose d method
CEPARMI	87.18	95.28	85	89.23	95.11
CEDAR	91.23	96.45	89	90.24	95.43
MNIST	88.73	95.92	91	91.27	97.63
Scanned Images	85.15	96.18	84	90.87	96.95

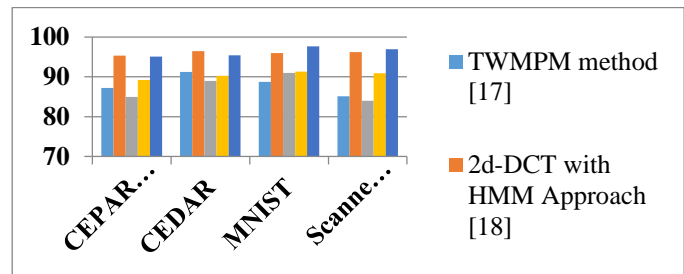


Fig. 11. Graphical representation of the % recognition of the proposed method and other existing methods

### V.CONCLUSION

In this paper a very large database of unconstrained handwritten digit images for experiment purpose was introduced. No author tested with such a large data base. It is proposed to make this database available for research purpose. 100000 handwritten characters collected from 1000 different individuals of different age groups and different areas. No author has attempted to such collection. The proposed method utilizes a small algorithms for preprocessing such as Binarization, normalization and thinning algorithms. The proposed method takes very less amount processor time for recognition of digit image. The main object of the proposed method is an efficient feature extraction method is derived for handwritten digit recognition. For thinning digit image, a well-known technique is used i.e. improved two-step thinning method. The overall percentage of the proposed method achieves 97.36%. The proposed method is evaluated by Ten Cycle Cross Validation (TCCV) approach to find the effectiveness of the proposed method No such method is available up to now utilizing the large set data base for testing.

### REFERENCES

- [1] J. pradeep, E. Srinivasan, and S. Himavathi, Neural network based handwritten character recognition system with feature extraction. International conference on computer, communication and electrical technology- ICCET 2011, 18th & 19th Mar. 2011.
- [2] W. Wu and Y. Bao, Online handwritten magnolia words recognition based on multiple classifiers, 2009.
- [3] C. L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, Handwritten digit recognition: investigation of normalization and feature extraction techniques, Pattern Recognition, vol. 37, iss. 2, pp. 265-279, 2004
- [4] C. Y. Suen, C. Nadal, R. Legault, T. A. Mai, and L. Lam, Computer recognition of unconstrained handwritten numerals, Proc. IEEE, vol. 80, iss. 7, pp. 1162-1180, 1992.
- [5] D. Gorgevik and D. Cakmakov, "Handwritten Digit Recognition by Combining SVM Classifiers," in The International Conference on Computer as a Tool (EUROCON), 2005.
- [6] M. D. Garris, J. L. Blue and G. T. Candela, "NIST form-based handprint recognition system," NIST, 1997.
- [7] P. J. Grother, "NIST special database 19 hand printed forms and characters database," National Institute of Standards and Technology, 1995.
- [8] X. Chen, X. Liu and Y. Jia, "Learning Handwritten Digit Recognition by the Max-Min Posterior Pseudo-Probabilities Method," in Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), 2007.
- [9] Xu, L., A. Krzyzak and C.Y. Suen (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. IEEE Trans. Syst. Man Cybernet. 22 (3), 418-435.

- [10] Y. LeCun and Y. Bengio. Convolutional networks for images, speech, and time-series. In M. A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*. MIT Press, 1995
- [11] B T Chen, Y S Chen and WH Hsu, "A Parameterized Fuzzy Processor and its Applications," *Fuzzy sets and systems*, Vol 59, pp. 149-172, 1993
- [12] Sehultze-Lam S, Harauz G, Beveridge TJ, "Participation of a cyanobacterial S-layer in fine grain mineral formation", 1992
- [13] Maher Ahmed, Rabab Ward, "A Rotation Invariant Rule-Based Thinning Algorithm for Character Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 24 Issue 12, December 2002, Page 1672-1678
- [14] Peter I. Rockett, 2005, An Improved Rotation- Invariant Thinning Algorithm, *IEEE, Trans. Patt.Anal.MachineIntll.*, Oct, Vol. 27.No.10. pp. 1671-1674.
- [15] Gonzlez R.C and Woods R. E, 2002 "Digital Image Processing text book" 2<sup>nd</sup> edition 2002
- [16] Nafiz Arica and Fatos T. Yarman-Vural, "An Overview of Character Recognition Focused on Off-Line Handwriting", *IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews*, Vol. 31, No. 2, May 2001
- [17] Zhijie Xu, Jianqin Zhang and Hengyou Wang, "Twin Minimax Probability Machine for Handwritten Digit Recognition", *International Journal of Hybrid Information Technology* Vol.8, No.2 (2015), pp.31-40
- [18] Syed Salman Ali, Muhammad Usman Ghani "Handwritten Digit Recognition using DCT and HMMs", *CONFERENCE PAPER* · JANUARY 2014
- [19] Sakshica, Dr. Kusum Gupta, "Handwritten Digit Recognition using various Neural Network Approaches", *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 4, Issue 2, February 2015
- [20] Ms. D. Devikanniga, Dr. M. Thangamani, "Handwritten Digit Recognition using Genetic Algorithm", *International Journal of Innovations & Advancement in Computer Science*, Volume 4, Special Issue May 2015

# QoS web service Security Access Control case study using HTTP Secured Socket Layer Approach

M.Swami Das  
Faculty of CSE, MREC  
Hyderabad,India, 500 100  
msdas.520@gmail.com

A.Govardhan  
Professor,SIT,JNTU Hyderabad  
India, 500 085  
govardhan\_cse@yahoo.co.in

D.Vijaya lakshmi  
Professor,Dept. of CSE  
MGIT,Hyderabad  
vijayadoddapaneni@yahoo.com

## Abstract

Web services are expected to play significant role in message communications over internet applications. Most of the future work is web security. Online shopping and web services are increasing at rapid rate. In this paper, we have presented the fundamental concepts related to web security threats, web server architectures, and web server protocols. QoS web service security is important concern in network communications and firewalls security. we have discussed various issues and challenges of web security including the fundamental concepts of network security encryption, decryption processes and network security hierarchies. In this paper, we also address various issues related to secure communications, and handshake protocol using lossless compression technique to improve the speed of web services. As a case study, it is also about implementing message authentication code using using HTTP secured socket layer approach.

## Categories and Subject Descriptors

C.2 [Computer-communication networks]: Security and protection, Data communication, OSI reference model, and firewalls.

## General Terms

Web services, Security, Quality of web services, and Access control.

## Keywords

Web Security, QoS web service, HTTP, Secure Socket Layer, Network Security, Network operations.

## 1 Introduction

Web services are loosely coupled, language neutral, platform independent, communication and coordination of messages passing between client and server processes. Web services are the functionalities offered by service providers to

end users. Web services are accessible using interfaces, applications over the web. Web services are self-contained, modular business applications that have internet oriented standard interfaces [1]. Web security is required for the organizations to communicate with the customers without any risk of data or information over transactions/web operations. The biggest on-line retailer is Amazon at USA with 44 million users using applications, and web services [5]. Recently attackers used man in the middle attack by changing DNS settings and redirecting information to other servers that affected 300,000 devices in countries like India, Italy and Thailand[4].

Cryptographic researchers proposed digital signatures procedures to protect digital data. Ron Rivest, Adi Shamir and L.Adleman created the first cryptographic algorithm named RSA cryptographic algorithm. The first digital web based certificate was introduced by International Telecommunication Unions Telecommunication standards in 1998, later secure socket layer technology was developed in 1990, and successor Transport layer security in 1999 [26].

It is essential and more significant to provide web security for financial web sites. Web services most widely in Government, Businesses and Individuals with use of internet, and some of the key terms used in web security/network security are: integrity, confidentiality, denial of service, and authentication integrity defined by the property that data has not been modified by non permitted users. Confidentiality means information not available to unauthorized users. Denial of service is a malicious attempt to make server or network not available to the user. Attack occurs when multiple systems flood the bandwidth or resources of under attack system, usually one or more web servers[14]. Authentication makes sure that a message is coming from true source[3]. Three major roles of web server architecture as Service provider that makes service available over Internet, Service requester is the end user requesting services over the Internet; Service registry is the logical centralized directory of services. This is used to provide/publish the services

## 1.1 Web service security by XML and SOAP

Web services security is provided by XML and SOAP specification. Three specific issues with the web services are Confidentiality, authentication and Network security

### 1.1.1 Confidentiality

Confidentiality is to make sure that private or confidential information is not accessible to unauthorized individuals.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

ICEMIS '15, September 24-26, 2015, Istanbul, Turkey  
©2015 ACM. ISBN 978-1-4503-3418-1/15/09...\$15.00  
DOI: <http://dx.doi.org/10.1145/2832987.2833072>

### 1.1.2 Authentication

Authentication is a process of assurance that the communication is between true source and intended destination

### 1.1.3 Network security

Network security is the process of using cryptographic techniques in network protocols and network applications during data communications over web services[3].

## 1.2 Web service security protocol stacks

Web service protocol stack used for web security it has four major layers such as service transport, XML messaging, Service description and service discovery.

### 1.2.1 Service transport

Transport layer security (TLS) can make offered point to point security, but not end to end. It is required TLS ensures message reliability and confidentiality. This layer includes Hyper Text Transfer Protocol (HTTP), Simple Mail Transfer Protocol (SMTP), File Transfer Protocol (FTP), and Blocks Extensible Exchange Protocol.

### 1.2.2 XML Messaging

This layer responsible for encoding messages in a common XML format. Messages can be understood at either end. For example XML RPC, and SOAP.

### 1.2.3 Service description

This layer is responsible for relating the public interface to a specific web service. Service handled by Web service Description Language (WSDL)

### 1.2.4 Service Discovery

This layer is responsible for centralizing the services into a common registry and providing functionality. For example UDDI [2]. The rest of this paper is organized as follows Section 2: Related work, Section: 3 Issues and challenges, Section 4: Web security and Network security, Section: 5 Discussions and Interpretations and Section 6. Conclusion.

## 2 RELATED WORK

Yang Kuihe etl, Proposed a VPN based SSL to provide data security in tunnel when data transferred between client and server. Intelligent IC card is used to authenticate identity, and cipher text is used to interpret the data. Introducing SSL proxy server behind the enterprise firewall. Internet security using SSN VPN system is discussed. Introducing SSL proxy server behind the enterprise firewall, provides Internet security over communication using SSL VPN system architecture technique[25].

The SOAP messages ensure integrity and authentication during the data transmission. Web services require partial signing of SOAP request which is achieved using XML signature by WSDL documents and operations as suggested by Padmanabhuni and Adarkar etl[9,12]. Web service security is critical task for message invocations by web servers. SOAP uses XML encryption, XML digital signature, and SSL/TLS methods.

Web service standards SOAP level security authentication and authorization management. Web security is defined as attach signature and encryption header to SOAP messages. it describes security tokens. Web security policy is defined as set of specifications that describe rules, constraints and other business policies on intermediaries and end points. (Example. Encryption algorithms). Web security trust describes a

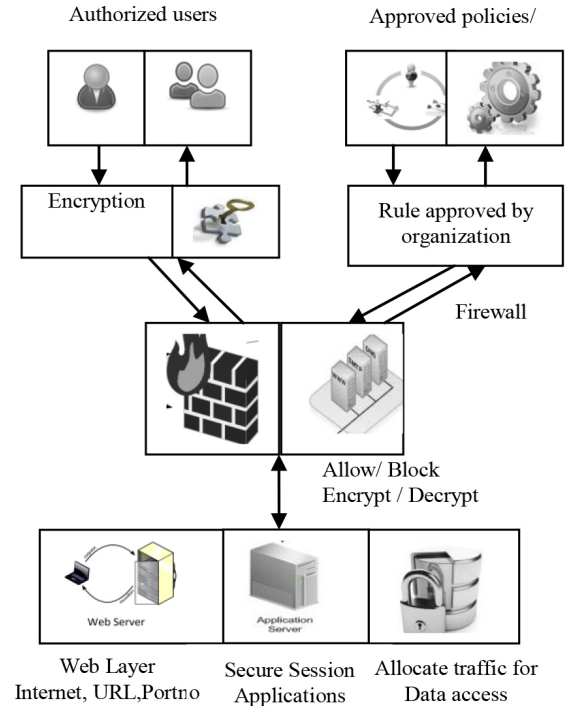


Figure 1. Firewall Architecture

frame work to design a model that enables web services to securely inter-operate request, issues, and exchange security operations.

Web security Privacy: A model web service and requesters state privacy preferences and organizational privacy instructions[9]. Web service securities by standards are firewalls, XML networks, EAR application Server technologies. Firewall is a network security system that monitors and controls the incoming and outgoing network traffic based on rule set/ policies. A firewall establishes a bridge between a trusted, secured internal network to other network[11]. Firewall architecture is shown in Figure.1.

### 2.1 XML Firewall

Web services environment, malicious attacks and DoS attacks are new challenges. Firewall allows Service providers residing in a network to be invoked from outside the network, keeping a high security[9]. HTTP protocol is not suitable for creating public key infrastructure. The prototype is used by application behind a firewall[10].

### 2.2 XML networks

Web service management vendors develop network based solution for web service applications that use peer to peer networks. (Blue Titan Network, Flamenco Network) provide better QoS web services by providing security to various networks endpoints service consumers and service providers. Extension of EAI and application Server Technologies: Enterprise application integration (EAI) product is enabling web based security (like SSL) security infrastructure by Vendors like Cisco, Dlink and others enterprise management vendors are meant for providing Web service standard

solutions to protect enterprise, financial, banking applications[10].

Fang Qi .etl, proposed Automatic Detecting Security Indicator (ADSI) for preventing Web spoofing on a confidential computer which is a harmless environment. if web browser is not trusted, attackers manipulate or steal user passwords, attack on web contents. web content to web server is authenticated with SSL protection. ADSI creates random indicator which is used to identify authenticated web page and detect bogus pages, with URL, screening data[23].

Fengying Wang, etl proposed web based information system that provides user identity and authentication. Web based system by digital certificate information service CA (center agent).After the authentication file is checked and considered correct, CA will send a user Digital certificate by using the RA (register agent). This certificate contains users personal information , public key information and Signature from CA. User needs to encrypt the certificate using the secret key and then send it to administrator,who uses users public key to decrypt and checks the identity. Web browsers and web services provide users authentication and encrypted channelSSL . SSL encryption technology and data package is supposed to achieve high secret communication[18].

Jin-Ha Kim.,etl proposed (Round Robin, ssl with session and ssl with buffering (backend forward) to analyze the performance of SSL enabled application servers. A back end forwarding scheme for getting better server performance through load balancing using ssl with buffer scheme in user level communication, and minimizing intra-cluster communication overhead [19].

Jingli Zhou., etl proposed an Asymmetric SSL tunnel solution which improves VPN throughput. SSL tunnel is the overlay networking technique for creating a SSL VPN on top of Internet ( or IP based network). Before passing SSL tunnel, IP packet is compressed and decrypted in advance using SSL/TLS protocol[20].

Seyed Hasan M., etl proposed Fuzzy Secure Socket layer (FSSL). This protocol is the combination of fuzzy controller and SSL protocol. Functionality is allotted to transport layer like TCP with FSSL protocol. This improves performance in SSL blocks in Cipher Block Chaining mode(CBC) after the MAC is generated. The data is fixed blocks of always eight bytes length. This will lead to vulnerability of cipher text length within easily discloses the length of plaintext. The cipher text is converted to Fuzzy by Fuzzy controller and cipher text length cannot disclose the length of fuzzy data[21].

Lin-Shung Huang., etc introduced a new method for detecting SSL man -in- the middle attacks against website users, feasibility detecting man-in-the middle attack mechanisms over of SSL connections at the top web sites by checking certificates as number of CA certificates. Trace of pervasive malware in the wild that intercepted SSL connections. Multiple available defenses for better protection [22].

S. Bregni, etl. proposed the Secure Socket Layer (SSL) protocol which provides applications by introducing SSL accelerating cards in network nodes security, confidentiality, message integrity and end to end SSL Security, performance improved by minimizing the total cost, end to end delay for completion of distribution[24]. Stefano Bregni.,etl proposed

a solution for SSL to minimize the cost by tabu search approach, which improves the confidentiality, method integration and Authentication [15].

### 3 ISSUES AND CHALLENGES

The following are the list of issues /challenges in Web security/network security. Digital certificates are designed to establish credentials of the people doing business or communicating in the web described in Table.1.

**Table 1. Issues related to web security**

Safety check or vulnerable attacks	Solution
Created fake certificates on trusted companies	Digital certificates authenticated by online checking
Installing Plug-in players Free softwares protected antivirus	Use firewall and Authorized or licensed antivirus software
Various attacks due to entrusted system	Provide security by authentication
Cryptographic methods not available in old browsers	Use new Browsers use SSL security protocols Monitor web site history

Web security developers provide secured connections with server digital certificates, that are useful for on-line transaction to identify the trusted systems.[26] Mal-ware, Denial of service attacks to modem routers against other systems by unknown users by stealing personal information and credentials to access certain web sites. Hackers used stolen laptops/equipment to hack web data where there is vulnerability like private wireless network or wireless network unsecured with no password is immediately accessible to hackers. Hacker used wireless antenna and software, nearby buildings and capture/ steal information like passwords, email-messages, and any data transmitted over the network when a network is not secured.[5]. Brute-force attack is the password cracking method, trying all the solutions seeking one that fits[11].Stealing the login password controlling the devices by malicious scripts and malicious DNS servers, attacked on DSL modems.

Manufactures of Modems/Routers are supposed (like D-Link, Cisco etc) to give a better solution to protect the information for Small offices, homes that use broadband connection devices. If the security devices are not configured properly, the people who want to connect the network and use it, lead to all sorts of issues such as attackers planting malicious DNS server in the device configuration in LAN, DHCP services.

Secure Session Layer: to request URL, XML format formatted router. This will extract the router hardware and firmware, and connect to portno, using SSL to request URL. Admin credentials are not checked by this script. Linsys

**Table 2. Discussions and Interpretations of web security Hacker send phishing mails to web users**

Web services	Solutions or
Problem in domain or threats	Safety precautions
Mail respond with Cyber attacks by phishing mails	Phishing and cyber Do not respond mails with personal data
Credentials stolen from email, accounts	Beware of web browsers environment network

router is aware of vulnerability of CGI script running on these routers [7].

In recent one used Wi-Fi network in traveling. Most virus to computers come from internet resources, internet access points, because wireless access point is least protected by encryption password. If virus are trying to propagate and looks access point, and strongly protected Wi-Fi point on locations as shops, airports. When router is effected malicious virus. Router firmware easily decodes the root password with target sites various domains targeted like face book, Ask, Bing,Google, LinkedIn,Slide Share. Solution is to disable SNMP, ASAP [7].

Routers are poorly protected by default password (Admin). If DNS is to be found malicious wires then that leads to change the DNS settings and cyber hackers steal credentials of personal data like credit card details. When web server applications run with brute force attack HTTP request to malicious DNS server IP address, except navigation temporary files, no files are created in virtual machines, no persistent techniques can be found. Modified DNS settings clone trusted websites. Network router is exposed to attack through mobile phones. Hackers try to send junk mails to the user to give feedback, when the user is trying to send the details of personal information/credentials. The hackers steal the user credentials. Example. Online banking services. The hackers send the junk mails asking the user to send personal details like credit card account no, cvv, pin etc..

#### 4 WEB SECURITY NETWORK SECURITY

Web Service Security: Three types of digital certificate are domain validated certificate, organizational validation certificate and Extended validation certification.

Domain validated Certificate: trusted domain name of owner. Organizational Validation Certificate: validation of organization by DNS names. Extended validation certification: Certificate Agent must meet minimum validation criteria. Organizations, application vendors, Browser makers issue extended validation certificate[26]. Web services standards worked at w3C, OASIS, IETF and other bodies to enable faster inventions of web services and security. A web service provides a flexible set of mechanism to design a range of security protocols. It is essential to design a non-vulnerable protocols for web services security. Web services specifications goals are to provide multiple security token formats, multiple trust domains, multiple signature formats, multiple encryption methodologies, and end to end message

**Table 3. Web service problems with Various attacks**

Web services Problem or threats	Solutions or Safety precautions
Denial of service attack inserting body of the table Cyber attacks by	Use firewall and secured network Do not respond
domain vulnerability and weak security web sites	Provide security by Firewall and Authentication
NTP server Vulnerabilities if visited web site is malicious web site	Beware of web browsers environment network Use anti virus software with firewall
Due to cloud client leaks the sensitive information	Use Authentication by Encryption process
Attackers execute arbitrary weak secret key due to vulnerability	Provide firewall and use strong security policies
Request invocations Content stealing by hackers	provide more security with Secured session layer OTP passwords

content security[12].

#### 4.1 Network Security Encryption and Decryption process

Cryptography is a method of converting plain text into cipher text. Encryption and decryption process shown in Figure.2. Cryptography classified into three types: Secret Key cryptography, Public Key Cryptography and Hash functions

Secret Key cryptography (SKC): It uses a single key for both encryption and decryption

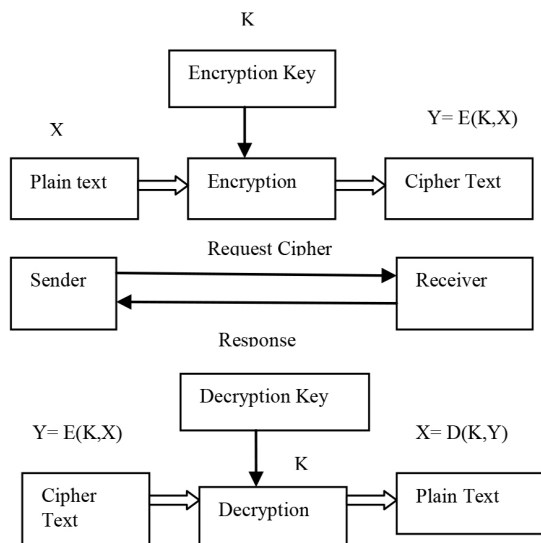
Public Key Cryptography (PKC): It uses one key for encryption and another key for decryption.

Hash Functions: It uses a mathematical transformation function to encrypt the message

Plain text: Actual information is send to web server by files, mail, images, web data processing(i.e. sequence of character etc.).Example. Plain Text= X

EncryptionDecryption key: The number of bits used for encryption or decryption process. Depending on the application the Key is represented by Secret key, public key and Hash functions. For example in Banking transactions user authentication and by using this decryption key confirmed application authorized cleint. Ex. Encryption Key= Decryption key= K

Cipher text: The decoded information by encryption key. The information will be represented by encryption key. (This code is not understood by users) and this information will be transferred over the network to provide security l to



**Figure 2. Encryption and decryption process in networked communication**

the other end of web server/application server. Example. Cipher Text  $Y = E(K, X)$

Sender: sender sends the information after encryption process that converts plain text into cipher text. This cipher text Sender will send cipher text  $Y = E(K, X)$

Receiver: Receiver receives the cipher text (i.e. encoded information) and decrypts key the information that converts cipher text into plaintext. Example Plain Text  $X = D(K, Y)$

## 4.2 ISO/OSI Reference model Fundamentals

ISO OSI (open system Interconnection) reference model. The reference model proposes the following layers (or abstractions) [27, 28].

**Physical Layer (Layer1):** Transmits raw bits the physical link connecting to network nodes. Physical signal is transmitted over hardware. It deals with electrical, mechanical, frequency bandwidth etc. Example. Concerns with bit rate, point to point, multipoint, Network topology. Example Blue tooth.

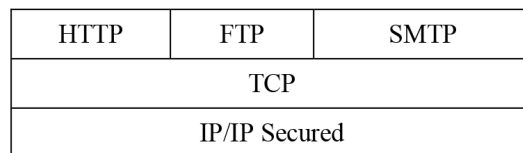
**Data link layer (Layer2):** It recognizes the frame boundaries. It deals with framing, error control, and Flow control. Ex. Point to Point protocol.

**Network layer (Layer3):** Controls the operation on the subnet. Receives the frames from Data link layer and converts into packets. Example. Internet protocol, IPSec.

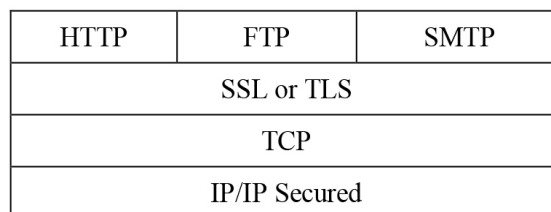
**Transport Layer (Layer4):** This Layer determines the type of service, transport connection, error free point to point channel the delivers the stream of bytes. It deals with reliability, flow controlling and multiplexing. Example TCP, UDP.

**Session Layer (Layer 5):** Allows establishing sessions between various services. Keeps track of token management and synchronization. Communication sessions, Request and response. Example. RPC.

**Presentation layer (layer6):** Concerned with syntax and



**Figure 3. Network Level Security**



**Figure 4. Transport Level Security**

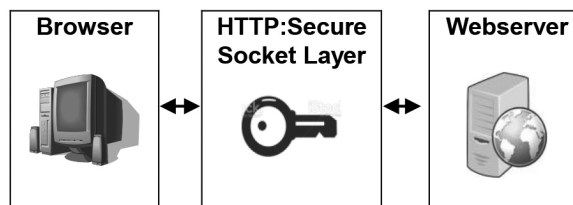
semantics of transmitted information. This gives data structures, delivery information to Application layer. ASCII code, Serializable, XML format, Data conversation, compression, Encryption and decryption are handled in this layer. Example. Telnet.

**Application layer (Layer7):** This layer is concerned with interface that is responsible for display information. Host communicate network interface responsible for display information. P2P, client server model. File transfer, Email, News, HTTP, and DNS.

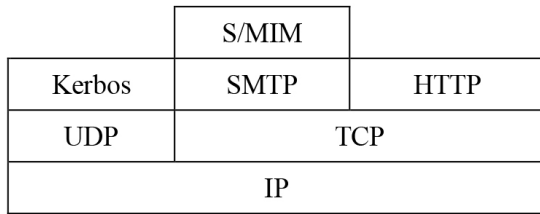
## 4.3 Security in TCP/IP Protocol stack

Security provided in each layer of Network protocol hierarchies Network security provided in three levels as depicted in Figure.3., Figure.4 and Figure.5 We can extend the web security service by providing encryption that uses pre-images in poly to one trapdoor function methodology in web server. The disadvantage of earlier method is hard to invert decryption function in web server and not injunctive way of communication. The advantage proposed approach is easy to invert to recover for pre-images unique by injunctive trapdoor function, helps to ensure authentication. [29]

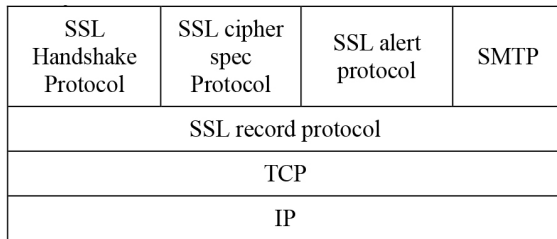
The standards for security in web services to be extended standard mechanism like encryption, digital signature, public key to handle XML and web services. W3C defines a set of services as software systems designed to support inter-operable system to system interactions over a network. W3C web services API implemented in web service and accessed over networks using internetwork.



**Figure 5. HTTP Secured Socket Layer Protocol**



**Figure 6. Transport Level Security**

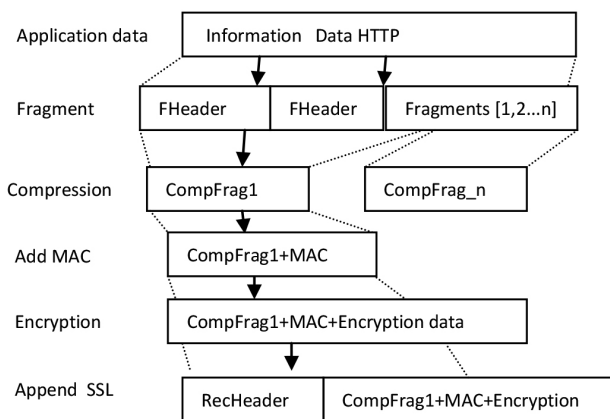


**Figure 7. SSL Protocol stack**

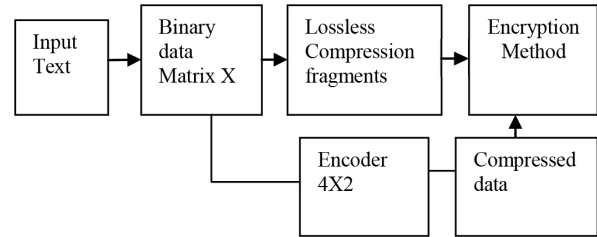
Web services are accessed by sending SOAP messages to endpoints. This is handled by transport layer security protocol such as HTTP, SSL, and TLS others. This ensures secured peer to peer messages. Web based security standards mapping to XML message security. All protocols are used to carry security data as part of XML document. The XML document is critical part of security requirement of web services.[9]

#### 4.4 Secure Socket Layer and Transport Layer security

HTTP Secured Socket Layer Protocol: HTTP over Secured Socket Layer combination to secure communication between browser and web server systems is shown in Figure.6. SSL Secured Socket Layer Protocol is transient, peer to peer communication, SSL protocol stack described in Figure.7, link associated with SSL session Record Protocol operation is shown in figure 8. These SSL sessions are associations between client & server by handshake protocol, with defined set of cryptographic parameters that may be shared by multiple SSL connections[3,12,13]. HTTP protocol stack



**Figure 8. SSL Record Protocol Operation**



**Figure 9. Lossless compression using ENCODER**

provides transfer information for web services interaction which can operate on top of SSL. Three layers are defined as part of SSL such as Hand shake protocol, The change of cipher spec protocol and the Alert protocol. These protocols are used in management of SSL exchange.

##### 4.4.1 HTTP SSL record protocol architecture

Two important concepts are SSL session, and SSL connection. Connection: SSL connection that provides protocol suite is similar to ISO/OSI reference model. SSL connection is on peer to peer network, and sessions are created by the Handshake protocol. Each session is identified by session state, session id, compression method and cipher spec. HTTP secured socket layer approach is described in algorithm.2. Established connection between end to end communication is described in Algorithm.3. Web security access control algorithm is described in Algorithm.4.

The secret key that used in MAC operation on data sent by client client\_write\_mac: The secret key is used in MAC operation on data sent by server and decrypted by the client client\_write\_key. As a case study example Message Authentication Code is implemented in Java program and testing results tomcat web server. The corresponding outputs are shown in Figure.10, Figure.11 and Figure.12. The conventional encryption key for data encryption is decrypted by the client and server. initiation\_vector: Block cipher in CBC model is used and initialization vector is the first initiated by SSL hand shake protocol. seq\_number: each party maintains separated sequence number for transmitted of number for transmitted and received message.

##### 4.4.2 HTTP SSL Message

Message: Message is information as sequence of bytes. Fragments: each message is fragmented into blocks of fixed size. Compress: High level protocol is used for this segment. Compression is done by various techniques SSL-V3, TLS. Encryption algorithm: High security encryption algorithm is used. Proposed compression algorithm is described in Algorithm.1. Final step is SSL record protocol. Is a header consists of the following fields. ContentType: The higher layer protocol 8 bits are used to process the enclosed fragmentation. Major Version: Indicates major version of SSL consist of 8 bits in use for SSL-v. value 3 Minor version: Minor version consists of 8 bits SSL-v3 value is 0. Compression length: 16 bits length of plain text fragment or compressed fragment if compression is used. This is maximum value. The content type is defined by Change cipher spec protocol, Alert protocol and Hand shake protocol.



Figure 10. Client user name and password

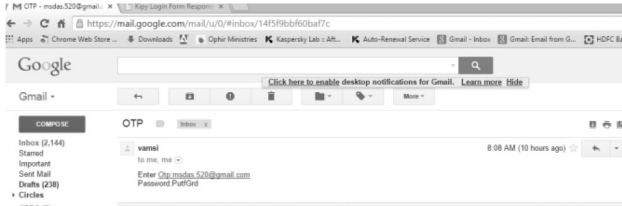


Figure 11. Server sends Random key to client

#### 4.4.2.1 Change Cipherspec protocol

One of the SSL specification protocol that uses SSLrecord protocol. Message state with cipher suite to be used on each connection.

#### 4.4.2.2 Alert protocol

This protocol is used for conveying SSL related alerts to peer entity. Application uses SSL to alert message is compressed and encrypted as specified with current state. Each message has two types of warning(1) or fatal\_error(2) to convey security. Alert SSL specification has unexpected message in appropriate message received, bad\_record\_mac is incorrect MAC received, hand shake failure is security parameters unable to accept, illegal parameters are handshake message was out of range , notify, and certificate messages.

#### 4.4.2.3 Handshake protocol

Most important part of SSL protocol is the hand shaking protocol. This allows the server and client to authenticate each other to negotiate an encryption and MAC algorithm. The cryptographic keys are to be used to protect data in an SSL record. The handshake protocol is invoked before any application data is transferred.

**Phase 1:** Establishes security, protocol versions, session\_id, cipher\_suite, compression\_methods, initial random numbers. **Phase 2:** Server may send certificate key exchange and request certificate server signal of given message. **Phase 3:** Client sends certificate if requested client sends key exchanged. Client may send certificate verification. **Phase 4:** Change cipher suite and finish hand shake protocol [3].

#### 4.4.3 Proposed Model

The problems with SSL in Handshake protocol Encryption and Decryption in Figure.2 are they have slow communication. And another issue is to Monitor every router with any malware/suspicious attacks. To give solution for these problems the proposed architecture as described in Figure. 9. As a case study Message Authentication code in Handshake protocol Phase 3.(key exchange) we implemented sample

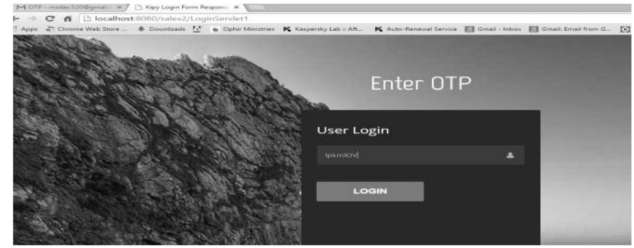


Figure 12. Client open the application using random key mailed to client

program in Java/J2EE language and executed tomcat web server, and results shown in Figure.10, Client gives user name and password, Server receives the client request and sends Random key to client mail id for validation authentication described in Figure.11, and finally client process to access web services by open secret key is shown in Figure.12

#### Algorithm 1 Compression algorithm

- 1: **procedure** ENCODING\_DECODING
- 2:   *begin*
- 3:   *Read Initialization variable Plaintext(binary)=X*
- 4:   *i, n, k, Xreduct(n,k), Xoriginal(n,m)*
- 5:   *Read plain text of binary data*
- 6:   *X= where X is a matrix hold binary data*
- 7:   *x =*

$$\begin{pmatrix} a_{11} & a_{21} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{k1} & a_{k2} & \dots & a_{kn} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$
- 8:   *Divide binary data 8 fragments with each fragment*
- 9:   *Apply reduction by Encoder 4 bits into 2bits*
- 10:   *output Xreduction(n,m)*
- 11:   *Apply Encryption algorithm*
- 12:   *Convert the information into cipher text*
- 13:   *This cipher text transmitted by using HTTP SSL*
- 14:   *Receiver receives the cipher text*
- 15:   *Use Decoder input 2 bits into 4 bits output*
- 16:   *verify the Matrix X*
- 17:   *with each block of binary 8 bits with check sum.*
- 18:   *This decoded data same as Xoriginal data*
- 19:   *end*

## 5 DISCUSSIONS AND INTERPRETATIONS

Secure Socket Layer provides security services between TCP and applications that use TCP. Internet standards TLS, SSL/TLS provide confidentiality using symmetric encryption and message integrity and Authentication code. This protocol as a case study will be used to Secure Electronic Transactions (SET) to protect credit card transactions on web service applications. We have implemented message authentication code by random message to client from Server. The

---

**Algorithm 2** HTTP Secure socket layer

---

```
1: procedure HTTP_SSL
2: SSL required to make to use of TCP , reliable end to end
  secure services
3: input :Web service URL request
4: Ouput : Secured HTTP Session layer
5: Read request of web service
6:   begin
7: Invoke HandShake Protocol
8: Connect: logical link between peer to peer network
9: Session: Created by handshake protocol. Set of crypto-
  graphic parameters
10: Session multiple secured connections are not used.
11: Session established with both read and write.
12: During handshaking protocol session is created success-
  fully
13: pending the new state if already the state is running.
14: With both read and write Session identified with state.
15: Compress data prior encryption
16: Invoke Encryption algorithm
17: Flag is used indicate new connection.
18:   end
```

---

---

**Algorithm 3** To Establish connection

---

```
1: procedure CONNECTION
2: A connection state is defined by the following Parame-
  ters:
3: Connection(clientid,clientMAC,ClientSecretKey,serverid)
4:   begin
5: Client sends request to Web server Invoke HandShake
  Protocol.
6: Each connection creates secure session layer and sets the
  flag.
7: If already connection is established in read/write mode
  do not open new connection else
8: Invoke Handshake protocol Sequence of each request is
  encrypted
9: Decrypt messages at receiver
10: Change CipherSpecProtocol: signal for coordinating
  client server.
11: Alert protocol: Application is allotting compression and
  close_notify alert for each message.
12:   end
```

---

---

**Algorithm 4** Web Security by access control to Encryption

---

```
1: procedure WS_ACESSCONTROL
2: HTTP protocol is used for web transactions into appli-
  cation passed to SSL.
3: HTTP: connection initiation
4: Input: web service request message
5: Output: Secure access control by encryption
6:   begin
7: read HTTP request by SSL to Authorized web server by
  invoking HandShake Protocol
8: Data is encrypted with user details
9: send the request to Web server
10: User data is verified with the data with existing data of
  concerned web server.
11: If(HTTPrequest is successful) then connection is estab-
  lished under service access policy else connection is
  closed with notification
12: if connection is established enable decryption of data at
  the web server
13: Message communication is accomplished by SSL en-
  cryption method.
14: close the connection
15: Connection closure if connection is closed in HTTP
  record
16: TLS level exchange close_notify alert then close TCP
  connection
17: handle TCP close before alert exchange send or com-
  pleted
18:   end
```

---

discussions and interpretations for web securities, precautions and remedies are described in table.2. and table.3. The table provides the information related to web security or network security problems and proposed solutions/precautions to meet network securities QoS parameters such as confidentiality, integrity, data authentication, and availability of information to trusted users from web service systems by merchants by digital certificates and required policy constraints.

## 6 CONCLUSION

Web services are expected to play increasingly important role for message communications over internet applications. Most of future work is web security. Online shopping and web services are increasing in the world. In this paper we described the fundamental concepts related to web security threats, web server architectures, Encryption, decryption and web server protocols. QoS web service security is important concern in network communications. Firewalls security, various issues and challenges of web security. As a case study we presented comprehensive study of HTTP Secured Socket Layer Protocol. In this paper we address the issue of secure communications in web services using HTTP secure socket layer approach. The key issue of communications handshake protocol is using lossless compression technique by encoder to improve the speed of web services during SSL communications between client and server. As a case study using HTTP secured socket layer approach implemented Message Authentication code implemented in Java/J2EE and tomcat

web server results are shown from Figure.10, Figure.11 and Figure 12. In future we propose to extend this paper to E-Commerce, On-line financial transactions, and this security concepts used for designing and developing firewalls which will protect web services applications.

## 7 References

- [1] Gustavo Alonso, Fabio Casati, Harumi Kuno: Web Services: Concepts, Architectures and Applications and Springer publications, 2003, 124-130
- [2] Sanjiva Weerawarana, Francisco Curbera, ony Storey: Web Services Platform Architecture: SOAP, WSDL, WS-Policy, WS-Addressing, WS-BPEL, WS-Reliable Messaging and O'Reilly Publications, 2005
- [3] William Stallings: Cryptography and Network Security, 5th Edition and Prentice Hall India
- [4] [www.livehacking.com](http://www.livehacking.com)
- [5] <http://www.cnet.com/news/the-unvarnished-truth-about-unsecured-wi-fi>
- [6] <http://www.cnet.com/news>
- [7] <https://forums.spybot.info>
- [8] <http://www.cisco.com>
- [9] Khaled M Khan: Managing Web Service Quality: Measuring Outcomes and Effectiveness and Information Science Reference, 290-292
- [10] Zoran Stojanovic and Ajantha Dahanayake: Service Oriented Software System Engineering Challenges and Practice
- [11] <https://en.wikipedia.org>
- [12] [www.cs.ucsb.edu](http://www.cs.ucsb.edu)
- [13] [mercury.webster.edu](http://mercury.webster.edu)
- [14] [www.incapsula.com/ddos/ddos-attacks/denial-of-service.html](http://www.incapsula.com/ddos/ddos-attacks/denial-of-service.html)
- [15] S.Bregni, P.Giacomazzi, A. Poli: Cost-Performance Optimization of SSL-Based Secure Distributed Infrastructures. IEEE Latin America Transactions, Vol. 9, No. 4, July 2011, 550-556, DOI: 10.1109/TLA.2011.5993742
- [16] <http://homepages.uel.ac.uk/>
- [17] Mohammed A. Alnatheer: Secure Socket Layer (SSL) Impact on Web Server Performance, Journal of Advances in Computer Networks, Vol. 2, No. 3, September 2014, 211-217. DOI: 10.7763/JACN.2014.V2.114
- [18] Fengying Wang, Caihong Li, Lei Zhao, Xiumei Li :A comprehensive security policy research on web information system Automation and Logistics, 2009. IEEE International Conference ICAL2009. 1776 - 1780, DOI: 10.1109/ICAL.2009.5262671
- [19] Jin-Ha Kim, Gyu Sang Choi, Das, C.R.: A Load Balancing Scheme for Cluster-based Secure Network Servers IEEE International Cluster Computing 2005. 1 - 10, DOI: 10.1109/CLUSTER.2005.347056
- [20] Jingli Zhou, Hongtao Xia, Xiaofeng Wang, Jifeng Yu :A New VPN Solution Based on Asymmetrical SSL Tunnels FCST '06. Japan-China Joint Workshop on Frontier of Computer Science and Technology, 2006. 71-78, DOI: 10.1109/FCST.2006.4
- [21] Mortazavi S.H., Yazdani. M., Jalilzadeh F., Avadhani. P.S : A novel secure protocol called FSSL using fuzzy controller for Web security IEEE International Conference on Systems, Man and Cybernetics (SMC), 2014, 1192 - 1197, DOI: 10.1109/SMC.2014.6974076
- [22] Lin Shung Huang, Rice.A., Ellingsen.E., Jackson.C.: Analyzing Forged SSL Certificates in the Wild IEEE Symposium on Security and Privacy (SP), 2014, 83 - 97, DOI: 10.1109/SP.2014.13
- [23] Fang Qi; Zhe Tang; Guojun Wang : Attacks vs. Countermeasures of SSL Protected Trust Model The 9th International Conference for Young Computer Scientists, 2008. ICYCS 2008. 1986 - 1991, DOI: 10.1109/ICYCS.2008.433
- [24] Bregni.S., Giacomazzi. P., Poli. A. :Cost-Performance Optimization of SSL-Based Secure Distributed Infrastructures IEEE (Revista IEEE America Latina), 2011, Volume: 9, Issue: 4, 550 - 556, DOI: 10.1109/TLA.2011.5993742
- [25] Yang Kuihe; Chu Xin: Implementation of Improved VPN Based on SSL, 8th International Conference on Electronic Measurement and Instruments, 2007. ICEMI '07, 2-15 - 2-19, DOI: 10.1109/ICEMI.2007.4350641
- [26] Leavitt, N. Computer: Internet Security under Attack: The Undermining of Digital Certificates Volume: 44, Issue: 12 Pages: 17 - 20, DOI: 10.1109/MC.2011.367
- [27] .A.S.Tanenbaum: Introduction to Computer Networks Computer Networks, 3rd Edition, Pearson publications., 2005, 41-45
- [28] Tanenbaum, and Wetherall: Computer Networks 5th Edition, Pearson publications, 2010, 12-15
- [29] Mihir Bellare, Shai Halevi, Amit Sahai, and Salil Vadhan: Many-to-One Trapdoor Functions and Their Relation to Public-Key Cryptosystems, CRYPTO98, Springer-Verlag Berlin Heidelberg 1998, 283-299

# QoS of Web Services Architecture

M.Swami Das  
Faculty of CSE, MREC  
Hyderabad, India  
msdas.520@gmail.com

A.Govardhan  
Professor, SIT, JNTU Hyderabad  
India, 500 085  
govardhan\_cse@yahoo.co.in

D.Vijaya lakshmi  
Professor, Dept. of CSE  
MGIT, Hyderabad  
vijayadoddapaneni@yahoo.com

## Abstract

This paper, web services architectures gives a compatible and scalable structure for web service interactions. The rapid development of web services and applications in various domains such as B2B, e-commerce, banking has led to the best quest for design in a QoS of web service architecture that can meet industry standards. Organizations to design and develop a best suitable QoS web service architecture to meet the industry demands various QoS web service parameters, reliable, quick services required for access the web services. In this paper we proposed a modified 3-tier architecture with a new component quality service manager that can fall in the core layer(i.e middle ware technology). This new component is concerned with the most important QoS parameters such as availability, reliability, adaptability, performance, response time, security and integrity. As a case study we have taken multimedia application and infrastructural components in multimedia applications. The improved quality parameters such as bandwidth, and access time, throughput by using round robin algorithm in the web server interface applications/services. We proposed this new architecture that would help the organizations in best possible QoS service architecture solutions

## Categories and Subject Descriptors

C.4 [PERFORMANCE OF SYSTEMS]: Design studies, Measurement techniques, Modeling techniques, Performance attributes, Reliability, availability, service ability

## General Terms

Architecture, Web services, Quality of web services

## Keywords

QoS, Web services, SOAP, UDDI, Middleware, XML, SOAP, Service Oriented Architecture

## 1 INTRODUCTION

Web services for Service-Oriented Architecture(SOA) give a compatible and scalable structure for web service interactions[1]. Web services are built upon protocols and components such as Simple Object Access Protocol(SOAP), middleware platforms using Object oriented approaches, Extensible Markup Language(XML), Web Service Description Language(WSDL). Universal Description Discovery and Integration(UDDI) open standards over internet standards for application services.

Web service applications that allows the users to access the services from web servers using internet communication standard protocols. Web services provided by organizations as service (internal or external services) to end users. These services can be implemented by (SOAP)[2,3]. Service-oriented architecture represents a method of building reliable, distributed systems that communicate functionality as services[4]. These services from various vendors publish application using Web Service Business Process Execution Language (WS BPEL)[5,6]. Hyper Text Markup Language (HTML) is used to design the front- end server applications like user screens. Extensible Markup Language is used to capture, store and exchange of information between systems. Service Object Access Protocol (SOAP) is used to provide registry of applications in the back end database server. Web server components such as UDDI, WSDL, IDL, CORBA, middleware applications. Organizations used to design and develop software components that provide best functionalities and QoS standards (i.e. reliable secured, performance, cost effective solutions). New web service applications which can be adoptable to industry requirements[19].

### 1.1 Web service

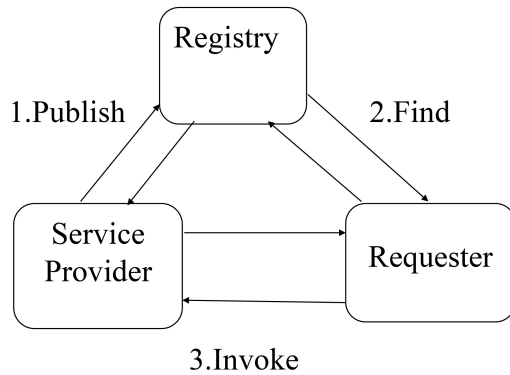
Web services is an interface, and a collection of operations or application programs that are used by clients over internet. Web Services are APIs on internet that allow external applications to interact with system to invoke operations[18,19].

### 1.2 Service Oriented Architecture

Service Oriented Architecture is a set of design methods, that are useful to develop system components and interactions, enables loosely coupled services that can be dynamically discovered for communication and coordination over internet[18]. Web services built upon top of open stan-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

ICEMIS '15, September 24-26, 2015, Istanbul, Turkey  
© 2015 ACM. ISBN 978-1-4503-3418-1/15/09...\$15.00  
DOI: <http://dx.doi.org/10.1145/2832987.2833079>



**Figure 1. major roles/components of Service Oriented Architecture**

dards architecture shown in Figure.1 that has three major roles/components. Service provider that makes service available over internet. Service requesters are the end users requesting HTTP request using internet to utilize the services. Service registry is the logically centralized directory of services. This is useful to provide and publish the services[20].

Service Oriented Architecture Protocol (SOAP) is designed and developed to enable client and server asynchronous interactions request,reply. SOAP is based on HTTP protocol. It uses protocols such as, SMTP, TCP, or UDP. SOAP. Most of API has been implemented in Java, Perl, Java Script, Python, .Net, C, C++, C# and Visual basic. REST (REpresentational State Transfer) is a web service development architecture used for better communication with clients with Internet[18].

### 1.3 XML Extensible Markup Language

XML is a markup language, which has an easy, flexible, format derived from Standard General Markup Language(SGML). This was designed to meet challenges of large scale organizations in electronic publishing. XML is a tool for information management and exchange. XML SOAP is used to encapsulate the message over HTTP[21].

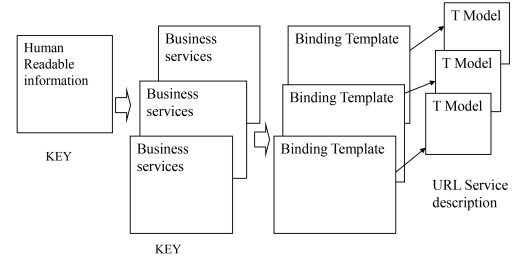
### 1.4 WSDL

Web Service description language is the interface description (Uniform Resource Locator URL i.e. address of the component) for system to access the resources on web server. WSDL definitions are provided by GUI functions in web applications.Example <http://www.jntuh.ac.in> URL is used to access the JNTUH website.

### 1.5 UDDI

Universal Description,Discovery and Integration (UDDI), Web Services Description Language (WSDL) documents are accessed via Uniform Resource Identifies(URI). Client and servers processes communicate either directly or indirectly to access UDDI service. UDDI is a directory service where businesses can be recorded for searching web services.

Human readable businessEntity describes the organization that provides these web services. It gives name, address, activities. For example.[www.jntu.ac.in](http://www.jntu.ac.in) businessServices: It



**Figure 2. UDDI data structure**

stores sequence of operational information as a set of instances of a web service. Example.such as its name and a description (for example Home, Administration, Academics, Research and Development Cell, etc)

bindingTemplate: It has instance of operation or function and service reference to service description,example bindingTemplate menu academics has four service descriptions. TModel: It gives service description usually by WSDL, stored outside the database and accessed via URL, name and its activities shown in Figure.3 Example.JNTUHyderbad database accessed by URL.

### 1.6 Middleware applications

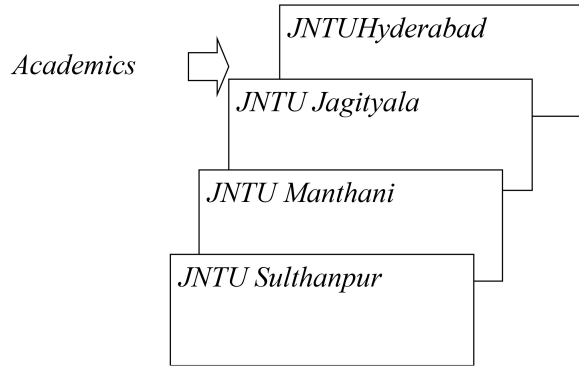
Middleware provides higher level programming representations for development of distributed system through layered architecture to provide heterogeneity in essential infrastructure to support inter-operability and portability. Interfaces developed in middleware platforms and programming languages that have specify signatures of services. Middleware providers access to different infrastructures, platforms and softwares. In Table.1. shows the categories of middleware technologies.

**Table 1. Categories of Middleware**

Major category	Sub category	Example
Distributed objects	Platform	JAVA RMI, CORBA
Distributed Component	Application Server	SUN EJB, CORBA
Public subscribe system	-	CORBA , Event Service
Message Queue	-	Web Sphere, Message Queue, JMS
Web Services system	Web Services Grid services	Apache Axis, The Globus Toolkit

Many applications involve remote invocations that are operated in different environment. The Table.1. describes different categories of middle ware technologies.

It is significant to the enterprise development and the service providers and software developers to maintain QoS ar-



**Figure 3. WSDL URL service description**

**Table 2. Communicate paradigm**

Inter process Communication	Remote Invocation	Indirect Communication
Message Passing	Request Reply	Group Communication
Socket	RPC	Public Subscribe Message Passing
Multicase system	RMI	Tuple space DSM

architecture that meet the client requirements response time, throughput, reliability, and other parameters. The quality standards are essential for designing a new novel architecture to maintain QoS web services. Discussed related work in section 2, Issues and challenges in section 3, Proposed model architecture discussed in section 4, Discussions and interpretations in section 5, and conclusion of the paper specified in section 6.

## 2 RELATED WORK

The evolution of architecture of software started with the monolith approach. A monolith is an architectural style or a software development pattern. Styles and patterns usually fit into different view type (a view type is a set, or category, of views that can be easily reorganized with each other). By evolutionary growth of development of architecture, software developers used this model for their requirements. By the rapid development of internet based applications software developers used client server architecture, this method is called 2-tier architecture. Client is a process (mobile device, or computer system) that communicates with server process( in this application services offered by clients available).The disadvantage of 2-tier architecture is tightly coupled business services and databases. If user want to modify anything in the business operations that will forcibly to change the databases applications. To overcome these problems the software developers used 3-tier architecture which is one of the most widely used architecture for web applications[20,28].

N-Tier Architecture: N-tier application architecture represents a model by which software designers can build flex-

ible and reusable applications, by segregating an application into tiers. N-tier model middleware is involved with developing, tuning, maintaining and evolving developers acquire the option of modifying or adding a specific layer or functionality, instead of modification the entire application[29].

Distributed objects: Distributed object architecture looking at the core architecture elements to design and develop sophisticated distributed systems solutions by layers style.(layers in Distributed systems are application and services,middleware, operating system, computer and network hardware)

Components can be related to objects in that they suggest problem-based abstractions for developing distributed systems. Interfaces or components provide solutions to the client problems. These interfaces also called API(Application Program Interface)

**Table 3. Communicating Entities**

System oriented Entities	Problem oriented Entities
Nodes process	Objects Components Web Services

Services: Web services is the most significant paradigm for rapid growth of Software. Applications are integrated into the World Wide Web(WWW). To design web applications that use communication paradigm and entities. The Table.2 describes the communication paradigm and in table.3 describes communication entities.

To understand the building blocks in web service shown Table 2,3. In designing web applications components used in various layers such as .(i) Architecture design (ii) communication protocols: Inter process, indirect, inter process communication, (iii)Request reply,(iv) method invocation (v),Group communications[18].

Gavin Shrivastava, et.al[22] proposes a method for a web architecture for adaptive location-based services for spatial domain web services. They proposed three tier architecture.

Client tier (allows interaction among the web page), middle tier (provides core services and functionalities) Personalization and visualization services allow the user sessions to handle user profiles, logs all relevant actions performed by clients. The personalization and visualization algorithms are also implemented within this service,that access spatial data sets and recorded interaction. Data sources also provides with spatial information, it has main component consists of a spatial DBMS, which stores spatial datasets.

C Aurrecoechea, et.al[23] proposed generalized frame approach for development of QoS architecture in distributed multimedia systems. They considered QoS assurance in distributed system platform. QoS assurance, complete data and media from remote server across the network points of delivery. Generalized frame work QoS architecture for research is to define a set of QoS configuration interfaces that formalize QoS in the end to end system and network. End to end QoS scenario continuous media flows in the applications area of telecommunication, computer communication and standard

communities using packet scheduling and flow control to improve data delivery and packet transmission.

D. Cotroneo M, et.al[23] proposed architecture that supports different classes of service by evaluating middle ware infrastructure, each with different quality attributes(QoS) concerning the network data delivery, availability, performance of network communications including web,video servers on demand services.

Marco Comuzzi, et.al[26] proposed negotiation broker to which both the consumer and the service provider approach by architecture notify the preferences on QoS attributes and negation strategies by specify value parameters. SOA architecture by service with price, availability and data quality as QoS attribute. Service invocation can be enriched by negotiation of the QoS attributes. SOA architecture by three different phases

Service publication provider who publishes as the service on architecture registry. Service selection consumer selects service by searching the service registry. Service invocation selected service has to be invoked by consumer by specified constraint on the QoS attribute a negotiation process between the consumer and the provider. They suggested architecture for QoS Negotiation (Web service Level Agreement Framework i.e WSLA). These approach selects QoS attributes during web services invocation.

M.Adel Serhani et.al[27] proposed a QoS broker- based architecture for web services architecture is to support the client in selecting web services based on user required QoS. The architecture is developed by four main participating roles the web service broker, the web service provider, the client and QoS- enabled UDDI registry services.

### 3 ISSUES AND CHALLENGES

The software developers design quality software which meets the requirements to overcome the problems of end users. Web services are to provide inter-operability across global internet, B2B applications, commerce etc. These software components must be reusable,efficient, cost effective, and must follow QoS parameters. Web applications interacting with server to access services using different methods, architectures, protocols. Heart of the web service is computing registry(service provider). Service provider will publish large number of function descriptions. All these provided by accessing service using HTTP protocols. Al-Masri et.al[10] suggested provided QoS of web services[8,9,12,13]

#### 3.1 Response time

The time taken to send and receive response from web server ( response time is measured in ms.)

$$RT = RCT - URT \quad (1)$$

where RT:Response time, RCT:Response Completion Time, URT:User Request Time

Software developers need to design a web service architecture that will minimize the response time.

#### 3.2 Availability

Number of successful invocations or total invocations are measured in (%).

$$AV = I/TI \quad (2)$$

where AV: Availability, I: Number of Invocations, TI: Total number of invocations

Availability of web services depends on middle ware, infrastructure, network performance and Service oriented architecture. We can improve the availability of service s by SOAP, UDDI[11]

#### 3.3 Throughput

Throughput is defined as average rate of successful messages received by communication channel per second (throughput is measured in invocationsseconds)

$$TP = TI/T \quad (3)$$

where TP:Throughput,TI:Total number of invocations, T:Period of time

Software kernel implementation is a high performance instruction set, which is required to improve throughput[30]

#### 3.4 Successability

Number of response over the number of requests (is measured in (%))

$$SUCC = Nresp/Nreq \quad (4)$$

where SUCC: Successability, Nresp: Number of response, Nreq:Number of requests

Software components are designed to improve successability

#### 3.5 Reliability

Reliability is measured as the number of failures over a period of time. It is derived from the unsuccessful invocations for a given period of measurement. ( reliability is measured in (%))

$$RE = Merror/Mtotal \quad (5)$$

where RE: Reliability, Merror: Error messages,Mtotal: Total Messages. Reliability to be improved by providing high availability of resources

#### 3.6 Compliance

Extends the WSDL document specification.(is measured in in (%)) Software developers provide insufficient information by WSDL file. WSDL is considered by input, output dependency, invocation sequence, hierarchical function description and concurrent specifications. WSDL are the standards,for the description of language written in XML and it is useful to locate the resource of a web service [14,15].

#### 3.7 Best Practices

The extension to which a Web service that follows, functionality, API, WS-I Basic Profile(measured in in (%)). Best practices like Enterprise Application Integration (EAI), Business to Business (B2B) Integration, SOAP etc.

#### 3.8 Latency

Time consumed for the server to process a given request is called latency (measured in ms)[6,10,16].

$$LAT = RT - Treq \quad (6)$$

where LAT: Latency, RT: Response time,Treq: Requested time. Latency performance depends on several implementations of software components like Middle ware technologies like Simple Object Access Protocol (SOAP), Java RMI, CORBA, UDDI, and Network Protocols.

### 3.9 Documentation

Measure of documentation (i.e. description of tags) in Web Service Description Language(WSDL) (measured in (%)).WSDL document is an XML document that defines a web service consuming of methods to invoke/access other web service applications. WSDL document is described among tags, XML specifications, and elements. WSDL gives the location of web services. This improves in the selection of service description[10,17].

### 3.10 WsRF

Web Service Relevancy Function: a rank for Web Service Quality (measured in (%)).This is used to measure quality of web services metrics (i to ix) to rank the web services described by Masri E[10] Service Classification 1. Platinum (High quality) 2. Gold 3. Silver 4. Bronze (Low quality) The categorization is based on the on the whole quality evaluation(rank) provided by Web Service Relevancy Function(WsRF). Using WsRF values attain for each Web services to a particular service group. The classification can be useful to distinguish between different services that offer the same functionality[10].

## 4 PROPOSED MODEL ARCHITECTURE

Web services are used to provide personalization, API customization, and support for multiple languages. They logically separate the layout Model ,View, and Controller in formats such as HTML, XML. Application centric services enable organizations to integrate applications, business process without constraints of proprietary infrastructure, platforms and operating system. These use open standards, including HTTP, XML, SOAP, WSD, UDDI. Application framework and web services. XML based web service architecture which allows programs written in different languages on different platforms to communicate with each other in a standard way as application frame[18,25].

**Functional Requirements:** In software development, functional requirement defines a function (services or operations) of the system and its components, described as set of inputs, performance and outputs.

**Non Functional Requirements:** Non functional requirements are the resources that specify to assess the operation of the system. For example network bandwidth. Adaptability: modify/update in system configurations and resource availability has been recognized as a further important aspect of service quality. Reliability and security issues are important in the design of most of the computer systems.

Performance is measured by quality of service, it was originally defined in terms of responsiveness and computational throughput, availability, but it has been redefined in terms of ability to meet timeliness, guarantees. Availability depends on communication resource that has guaranteed computing and communication resource.

We proposed a new QoS web service architecture by extending 3-tier Architecture adding new component Quality service manager depicted in Figure in 4.

### 4.1 Application Layer

End-users request applications by through this layer. Clients may use devices, web users and application users.

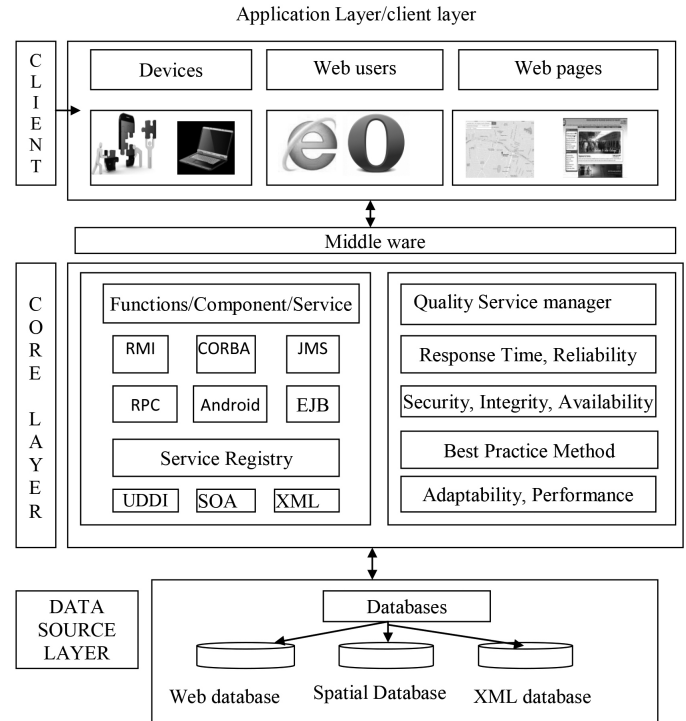


Figure 4. QoS of Web Service Architecture

### 4.2 Middleware

This layer is the core of the QoS architecture. The main concept of this layer to provide business functionality to the user. It has two parts such as Quality service manager, and Service manager.

#### 4.2.1 Quality service manager

Is a software component to select the QoS web service parameters that provide efficient, reliable, and secured services/components

#### 4.2.2 Service manager

Service Manger has various parameters, response time, reliability, availability, security, best practices and performance.

The performance of the services is measured by considering QoS parameters in functional and non functional requirements. To design an efficient Architecture of QoS services by extending three tier architecture[18].We proposed new Algorithm.1 in Quality Service Manager which will improve the QoS web service architecture with QoS web service parameters. We represented in Figure 5 as infrastructural components of multimedia application as a case study, using round robin algorithm in the Web Server interface applications will be improved the performance.

Multimedia applications run in networks of personal computers, they need resources from the work stations running applications(processor,buffer capacity,bus cycles) and network (physical transmission,routers,switches gateways). Workstations may support multimedia stream of data within individual/organization applications.

For multimedia application (non functional requirements, network communication, high bandwidth), and functional

### Algorithm 1 Quality Service Manager

- 1: **procedure** RANKING OF WEB SERVICES
- 2: Input: QWS data parameters related attributes  $D_i = \{A_1, A_2, A_3, A_k, A_i\}$
- 3: output: Component specify their QoS requirement from  $D_i$  to QoS manager to measure Rank of web service or (web service relevancy metrics quality metrics) i.e Service Classification Outcome =  $\{A, B, C, D\}$
- 4: *begin*
- 5: Read the QoS manager parameter  $D_i$  instances
- 6: Apply Classification Procedure to measure the Rank of Web services
- 7: Rank ClassificationMethod( $D_i, O$ )
- 8: If ( tuples will be equal required the rank after step 9 to step 11)
- 9: { Reserve these parameters as resource parameters;
- 10: Application will be proceeded with these parameters;
- 11: Software design notifies QoS manager will be used these parameters }
- 12: else
- 13: { Do not allow the application to be proceed
- 14: Negotiate or application refinement of QoS Manager parameters that will
- 15: Meet the required rank or service classification;
- 16: Give solution to improve the QoS parameters of different type of category;
- 17: To improve the efficiency by computer networks and operating system related parameters or Web services and software related parameters }
- 18: Repeat the steps step 7 to 11 until QWS data  $D_i$  instance (or tuples ) belonging to service classification.
- 19: *end*

Multimedia Service	Data rate (approximate)	Size bits	Size in frequency
Telephone speech	64 kbps	8 bits	8000/sec
CD Quality sound	1.4 Mbps	16 bits	44,000/sec
Standard TV video (uncompressed)	120 Mbps	Upto 640X480 Pixels 16 bits	24/sec
Standard TV video (MPEG-1) compression	1000-3000 mpbs	Variable	24/sec
HDTV video(Uncompressed)	1000-3000 mpbs	Upto 1920X1080 Pixels X 24 bits	24- 60 /sec
HDTV(MPEG-2/MPEG-4 Compression)	6-20 mpbs	Variable	24-60 /sec

Figure 6. Characteristics of Multimedia data stream

Component	In/out	Bandwidth	Latency	Loss rate	Resource required
Web camera	Out	20 frames / second raw video data	--	0	--
Service 1	In	20 frame second, raw video	Inactive	Low	20 ms CPU each 200 ms 20 mbytes RAM
Mixer	In	2X88 Kbps audio	Inactive	Very low	2 ms
	Out	1X88 kbps			2 mbytes RAM
Service 2	In	Various	In active	Low	10 ms CPU each 200 ms
	Out	100 frame s/ second			10 m bytes RAM
Service 3	Network connection	In/out	MPEG-1 stream , approximately 1.5 mbps	In active low	2.5 mbps , low loss stream protocol
Service 4	Network connection	Audio 88 Kbps	In active	Very low	88 kbps, very low stream protocol.

Figure 7. QoS specifications for components of the application shown in Figure.5.

requirements (business services, multitasking shared networks). The central process will perform tasks in the server process, in round robin algorithm scheduling for best effort to put among all the activities running in systems. Networks from different sources network topology, Ethernet, message protocols concepts manage the QoS web service parameters.

The key feature is the resource allocation scheme demanding availability of resources for more quickly completing the task. Round robin approach proposed is shown in Figure.5, best effort method for sharing the processor cycles and network bandwidth. This architecture given late delivery value is less. In order to achieve timely delivery, application need guarantees that the resource will be allocated and scheduled properly.

## 5 DISUSSIONS AND INTERPRETATIONS

From this study, QoS web services architectures, we assume the data as for Figure.6. It gives the characteristics of multimedia data stream the following observations are noted. Telephone speech is less data rate (i.e.bandwidth required or data transfer rate 64 kbps), and time taken to process information frequency size is 8000 seconds. HDTV(MPEG-2/MPEG-4 Compression) multimedia services requires the highest data transfer rate 6-20 Mbps, and time taken to process information frequency size is 24 to 60 seconds. The average case Standard TV video (uncompressed) requires data rate is 120 Mbps, and time taken to process information in frequency is 24 seconds. we proposed a Quality service man-

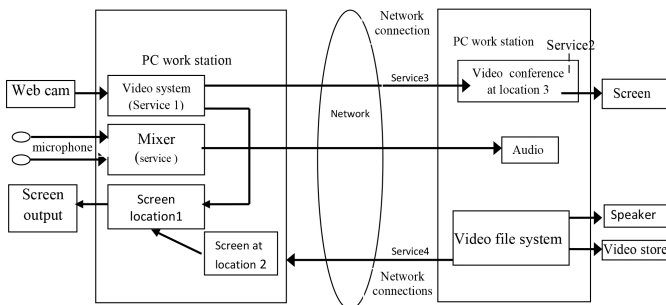


Figure 5. Infrastructural components in multimedia application

ager in the middleware of 3-tier architecture. This architecture suggests Algorithm.1, that gives the QoS parameters, such as response time, reliability and other parameters in the development of web applications.

## 6 CONCLUSION

In this paper we proposed a QoS web service architecture. This is a new model of QoS web service architecture for the rapid development of web services and applications in various domains such as B2B,e-commerce etc., organizations to design and develop the best suitable QoS Web service Architecture to meet the industry demands on various QoS web service parameters, which are reliable, quick services that required for access the web services. In this paper we presented web services fundamental concepts, HTTP,XML,UDDI,SOAP, issues in web services and discussed various architectures and implementation methods. To measure the performance of the services, we proposed a new QoS web service architecture that will meet the demands of organizations and individuals to design and develop QoS web service framework, ranking of web service algorithm which will be applicable for any platform and middleware technologies.

As a case study we have taken multimedia application and infrastructural components in multimedia application and improved quality parameters bandwidth, and access time, throughput by using round robin algorithm in the web server interface applications/services in Figure.7 The new proposed architecture will help the organizations to design and develop solutions using this new QoS service Architecture management and allocation of resource to give guarantees as quality of service management. In future the scope of this paper will extend the work to measure the various QoS web services [10]. Data set QWS from WSDL crawler and other web services data set to measure the performance of web services using various classification methods (i.e data mining techniques).

## 7 References

- [1] Tao Yu,Yue Zhang and Kwei-Jay Lin :Quality of service (qos) in web services: model, architecture and algorithms by University of California, Irvine, CA,ACM, Volume 1 Issue 1, May 2007,Article No. 6,1-26
- [2] [http://en.wikipedia.org/wiki/Serviceoriented\\_architecture](http://en.wikipedia.org/wiki/Serviceoriented_architecture)
- [3] <http://www.webopedia.com/TERM/W/Service.html>
- [4] <http://www.servicearchitecture.com>
- [5] Shrivastava, S.; Sharma, A. : An approach for QoS based fault reconfiguration in service oriented architecture, Information Systems and Computer Networks (ISCON), 2013 International Conference on Year: 2013,Pages: 180 - 184, DOI: 10.1109/ICISCON
- [6] <https://www.oasisopen.org/committees/download.php/10347/wsbpel-specification-draft-120204.htm>
- [7] Marco Crasso,Alejandro Zunino,Marcelo Campo :Easy web service discovery: A query-by-example approach., Elsevier, Science direct 144-163
- [8] <http://www.webperformancetoday.com/2011/06/30/revisiting-the-performance-equation/>
- [9] <http://www.lovelytool.com/>
- [10] QWS dataset <http://www.uoguelph.ca/qmahmoud/qws/>
- [11] <https://www.cs.kent.ac.uk/events/conf/2002/wads/Proceedings/cotroneo.pdf>
- [12] Zibin Zheng, Michael R. Lyu, QoS Management of Web Services, Springer Publications, 2013, 9-37
- [13] <https://www.oasis-open.org>
- [14] Tsai.W.T., Paul. R., Yamin Wang, Chun Fan and Dong Wang: Extending WSDL to facilitate Web services testing 7th IEEE International Symposium on High Assurance Systems Engineering, 2002, 71 - 172, DOI: 10.1109/HASE.2002.11
- [15] <http://www.w3schools.com/webservices>
- [16] Endo.T.,Kaneda.K.Taura.K., Yonezawa. A.High performance LU factorization for non-dedicated clusters IEEE International Symposium on Cluster Computing and the Grid, 2004. CCGrid 2004.678-685,DOI: 10.1109/CCGrid.2004.1336698
- [17] Ben Shil.A,Ben Ahmed.M.: Additional Functionalities to SOAP,WSDL and UDDI for a Better Web Services Administration 2nd. ICTTA '06. Volume: 1, 572-577, DOI: 10.1109/ICTTA.2006.1684434
- [18] George coloulouries, Jean Dollimore, Tim Kindberg and Gordon Blair : Distributed systems concept, 5th Edition,pearson publications
- [19] McArdle, Andrea Ballatore, Ali Tahir, Michela Bertolotto :An open-source web architecture for adaptive location-based services ,the International Archives of the Photogrammetry, RSSIS, Vol. 38, Part II
- [20] M.Swami Das, A. Govardhan, and D.Vijayalakshmi :A survey on web applications, QoS of web services and classification of web services by International conference on Computer and Communication Technologies (IC3T 2K14), CMR Technical Campus,112-116
- [21] <http://www.w3.org/standards/xml>
- [22] Gavin McArdle, Andrea Ballatore, Ali Tahir, Michela Bertolotto :An open-source web architecture for adaptive location-based services. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part II, IEEE, 296-301
- [23] Cristina Aurrecochea,Andrew T.Cambell, Linda Hauw : A survey of Qos architectures, Springer publications,138-151
- [24] <https://www.cs.kent.ac.uk/events/conf/2002/wads/.../cotroneo.pdf>
- [25] Gunjan Samtani, Dimple Sadhwan :Web Services and Application Frameworks (.NET and J2EE)
- [26] Marco Comuzzi and Barbara Pernici :An Architecture

for Flexible Web Service QoS Negotiation, EDOC Enterprise Computing Conference, 2005, Ninth IEEE international Conference.7079

- [27] M.Adel Serhani, Rachida Dssouli, Abdelhakim Hafid and Houari Sahraoui : A QoS broker based architecture for efficient web services selection, Ninth IEEE International,EDOC Enterprise Computing Conference, vol.1. 2005.113 - 120
- [28] <http://www.redbooks.ibm.com/redbooks/pdfs/sg246303.pdf>
- [29] [http://en.wikipedia.org/wiki/Multitier\\_architecture](http://en.wikipedia.org/wiki/Multitier_architecture)
- [30] Montgomery.D.Akoglu.A.: Methodology and Toolset for ASIP Design and Development Targeting Cryptography-Based Applications IEEE International Conference on ASAP.2007.,365 - 370, DOI: 10.1109/ASAP.2007.4459291

## QoS Web Service Security Dynamic Intruder Detection System for HTTP SSL services

M.Swami Das<sup>1</sup>, A.Govardhan<sup>2</sup>, D.Vijaya lakshmi<sup>3</sup>

Assoc. Professor, CSE, MREC<sup>1</sup>, Professor, SIT,JNTU Hyderabad<sup>2</sup>, Professor, Dept. of CSE, MGIT  
Hyderabad, India<sup>3</sup>

**Abstract:** Web services are expected to play significant role for message communications over internet applications. Most of the future work is web security. Online shopping and web services are increasing at rapid rate. In this paper we presented the fundamental concepts related to Network security, web security threats. QoS web service security intrusion detection is important concern in network communications and firewalls security; we discussed various issues and challenges related to web security. The fundamental concepts network security XML firewall, XML networks. We proposed a novel Dynamic Intruder Detection System (DIDA) is safe guard against SSL secured transactions over message communications in intermediate routers that enable services to sender and receiver use Secured Session Layer protocol messages. This can be into three stages 1) Sensor 2) Analyzer and 3)User Interface..

**Keywords:** Web Security, QoS web service, HTTP, Intruder detection, Secure Socket Layer, Network Security

### 1.Introduction

Intrusion detection system is a device or software application that monitors malicious attacks or network traffic if any policy violations [1]. Web services applications communicate and coordinate message passing between client and server. A web service provides functionality and services to the web users. The users to communicate in network channels, the hacker or Intruder tries to operate various attacks such as, DDOS attack, side channel attack, authentication attack, man in the middle attack, cloud computing attacks and steal sensitive information. Hacker execute arbitrary or malicious code in the system due to vulnerability, weak security and no Intruder Detection and Monitoring system.[2]. In Intrusion detection system is a device or software application that monitors malicious attacks or network traffic if any policy violations [1]. A web service provides functionality and services to the web users. The users to communicate in network channels, the hacker or Intruder tries to operate various attacks such as, DDOS attack, side channel attack, authentication attack, man in the middle attack, cloud computing attacks and steal sensitive information. Hacker execute arbitrary or malicious code in the system due to vulnerability, weak security and no Intruder Detection and

Monitoring system.[2]. In recently ISRO website homepage hacked by hackers, other examples related to Government and other web sites discussed[3,4]. It is essential to provide Intruder Detection and monitoring system for Government Institutions, Diplomatic offices, Energy, oil and gas companies, Research Institutions, private equity firms, and activist. Frequently to monitor and control the valid and authorized data operations over the network.

Web security has three important concepts confidentiality, integrity and availability. Confidentiality means Information not available to unauthorized users. Integrity defined by the property that data has not been modified by unauthorized users, and availability means web services are accessible to authorized users with access restrictions.[5] Intrusion: attempting to attack into or misuse the system from outside network or legitimate users of the network, intrusion can be a physical, system or remote intrusion. Automatic Intrusion detection system sensor, Analyzer and user interface. Intrusion Detection systems can be classified as i)Anomaly detection ii) signature based misuse iii) host based iv) network based v) stack based The rest of this paper is organized as follows Section 2: Related work, Section: 3 Issues and challenges, Section 4: Web security and Network

security, Section: 5 Discussions and Interpretations and Section 6. Conclusion.

## 2. Related Work

Zhiwen Bai etl, Proposed DTAD, a dynamic taint analysis detector aiming to protect malicious attacks and vulnerabilities. Attacker process is detected and precision intrusion, signature of collection of virtual systems and comparing network data and log files used to identify the attacks.[6]

Jiang Du etl, studied man in the middle attacker use ARP deception for both sides communication. Man in the middle will generate own public, private and self digital certificate, and this is interactive process validated by Service provider.[7]

Taro Ishitaki,etl proposed intrusion detection system using Neural network, Fuzzy logic, Probabilistic reasoning, Genetic algorithms capable for finding pattern behavior to detect normal and attack conditions[8].

The SOAP messages to ensure integrity and authentication during the data transmission. Web services require partial signing of SOAP request which is achieved using XML signature by WSDL documents and operations as suggested by Padmanabhuni and Adarkar etl[9,10]. Web service security is critical task for message invocations by web servers. SOAP uses XML encryption, XML digital signature, SSL/TLS methods. XML message security is achieved by service oriented security functionality.

Web service standards SOAP level security authentication, authorization management. Web security is defined as attach signature and encryption header to SOAP messages. It describes security tokens. Web security policy is defined as set of specifications that describe rules, constraints and other business policies on intermediaries and end points. (Example. Encryption algorithms).Web security trust describes a frame work to design a model that enables web services to securely inter-operate request, issues, and exchange security operations.

### 2.1 XML Firewall

Web services environment, malicious attacks and DoS attacks are new challenges. Firewall allows to the Service providers residing in a network to be invoked from outside the network, and keeping a high security[9]. HTTP protocol is

not suitable for creating public key infrastructure. The prototype is used by application behind a firewall.[11]

### 2.2 XML networks

web service management vendors develop network based solution for web service applications to provide better QoS web services with security to various networks endpoints service consumers and service providers.[11].

Fang Qi .etl, .proposed Automatic Detecting Security Indicator (ADSI) for preventing Web spoofing on a confidential computer which is a harmless environment. It creates a random indicator to identify and detect bogus pages with URL screening data.[13]

Jaing Du,.etl analyzed as a case study secured socket layer man in the middle attack based on SSL certification interaction. Attacker place computer gives a vital role two communication processes. [14]

Lin-Shung Huang., etc introduced a new method for detecting SSL man -in- the middle attacks against website users, over of SSL connections at the top web sites by checking certificates as number of CA certificates. Trace any malware in SSL connections for identify and provide better protection. [15]

## 3 Issues and challenges

The following are the list of issues /challenges in Web security/network security. Digital certificates are designed to establish credentials of the people use Router configurations with weak vulnerabilities and security policies described in Table.1. Web security developers provide secured operations and safety steps necessary to identify trusted systems. [16]

Table.1. Router or firewall configurations with weak or vulnerabilities

Web services Solutions or threats	Problem in domain or Safety precautions
Web service has arbitrary disclosure policy	Provide strong policies to web services
Passwords stored in browser	Do not save passwords in browser history
Institutions, organizations malicious code attacks, virus	Web security , Frequently monitor network operations. Use SSL security

Malware, Denial of service attacks to modems / routers against other systems by unknown users by stealing personal information and credentials to access certain web sites.

Hackers used stolen laptops/equipment to hack web data where there is vulnerability like private wireless network or wireless network is unsecured with no password is immediately accessible to hackers. Hacker used wireless antenna and software nearby buildings and capture/ steal information like passwords, email-messages, and any data transmitted over the network when a network is not secured. [5].Hacker will use some tools described in table.2. Brute-force attack is the password cracking method, trying all the solutions seeking one fits[11].Stealing the login password controlling the devices by malicious scripts and malicious DNS servers attacked on DSL modems.

### 3.1 Man-in-the-Middle (MIM) attacks

This attack where the attacker secretly relays and possibly alters the communication between two parties who believe they are openly communicating with each other. Attacker intercept all relevant messages by passing between victims and adding extra information. Attackers trying to access the services using fake address, fake certifications. Examples of MIM attacks One provides free Wi-Fi service with malicious software.

#### 3.1.1 ARP Cache Poisoning

Sender and receiver over message communication, PC sends IP packets broad cast to all systems in subnet. ARP(address resolution protocol is not secured protocol).

#### 3.1.2 DNS Spoofing

DNS cache poisoning is a computer hacking attack, where by data is communicated into a Domain Name System (DNS) resolver's cache, causing the name server to return an incorrect IP address, diverting traffic to the attacker's computer (or any other computer).Attackers creating a fake web site by redirecting data to shadow servers.

3.1.3 Session Hijacking: Client to server when session established, the hacker capture cookies information and diverting the session communications to un-trusted systems

#### 3.1.4 Session hijacking attack

Communication over TCP connections. Session normally consists of string of variables used in URL stealing and predicting valid session token to gain unauthorized access to the

web server [17]

**Table.2.Tools and software's used to steal the data**

Web services Solutions or threats	Problem in domain or Safety precautions
Suspicious downloads or plugins	Use firewall in secure network
Terminals with chip card vulnerabilities	Alert any where service by authentication and secret key.

## 4.Web Security Network security

Web Service Security: Three types of digital certificates are domain validated certificate, organizational validation certificate and Extended validation certification. Domain validated Certificate: trusted domain name of owner. Organizational Validation Certificate: validation of organization by DNS names. Extended validation certification: Certificate Agent must meet minimum validation criteria. Organizations, application vendors, Browser makers issue extended validation certificate.[4] Web services standards worked at w3C, OASIS, IETF and other bodies to enable faster inventions of web services and security. A web service provides a flexible set of mechanism to design a range of security protocols. It is essential to design non-vulnerable protocols for web services security. Web services specifications goals to provide multiple security token formats, multiple trust domains, multiple signature formats, multiple encryption methodologies, and end to end message content security.[12]

### 4.1 Intruder Identification and Detection System

#### 4.1.1 Various Attacks

Unauthorized system used to attack on router or servers using various attacks (DDOS attack, side channel attack, Man in the middle attack , Authentication attack and cloud computing attacks) methods practiced due to various reasons like, not secured web site, malicious code, denying encrypt , weak secret keys, vulnerabilities in content security, and policy constraints. In Figure.1. shows the intruder attacks on router.

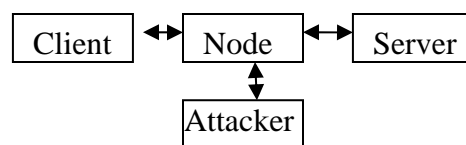


Figure.1. Intruder attacks on services

#### 4.1.1.1 Denial of Service attacks

DDOS attack were launched from distributed attacking hosts. This is launched two phases. First an attacker builds a network which is distributed and consists of thousand of compromised computers are called(Zombies, attacking hosts). The attacker hosts flood of tremendous volume of traffic towards victims either under command or automatically [32].

#### 4.1.1.2 Attack to change DNS settings

Attackers directly targeting DNS server two ways Cybersquatting aim is to steal the Victims identity and or divert traffic from victims website. Name jacking or theft : by appropriate the domain name (updating the holders field or taking control) by technical means to divert the traffic such as modifying the name of hosting the site.[18]

#### 4.1.1.3 Authentication attack

This type of attack targets and attempt to take advantage of following Brute force : allow attacker to guess persons username, other credentials by using Automated trail and error Insufficient Authentication: Allows an attached to access a web site sensible information without having to properly authenticate in web site. Sending phishing mail to user to steal sensitive information[19]

## 4.2 Intruder Detection System

Intruder Detection System has two type namely Network Intrusion Detection System and Host based Intrusion Detection System

### 4.2.1 Network based Intrusion Detection System

It deals with traffic accounting and network flow information. This system is implementing in Routers and switches Input and Output HTTP / TCP data, and testing various functions like port scanning , Reassembling, decoding, detecting virus, protocol violations.

### 4.2.2 Host based Intrusion Detection System

It deals with Analyzing logging facility for almost all failed or success services. The system is implementing in Routers or Firewall to access authorized client. It calculates the cryptographic checks of files, including owner ,group changes, and also checks system integrity.[20]

Web services accessed by sending SOAP messages to endpoints. This is handed by transport layer security protocol such as HTTP,

SSL, and TLS others. This ensures secured peer to peer messages. Web based security standards mapping to XML message security. All protocols use to carry security data as part of XML document. The XML document is critical part of security requirement of web services. [9]

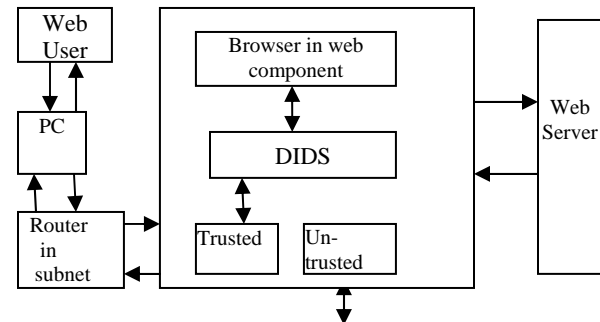


Figure.2. Dynamic Intrusion Detection System

## 4.3 Secure Socket Layer and Transport Layer

Security HTTP Secured Socket Layer Protocol: HTTP over Secured Socket Layer combination to secure communication between browser and web server systems. SSL Secured Socket Layer Protocol is transient, peer to peer communication, SSL protocol stack link associated with SSL session Record Protocol operation. These SSL sessions in association

between client & server by handshake protocol, with defined set of cryptographic parameters that may be shared by multiple SSL connections. [3, 12, 13]. HTTP protocol stack provides transfer information for web services interaction can operate on top of SSL. Three layers are defined as part of SSL such as Hand shake protocol, The change of cipher spec protocol and the Alert protocol. These protocols are used in management of SSL exchange.

## 4.4 Proposed Model: Dynamic Intruder Detection system

The Automated Intruder Detection System shown in Figure. 2. It will detect the unauthorized or hacker requests by invoking a procedure Intrusion Detection System in four subpaths, that are user requests to subnet router point to point in Transport layer, browser in the Intrusion detection system detection system invokes a procedure to check, Certification, digital signature of trusted client. If trusted request as a result then it inserts the process for further processing into Deque. The deque holds a batch of trusted services routed to next hop via point to

point protocol. In subpath3 browser contents security not known to attacker by padlock security. The forth sub-path content in the web server connecting a session request for web services. The algorithm:1,2 and 3 depicted table. Intrusion Detection System Message Format alert:(messageid; create time;nt pstamp;date;time; source;node;address;message; flag)

#### 4.4.1 Components of DIDS

Dynamic Intrusion Detection System has three components are sensor, Analyzer and User interface. Overall network security maintains a security state. when threat occurs by executing an event, the system will check the context of the event and data by following

##### 4.4.1.1 Sensor

Sensors are responsible for collecting data. Example network packet, log files, and system call traces, sensors collect and forward to the analyzer

##### 4.4.1.2 Analyzer

Analyzer receives input from one or more sensors from the system. Control the behavior of the system.

##### 4.4.1.3 User Interface

The user interface to DIDS that enables a user to view output from the system or control the behavior of the system. System component as manager or console component.[20]

#### **Algorithm: 1** Initialization of Request Procedure: DYNAMIC INTRUDER DETECTION SYSTEM

*Input :* Sensor/ node send a service request

*Output:* Trusted service or Un-trusted service  
begin

1. Establish connection between sender and receiver
2. User system to web server consists of four sub paths
3. subpath1: user request to router in subnet(trusted system)
4. Subpath2: Web browser in router checks the procedure using Dynamic Intruder detection system.
5. Identify the request process trusted request pushed onto Deque and un-trusted requests rejected and access restricted.
6. Subpath3: Web browser content and security sign which are not known to the attacker. Ex icon with Padlock security sign unknown to the attacker.
7. Subpath4: Web content to Web server: Connecting via subnet routers with trusted

systems the request connection established between sender to receiver  
using SSL handshake protocol  
end

#### **Algorithm.2: Connection Establishment.**

#### Procedure PROCESS DETECTING TRUSTED REQUEST

*Input :* Web service request

*Output :* Secured HTTP Session layer

Begin

1. Read DIDS
2. Client sends a request to Web server by invoking HandShake protocol using cryptographic parameters (clientid, clientMAC, ClientSecretKey, serverid)
3. During handshaking protocol session is created successfully by resuming the new state if already the state is running. With Session identified by its state, prior to encryption algorithms.
4. Each connection creates a secure session layer and sets the flag. Here flag indicated the connection.
5. The request process is checking by Certificate Agent, and Digital certificate.
6. Detecting trusted service or un-trusted service. if request is trusted service then insert the process in Deque for further processing communication to next hop if connection request is un-trusted requests are denied/ rejected  
end

#### Algorithm.3 Closing the Connection

#### Procedure WS SECURED CONNCTION

*Input:* web service request message

*Output:* Secure access control by encryption

Begin

1. Read PROCESS DETECTING TRUSTED REQUEST
2. User data is verified with the data with existing data of concerned web server.
3. If(HTTPPrequest is successful) then Connection is established, under service access policy else Connection is closed with notification
4. if connection is established enable decryption of data at the web server
5. Message communication is accomplished by SSL encryption method.
6. close the connection
7. Connection closure if connection is closed in HTTP record
8. TLS level exchange close notify alert then close TCP connection
9. handle TCP close before alert exchange send or completed

End

## 5 Discussions and Interpretations

Secure Socket Layer provides security services between TCP and applications that use TCP. Internet standards TLS, SSL/TLS provide confidentiality using symmetric encryption and message integrity and Authentication code. This DIDS (Dynamic Intruder Detection System) protecting the man-in-the-middle attacks and deny services. And allows trusted services forwarded to next hop to reach peer entity web server and Web service applications. The discussions and interpretations for web securities, precautions and remedies are described in table.2 provides the information related to web security/network security problems and proposed solutions/precautions to meet network securities QoS parameters such as confidentiality, integrity, data authentication, and availability of information to trusted users from web service systems to detect various attacks. The genuine merchants by Digital certificates and required policy constraints to validate authentication process in DIDS system architecture.

## 6 Conclusions

Web services are expected to play increasing important role for message communications over internet applications. Most of future work is web security. Online shopping and web services are increasing in the world. In this paper we described the fundamental concepts related to web security threats, web server architectures, web server protocols. QoS web service security is important concern in network communications. Firewalls security, various issues and challenges of web security. Discussed fundamental concepts, network security encryption and decryption process, and Network security hierarchies.

We proposed a novel Dynamic Intruder Detection System(AIDA) is safe guard against SSL secured transactions over message communications to intermediate routers that enable services to sender and receiver use Secured Session Layer protocol messages. This can be into three stages 1) Weak security assumption 2) Intruder attacks on browser 3)Trusted system detect service and safe guard information. As a case study we proposed the architecture of system in Figure .2.In future we can extend this paper to E-Commerce, Online Financial transactions, and this security concepts

used for designing and developing Firewalls which will protect web services applications.

## References

- [1] <https://en.wikipedia.org/>
- [2] [www.livehacking.com](http://www.livehacking.com)
- [3] <http://www.thehindu.com/news/national>
- [4] <http://www.ndtv.com/topic/websites-hacked>
- [5] M.Swami Das, A.Govardhan, and D.Vijya lakshmi: QoS web service Security Access Control case study using HTTP Secured Socket Layer Approach ICEMIS 15, September 24-26, 2015, Istanbul, Turkey 2015 ACM. ISBN 978-1-4503-3418-1/15/09
- [6] Zhiwen Bai,Liming Wang, Jinglin Chen,Jain Liu,Xiyang Liu on " DTAD A Dynamic Taint Analysis Detector for Information Security",IEEE, Web age Information system 2008, pp,591-597
- [7] Jaing Du,Xing Li and Hua Huang :A study of man in the middle attack based on SSL certificate interaction",IEEE,ICIMCCC 2011, pp 445-448
- [8] Taro Ishitaki, Donald , Yi Liu, Tetsuya Oda, Leonard Barolli, and Kazunori Uchida : Application of Neural Networks for Intrusion Detection in Tor Networks.IEEE, ICAINAW 2015, pp 67-72
- [9] <https://events.ccc.de/congress/2005/fahrplan/638-22c3 ids.pdf>
- [10] [www.cs.ucsb.edu](http://www.cs.ucsb.edu)
- [11]Service-oriented Software System Engineering: Challenges and Practices by Z Stojanovi, Ajantha D
- [12]Service-oriented Software System Engineering: Challenges and Practices by Z Stojanovi, Ajantha D
- [13] Fang Qi, Zhe Tang, Guojun Wang on" Attacks vs. Countermeasures of SSL Protected Trust Model", IEEE confernece 2008, pp1886-1991
- [14] Jin-Ha Kim, Gyu Sang Choi and Chita R. Das :A Load Balancing Scheme for Cluster-based Secure Network Servers,IEEE
- [15] Lin-Shung Huang, Alex Ricey, Erling Ellingseny,Collin Jackson :Analyzing Forged SSL Certificates in the Wild,2014 IEEE Symposium on Security and Privacy,pp.83-97
- [16] Neal Leavitt :Internet Security under Attack: The Undermining of Digital Certificates",Technology news in IEEE 2011,pp17-20
- [17] <https://en.wikipedia.org>
- [18] <https://www.afnic.fr/> DNS Types of attack and security techniques
- [19][http://www1.ibm.com/support/knowledgecenter/SB2MG.6.0/com.ibm.ips.doc/concepts/wap auth entication.htm](http://www1.ibm.com/support/knowledgecenter/SB2MG.6.0/com.ibm.ips.doc/concepts/wap_auth entication.htm)
- [20] <https://s2.ist.psu.edu/paper/ddos-chap-gu-june-07.pdf>

# **Optimisation of Characteristics Influence the Value of a Supplier Management ERP Application**

**K.Ramakrishna Reddy, Dr. G. Ravindra Babu, Dr. S. Krishna Mohan Rao**

Assistant Professor, Department of Computer Science Engineering, Malla Reddy Engineering College(Autonomous)  
Hyderabad, Telangana, India.

Professor, Department of Computer Science Engineering, Trinity College of Engineering & Technology, Karimnagar,  
Telangana, India.

Professor, Department of Computer Science Engineering, Siddhartha Institute of Engineering & Technology,  
Hyderabad, Telangana, India.

**ABSTRACT:** According to the earlier researches there is no evidence that the information systems (IS) have any tangible benefits. But, it is very clear that IS can identify less costly sources of inputs for firms by making it cheaper and easier to evaluate prospective supplier offers. According to this, adopting IS should denote a value of information. We now develop a model which relates this potential value to various characteristics. The validity of our model is tested using a random sample of 310 medium and large size manufacturing organizations. The results are strongly supportive of the model, and hence provide support for the hypothesis that IS provides tangible benefits and also has a positive impact on performance of the organization.

**KEYWORDS:** Information Systems, Decision support systems, ERP, Linear regression, organizational characteristics, value of information.

## **I. INTRODUCTION**

Part of the lore of contemporary business is that information systems (IS) are utilized to achieve competitive advantage by lowering costs, speeding the processing of potential input suppliers' offers, enhancing differentiation and changing competitive scope [24], IS applications increase the speed with which corporate functions are performed [13]. Information systems for the manufacturing industry can have a major impact on organizational performance. It has been argued that adopting IS in manufacturing companies can generate substantial reductions in inventory holding costs by minimizing the amount of inventory, reducing raw material costs by enabling the buyer to obtain the best possible conditions, reducing working hours, and lowering response time to market demands [12,18,28,33]. These manufacturing information systems are known as Enterprise Resource Planning (ERP).

Most previous studies attempting to determine the relationship between investment in IS and the performance level of the organization failed to detect a positive relationship. One group of studies compares investment in IS against firm performance, as measured by stock returns [11,14,29]. A second set analyzes industry-level statistics and finds that the marginal dollar of IS investment tends to add less than \$1 to the value of output [6]. A third group, using a relatively small sample of firm level information, fails to detect a significant positive relation between IS spending and productivity [15,17,31]. While these groups of results could indicate the absence of a significant relation between IS and firm performance, each group of results is subject to criticism. First, to the extent that financial investors employ rational expectations, anticipated changes in IS investment (or any other type of investment, for that matter) will have no impact on stock returns; only unexpected changes in IS investment should generate noticeable gains. Thus, the lack of a significant relation between IS investment and stock returns is not evidence that IS investment does not enhance firms' productivity. Second, if firms within an industry invest in IS to varying degrees, it is plausible that big investors will do well partly at the expense of small investors. If this is so, one cannot draw any inferences about the impact of IS investment on firm performance from industry-level statistics. Third, even if one focuses on firm-level data, a significant positive relation can be masked if the performance variables are suspect or the sample is not sufficiently large. Brynjolfsson [25] has argued that aspects often attributed to IS "are precisely the aspects of output measurement

that are poorly accounted for in productivity statistics...". Failing to correct for this mis-measurement can yield underestimates of the benefits from using IS.

One obvious resolution to these difficulties is to use large samples with firm level data, where evidence of benefits is based on the expert assessment of someone in the firm.<sup>2</sup> Further, such an approach should take organizational characteristics into account, since organizations with different characteristics may gain different benefits from using similar individual IS applications [14,19,26]. Porter et al. argues that "competitive advantage cannot be understood by looking at a firm as a whole. It stems from many discrete activities a firm performs in designing, producing, marketing, delivering, and supporting its products. Each of these activities can contribute to a firm's relative cost position and create a basis for differentiation". Porter also indicates role of IS in supporting organization's activities: "Information systems technology is particularly pervasive in value chain, since every value activity creates and uses information." Porter's theory suggests that the impact of the IS should be measured for each activity area by itself, and not for the entire organization as a whole. Barua et al. [20, p. 6] expanded this theory by suggesting that "the primary economic impacts or contributions (to performance) of information technologies (if any) can be measured at lower operational levels in an enterprise, at or near the site where the technology is implemented". That is to say, the contribution of IS to organizational performance should be measured at the lower operational level and not at the aggregated (highest) level of the organization. Barua et al. suggest that "[t]hese effects may then be traced through a chain of relationships within the organizational hierarchy to reveal higher order impacts (if any) on enterprise performance". For example, a direct savings in inventory holding costs (the lower operational level) will contribute to total cost reduction (a higher level), and eventually to organizational performance.

Following this approach, together with Porter's theory, we will look for the benefit IS can provide to a manufacturing organization at the lower level of the organization - the benefit provided by an individual IS application to an individual activity centre. The data we utilize in this study was obtained from surveys of senior managers at 310 manufacturing organizations in Israel. These managers were intimately aware of the benefits gained by using IS for different functional areas in their organization. Each respondent was asked to rate the importance of IS, on a scale of 1 to 7. Respondents were also asked a series of questions about the organization's characteristics, including number of suppliers, relative share of raw materials cost in the cost of the final product, number of customers, average lead time to customers, number of products, number of production lines, volume of sales, and number of employees in the firm. In order to better understand the benefit provided by IS to the lower level of the organization, we follow the approach suggested by Mukhopadhyay et al. [19]. They state (p.7) "for example, in a manufacturing environment, investment in material requirement planning (MRP) systems may improve overall capacity utilization (first-order effect), which, in turn, may lead to a higher return on investment (second-order effect)". To accurately predict the first-order effect, however, more must be understood about the particular manufacturing environment. To this end, we seek to identify the relation between perceived IS benefits and various structural characteristics of the firm.

Since we want to link the benefit from using IS applications to the performance of the organization, we use organizational characteristics that were found to have an impact on the organizational performance [3] as the independent variables for our analysis. We found that these characteristics have a significant impact on the benefit the organization gains from using an individual IS application. By considering both the impact of the certain organizational characteristics on the organizational performance, on the one hand, and the impact of the same variables on the benefit the organization derives from using an individual IS application, on the other hand, we can relate the impact of the individual IS application to the organizational performance. These findings support and expand the findings presented by Mukhopadhyay et al. [19] and Mukhopadhyay et al. [20], Following Porter's approach. Barua et al. [20], and Mukhopadhyay et al. [19,20], we examined the impact of an individual IS application that supports an individual activity area on the performance of the organization as a function of the organizational characteristics.

## II. CONCEPTUAL FRAMEWORK

The MRP application is an important and the most commonly used CIM/MMIS (manufacturing management information system) application [26]. This application can help to plan the ordering of raw materials as close to the time they are needed as possible. This will reduce inventory holding costs on the one hand, and will increase the overall capacity utilization by assuring the availability of raw materials when needed, on the other hand. Yet there are some cases where this will not work. A process plant that produces products such as chemicals, milk, and cans usually

produces these products over and over again. The manufacturer usually dedicates each production line to the manufacture of one product. This type of organization makes sure they have raw materials available by keeping minimum levels of raw material in inventory. Hence, for this type of manufacturer, the MRP system would be of little value. Indeed, a recent study Johansen et al. [29, p. 971] suggested that “CIM is not a generic concept, and a common approach does not work in all cases. Understanding the product process, and plant characteristics associated with successful implementations will assist in evaluating and formulating more effective integration and standardization policies, a poor fit with the firm’s characteristics and/or its product markets could easily lead to failure”.

Mukhopadhyay et al. [19] found that the benefit information systems contribute to the organization is provided by individual applications. Mukhopadhyay et al. [20] found individual IS applications increase the volume and the quality of the output in the mail industry. Based on the Barua et al. [20] study and on Porter’s theory, and the two studies presented by Mukhopadhyay et al. [19, 20], it seems that the benefit a manufacturing organization may gain by using IS should be measured at the lower level of the organization (individual activity center) and should be provided by an individual IS application as opposed to the entire IS applications portfolio. In turn, these benefits are likely to depend on the manufacturing organizational characteristics of the activity area, such as the average number of purchase orders per month for the purchasing activity area, supply time to customers for the sales activity area, etc. [14], Indeed, Bartezzaghi and Francesco [18, p. 46] argue that performance so defined depends on a series of production system structural characteristics (that come from design and management choice of the system itself, besides, of course, from a series of technological and environmental constraints). There are some significant parameters that, taken together, describe the functioning characteristics of the production system. These parameters, which can be called the operating conditions, are, for example, time variables, such as lead time and throughput time; physical measurements, such as lot size; relative parameters such as capacity utilization, percentage of defects, manpower efficiency, etc.

An important part of the MRP application is selecting the optimal supplier(s) of the raw materials the firm uses to produce each of its products. The attributes of a given supplier’s offer might include unit price, quality, delay until delivery (i.e., lead time), special conditions (e.g., volume discounts), and so on. Evaluating a given prospective supplier’s offer can entail computing the implied cost of an order under a variety of scenarios. Comparing two different offers requires analyzing the offers under several contingencies. It follows that identifying the optimal offer out of a set of many offers will commonly require the application of “sensitivity analysis”, whereby the manager considers a variety of possibilities. Clearly, there is a tangible cost associated with considering further information, be it from the further evaluation of existing offers or from the consideration of an additional offer.

Because the manager faces a substantial computational burden in identifying the implied unit cost associated with any supplier, this implied cost cannot be known ex ante. Thus, identifying the optimal offer is a problem of searching for the most attractive opportunity under uncertainty with respect to implied cost. In this context, adopting a decision support IS application has obvious potential value, since it facilitates the evaluation of any potential supplier’s package. That is, it allows the firm to evaluate an additional offer at substantially smaller cost. The value of information associated with the IS application can be traced to the expected cost savings due to the additional searches that can be conducted. That is to say, by using a decision support system a manufacturing organization can identify a less expensive offer for the needed raw materials, and the cost of the search will also be less expensive than looking for the opportunity without using an IS application.

It is well known that marginal cost - the extra cost of producing an extra unit of output - is tied to the input prices faced by the firm.<sup>4</sup> Based on its marginal cost and the demand curve the firm perceives that it faces, the firm determines a profit maximizing level of production. Because both costs and the profit maximizing level of output are ultimately tied to the various input prices, so is maximal profit. In particular, a reduction in the price of any raw materials will typically raise maximal profits. Because profits are a declining function of input prices, the firm has an obvious incentive to seek the best raw materials offer it can find. However, since evaluating an offer is costly, the firm has a disincentive to search. The trade-off between lowering the expected price of raw materials, and so raising expected profits from production on one hand, and bearing increased search costs on other hand, frames firm’s search problem.

The opportunity cost of searching for alternative sources of raw materials will depend on the difficulty of processing the order; in turn, this is likely to be larger the greater is the potential complexity in suppliers’ offers. Such complexity could occur for a variety of reasons. Surely the number of suppliers or the number of orders per period of time is related to the potential complexity, since either of these attributes would raise the number of potential offers to be

# International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 2 , February 2016

evaluated, and so would increase the number of comparisons that must be made. But it is also true that the range of the required sensitivity analysis is linked to the potential differences in suppliers' offers. While it is easy to see how such differences can exist if there are differences in potential supplier's prices, they can also be linked to the possibility of volume discounts. Such possibilities complicate the problem of identifying the optimal input combination, and the associated profit maximizing output, particularly as different scenarios are contemplated. Differences can also occur if there are differences in lead times, since there may be a trade-off between the possibility of obtaining a shipment of inputs closer to the time they are needed against an earlier arriving shipment that contains a larger quantity of the inputs (which increases the inventory holding cost), but is obtained at a volume discount. Thus, we regard the search cost as a function of such features as the number of suppliers, the number of orders, and heterogeneities in suppliers' offers (including prices, volume discounts, lead times, and so on).

The likely benefit associated with a search for alternative sources of raw materials is linked to the significance of raw materials in the firms' total costs. In turn, this can be measured by inspecting the relative share of raw materials in the cost of the final product. For example, consider an organization that has a very large number of suppliers of raw materials, where differences prevail among the suppliers (differences in prices, lead time, etc.). The purchasing manager has to examine a large quantity of data in order to compare the suppliers and make the best purchase with respect to price, lead time, etc. The manager can reduce purchasing cost by 10-15% by using a computerized decision support system to support such decisions [28]. Given this figure, it may be justifiable to implement a sophisticated purchase order IS application to reduce the cost of raw materials. However, if the share of raw material cost relative to the total production costs is low, then even if the firm makes the best decision concerning suppliers, it will have little impact on the organizational performance in terms of cost reduction. For instance, if the material cost is 5% of the total cost of the product, the savings will amount to 0.75% ( $15\% \times 5\%$ ) of the cost of the finished product. In this case, a decision support system that helps the purchasing manager to compare and evaluate more suppliers in order to get the best purchase order (the lowest cost) is not particularly valuable.

The firm's size can also affect the value it places on the ability to consider additional potential offers at lower cost. In general, one expects that the larger the firm the smaller are its unit costs. Correspondingly, firms with lower sales levels are more likely to have larger unit costs and hence a greater potential for cost savings. On the other hand, to the extent that firms can exploit network externalities or scale economies, larger firms stand to gain more from employing IS [8]. Either way, the implication is that firm size may matter.

Before proceeding to the empirical analysis, we briefly discuss two complications of the analysis. In the process we envision, the firm faces a list of potential suppliers. It may have some past experience with some of these suppliers, either because it has ordered from them or because it has considered offers from them. Nevertheless, market conditions are subject to sufficient flux so that this year's offers cannot be assumed identical to last year's, and so each potential supplier's current offer must be analysed. In practice, these offers are likely to possess a great number of characteristics. They may differ in terms of the price charged, in terms of the possibility of volume discounts, in terms of delivery time, or in terms of quality. Their features may prove particularly advantageous in certain market conditions, but less so in others. Despite this apparent complexity, the firm's optimal strategy is to assign a number to each potential offer, reflecting the anticipated gains from analysing that offer, and to analyse offers in decreasing order of these indices [32]. This assignment will generally be based on the various contingencies. It is also tied to the values associated with other offers, so that the general solution involves a fairly complicated dynamic programming problem. Indeed, Weitzman [32] remarks "[t]he actual computation is likely to be a combinatory task of unwieldy proportions unless the number (of offers) is very small". In this interpretation, there is a set-up cost associated with search, which depends on the number of offers being considered. It follows that there is an optimal number of offers to include in the list to be considered, with the usual property that the incremental expected net gain from the last offer placed in the list is non-negative, while including one more offer would generate a negative expected net gain. Plainly, a change that facilitates more rapid and more elaborate computation, such as the adoption of IS, would lower this set-up cost, thereby increasing the optimal number of offers to be considered. In turn, this would lower the reservation price for each offer and raise expected payoff from search [32].

The second complication is that potential suppliers' offers are likely to be correlated if these offers are influenced by some common characteristics. For example, firms supplying paper products to an office will be likely to base their offers on the current and anticipated price of wood pulp; supply shocks in the timber industry would plausibly influence

these paper suppliers in similar ways. To the extent this is true, considering one potential supplier's offer may convey information about the probable outcome of analyzing a second potential offer. This generates an extra benefit from analyzing an offer, but does not really change our fundamental result, that lowering search cost raises the optimal number of offers to process and reduces the expected cost of production [27].

### III. DATA ANALYSIS

The research is based on a sample of 310 manufacturing organizations in Israel. Organizations were randomly selected from each major industrial sector in the Annual Survey of Manufacturing Companies published by the Central Bureau of Statistics in Israel. The companies in the sample varied widely in their characteristics. Table 1 provides information on the minimum, maximum, median, and mean values in our sample for six key characteristics, as well the sample standard deviation. Structured questionnaires were filled out through interviews.

The questionnaire was divided into two sections. The first section included questions about the organization's characteristics (e.g., number of suppliers, relative share of raw materials cost in the cost of the final product, number of customers, average lead time to customers, number of products, number of production lines, volume of sales, number of employees). In the second section the respondent was asked to assess the benefit the organization derived by using several specific IS applications. All the organizations that participated in the study used ERP software packages for managing manufacturing organizations. Hence, all of them had the same idea regarding the possibilities and use of each application. Cases that were reported by the interviewer as different than the meaning of the study were pulled out of the sample. Most of the companies that participated had the same use of the purchasing application and only six organizations were pulled out of the sample.

**Table 1**  
**Distribution of the respondent organizations' characteristics.**

	Min.	Max.	Median	Mean	St. Dev.
Volume of 1990 sales (in million \$ US)	1	400	33	41.78	39.52
Number of employees	10	2400	100	200	321
Number of suppliers	1	5000	45	177	526
Relative share of raw materials in the cost of	2%	85%	44%	43.65%	15.31%
Average number of orders per month	1	3.000	47.5	85.7	249
Average lead time of raw materials (in days)	1	160	37.3	38	27.93

A structured interview was conducted with a senior manager in each organization. The person selected was a user of IS, well aware of the benefits gained by using IS for different functional areas in the organization. The majority of the interviewees were the presidents of the organizations (57%). The other was senior managers of the organization (e.g., Production VP, Finance VP). The interviewers explained each question to the interviewees. Usually the interviewee knew the answer. When the interviewee did not, he/she called the person in charge of the topic and got the answer. An ANOVA analysis revealed no significant difference in the respondents' estimates of the benefit derived by using the information systems as a function of their role in the organization.

To avoid bias due to differences among interviewers, a principal researcher trained each interviewer. First, the candidate interviewer observed the researcher conduct two or three interviews. Next, the interviewer conducted two or three interviews in the presence of the researcher and only afterward began interviewing independently.

During a pilot test in several organizations we found that we could not determine the real benefit the organization gains by using IS because this information was either unavailable or not released by the interviewee. An alternative is to focus on the perceived value of IS, which has been argued to be "very instrumental in assessing the value of an IS that supports unstructured decisions where the outcomes are somewhat intangible or planned for long range" [1], for these reasons we decided to use the perceived value of the benefit. Each respondent was asked: "How would you rank the benefit derived by your organization from each of the following applications on a scale of: 1 2 3 4 5 6 7, 1 indicating very low benefit, and 7 indicating very high benefit?" Ranking the benefit on semantic scales (usually, but not necessarily, ranging from 1 to 7) is a well-tested method for investigating the perceived benefit [1].

Using perceived benefits, as opposed to measured benefits, does have some problems. Because the measure is subjective, there is a concern that responses may not be comparable across firms. That is to say, a dollar impact that one

# International Journal of Advanced Research in Science, Engineering and Technology

Vol. 3, Issue 2 , February 2016

manager regards as significant may strike another manager as merely average. There is also the concern that respondents may not respond truthfully, since they have no disincentive to lie. In particular, the results may be biased if the respondent senses the relationship being examined, and subconsciously provides responses that overstate that relation. While these concerns need to be taken seriously, it is important to bear in mind that it is systematic relations between any biases in response and the explanatory variable in the analysis, and not simply the presence of any errors, that would contaminate the results. The fact that subjective, as opposed to objective, measures of benefits are being used really means that our analysis uses a proxy for the dependent variable of interest. The first of the concerns identified above simply points to the presence of measurement errors in the dependent variable, and does not imply any systematic link between those measurement errors and firm size, number of suppliers, or differences amongst suppliers' features. The second concern speaks directly to the possibility that responses could be biased in a way that varies with the explanatory variable, but here again the concern is overstated.

The fear that respondents could deduce the desired relation from the questions asked is unconvincing, since each questionnaire contained a large number of questions.<sup>5</sup> The dependent variable in our analysis is the perceived value to the organization from using IS for the Suppliers and Purchase Order application. This application manages all the details regarding the organization's suppliers and each raw material purchase order. Specifically this application governs the procurement of raw materials. The information provided by this application helps the organization in planning the purchase of raw materials according to the production plan and the anticipated needs for raw materials. This application also provides the purchasing managers with information regarding previous purchase orders. Such information enables the purchasing managers to conduct a more detailed sensitivity analysis, including a comparison of the benefit derived by quantity discount versus the added holding costs of the inventory for a longer time. It also allows the firm to better negotiate with its suppliers. By using this information the organization can obtain the best prices for raw materials and receive supplies close to the date they are needed.

These potential benefits should be larger the greater the potential for IS to lower search costs or the greater the potential gain from lowering unit costs. The potential reductions in search costs are greater the larger the initial search cost, which should be correlated with the complexity of prospective suppliers' offers. Ceteris paribus, we would expect a greater potential for firms to gain from adopting IS (1) the larger the number of potential suppliers they deal with; (2) the larger the number of orders they make in a typical month; (3) if there are differences between potential suppliers' offers, either in terms of price, lead time, or quality; or (4) if some of the potential suppliers offer price discounts based on quantity. The potential gain from lowering unit costs will be greater the larger the initial share of raw materials in the firm's overall costs. It will also be greater the larger the firm's potential to expand production following a reduction in unit cost. Economic theory predicts that smaller firms (which tend to be higher cost firms) will generally realize a larger increase in equilibrium output and profits for a given reduction in unit costs. Thus, we would expect the potential gain from adopting IS to be positively correlated with the share of raw materials in the firm's costs, and negatively correlated with the firm's size.

Based on the above description, we used the following variables in our empirical analysis:

- (1) the relative share of cost of raw material in the overall cost of the final product;
- (2) a variable reflecting differences among the suppliers' lead times;
- (3) a variable reflecting differences between suppliers' prices;
- (4) a variable reflecting differences between the quality of suppliers' products;
- (5) a variable reflecting possibility of obtaining raw material price discounts as a function of ordered quantity;
- (6) the average lead time of all the raw materials the organization uses (in days);
- (7) the number of suppliers the firm maintains contact with;
- (8) the average number of purchase orders per month;
- (9) a variable reflecting the firm's size (annual sales, in US \$).

Most of these variables are self-explanatory. The variables reflecting differences between potential suppliers and the firm's size require further discussion. Suppliers can differ in terms of supply lead time, price, and quality. When there are differences among the suppliers regarding these attributes, more-combinations have to be checked, and a greater amount of information is needed to choose the best supplier for a specific purchase order. We capture these effects by using dummy variables. The respondents were asked questions like: "Are there differences in lead time among the suppliers? YES/NO." The dummy variable reflecting differences among the suppliers' lead time takes the value 1 if the respondent indicated there were differences and 0 otherwise. The dummy variable reflecting differences among the

suppliers' prices takes the value 1 if there are differences in the suppliers' prices, and 0 otherwise. The dummy variable reflecting differences among the suppliers' qualities takes the value 1 if there are differences in the quality of products from the firm's suppliers, and 0 otherwise. Finally, the dummy variable reflecting the possibility of obtaining raw material price discounts as a function of ordered quantity takes the value 1 if some of the firm's suppliers offer quantity discounts, and 0 otherwise.

There are several potential variables that might reflect firm size, including the number of employees per establishment, the number of employees per firm, value added per establishment, sales per establishment, and the firm's total sales [10]. At first blush, these variables would appear to provide a roughly equivalent picture of firm size. It has been observed, however, that IS tends to induce firms to substitute skilled labour for unskilled labour, with the net effect on employment being indeterminate [7]. Thus, we prefer not to use a measure that is tied to the number of employees. Our model suggests firm size will matter to the extent that smaller firms have a greater potential for cost savings; so that one expects smaller firms to perceive larger benefits. On the other hand, it has been argued that IS may enhance potential networking externalities, which would tend to favour larger firms. Under either scenario, firms with smaller sales differ in a predictable way from firms with larger sales. But having larger sales does not imply a larger value added, since a small firm with a large margin might easily have a greater value added than a big firm with a small margin. Thus, we prefer not to use a measure that is related to value added. For these reasons, we adopt the firm's annual sales, in US dollars, as our measure of firm size.

### IV. RESULT ANALYSIS

We now turn to the task of statistically identifying the relation between perceived benefits for firm  $i$  ( $Y_i$ ) and relevant organizational characteristics. In this analysis, we hypothesize that perceived benefits can be explained by the following variables:

- $X_{1i}$  = the relative share of raw materials in firm  $i$ 's overall cost,
- $X_{2i}$  = 1 if and only if there are differences in lead time among firm  $i$ 's suppliers,
- $X_{3i}$  = 1 if and only if there are price differences among firm  $i$ 's suppliers,
- $X_{4i}$  = 1 if and only if there are quality differences among firm  $i$ 's suppliers,
- $X_{5i}$  = 1 if and only if some of firm  $i$ 's suppliers offer price discounts as a function of ordered quantity.
- $X_{6i}$  = the average lead time of raw materials for firm  $i$ ,
- $X_{7i}$  = the number of suppliers firm  $i$  deals with,
- $X_{8i}$  = the average number of purchase orders firm  $i$  makes per month,
- $X_{9i}$  = firm  $i$ 's sales, in US \$.

Specifically, we consider the following regression equation:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + \eta_i,$$

where  $\eta_i$  is a disturbance term. For now, we assume that the disturbance term satisfies the assumptions in the Gauss – Markov theorem, namely that it is identically and independently distributed across agents, and that this distribution is symmetric about a mean of zero. We discuss the implication of relaxing the assumed homoscedasticity below. We analyzed this regression using ordinary least squares (OLS).<sup>6</sup>

The formal hypotheses we wish to test, and the associated alternative hypotheses are:

$$H_0: \beta_k = 0 \quad (\text{regressor } k \text{ does not matter});$$

$$H_a: \beta_k > 0 \quad \text{for } k = 1, \dots, 8, \text{ or } \beta_9 < 0.$$

The alternative hypotheses are based on the discussion from section 2.

These hypotheses can be tested by comparing the  $J$ -statistic for  $f_{5k}$  against the critical point on a Student's  $t$ -distribution. Because each alternative hypothesis is one-sided, the null hypothesis of no impact in favour of the alternative hypothesis suggested by our model in each case.

**Table 2**  
**OLS regression results.**

Independent variable	Point estimate	f-statistic
Intercept	0.0377	0.09
Relative share of cost of raw materials in	0.05129	10.70 <sup>a</sup>
Lead time differences among the suppliers	0.4615	2.80 <sup>3</sup>
Price differences among the suppliers	0.7163	3.21 <sup>a</sup>
Quality differences among the suppliers	0.2756	1.64 <sup>b</sup>
Price as a function of ordered quantity	0.7597	3.94 <sup>a</sup>
Average lead time of raw materials	0.01528	5.61 <sup>a</sup>
Number of suppliers	0.00028	2.06 <sup>b</sup>
Number of orders per month	0.00117	3.71 <sup>a</sup>
Sales	-0.00440	-2.17 <sup>b</sup>
R – squared = 48.3%		
Adjusted R – squared = 46.5%		

The coefficient on sales deserves particular discussion. Since this coefficient has a negative sign, the implication is that smaller firms perceive larger benefits than larger firms from applying IS to the management of suppliers' orders. This could occur because small firms have larger unit costs than big firms, and so have a larger potential to reduce their raw materials costs. An alternative explanation, suggested by Brynjolfsson [25], is that IS reduces the transactions costs firms must bear when they choose to increase their dealings with outside sources. Ceteris paribus, this will favour firms with smaller internal production, and be less favourable for firms that have large levels of internal production, i.e., firms that are heavily vertically integrated. To the extent that vertical integration has generated a competitive advantage in recent times, one might expect larger firms to have a larger degree of vertical integration. In such an event, smaller firms might perceive a larger potential benefit from using IS to manage the supplier order application. A hypothesis that our results are not supportive of, is that IS favors larger firms.

While these results unequivocally indicate each variable has a statistically important effect on perceived benefits, this does not imply the effect is economically significant. To assess economic importance we must consider the elasticity of perceived benefits with respect to each independent variable. To construct each of these elasticities, we first calculate the average values for perceived benefits,  $K_a$ , and each of the independent

variables,  $X_{ka}$ ,  $k = 1, \dots, 9$ . The elasticity of perceived benefits with respect to the  $k$ th independent variable is:

$$e^* = \beta_k X_{ka} / Y_a.$$

Observe that  $e^*$  may be interpreted as the percentage change in perceived benefits due to a one percent change in the  $k$ th independent variable.<sup>9</sup> Estimates of these elasticity's are given in table 3.

**Table 3**  
**Implied Elasticities for suppliers and purchase order application.**

Independent variable	Point estimate	Elasticity
Relative share of cost of raw materials in	0.05129	0.4706
Lead time differences among the suppliers	0.4615	0.0588
Price differences among the suppliers	0.7163	0.1243
Quality differences among the suppliers	0.2756	0.0792
Price as a function of ordered quantity	0.7597	0.1282
Average lead time of raw materials	0.01528	0.1290
Number of suppliers	0.00028	0.0104
Number of orders per month	0.00117	0.0302
Sales	-0.00440	-0.0386

We note that the independent variable with the largest elasticity is the relative share of raw materials in the overall cost of the final product. This is comforting, since one expects that potential benefits from using an IS application to manage supplier orders would be positively and strongly related to the potential cost savings, which in turn is linked to

raw materials' relative share in costs. It is interesting that the three variables with the next largest impact are "average lead time", "price as a function of ordered quantity", and "price differences among suppliers". This is completely consistent with the trade off we described at the outset, where the firm may have to compare "just in time" delivery of a shipment of raw materials that just meets its current needs against a larger shipment that brings a quantity discount, but forces the firm to bear additional inventory holding costs. The comparison of such possibilities is precisely the sort of "what if" sensitivity analysis that we argue can be facilitated by applying IS. Such comparisons are all the more complex when the firm's potential suppliers offer different prices as well.

One potential criticism of the regression analysis presented above is that it does not allow for idiosyncratic differences across firms in different industries. In particular, it is conceivable that firms in industries that manufacture electronic components will assess different benefits to IS as compared to firms in food processing industries. This could occur because there are intrinsic differences in the benefits from IS in different industries. Alternatively, firms in a given industry may hold similar subjective views towards the potential for IS to lower their costs, and these attitudes could differ from firms in some other industry, whether there are any true differences or not. Idiosyncratic industry effects can show up as systematic differences in perceived benefits, i.e., they may cause the value of the intercept  $\beta_0$  in the regression equation above to differ across industries. Industry specific dummy variables would capture such idiosyncratic differences. Alternatively, the variance of the disturbance term  $\epsilon_i$  could be tied to the industry in which firm  $i$  participates. While OLS still produces unbiased coefficient estimates in such a setting, the associated estimates of the covariance structure are biased. In this case, a weighted least squares approach that allows the variance of the disturbance term to differ across industries would be appropriate.

As a robustness check of our results, we ran a second regression that allowed for both these types of industry specific effects. In this regression, we replaced the common intercept term with dummy variables for each industry. We also allowed for different variances in different industries, and used the information on these different variances to perform weighted least squares. Our survey asked the respondent to identify the industry the firm participated in from the following list: wood, metal, food, textile, rubber, chemistry, paper, electronic, or "other". This allows us to define dummy variables for each of these ten categories. For example, the dummy variable for wood is coded as "1" for every firm that indicated it was in wood and "0" for all other firms. The results of this regression are reported in table 4.

There are several features we wish to point out here. First, the industry specific effects do vary across industries, although the statistical significance of this variation is borderline.<sup>10</sup> Second, the goodness of fit measures is roughly similar to the ones reported above. While there is an increase in  $R$ -squared, this is to be expected because of the inclusion of extra regressors. Finally, most of the coefficients on the other variables are quite similar to their analogs in the original regression, as reported in table 2, although the estimated coefficient on quality differences is no longer significant. We conclude that the results we reported above remain significant when industry specific effects are allowed for. These findings are consistent with findings presented by Ragowsky and Adams. They compared the impact of organizational characteristics on the benefit the organization gains from using IS applications with the impact of the sector (wood, metal, etc.) the organization belongs to on this benefit. They found that specific organizational characteristics have strong impacts on the benefits that organizations gain from using IS applications and that sectorial classification was less important in that respect.

On the whole, our empirical results are corroborative of our discussion in section 2. In addition, they demonstrate a link between organizational characteristics and the value of an IS application for manufacturing organizations.

**Table 4**  
**Weighted least squares regression results, industry dummies included.**

Independent variable	Coeff.	St. Dev.	t – statistic
Intercept	2.4681	0.5165	4.78 <sup>a</sup>
Relative share of raw materials in overall cost of the final product	0.0538	0.0049	10.95 <sup>a</sup>
Lead time differences among the suppliers	0.4129	0.1701	2.43 <sup>a</sup>

Price differences among the suppliers	0.6086	0.2243	2.71 <sup>a</sup>
Quality differences among the suppliers	0.1755	0.1720	1.02
Price as a function of ordered quantity	0.6972	0.1945	3.49 <sup>a</sup>
Average lead time of raw materials	0.0121	0.0029	4.12 <sup>a</sup>
Number of suppliers	0.00026	0.00014	1.87 <sup>b</sup>
Number of orders per month	0.00116	0.00032	3.66 <sup>a</sup>
Sales	-0.0038	0.0020	— 1.85 <sup>b</sup>
Wood	0.1405	0.6455	0.2178
Metal	-0.2061	0.5070	-0.407
Food	-0.8075	0.5271	-1.53
Textile	-0.4237	0.5372	-0.789
Rubber	0.1261	0.5306	0.238
Chemistry	-0.4035	0.5374	-0.751
Paper	-0.3274	0.5656	-0.579
Electronic	0.0310	0.5179	0.060
Construction	-0.5399	-0.6156	-0.877
$R$ – squared = 50.1%			
Adjusted $R$ – squared = 46.6%			

a Estimate significant at 1% level (one-sided test).

b Estimate significant at 5% level (one-sided test).

## V. DISCUSSION AND CONCLUSION

In this paper we have posed and tested a model describing the potential for IS to benefit manufacturing firms by allowing them to analyze a greater number of offers from raw materials suppliers. Interpreting this as a reduction in the cost of making an additional search for a lower cost supplier, we characterize the value of information induced by adopting IS. Our model predicts that the benefits from IS adoption should be positively correlated with the number of potential suppliers, the average number of orders in a typical month; the presence of differences between potential suppliers' offers, either in terms of price, lead time, or quality; the presence of price discounts based on quantity and the share of raw materials in overall costs, and negatively correlated with the firm's size. Using a sample of 310 manufacturing firms, we were able to empirically confirm each of these predictions.

The benefit a manufacturing organization may gain from an IS application is a crucial factor when considering any investment in IS. Most previous studies sought a relationship between the entire IS application portfolio of the organization and the performance of the entire organization. These studies generally fail to detect a positive relationship between the investment in IS and the performance of the organization. We expanded the findings presented by Barua et al. [20] and the two studies presented by Mukhopadhyay et al. [19] and Mukhopadhyay et al. [20]. While the last two studies focused on the impact of individual IS applications on the output and the quality of the process, we focused on the benefit an organization may gain by using an individual IS application, namely applying IS to lower the cost of searching for a cheaper source of raw materials, and the organizational characteristics that impact this benefit. The analyses performed in this study found a relationship behind the benefit a manufacturing organization may gain from using a specific application, as a function of organizational characteristics such as the presence of significant differences among the unit's suppliers, the relative importance of raw materials, the number of suppliers the firm uses, the number of orders it makes, and firm size.

We found that these characteristics have a significant impact on the benefit the organization gains from using an individual IS application. As mentioned in section 2, by using a decision support system, a manufacturing organization can identify a less expensive offer for the needed raw materials. This, of course, will reduce the overall costs of the organization and will contribute to the organizational performance. Yet, the contribution will be more or less significant based on the organizational characteristics (e.g., the value of raw materials out of the total costs of the product). Hence, the decision of what applications to implement should be based on the organizational characteristics.

Obviously there are many more individual IS applications that can be used by manufacturing organizations to support the organization's performance either by lowering costs or by increasing sales. The purpose of this study is to present

and examine a new approach to relate the benefit derived from using IS to the performance of the organization and not to analyze all the possible applications. Our findings indicate that the contribution of IS to the manufacturing organization's performance depends on various organizational characteristics. By using the approach presented in this study, other studies can develop similar models in order to relate other individual IS application to the performance of manufacturing organizations. For example, as we described above, the MRP application can be useful for inventory levels reduction, but not every organization may need it. Some organizations (e.g., process manufacturing plant) will not gain much benefit from using it, and other organizations (e.g., automotive suppliers) may have to pay a penalty if relying on the MRP recommendations. The customers' management application may be more beneficial the more complicated the product is and the more it is adjusted to the customer's specification.

By using the approach presented in this paper, IS managers can better identify the best way to invest in organizational information systems. The business value of IS should be identified for each activity area by itself, and determined for each application rather than an entire portfolio of applications. In addition, the business value of a particular application depends upon how the functionality of the application relates to the organizational characteristics that are related to the particular activity area that is supported by the individual application. These findings suggest that decisions on IS investment should be made at the application level.

#### REFERENCES

- [1] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. Washington, D.C.: American Psychiatric Press, 1987.
- [2] Anthony, W. A. *The principles of psychiatric rehabilitation*. Baltimore, MD: University Park Press, 1980.
- [3] Anthony, W. A., & Blanch, A. Supported employment for persons who are psychiatrically disabled: An historical and conceptual perspective. *Psychosocial Rehabilitation Journal*, XI(3), 5-23, 1987.
- [4] Anthony, W. A., Cohen, M. R., & Danley, K. S. The psychiatric rehabilitation model applied to vocational rehabilitation. In J. A. Ciardiello & M. D. Bell (Eds.), *Vocational rehabilitation of persons with prolonged psychiatric disorders* (pp. 59-80). Baltimore, MD: Johns Hopkins University Press, 1988.
- [5] Arns, P. G., & Linney, J. A. Work, self, and life satisfaction for persons with severe and persistent mental disorders. *Psychosocial Rehabilitation Journal*, 17(2), 63-79, 1993.
- [6] Bond, G. R. Supported work as a modification of the transitional employment model for clients with psychiatric disabilities. *Psychosocial Rehabilitation Journal*, 11(2), 55-73, 1987.
- [7] Bond, G. R. Vocational rehabilitation. In R. P. Liberman (Ed.). *Handbook of Psychiatric Rehabilitation* (pp. 244-263). New York: Macmillan Press, 1992.
- [8] Bond, G. R., Drake, R. E., Mueser, K. T., & Becker, D. R. An update on supported employment for people with severe mental illness. *Psychiatric Services*, 48(3), 335-346, 1997.
- [9] Bond, G., Witheridge, T., Dincin, J., Wasmer, D., Webb, J., & De Graaf-Kaser, R. Assertive community treatment for frequent users of psychiatric hospitals in a large city: A controlled study. *American Journal of Community Psychology*, 18, 875-893, 1990.
- [10] Bybee, D., Mowbray, C. T., & McCrohan, N. Towards zero exclusion in vocational opportunities for persons with psychiatric disabilities: Prediction of service receipt in a hybrid vocational/case management service program. *Psychosocial Rehabilitation Journal*, 18(4), 73-93.
- [11] Cook, J. A., Jonikas, J. A., & Solomon, M. L. Models of vocational rehabilitation for youth and adults with severe mental illness. *American Rehabilitation*, 77(1), 6-11, 32, 1991.
- [12] Derogatis, L., & Melisaratos, N. The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13, 595-605, 1983.
- [13] Drake, R. E., Mclugo, G. J., Becker, D. R., Anthony, W. A., & Clark, R. E. *Journal of Consulting and Clinical Psychology*, 64(2), 391-399, 1996.
- [14] Griffith, R. D. P. A standardized assessment of work behavior of psychiatric patients. *Journal of Psychiatry*, 123, 166-771, 1973.
- [15] Grossman, J., & Tierney, J. P. The fallibility of comparison groups. *Evaluation Review*, 17(5), 556-571, 1993.
- [16] N. Ahituv, Assessing the value of information: Problems and approaches, in: Proceedings of the 10th International Conference on Information Systems (Boston, 1989).
- [17] J.Y. Bakos and M.E. Treacy, Information technology and corporate strategy: A research perspective, *MISQ* 19 107-119, 1986.
- [18] E. Bartessaghi and T. Francesco, The impact of just-in-time on production system performance: An analytical framework, *International Journal of Production Management*, 40-62, 1989.
- [19] A.C. Barua, C. Kniebel and T. Mukhopadhyay. Information technology and business value: An analytic and empirical investigation, *Information Systems Research*, 3-23, 1995.
- [20] A. Barua and A.B. Whinston, An information economics approach to analyzing information systems for cooperative decision making, in: Proceedings of the International Conference on Information Systems (1991).
- [21] E. Brynjolfsson and L. Hitt, Paradox lost? Firm-level evidence on the returns to information systems spending, *Management Science* 42 541-558, 1996.

Volume: 4, Issue: 2  
April - September 2015

ISSN: 2248 - 9584

international journal of  
advances in computer,  
electrical & electronics  
engineering



Editor-in-Chief  
**Dr. Vaka Murali Mohan**



Published by  
**BHAVANA RESEARCH CENTER**

# Case Management Teams Through Integrating Vocational Services

**K.Ramakrishna Reddy<sup>\*1</sup>, G. Ravindra Babu<sup>2</sup> and S. Krishna Mohan Rao<sup>3</sup>**

1. Department of CSE, Malla Reddy Engineering College, Hyderabad, Telangana, India.

2. Department of CSE, Trinity College of Engineering & Techn., Karimnagar, Telangana, India

3. Department of CSE, Siddhartha Institute of Engineering & Techn., Hyderabad, Telangana, India

## KEYWORDS

Case  
Management;

Psychiatric  
Disabilities;

Vocational  
Rehabilitation;

Service  
Integration;  
Mental Health  
Services;

Assertive  
Community  
Treatment;

**Abstract:** Recent innovations to improve employment rates among persons with psychiatric disabilities include “hybrid case management/employment services.” Project WINS was a research/ demonstration project which integrated specialized vocational services into case management teams. In this report, client outcomes of WINS involvement are evaluated, using a quasiexperimental, longitudinal design. On almost all the work-related variables, participants in the immediate and delayed treatment conditions displayed better outcomes than those in the control condition, as did individuals receiving moderate or substantial service versus no/ minimal services. To address possible selection bias due to the quasi experimental nature of the design, further analyses used baseline differences across conditions and participation levels as covariates. Results of multivariate analyses showed some anomalous findings regarding significant positive effects for the delayed, but not the immediate treatment condition versus the no-treatment control group. However, in similar analyses involving participation level as the independent variable, a moderate or substantial amount of service increased the odds of working by almost five times and also positively affected three other work-related variables. While limitations of this quasi experimental design are noted, the results appear promising enough to support replications of WINS.

## 1. INTRODUCTION

The significance of work in the lives of individuals with a psychiatric disability and its therapeutic potential has been widely recognized by professionals and advocates alike. Over the past 30-40 years, a variety of new interventions have focused on vocational training and/or employment opportunities for individuals with severe and persistent mental illness. While demonstrating some positive effects on clients’ vocational outcomes, this first wave of rehabilitation programs has been typified by freestanding, isolated models (e.g., clubhouse, Fairweather Lodge, etc.) not

frequently available in most geographic locations or to most persons with severe mental illness. Also, as noted in several past reviews, the extent to which such vocational programming significantly affects independent, competitive employment rates is problematic (Bond, 1987, 1992; Lehman, 1995).

According to Cook, Jonikas, and Solomon (1991), we now confront the challenge of a “second wave” of vocational issues—developing and expanding vocational programming broadly for all clients. How such widespread availability can occur, however, has not been established. As is the case with other systems (e.g., substance abuse), access to services provided through vocational rehabilitation (VR) agencies is difficult for most persons with psychiatric disabilities. Additionally, individuals with psychiatric disabilities have less placement success in VR than persons with other disabilities (Rimmerman, Botuckl & Levy,

### \* MR. K.RAMAKRISHNA REDDY

Assistant Professor  
Department of CSE  
Malla Reddy Engineering College  
Hyderabad, Telangana, India  
Ph. No: 91-9603016200  
E-mail: ramakrishnareddy524@gmail.com

1995), and VR services do not seem to demonstrate much sustained impact on competitive employment for them (Bond, Drake, Mueser & Becker, 1997).

We now recognize that it is not appropriate or feasible to expect mental health consumers to negotiate unassisted through separate service systems—be they health, housing, or employment supports. Case management services are supposed to do this. However, most case managers have large caseloads and too little time to deal with services other than avoiding hospitalization and meeting clients' basic needs. Many case managers also lack knowledge about rehabilitation interventions or methods to support disabled workers in their jobs (Anthony & Blanch, 1987).

If we are to move beyond the very low employment rates (15% according to Anthony and Blanch, 1987) experienced by psychiatric consumers, we need models of how to integrate vocational services within existing mental health service systems. Several recent reviews indicate that coordinating vocational approaches with clinical care can significantly increase vocational outcomes, especially competitive employment rates (Bond *et al.*, 1997; Lehman, 1995). Russert and Frey (1991) describe how the Program of Assertive Community Treatment (ACT) has addressed this need. Research funded by the National Institute on Disability and Rehabilitation Research has identified other models of "hybrid case management/employment services." Through a variety of structures and settings, these programs utilize innovative ways to integrate vocational services with case management. The programs include an assortment of case management models besides ACT and are therefore applicable to a wider array of service sites than that described by Russert and Frey (1991). The program similarities involve taking an assertive approach to motivating clients toward vocational pursuits, using individually tailored vocational interventions, and providing indefinite follow-along services. Mowbray, Leff, Warren, McCrohan, and Bybee (1997) described six of these programs identified as exemplars. Results have recently been reported from a true experimental design evaluating one of these programs,

Individual Placement and Support (IPS) in New Hampshire. This research found that in comparison to those in a brokered model (offering preemployment training and individual placements with support), clients in the hybrid/integrated program were more likely to be employed and, when employed, earned more wages and worked more hours over an 18-month period (Drake, McHugo, Becker, Anthony, & Clark, 1996).

Project WINS in Grand Rapids, Michigan, was another of the six models of hybrid integrated case management /vocational services. Its purpose was to add resources to expand the vocational focus of existing case management teams in a large, suburban community mental health service system. Project WINS accomplished vocational integration in two ways: (1) vocational specialists were assigned to and worked closely with case management teams, and (2) consumers were employed as peer support specialists to function as case manager extenders, job coaches, role models, and lay counselors. Project WINS was funded as a 3-year demonstration project through the Substance Abuse/Mental Health Services Administration's Center for Mental Health Services. The present report presents the results from a quasiexperimental research design which evaluated client outcomes from the Project WINS demonstration program over an 18-month period. We also contrast our results to a recent review of research on supported employment by Bond *et al.* (1997) and address issues and criticisms which that review raises.

## **BACKGROUND**

Project WINS was sited in Kent County, Michigan, which operates a full management community mental health (CMH) board and provides a comprehensive array of services for persons with severe and persistent mental illness (PSMI). WINS services were housed in two large case management agencies that provide team-based case management services to all PSMI clients of the CMH board, serving a total of more than 1,000 persons, through either Assertive Community Treatment (ACT) or less intensive Comprehensive Community Treatment (CCT) team models.

## WINS SERVICE MODEL

The principles guiding Project WINS operations were based on the Choose-Get-Keep model of vocational rehabilitation, originally developed at Boston University Center for Psychiatric Rehabilitation (Anthony, 1980; Anthony, Cohen, & Danley, 1988). In this model, staff focus on assisting individuals to choose a job or career of interest, to get a job congruent with their choices, and to keep that job so long as the participant finds it desirable. WINS also emphasized self-determination and minimal exclusion—with eligibility based primarily on clients' self selection rather than staff determination of work potential. The major component of Project WINS' operations was the services provided by individual vocational specialists (VS). Seven of eight case management teams (ACT or CCT model) had assigned VS, who provided *direct* services to a rotating caseload of about 20 of the team's clients. Direct services of VS to their clients included vocational planning, assessment, job preparation and choice, job acquisition, and job maintenance. Vocational specialists' activities also included job development and placement efforts on behalf of these clients. Additionally, VS provided *indirect* services to their assigned teams, by attending team meetings and by providing case consultation to case managers on the teams, and sharing information with them about vocational training and opportunities, employment rights of persons with disabilities, etc. The VS were recruited from positions in either rehabilitation or community mental health agencies; all were required to have significant prior experience in working with the other service system. In addition to working with a VS, clients were offered the opportunity to work with a peer support specialist (PSS), another consumer also served by the case management teams; PSS worked as case manager extenders and as role models for assigned clients (see Mowbray, Leff, *et al.*, 1996, and Mowbray, Moxley, *et al.*, 1996, 1998, for an extensive description of the role of the PSS in WINS).

### Eligibility and Entry into WINS

Clients who were unemployed or already working were eligible for WINS services. Exclusion criteria for WINS were (1) clients who were uninterested or unable to

work (including those choosing to be full-time parents or homemakers or those who considered themselves retired), (2) clients in long-term psychiatric care settings, and (3) those whose involvement in the labor force was precluded by severe medical or substance abuse disability or by acute psychiatric symptomatology. Prior to the initiation of WINS services, case managers from the participating teams were asked to identify *all* clients who *could* potentially be served through WINS, utilizing the eligibility criteria. From a total of 690 clients served by the teams, case managers identified 147 clients who fit the exclusion criteria (or 21.3%). Clients classified as ineligible for WINS could be reconsidered at any time.

Clients became direct recipients of WINS services either through self-referral or referral to the assigned VS by case managers on the team. There were always more referrals than WINS staff could accommodate, and an ongoing waiting list was maintained. Each VS was responsible for prioritizing referrals received by discussing each individual's needs and work motivation with the treatment team.

Initial contact with prioritized clients occurred either through the case manager or through direct initiation by the VS. Clients who then confirmed that they wished to receive direct services from a WINS vocational specialist were scheduled to meet with the VS and to begin assessments of their work history, skills, training and support needs, interests, and vocational preferences. The client and the VS then developed an intervention plan that specified vocational goals, anticipated problems, kinds of help needed, and planned intervention activities (both those to be initiated by the VS and those to be followed up by the client).

All WINS services were individualized to address the specific needs and goals of each client. The main intent was to procure jobs of the client's choice in the community, in independent settings; however, more traditional VR placements could be arranged, as appropriate and/or requested by the client. Access to other vocational and related services was also available through WINS: vocational training programs, through VR, community education, or educational programs (GED, community college or

university, or volunteer activity). Community jobs could be transitional or permanent, dependent on the client's needs and preferences. A priority was placed on presenting vocational options in a way that fostered client self-determination, encouraging clients to make decisions about their own vocational endeavors and to exercise choice about the support and assistance they required to attain their goals. The specialized direct services that WINS provided to individual clients were intended to be time-limited, following the process outlined in each client's service plan. Follow-along vocational services were to be provided by the client's case management team, with consultation from the WINS staff available at any time. (Additional information on program operations can be found in Mowbray *et al.*, 1994; Mowbray, McCrohan, & Bybee, 1995; and Bybee, Mowbray, & McCrohan, 1995).

#### **METHOD: Research Design**

A quasi experimental research design was utilized to determine the effectiveness of the WINS vocational service enhancement. Since case management services were provided on a team basis at both agencies and since WINS provided direct services to assigned clients as well as indirect services to the team, individual clients could not be randomly assigned to WINS versus the usual treatment without the likelihood of confounding. Thus, assignment to treatment condition was done by teams rather than by individuals; teams were assigned to one of three conditions: immediate treatment, delayed treatment, and control. This design was selected to maximize the likelihood of detecting change by including three time points and three sets of contrasts to examine. Of the eight available case management teams at the two agencies, one CCT was selected to be the control group. The other seven teams received direct and indirect WINS services according to the following schedule: after a baseline data collection, WINS services were immediately available to three ACT teams at the first case management agency and to one ACT and one CCT team at the other agency. Approximately 9 months later, after a second data collection, WINS services were expanded to two more teams (one each, ACT and CCT) at the second

agency. Referrals to WINS from these five ACT and two CCT teams then continued throughout the duration of the research demonstration period, for about another 9 months. The control condition (the CCT team at the second agency) never received any specialized vocational services. Team assignments to the treatment conditions could not be random, for logistical reasons: The first agency had already begun implementing some less intensive vocational services on some of its teams; therefore its teams could not realistically be assigned to the control or the delayed-treatment conditions. Also, one team strongly asserted a preference not to implement vocational services during the intervention period because of management and staffing issues, so of necessity it became the control team. Finally, the CMH agency was not willing to have more than one of the teams constitute a control condition due to the desirability of initiating vocational services agency-wide.

#### **RESEARCH SAMPLE**

Because of the previously described factors, plus the fact that not all eligible clients could be interviewed due to funding constraints, selection and establishment of the research sample for WINS was somewhat complicated. The research sample consists of samples of clients who were interviewed from the teams which received immediate or delayed treatment, plus a sample of clients from the control condition. Clients from the control condition were not interviewed; however, for them and for the interview sample, data from records and case managers were collected.

Using a power analysis to determine desirable sample-size and incorporating expectations about attrition, a sample of 437 was randomly selected from the pool of eligible on the treatment teams into the interview study, and contacts were attempted; 304 (69.6%) agreed to participate. The remainder either refused ( $N = 73$ ) or were in crisis, unreachable, or ambivalent and could never be scheduled ( $N = 85$ ). There were some significant differences between those who did versus did not consent to interview participation: consenters were less likely to be minority, had higher education levels, were more likely to live independently and had

higher functioning levels, on average (Mowbray *et al.*, 1995). Overall, these differences suggest that consenters (those participating in the interview) were more employable. Of the 304 consenters interviewed, 279 interviews were rated as generally reliable; this established the group of 279 as the interview sample.

Individuals in the interview sample were interviewed at baseline, then at 9 and 18 months; they were also subject to case manager ratings and record reviews. Although all case management team clients were potentially eligible for referral to WINS (except for the control team), interviewing and other data collection was done only for these 279 clients. It should also be remembered that the interview sample of 279 consisted of individuals *potentially* eligible for WINS services. Clients did not have to be in the interview sample in order to receive WINS service; WINS staff also served clients *not* in the interview sample, e.g., those not randomly selected from the eligible pool of 543, those who did not consent to the interview, and new clients entering the agency.<sup>5</sup> The outcome evaluation results presented in this report, however, are based on data from only the 279 clients from the interview sample (from the immediate and delayed-treatment conditions) plus 124 selected for the research sample from the no-treatment control group, for a total of 403.

#### **INTERVIEW SAMPLE CHARACTERISTICS**

About 60% of the interview sample were male, and 80% were White. Minority representation included African Americans (16.8%), American Indians (1.1%), Hispanics (1.1%), Asians (0.7%), and mixed heritage (.4%). The average age was 37 years. The majority were single and had never been married. Most (74%) of the participants were high school graduates ( $M = 12.1$  years). Living arrangements varied: 38% lived alone, 26% in supervised settings, 23% with family, and 14% with friends or roommates. The most common diagnosis was schizophrenia (68%), followed by major affective disorder (21%), with 11% in other categories. Only one third had been hospitalized for mental illness in the previous 2 years. Mean age at onset of mental illness was 19.5 years ( $SD = 12.7$ ). The average Global Assessment of Functioning Scale (GAF Scale, Axis 5 of the DSM-III) was

53 on a continuum of mental health-illness ranging from 1 to 90, which falls into a range indicating moderate symptoms or moderate difficulty with social, occupational, or school functioning (American Psychiatric Association, 1987). Nearly all were taking prescribed psychotropic medications (93.8%). Nearly all the individuals also had a work history, mostly in the previous 2 years. Forty percent were employed at the time of the baseline interview. However, the types of jobs recently held were typically in the unskilled-semiskilled category and provided very low wages. Nearly 90% reported that they wanted to work, and 71% expected to be working in a year. (See McCrohan, Mowbray, Bybee, & Harris, 1994, for a more extensive description of the employment histories, expectations, and desires of the interview sample.)

#### **MEASUREMENTS**

Data were gathered from several sources: (1) one-hour-long, in-person interviews, conducted at baseline by trained undergraduate students in human services fields, (2) reviews of mental health case records, conducted by trained data collectors using structured protocols, (3) checklists indicating clients' vocational and related activities, completed by case managers at the second and third data collection points, and (4) daily logs of WINS services maintained by the vocational specialists.

#### **INTERVIEW**

Interviews gathered information about individuals' living situations (type and tenure), marital status, and employment history (baseline employment status; regency, duration, and pay level of employment; number of jobs held since age 15). Individuals were asked if they wanted to work and if they expected to be working in 1 and 5 years.

Other self-report measures were utilized at baseline and at the 9- and 18-month follow-up interviews: the Work Behavior and Attitudes Scale was developed for the interview as an adaptation of an observational measure of work behavior developed by Griffith (1973). Symptoms were assessed through the Symptom Checklist-10 (SCL-10), 10 items from the Brief Symptom Inventory (Derogates & Melisaratos, 1983). Community functioning was reported with the Areas of Difficulty Checklist (Bond, ... , & ..., 1990, adapted for this study). Self-perception

measures included Pearlin Self-Mastery (Pearlin & Schooler, 1978), Purpose in Life Scale (Reker & Cousins, 1979), and subscales and items from Lehman's Quality of Life Scale (Lehman, 1988, 1991). All scales utilized in the analyses had Cronbach alpha reliabilities of .66 and above (Mowbray *et al.*, 1994).

### **CASE RECORD REVIEWS**

Case records provided basic demographic information (age, race, sex), age at onset of mental illness, most recent psychiatric diagnosis, hospitalizations during the previous 2 years, and two ratings of individuals' functioning: Global Assessment of Functioning (GAF) ratings and Community Living Adaptation Scale (CLAS) ratings (used at this CMH by case managers quarterly to assess community living ability, including residential status, daily structure, financial management, etc.).

### **CASE MANAGER REPORTS**

At the 9- and 18-month follow-up time points, case managers were asked to complete a checklist summarizing clients' vocational and related activities in the last 6 months (CVAs).

### **SERVICES**

Service activity logs (SALs) were completed daily by vocational specialists and peer support specialists on the clients served directly by WINS. Information from the SALs included direct contacts and any other service activity on behalf of clients on their caseloads, and amounts of staff and client time for each client service activity. "Substantial" WINS service was defined as total VS service totaling 6 hr or more. "Moderate" service consisted of total contact time from 1 hr to 6 hr. The "minimal" service category involved clients who were referred to WINS, but received fewer than 60 min of staff contact. Those receiving no service were clients in the control condition, as well as those on the treatment teams who were never referred to WINS; usually this was because the clients on the team were institutionalized or otherwise not available for referral or not interested (either because they were already satisfied with their vocational activity, or they were not motivated toward vocational pursuits at the time).

Additional data on the types of jobs which WINS clients obtained while receiving

project services were compiled through job data sheets completed by VS at the end of the project on WINS clients served.

The job data gathered included the type of activity (whether it was a paid job, training, volunteer work, school, or day treatment activity); the setting of paid jobs (competitive, sheltered, affiliated with a rehabilitation agency); a description of the job title, duties, and place of employment; the dates of employment; and reasons for termination of the jobs. The job data sheets only describe jobs held by the clients during the WINS demonstration period and jobs of which the VS were aware. The data on job duties and place of employment were used to code jobs into activity categories according to the Department of Labor's Dictionary of Occupational Titles (1991), as well as prestige categories according to Hollingshead's prestige scale (Hollingshead & Redlich, 1958).

### **FOLLOW UP**

Research assistants were given training in effective longitudinal tracking and interviewing methods (Ribisl *et al.*, 1996). They utilized extensive contact information which had been collected from participants at the baseline interview. At the second and third interview time points, 225 (80.6%) and 233 (83.5%), respectively, were reinterviewed. Data were available on all participants from case manager ratings (CVAs) and service activity logs (SALs).

### **DATA ANALYSIS PLAN**

Logistic regression analysis was used to examine the effects of condition and service amount on Time 3 outcomes. The outcomes examined included several dichotomous behavioral outcomes related to work or schooling: any paid work (competitive or sheltered workshop or enclave), work for more than 10 hours per week, competitive work, seeking work, volunteering, educational activities, prevocational classes, day treatment, and any productive activity (paid work, volunteering, or educational activity). Time 3 outcomes were selected as the target of analysis because the greater amount of time and consequently greater potential treatment exposure was expected to increase the likelihood of observing effects.

To examine the effects of condition and service amount on Time 3 outcomes, a two-

step analysis strategy was employed. The first step involved comparison of all research participants in the three conditions on work-related outcomes. Because the design was quasiexperimental, with teams assigned to condition on a nonrandom basis, efforts were made to adjust for preexisting condition differences. Control variables included those showing Time 1 differences among conditions and Time 1 variables significantly predictive of Time 3 work outcomes. The second analysis step involved only individuals who could have received WINS services—those in the immediate and delayed conditions—and focused on the relationship between Time 3 outcomes and amount of WINS service received. Because individuals self-selected into WINS services to some extent, we could not assume that amount of services an individual received was independent of their preexisting likelihood of vocational success. It was possible that, without WINS, the individuals who sought vocational services and received them in substantial amounts would have been more successful than those who did not pursue vocational help. To adjust for the effects of such preexisting differences, characteristics found to predict postintervention work success, along with Time 1 variables that covaried with amount of WINS service received, were used as control variables in this analysis.

Other analyses were conducted with the interview sample data to examine changes in self-perceptions and the effects of condition and service amount on self-perceptions. Doubly multivariate repeated measures analysis of variance was used to examine these changes from baseline to Time 2 and Time 3. All the analyses results described below are significant at  $p < .05$  at least (unless otherwise noted as marginal, where  $p < .10$ ).

### **RESULTS : Descriptive Results**

Table 1 provides data describing the amount of WINS service received by each of the two treatment groups. The immediate group received more services overall than the delayed group, as would be expected since these clients had a longer period of time in which they were eligible to receive WINS service. However, the two treatment groups were about equal in the percentages receiving

a substantial amount of service. Individuals in the control group received no services.

Table 2 provides descriptive information on several outcome variables at Time 3, examined by condition and service amount. In terms of condition, on virtually all outcomes, greater percentages in immediate and delayed-treatment conditions demonstrated successful outcomes. For example, 69% of those in the delayed-treatment group engaged in some productive activity (i.e., work, education, or volunteer work) at Time 3, compared to 59% of those in the immediate-treatment group and 44% of those in the no-treatment group. Regarding service level, greater percentages of those who received moderate or substantial service levels demonstrated successful outcomes compared to those who received no service or minimal service. For example, 97% of those receiving substantial service were engaged in productive activity at Time 3, compared to 88% of those who received moderate service, 56% of those who received minimal service, and 54% of those who received no service.

### **MULTI VARIATE ANALYSIS**

Two sets of logistic regression analyses were conducted, one each to examine the effects of condition and of service amount. Because WINS utilized a quasiexperimental design, and consequently there was no random assignment to either condition or service amount, preliminary analyses were conducted to identify covariates demonstrating significant associations with condition and service amount, so that their effects might be controlled in determining the effect of condition and service amount on outcomes. Preliminary analysis was also conducted to identify baseline predictors of work at Time 3. This was necessary to control for any preexisting differences that might exist between conditions and service amounts that would confound the relationships between condition and Time 3 outcomes, and service amount and Time 3 outcomes. While several Time 3 outcomes were examined, all were related to work and work-related activities, hence the control of covariates related to work at Time 3 was determined to be sufficient for all analyses.

### **CONDITION EFFECTS ON WORK OUTCOMES**

Baseline variables were tested to determine if they were associated with working at Time 3. Four variables demonstrated significant associations with working: race, entitlement status, baseline functioning, and working competitively at baseline. Whites were more likely to be working (53%) than members of minority groups (35%),  $X^2(1, N = 386) = 8.46, p < .01$ . Those receiving entitlements were less likely to be working (47%) than those not receiving entitlements (63%),  $N = 373) = 4.60, p < .05$ . Those working at Time 3 were higher functioning at baseline on the Community Living Adaptation Scale ( $f = 4.78, p < .001$ ). Finally, those involved in competitive work at baseline were more likely to be working at Time 3 (84%) than those not involved in competitive work (42%),  $^{\wedge}(1, N = 372) = 34.06, p < .0001$ .

**Table 1: WINS Service Amount by Condition (N = 403)**

WINS service amount	Condition: exposure to WINS		
	Immediate (18 months exposure)	Delayed (9 months exposure)	Control (no exposure)
None (no documented contact with staff)	82 58.2%	97 70.3%	124 100.0%
Minimal (less than 60 min total staff contact)	16 11.3%	10 7.2%	0
Moderate (total staff contact 60-359 min)	26 18.4%	14 10.1%	0
Substantial (360 min or more total staff contact)	17 12.1%	17 12.3%	0
Total	141	138	124

Baseline covariates of condition were also explored as potential control variables in the analysis. Of the variables collected on all three conditions, five variables were found to be significantly associated with condition, including the four already found to be predictive of working at Time 3 (race,

entitlement, baseline functioning [CLAS], and competitive work). The delayed and control conditions had significantly lower proportions of minority clients (14% and 19%, respectively) than the immediate service condition (26%);  $r(2, N = 396) = 6.30, p < .05$ . A higher proportion of individuals in the delayed and control groups were receiving entitlements (91% and 92%, respectively) than in the immediate service group (74%);  $Jc(2, N = 382) = 19.81, p < .0001$ . The proportion in the control condition who were engaged in competitive work (7%) was much lower than in the immediate (19%) and delayed (20%) conditions,  $^{\wedge}(2, N = 381) = 9.08, p < .01$ . On community functioning, the immediate group was significantly higher ( $M = 3.60$ ) than the control group ( $M = 3.32$ ) in CLAS scores, with the delayed group indistinguishable from the other two ( $M = 3.42$ );  $F(2, 375) = 6.09, p < .01$ . Nearly twice as many in the control group (24%) were in day treatment compared with the immediate service group (12%); the delayed group (18%) was not different from the other two,  $^{\wedge}(2, N = 380) = 6.08, p < .05$ . However, in multivariate analysis used to determine which set of variables made independent contributions to account for variance between pairs of conditions, day treatment was not found to make an independent contribution, and was not retained in the logistic regression model examining the influence of condition.

**Table 2. Time 3 Outcome Variables by Condition and Amount of WINS Service Received**

Time 3 outcome variable	By condition			By amount of WINS service (immediate and delayed conditions only)					n
	Imme diate	De laye d	Con trol	n	No ne	Mini mal	Mod erate	Subst antial	
Any paid work	50%	61%	37%	390	45%	40%	85%	85%	262
Work & 10 hr/w week	37%	38%	14%	229	31%	23%	55%	67%	229
Com petiti	21%	27%	12%	39	19%	8%	33%	50%	26

ve work				3					9
Seeking work	11%	3%	2%	35%	3%	5%	15%	19%	23%
Volunteering	5%	5%	3%	37%	2%	9%	8%	12%	25%
Educational activities	15%	17%	7%	38%	13%	21%	15%	27%	26%
Prevocational classes	17%	15%	18%	37%	19%	9%	15%	6%	25%
Day treatment	12%	23%	30%	24%	28%	6%	3%	7%	17%
Any productive activity*	59%	69%	44%	39%	54%	56%	88%	97%	26%

The results of the logistic regression analysis examining condition effect on the outcome of productive activity are given in Table 3. After controlling for race, entitlement status, functioning (CLAS), and competitive work at baseline, the delayed condition demonstrated a significant effect, more than doubling the odds that an individual was engaged in productive activity at Time 3 compared with the control condition. However, the immediate versus the control condition effect was not significant.

For the remaining outcome variables examined, odds ratios and model chi-square improvement are reported in Table 4. Six variables, in addition to productive activity, demonstrated significant condition effects. In two cases (working and competitive work), the delayed condition showed an effect, while the immediate condition did not. In two additional cases (day treatment and seeking work), the immediate condition showed an effect, while the delayed condition did not. In the case of educational activities, both immediate and delayed conditions showed an effect, with the delayed condition

demonstrating the stronger effect. In the case of working more than 10 hours per week, both immediate and delayed conditions showed effects, with the delayed condition demonstrating a slightly stronger effect.

**Table 3: Condition Effects on Productive Activity at Time 3 (*N* = 365)**

Covariates	Model 1: Baseline covariates of work at T3				Model 2: Condition added			
	<i>B</i>	<i>SE</i>	<i>df</i>	Odds ratio	<i>B</i>	<i>SE</i>	<i>df</i>	Odds ratio
Race	-.018	.28	1	0.84	-.015	.28	1	0.86
Receiving entitlements	0.19	.34	1	1.20	0.22	.35	1	1.25
Functioning level (CLAS)	0.52	.18	1	1.69**	0.52	.19	1	1.69**
Competitive work	1.76	.46	1	5.34***	1.67	.46	1	5.33***
Condition							2	
Immediate vs. control					0.33	.28	1	1.40
Delayed vs. control					0.84	.28	1	2.31**
Constant	-1.70	.80	1		-1.76	.83	1	
Model $\chi^2$			40				9.34	
-2 log likelihood			458.54				449.21	

$p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ ; \*\*\*\* $p < .0001$ .

**Table 4. Condition Effects on Other Time 3 Outcomes**

Time 3 outcome variables"	Odds ratio Immediate vs. control	Delayed vs. control	Model chi-square
Any paid work ( <i>n</i> = 365)	1.22	2.20**	8.72**
Paid work & 10	2.45*	3.31**	10.99*

hr/week ( <i>n</i> = 306)			
Competitive work ( <i>n</i> = 352)	0.88	2.45*	7.94*
Seeking work ( <i>n</i> = 329)	7.37**	2.08	10.02* *
Volunteering ( <i>n</i> = 355)	1.19	1.51	NS
Educational activities ( <i>n</i> = 361)	2.46*	3.23**	8.26*
Prevocational classes ( <i>n</i> = 355)	1.34	0.91	NS
Day treatment ( <i>n</i> = 231)	0.37*	0.87	5.10 <sup>t</sup>
Any productive activity: paid work, volunteering, education ( <i>n</i> = 365)	1.40	2.31**	9.34**

It had been expected that if there were differences between the two treatment conditions (immediate and delayed), the results would favor the immediate condition, since that group had the opportunity to receive WINS services for a longer period of time. Since the results showed a different pattern (more significant results with the delayed treatment condition), post hoc exploratory analyses were undertaken to explain these results, through a logistic regression comparing clients in the immediate condition at T2 (9 months) and clients in the delayed condition (who at that time had received no WINS services) with those in the control condition, utilizing covariates, as in the previous logistic regressions on T3 outcomes. It was expected that if there were any effects, they would be limited to the immediate treatment condition (since the delayed condition had not yet been assigned specialized WINS services). Results of this set of logistic regressions are shown in Table 5. Only four regressions were undertaken, since only four variables were significantly related to condition in simple bivariate analyses.

Condition is a significant predictor of the vocational outcome variable for two of these four variables: productive activity and any paid work. In these two analyses, both the immediate and delayed conditions showed an effect, with the delayed condition demonstrating a somewhat stronger effect.

### SERVICE AMOUNT EFFECT ON WORK OUTCOMES

Preliminary analyses again sought to determine two types of covariates, to be controlled in order to determine the effect of service amount on later outcomes: those associated with working at Time 3 and those associated with WINS service involvement. Although baseline predictors of work at Time 3 had already been examined for the analysis of condition effects, these needed to be reexamined in this analysis because the analysis of service amount effects was limited only to the immediate and delayed conditions. Nonetheless, the same set of baseline variables was found to be associated with working at Time 3 in this reduced sample: race,  $\chi^2(1, N = 264) = 7.78, p < .01$ ; entitlement status,  $\chi^2(1, N = 251) = 5.16, p < .05$ ; competitive work at baseline,  $\chi^2(1, N = 250) = 20.05, p < .001$ ; and baseline functioning CLAS score  $t = 3.97, p < .001$ . One additional variable, age, was found to be significantly associated with working at Time 3, and was controlled in the analysis of service amount effects. Those employed at Time 3 were younger than those not employed (36.01 vs. 38.96;  $t = 2.71, p < .01$ ).

The examination of variables associated with service amount identified three significant variables: length of time in current residence (1 year or less vs. more than 1 year), number of previous jobs (five or less vs. more than five), and expectations about working in the future. Those who were in their current residence 1 year or less were more likely to receive a moderate or substantial amount of service (33%) versus those who were in their current residence more than 1 year (22%),  $\chi^2(1, N = 278) = 4.84, p < .05$ . Those with five or fewer jobs were less likely to receive a moderate or substantial amount of service (14%) than those who had worked more jobs (32%),  $\chi^2(1, N = 274) = 9.61, p < .01$ . Individuals who, at baseline, expected

to be working in 1 year were more likely to receive substantial WINS service (32%) than those who did not (14%),  $\chi^2(1, N = 278) = 9.95, p < .01$ .

For the multivariate analysis, service amount was dichotomized into none/minimal versus moderate/substantial categories to address otherwise small cell sizes. The results of the logistic regression analysis examining the effect of service amount on productive activity are given in Table 6. After controlling for predictors of work (race, entitlements, baseline functioning [CLAS], working competitively at baseline, and age) and predictors of service amount (residential stability, limited work history, and 1-year work expectations), there was a significant effect of service amount on productive activity.

**Table 5. Condition Effects on Time 2 Outcome**

Time 2 outcome variables	Odds ratio		
	Immediate vs. control	Delayed vs. control	Model chi-square
Productive activity ( $n = 364$ )	1.71*	1.86*	5.97*
Any paid work ( $n = 370$ )	1.98*	2.23**	9.68**
Paid work £10 hr/week ( $n = 370$ )	1.96+	NS	NS
Competitive work ( $n = 358$ )	NS	NS	NS

Service participation effects on other outcomes can be found in Table 7. In addition to productive activity, service amount was a significant predictor of four other outcome variables. A moderate or substantial amount of service increased the odds of working by almost five times, more than doubled the likelihood working more than 10 hours per week, and increased the odds of work-related activity such as seeking work (by more than four times). A moderate or substantial amount of service also reduced the odds by 80% that an individual would be in a day treatment program at Time 3.

#### Cases Omitted from Outcome Analysis due to Missing Covariate Data

Missing covariate data required the removal of 43 individuals from the analysis of condition or service effects on productive activity. The proportion of omitted cases was higher in the immediate (14%) and delayed (11%) conditions compared with the control group (3%);  $\chi^2(2, N = 403) = 8.63, p < .01$ . Within the immediate and delayed conditions, a higher proportion of individuals who received no or minimal WINS service were omitted due to missing data (17%) than were those who received substantial service (7%);  $\chi^2(2, N = 279) = 4.37, p < .05$ . No differences between the omitted and included cases were found on sex, race, age, baseline work, entitlement status, baseline functioning level (CLAS), or diagnosis.

**Table 6. Effect of Receiving Substantial WINS Service on Productive Activity at Time 3 ( $N = 240$ )**

Covariates	Model 1: Baseline covariates of work at T3			Model 2: Baseline covariates of receipt added			Model 3: Amount of WINS service added		
	B	SE	df	Odds ratio	SE	df	B	SE	df
Race	0.51	.35	1	0.60	.35	0.60	0.50	.37	1
Receiving entitlements	0.57	.44	1	1.77	.45	1.60	0.47	.1	1.69
Functioning level (CLAS)	0.44	.23	1	1.55	.24	1.5	.25	.1	1.33
Competitive work	1.26	.48	1	3.53**	.50	2.80*	0.52	.1	2.33+
Age	0.04	.02	1	0.96*	.02	0.97*	0.03	.02	1
Residential stability (in current home > 1 year)									

Limited work history (^5 jobs)				0.20	.32	1.20	0.33	1.094
One-year work expectation				0.76	.321	2.14*	0.331	1.81+
Substantial WINS service received							1.95	7.01****
Constant	0.27	1.181		0.39	1.221		0.41	1.25-0.41
Model $\chi^2$		30.09****			6.711			21.53****
-2 log likelihood		284.24			277.53			256.00

**Table 7. Effect of Substantial WINS Service on Other Time 3 Outcomes**

Note. Analysis of WINS service amounts was limited to intervention conditions only (immediate and delayed),  $tp < .10$ ;  $*p < .05$ ;  $**p < .01$ ;  $***p < .001$ ;  $****p < .0001$ .

Time 3 outcome variables"	Odds ratio*	Model chi-square
Any paid work ( $n = 240$ )	4.92***	17.57****
Paid work £10 hr/week ( $n = 240$ )	2.15*	5.09*
Competitive work ( $n = 227$ )	1.80	NS
Seeking work ( $n = 209$ )	4.06*	5.39*
Volunteering ( $n = 233$ )	3.18+	2.69+
Educational activities ( $n = 236$ )	1.54	NS
Prevocational classes ( $n = 233$ )	0.74	NS
Day treatment ( $n = 158$ )	0.20*	4.92*
Any productive activity: paid work, volunteering, education ( $n = 240$ )	1.01****	21.53****

Note. For each logistic regression listed here, the control variables detailed in Table 6 were entered prior to the final variable indicating receipt of substantial WINS service. Odds ratios and model chi squares reflect the final step in each logistic regression.

"Values of  $n$  differ due to different amounts of missing data on each Time 3 outcome. 'Odds ratios are in comparison with the no-treatment control condition,  $tp < .10$ ;  $*p < .05$ ;  $**p < .01$ ;  $***p < .001$ ;  $****p < .0001$ .

#### Self-Reported Interview Scales

Doubly multivariate repeated-measures analysis of variance was conducted on the set of six self-report measures from the interview (Pearlin Self-Mastery Scale, Purpose in Life Scale, Lehman's Global Quality of Life, Work Behaviors and Attitudes, Symptom Checklist-10, Areas of Difficulty Checklist) to examine condition and service amount effects on these outcomes. A significant time effect was found, Wilks' lambda = .54;  $F(12,172) = 12.03$ ,  $p < .001$ ; however, there were no significant interactions of condition or service amount with time.

#### DISCUSSION: Interpreting the Outcome Results

Since this was a quasiexperimental design, the preexisting differences in the characteristics of participants by condition and by service model needed to be addressed in outcome analyses. Therefore, the significant preexisting differences between groups by condition and by service level on relevant baseline variables were included as covariates in multivariate analyses predicting work outcomes. In these analyses, a condition effect was significant in terms of productive activity and on live of eight other work-related outcomes. Service amount effects were also found to be significant for productive activity and four of the other variables.

We had expected that the immediate-treatment group would have more positive outcomes than the delayed condition (because of the longer time frame available to effect change) and that both conditions would be superior to the control condition. While the latter result was demonstrated, the former was not. That is, on most variables, the delayed condition actually had somewhat better outcomes. Post hoc analyses

undertaken to examine T2 differences revealed fewer condition effects, but outcomes for the delayed and immediate treatment groups were both significantly better than for the control group. The explanation for these findings is not immediately clear. We attempted to control statistically for baseline differences across the three conditions; however, statistical controls cannot be used to fully equate non-comparable groups (Grossman & Tierney, 1993; Porter & Raudenbush, 1987). At baseline, the clients on the teams in the immediate condition appeared to be somewhat better functioning (less likely to be receiving entitlements, with higher CLAS scores, and less likely to be in day treatment). Perhaps this meant that there was less likelihood of their showing significant improvements in work-related outcomes. Another possibility is a “compensatory rivalry” explanation—that is, teams in the delayed condition knew that they would have vocational specialists added to their teams at the end of 9 months. In anticipation, perhaps they prepared their clients for work, and/or wanted to demonstrate that they could get clients jobs on their own, for example, by pursuing more referrals to sheltered workshops, or other services available through the vocational rehabilitation agency. This could explain equivalent outcomes for the immediate and delayed groups at T2.

To explain the differences favoring the delayed group at T3, we might speculate that the jobs obtained by Project WINS’ clients in the immediate condition were short-term and/or unstable jobs, with little career potential. A similar conclusion was drawn by Bond *et al.* (1997). Thus, while individuals may have obtained jobs by T2, by T3 these jobs were lost. Furthermore, perhaps the vocational specialists were better at finding more stable jobs when they started working with the delayed-condition clients 9 months later. The logistic regressions partially confirm this explanation, finding that the immediate group was more likely to be “seeking jobs” at T3 than either the delayed or control conditions. Previous published results on this project also offer some confirmation for the suggested explanation: McCrohan *et al.* (1994) reported that many Project WINS clients obtained low-skill, high-turnover jobs.

On the self-report (interview) variables, although there were significant time effects, there were no significant interactions for time by condition or for time by service amount on outcomes like quality of life, symptoms, self-esteem, community functioning, etc. This is congruent with results reported by Drake *et al.* (1996) on a similar integrated vocational/clinical intervention and to the conclusions in the Bond *et al.* (1997) review that results “lend little support for the hypothesis that supported employment programs have a generalized effect on other outcomes” (p. 343). Arns and Linney (1993) posit that this somewhat counterintuitive result can be explained with the hypothesis that changes in self-perceptions occur as a result of getting and keeping a job. The short amount of time available for follow-up in Project WINS may not have been long enough to allow for such sequential changes to occur. On the other hand, the self-perception and quality-of-life measures may be insensitive to change.

### **Comparison to Other Research Results**

In comparison to results from research studies summarized in Bond *et al.*’s (1997) recent review, the outcomes from Project WINS are somewhat difficult to judge. The competitive employment rates for WINS clients were somewhat low (21-27%). However, if percentage with a schizophrenic diagnosis can be used as an indicator of severity (as suggested by Lehman, 1995), then WINS, with 68% of clients having a schizophrenic diagnosis, appears to have served clients that are more disabled than any of the other studies, and competitive employment rates might be expected to be correspondingly lower.

WINS also differed from other supported employment programs reviewed by Bond *et al.* (1997) in that some WINS clients preferred to start employment initiatives within the security and predictability of a sheltered workshop setting. Since WINS operated as much as possible on a “choice” principle, it respected desires of participants to start with sheltered work, but recontacted them periodically to ascertain their desires to move on to more independent employment. Thus, it does seem appropriate to include as positive outcomes any client-chosen, work-related outcome, including noncompetitive work. When “any work” is considered, WINS

active conditions achieved 50% and 61% employment rates, which compare quite favorably with other studies.

The Project WINS results can be assessed against some other topics raised by Bond *et al.* (1997). One is participation rate. In the present analysis, we report on service amounts by condition (Table 1). Those statistics indicate that only 26 of 103 participants in the immediate and delayed conditions (or 25%) received minimal services. (Note that the category "None" refers to those individuals never referred to WINS; this was appropriate, as WINS staff were not supposed to service all eligible clients.) Clients receiving minimal services were typically contacted by WINS program staff after a referral, but then did not show up for appointments or follow-up on agreed-upon vocational activities. Thus they should be equivalent to Bond *et al.*'s characterization of program dropouts. In their review, Bond *et al.* cite dropout rates of more than 40% as being common (p. 343). Our retention is better than this comparison figure.

Another limitation of many studies cited by Bond *et al.* is possible screening processes, which limit the generalizability of positive outcomes. In Project WINS, the exclusion criteria were intended to pose minimal limitations. This appears to be the case in that only 21.1% of all clients served by case management teams were screened out. Thus, our rate of inclusion seems appropriately high. Previous published reports on WINS indicated that the productivity of staff was also high (i.e., staff served about 80% of the maximum number expected, not taking into account startup, staff vacancies and leaves of absence, etc.; Bybee *et al.*, 1995).

WINS was conceptualized as a minimal-exclusion, client-determined model. That meant that other than for disability or client preference reasons, client characteristics should not differentiate those individuals from the public mental health system in this county who were referred/served versus not served by WINS. Previous analyses, however, indicated that this was not totally the case. That is, traditional labor market variables did successfully predict WINS participation

(participants were more likely to be those with a greater work history, younger ages, and more education). However, these variables accounted for only a modest percentage above a chance prediction (17.1%; Bybee *et al.*, 1995).

A final point of comparison concerns the types of jobs acquired by supported employment participants. The Bond *et al.* review indicated that most jobs obtained were unskilled and entry-level. This was, unfortunately, the case for many WINS participants. On the Hollingshead rating, two thirds of jobs obtained were unskilled, versus 30% that were clerical, sales, or minor professionals. According to Dictionary of Occupational Title coding, about 32% of WINS jobs were in benchwork, and 35% were in service industries. The jobs that could be obtained for participants were limited by a number of factors: the checkered work histories of participants and their resultant skill deficits in technology and computer equipment, as well as the difficulties in finding suitable jobs utilizing an individualized rehabilitation model. That is, Project WINS staff had to find or develop jobs for clients one by one, based on a client's preference, skills, etc. If the client left the job, the job slot was usually lost.

These results suggest two things: first, that there should be more emphasis on vocational training and/or higher educational opportunities, like supported education, being available for all individuals with psychiatric disabilities who are interested in jobs. Second, that integrated VR models need to have a tie-in to a larger job development resource. Then, rather than spending time finding or developing job slots that easily get lost, vocational specialists could tap into a diverse array of job possibilities and spend their time working with clients on selecting, getting, and keeping their jobs.

### **Methodological Limitations**

The research design limitations associated with Project WINS need to be acknowledged. In terms of outcome measures, it should be noted that status (dichotomous) variables were used, rather than continuous variables like number of weeks worked, wages earned, etc. Unfortunately, data reported by

case managers and project staff were not sufficiently detailed to allow construction of continuous variables.

The major shortcoming of the WINS research was its complicated design in which teams, not clients, were assigned to conditions and not all eligible clients could be interviewed. Statistical controls are likely to be the only method to address this issue. It was therefore encouraging to see positive results from our logistic regressions on work outcomes, examining effects of model assignment and of service amounts when appropriate covariates were utilized. However, the fact that effects for the delayed rather than the immediate condition are more robust is of significant concern.

The results need to be interpreted with caution. On several variables strongly predictive of vocational success, the control group looked worse at baseline. Although we employed powerful statistical methods to adjust for these differences, it is rarely possible to remove their effects entirely. Additionally, we could not adjust for differences implied by the staff choice that led to the control team being designated as such. It seems possible that a portion of the condition difference in work-related outcomes could be attributed to the negative impact of staff who requested that their team not implement vocational services. Finally, higher attrition from outcome analysis in the WINS service conditions may have artifactually strengthened the condition differences. Due to missing covariate data, a higher proportion of individuals in the immediate and delayed conditions (particularly those who received no or minimal WINS service) were omitted from the outcome comparison. In the control condition, very few cases were omitted due to missing data. In other words, through differential missing data, the WINS service conditions may have been "creamed" to a greater extent than the control condition, creating a bias that also may have worked in favor of finding positive WINS effects. Of course, the opposite selection bias might have occurred, wherein those cases lost to follow-up were unavailable because they had left the treatment system because of better functioning. Finally, we must note the fact that overall only about 28% of clients on the treatment teams received moderate or

substantial levels of WINS services, also limiting the generalizability of findings. Collectively, these caveats indicate that the results of this single study are not sufficient for unqualified conclusions about the effectiveness of integrating vocational services into case management teams. But the results are promising and suggest that this approach may be an effective way to deliver vocational services. Replication, with random assignment of a larger number of teams to conditions and with sufficient resources for repeated assessment of all clients eligible for WINS services, would allow stronger inferences and more definitive conclusions about the model's effects.

## CONCLUSION

Results from Project WINS offer additional evidence of the effectiveness of an integrated model for case management and vocational services. In addition to their effectiveness, such models have the potential for being more easily accommodated within existing community mental health programs and also more acceptable to managed behavioral health care providers (versus separate vocational services). WINS results also offer some challenges for improvement as well. While paid employment and overall productive activity was high, this was not true for competitive employment. Furthermore, the kinds of jobs obtained by WINS clients were mainly in lower status occupations, usually corresponding to employment situations which offer much boredom and sometimes noisome occupational conditions, but little in the way of stability, career advancement, or necessary fringe benefits. While at times starting with workshop employment or entry-level jobs may be the desire of the client, agencies need to ensure that better opportunities are also available beyond these beginnings, such as through education, training, or increased or more creative job development activities.

Project WINS experiences also indicated the need for improvements in research designs and methods. More standardized record-keeping systems and more valid management information systems might curtail resource limitations and allow expansion of data collection on larger cohorts and/or multiple sites to enhance generalizability and knowledge.

## REFERENCES

- [1] American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders*. Washington, D.C.: American Psychiatric Press.
- [2] Anthony, W. A. (1980). *The principles of psychiatric rehabilitation*. Baltimore, MD: University Park Press.
- [3] Anthony, W. A., & Blanch, A. (1987). Supported employment for persons who are psychiatrically disabled: An historical and conceptual perspective. *Psychosocial Rehabilitation Journal*, XI(3), 5-23.
- [4] Anthony, W. A., Cohen, M. R., & Danley, K. S. (1988). The psychiatric rehabilitation model applied to vocational rehabilitation. In J. A. Ciardiello & M. D. Bell (Eds.), *Vocational rehabilitation of persons with prolonged psychiatric disorders* (pp. 59-80). Baltimore, MD: Johns Hopkins University Press.
- [5] Arns, P. G., & Linney, J. A. (1993). Work, self, and life satisfaction for persons with severe and persistent mental disorders. *Psychosocial Rehabilitation Journal*, 17(2), 63-79.
- [6] Bond, G. R. (1987). Supported work as a modification of the transitional employment model for clients with psychiatric disabilities. *Psychosocial Rehabilitation Journal*, 11(2), 55-73.
- [7] Bond, G. R. (1992). Vocational rehabilitation. In R. P. Liberman (Ed.). *Handbook of Psychiatric Rehabilitation* (pp. 244-263). New York: Macmillan.
- [8] Bond, G. R., Drake, R. E., Mueser, K. T., & Becker, D. R. (1997). An update on supported employment for people with severe mental illness. *Psychiatric Services*, 48(3), 335-346.
- [9] Bond, G., Witheridge, T., Dincin, J., Wasmer, D., Webb, J., & De Graaf-Kaser, R. (1990). Assertive community treatment for frequent users of psychiatric hospitals in a large city: A controlled study. *American Journal of Community Psychology*, 18, 875-893.
- [10] Bybee, D., Mowbray, C. T., & McCrohan, N. (1995). Towards zero exclusion in vocational opportunities for persons with psychiatric disabilities: Prediction of service receipt in a hybrid vocational/case management service program. *Psychosocial Rehabilitation Journal*, 18(4), 73-93.
- [11] Cook, J. A., Jonikas, J. A., & Solomon, M. L. (1991). Models of vocational rehabilitation for youth and adults with severe mental illness. *American Rehabilitation*, 77(1), 6-11, 32.
- [12] Derogatis, L., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine*, 13, 595-605.
- [13] Drake, R. E., McIlugo, G. J., Becker, D. R., Anthony, W. A., & Clark, R. E. (1996). *Journal of Consulting and Clinical Psychology*, 64(2), 391-399.
- [14] Griffith, R. D. P. (1973). A standardized assessment of work behavior of psychiatric patients. *Journal of Psychiatry*, 123, 166-771.
- [15] Grossman, J., & Tierney, J. P. (1993). The fallibility of comparison groups. *Evaluation Review*, 17(5), 556-571.
- [16] N. Ahituv, Assessing the value of information: Problems and approaches, in: *Proceedings of the 10th International Conference on Information Systems* (Boston, 1989).
- [17] J.Y. Bakos M.E. Treacy, information technology and corporate strategy: A research perspective, *MISQ* 19(1986) 107-119.
- [18] E. Bartessaghi and T. Francesco, The impact of just-in-time on production system performance: An analytical framework, *International Journal of Production Management* 9 (1989) 40-62. published by july or aug 2014

Volume :5, Issue: 2  
July - December 2015

ISSN : 2229 - 3515

International Journal of  
**ADVANCES IN  
SOFT COMPUTING  
TECHNOLOGY**

Editor-in-Chief  
**Dr. Vaka Murali Mohan**



Published by

**BHAVANA RESEARCH CENTER**

# The New Economic Order Using Financial Information for Keeping Social Scoreehics, In Equity and Social Justice

**K.Ramakrishna Reddy<sup>\*1</sup>, G. Ravindra Babu<sup>2</sup> and S. Krishna Mohan Rao<sup>3</sup>**

1. Department of CSE, Malla Reddy Engineering College, Hyderabad, Telangana, India.
2. Department of CSE, Trinity College of Engineering & Techn., Karimnagar, Telangana, India
3. Department of CSE, Siddhartha Institute of Engineering & Techn., Hyderabad, Telangana, India

## KEYWORDS

Ethics;  
Equity;  
Financial Information;  
Social Justice;  
Social Scoreehics;  
Economic Order

**Abstract:** In the present world order unbridled forces of free market capitalism are frequently cited for much of the social injustice, inequity, and disparity of wealth between the rich and the poor. Although history's verdict in favor of the free markets could hardly be harsher or clearer, it is clear that after the initial wave of triumph, the free market paradigm has developed some cracks in its facade. What marks the trail of such sustained and pronounced move toward free markets in terms of ethics, morality, social welfare and social justice? How does one keep a social score in this seemingly relentless and irreversible move all over the world toward free market capitalism?. In this paper we shall attempt to address these and related questions. Drawing on concepts from organization theory and social philosophy and using publicly available financial information, we shall illustrate how, amidst the myriad and mixed noises, some sense of order and signal can be discerned in addressing issues of equity and social justice. Toward this end, first, we provide a broad contrast between two models of financial markets: the command model and the free market model and proceed to examine publicly available financial information and analyze the trends and patterns with graphical representations using publicly available data from Handbook of International Economic Statistics. Next, we explore the implications of financial performance measures for social welfare and social justice and discuss the social perils of free markets using the Mexican and Asian Financial crises as the focal points. Finally, we present a set of recommendations for smoother structural transition..

## 1. INTRODUCTION

Unbridled forces of free market capitalism are frequently cited for much of the social injustice, inequity, and disparity of wealth between the rich and the poor. The polarization is so pronounced that only two worlds seem to matter in today's world economy: the first

world and the third world. Some have even argued that the issue of whether there is such a thing as the second world has almost become an existentialist question.

History's verdict in favor of the free markets could hardly be harsher or clearer. Former strongholds of socialist economies including Poland, Czechoslovakia, and Hungary have not only embraced free market capitalism. However, it is clear that after the initial wave of triumph, the free market paradigm has developed some cracks in its facade. What marks the trail of such sustained and pronounced move toward free markets in terms of ethics, morality, social

### \* MR. K.RAMAKRISHNA REDDY

Assistant Professor  
Department of CSE  
Malla Reddy Engineering College  
Hyderabad, Telangana, India  
Ph. No: 91-9603016200  
E-mail: ramakrishnareddy524@gmail.com

welfare and social justice? How does one keep a social score in this seemingly relentless and irreversible move all over the world toward free market capitalism?

In this paper we shall attempt to address these and related questions. Drawing on concepts from organization theory and social philosophy and using publicly available financial information, we shall illustrate how, amidst the myriad and mixed noises, some sense of order and signal can be discerned in addressing issues of equity and social justice. This paper is organized into five sections. First, we provide a broad contrast between two models of financial markets: the command model and the free market model and proceed to examine publicly available financial information and analyze the trends and patterns! For this purpose, we present graphical representations in Figures 1-9 using data from the *Handbook of International Economic Statistics: Central Intelligence Agency, Directorate of Intelligence*. Next, we explore the implications of financial performance measures for social welfare and social justice using data from countries which have implemented market reforms. Third, we discuss the social perils of free markets using the Mexican and Asian Financial crises as the focal points. In Sections "Going it gradually: the incrementalism imperative" and "Mechanisms first, market integration later: the infrastructure imperative" we present a set of recommendations for smoother structural transition. The paper concludes with an end note based on the distinction between doing things right versus doing the right thing.

## **TWO MODELS, TWO MOTIVE POWERS: A CONTRAST**

At one extreme, it is argued that self-interest provides the noblest and self-correcting motive power for social development by offering rewards commensurate with one's effort. In a world increasingly characterized by individualistic ethic, such an argument seems to be gaining ground steadily. However, voices of dissent can be heard loud and clear.

Kenneth Lux (1990) provides a provocative analysis of how the sees of social strife and ecological damage can be traced to self-interest as advocated by Adam Smith

more than 200 years ago. In support of his argument, Lux quotes the following passage from Adam Smith's book, *The Wealth of Nations*:

As Lux has rightly pointed out, Adam Smith's mistake lies in the implication that in any economy based on exchange, self-interest is more effective than benevolence in promoting social good. On the contrary, the free market capitalist system can only be expected to promote the interests of the rich at the expense of the disadvantaged. Roth (2002) similarly describes capitalism as an economic system that is uncaring and cut-throat and argues for its replacement by what he calls "democratic socialism." Henry Demarest Lloyd argued in a similar vein: "The wealth created by a thousand men under the motive power of the self-interest of the capitalist is not, and cannot be, equal to the wealth that will be created by the same men under the motive power of cooperation or democracy".

The two motive powers ultimately find expression in market forms. As Bathala and Korukonda (2003) have argued, state-controlled and market-oriented economies are the two prominent and opposing economic structures. They differ greatly in political ideology, private property rights, and the structure and the openness of financial markets. The state-controlled economic system suppresses competitive spirit of individuals which would have a detrimental effect on entrepreneurship, innovation, and productive efficiency. Although state-controlled systems ostensibly seek equality and equity in the distribution of incomes and wealth among their citizens, the recent historical developments suggest that the countries which have pursued this path were not very successful in accomplishing the objective. Inequalities arose between ordinary citizens (peasants) and those in control (bureaucrats and polit-bureau members). The greed for control and power replaced the greed for accumulating private wealth. The governments took control over the production and distribution of goods and services and removed the public choice from the system. The public enterprises became tools for promoting the agenda and welfare of politicians. In spite of some limitations, in

market-oriented economies individual property rights are well defined, and individuals as well as private firms would have incentives to maximize their wealth. These features are conducive to promote competition, innovation, and economic growth. Private enterprises are separate from their governments and this separation enables the governments to exercise monitoring and regulating functions which is deficient in state-controlled economies. Additionally, capital market monitoring plays a major role in the efficient allocation of resources among competing needs and this also is a limitation in state-controlled economies.

### **IMPLICATIONS OF THE FREE MARKET CAPITALIZATION FOR SOCIAL ISSUES**

The data from *World Economic Outlook*, October 1998, International Monetary Fund, presented in Table I substantiate some of the above contentions. Table I contains annual growth rates in several macro economic variables for Advanced Economies (AE) and the Newly Industrialized Asian Economies (NIAE). The growth rates are for the two subperiods, 1980–1989 and 1990–1999. As can be seen, NIAE recorded higher growth rates of all variables, except consumer prices and unit labor costs, during both the subperiods. The inflation rates and the growth in unit labor costs in NIAE, although higher than in AE, are commensurate with productivity and consumption gains in those countries.

Several studies (e.g., Bennett and Johnson, 1979; Boardman and Vining, 1989; Boycko et al. 1993; Galal et al., 1994) have found evidence that (i) state-owned businesses are less profitable and less efficient than private businesses, and (ii) privatization improves efficiency. More recently, Gupta (2001) finds evidence that even partial privatizations in India have produced positive results in productivity and performance. The author attributes the improvements to “the role of stock market in monitoring managerial performance rather than to a change in owners’ objectives.”

Countering this evidence, some studies have found no evidence of deficiencies in efficiency in public enterprises. For example, Caves and Christensen (1980) find that the government-run railroads in Canada

performed as well as its private counterpart. They argue that it is the lack of effective competition and not the public ownership per se that produces inefficiencies in public enterprises. Aharoni (1986) argues that the dismal financial performance of state-owned enterprises (SOEs) could be due to the social and political demands on them even though they are as efficient as private businesses in producing the same products. These mixed findings from empirical studies make it imperative to explore the ideological and ethical issues pertaining to market reforms. Broad and abstract arguments sound plausible whether they arise from the right or left wing ideological camps, yet the bottom line is how real-life experience stacks up against theoretical and conceptual arguments. In contrast to the ethical egoism propagated by free market advocates, the theory of utilitarianism proposed by Jeremy Bentham argued for consequentialism as the criterion for deciding moral choices. Under this theory, morally correct rules are those which produce the greatest good for the greatest number of people. How does the free market system under this rule?

This is a task we turn to in this section of the paper. Using data from countries which have implemented market reforms, we will endeavor to flesh out implications of market reforms and free market capitalization in terms of their impact on social issues. We will then extend the argument using concepts from organization theory and elaborate on the policy implications.

### **STOCK MARKET LIBERALIZATION**

To examine this question, we look at financial and social data for various factors before and after stock market liberalization for select countries. The comparative data, presented in Figures 1–3, shows the change of (a) industrial production (b) gross domestic product (GDP), and (c) net foreign assets before and after stock market liberalization for select countries.

A review of Figure 1 reveals that industrial production has registered a significant increase only in the case of one country out of the four examined.

TABLE I: Comparative macro-economic data: advanced economies versus newly industrialized Asian economies

	Advan ced econo mies	Newly industri alized asian economy es	Advan ced econo mies	Newly industri alized asian economy es
Annual growth rates	1980- 1989	1980- 1989	1990- 1999	1990- 1999
Real GDP growth rate	2.9	7.8	2.3	5.2
Real GDP growth rate (Per Capita)	2.2	6.4	1.6	4.1
Gross fixed capital formatio n	3.1	7.7	2.8	5.9
Pnvate consum ption expendit ure	3.0	7.4	2.4	4.9
Public consum ption	2.5	6.1	1.3	4.2
Unemplo yment rate (%) - average	6.9	3.2	7.1	2.9
Growth in employ ment	1.2	25	0.8	1.91
Consum er prices	6.3	6.7	2.9	5.1
Hourly wages	6.4	11.7	4.3	10.7
Producti vity	2.3	5.8	3.0	6.8
Unit labor costs	4.1	4.6	1.2	2.7
Export volume	5.3	10.8	5.8	10.0
Import volume	5.2	9.8	5.5	9.1

Source: World Economic Outlook, October 1998, International Monetary Fund.

### Industrial Production before and after Stock Market Liberalization

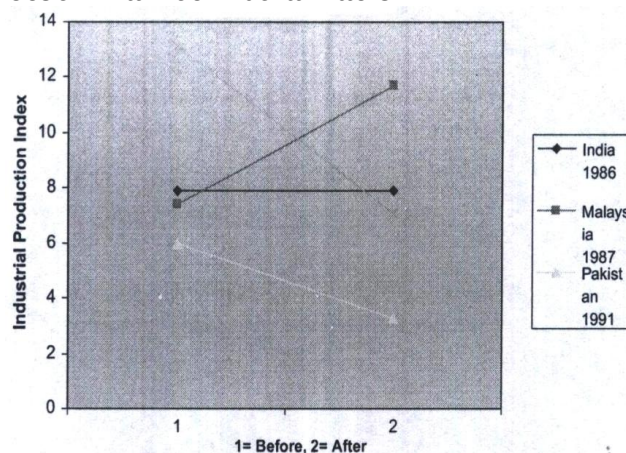


Figure 1. Industrial production before and after stock market liberalization.

### Gross Domestic Product and after Stock Market Librealization

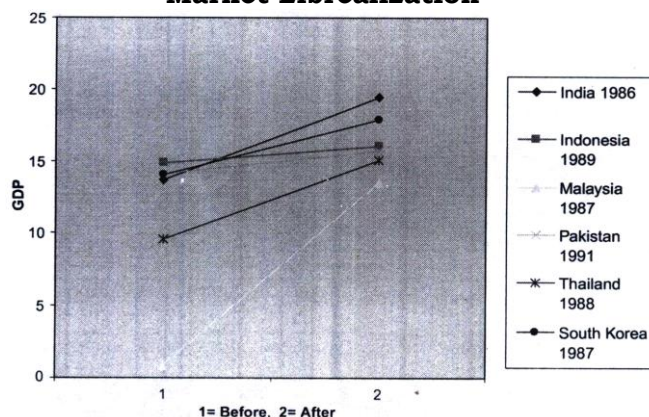


Figure 2. GDP before and after stock market liberalization.

### Net Foreign Assets before and after Stock Market Liberalization

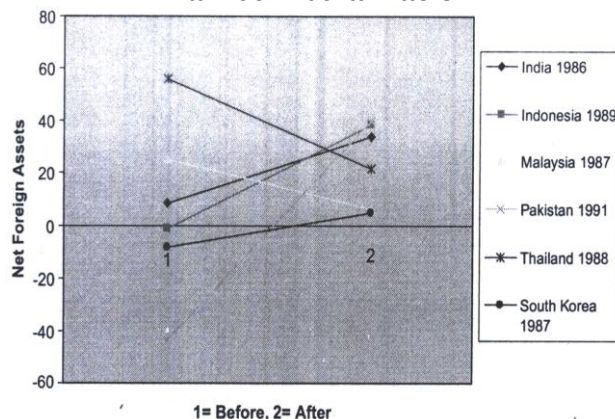


Figure 3. Net Foreign assets before and after stock market liberalization.

GDP, by contrast, has registered a dramatic increase in the case of all the countries examined. These mixed findings are quite perplexing and it appears that some of the turmoils faced by developing nations could be the result of GDP growth without the accompanying growth in industrial production. Net foreign assets have shown a dramatic increase in the case of India and Indonesia, while in the case of Malaysia they have registered an equally dramatic decline. An increase in net foreign assets is expected to occur virtually by definition; however, the fact that such an increase is selective reveals context- specificity of foreign investment.

### ILLUSARY BENEFITS FOR FREE MARKETS

Benefits from free markets have been ostensibly reported to accrue to the countries implementing privatizations and economic liberalizations. Superficially, this appears to be the case as can be seen from Figure 4.

Every newly industrialized economy identified in the figure has shown tremendous growth in exports during the period 1983–1996, with the steepest growth taking place between 1983 and 1994 in most cases. While this is laudable, what is perplexing is the ratio of exports-to-imports to Big 7 countries depicted in Figure 5. Initially, during the period of 1980–1990, every country identified in the figure with the exception of Taiwan and Mexico, has shown impressive improvements in the exports-to- imports ratio meaning exports have increased relative to imports. After that, starting 1990, every developing country with the exception of Mexico has lost ground to the Big 7 countries as their imports have outgrown the exports. This trend is quite disturbing as Big 7 countries could be encouraging market orientation in less developed countries as a means to push their exports to developing nations while at the same time carefully managing the level of imports from those countries. It will be a case in point to conduct additional in-depth research to examine these issues further and verify if our conjecture is true.

Another interesting aspect of moving toward market orientation is the potential for currency depreciations that are associated with market reforms and other expansionary

measures in developing economies (Mishkin, 1998). Their cost of imports would grow disproportionately relative to the revenue from exports as international transactions are typically denominated in US dollars or pounds sterling. This upsets the ratio of exports-to-imports, unless -Big 7 countries take in more imports from developing nations to offset the negative effects of currency depreciations in those countries. Figure 5 shows that was not the case.

### MATHHUE EFFECT IN ACTION

Matthew effect as proposed by Robert Merton (1968) is based on the following gospel:

#### Exports from Newly Industrialized Economies

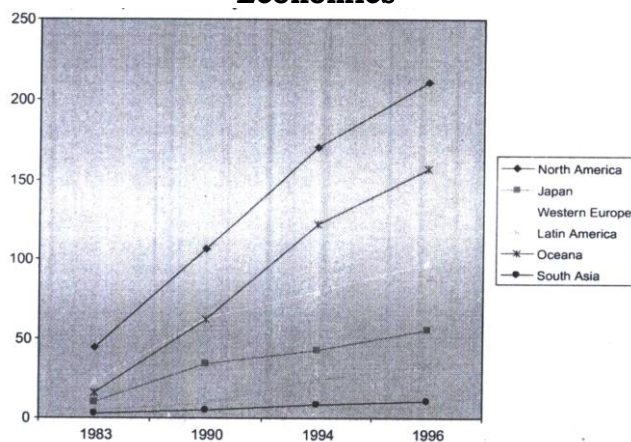


Figure 4. Exports from newly industrialized economies.

#### Ratios of Exports/Imports to Big 7

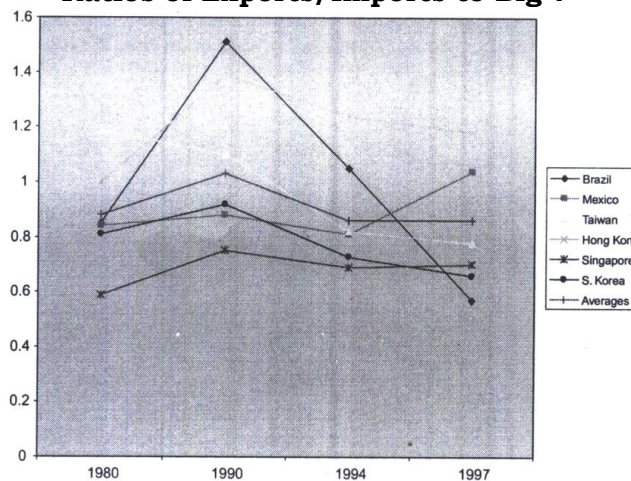


Figure 5. Ratios of exports/imports to Big 7 countries.

For unto everyone that hath shall be given, that he may have abundance; but from him that hath not, even that which he hath shall be taken away; and cast ye the useless servant into the outer darkness (Matthew, 25; 25, 28-30; Luke, 19; 24-26).

The picture with respect to employment provided in Figure 6 offers an interesting contrast to GDP described above. Employment index shows a drastic decline after stock market liberalization giving credence to the popular commonwealth argument about market liberalization leading to escalation of greed, self-interest and increase in wealth for the select few haves. This seems to be an overt manifestation of the Matthew effect.

#### **Employment before and after Stock Market Liberalization**

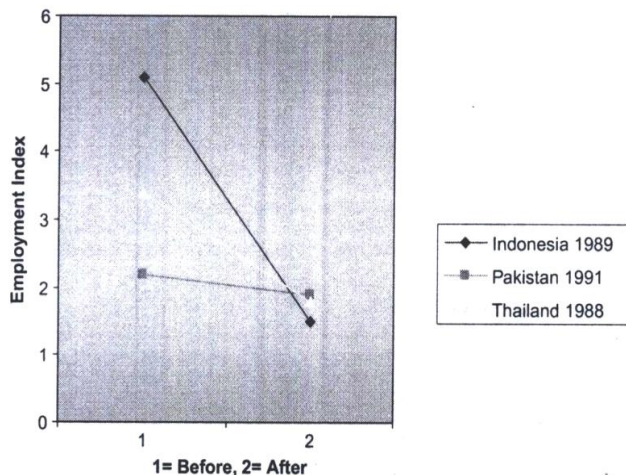


Figure 6. Employment before and after stock market liberalization.

#### **DISPARITY AND DOUBLE STANDARDS**

It is often pointed out that the U.S. is the world's largest debtor nation with consumption exceeding production and expenditure exceeding savings. This could be due to its heavy dependence on foreign oil, constant increases in import of other goods relative to exports, and lax domestic fiscal policies. We illustrate this taking the example

of oil production and consumption as depicted in Figure 7 (production trends) and Figure 8 (consumption trends). Together, these trends dramatically illustrate the validity of this argument when we examine the wide disparities in production and consumption of primary energy among the different countries examined the figures.

As can be seen, U.S.A. is among the lowest producers of energy with nearly flat production levels (measured in thousands of barrels of oil per day) during the period 1970—1996. However, during the same time frame, its dependence on foreign oil has shown a multifold increase. In spite of this increased thirst for oil from outside sources, oil prices in the U.S. have not shown significant increases commensurate with the level of consumption by the U.S. China, on the other hand, seems to be playing a fair game as it has nearly matched its need for increased energy with increases in domestic production. All other countries (Argentina, Brazil, Mexico, India, and U.K.) have shown relatively stable trends in both production and consumption of oil. Japan, of course, has increased its appetite for oil during the period examined; however, its increased dependence on oil is nowhere near the levels seen for U.S.

Further evidence of such disparities can be seen from the IMF policies toward developing countries. IMF has imposed on the debtor nations a number of structural reforms such as reduction of budget deficits, public sector expenditure. However, the issue of double standards in IMF conditionalities stares one in the face: it has not been able to impose them with equal force on the first-world countries. The most glaring example of this contradiction can be found in the case of the U.S. In spite of being the largest debtor nation of the world, the U.S. continues to consume more than what it is producing and to spend more than what it is saving. Further, critiques have been arguing that the U.S., as the most dominant country in the world, engages in arm-twisting tactics aimed at largely benefiting at the cost of those with little or no bargaining power. In other words, the implications of Figures 7 and 8 would seem to extend far beyond the production and consumption of oil.

### Primary Energy Production in Thousands of Barrels of Oil Per Day

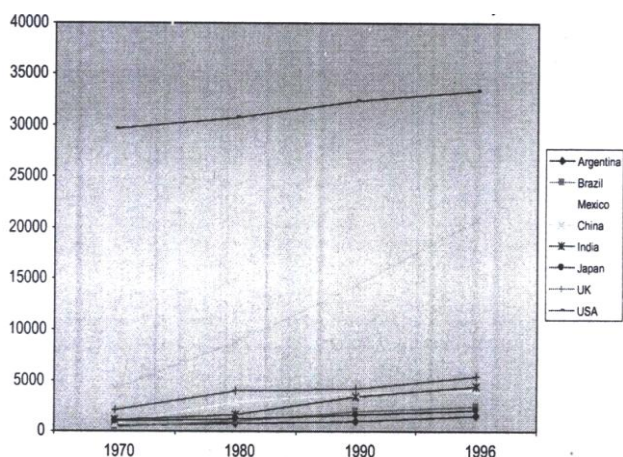


Figure 7. Primary energy production in thousands of barrels of oil per day by country.

### A QUESTION OF DEBT AND HUMAN DIGNITY

Without doubt, free market capitalist philosophies have brought in their wake an enormous burden of foreign debt to developing countries. Figure 9 illustrates this phenomenon.

### Primary Energy Consumption in Thousands of Barrels of Oil per Day

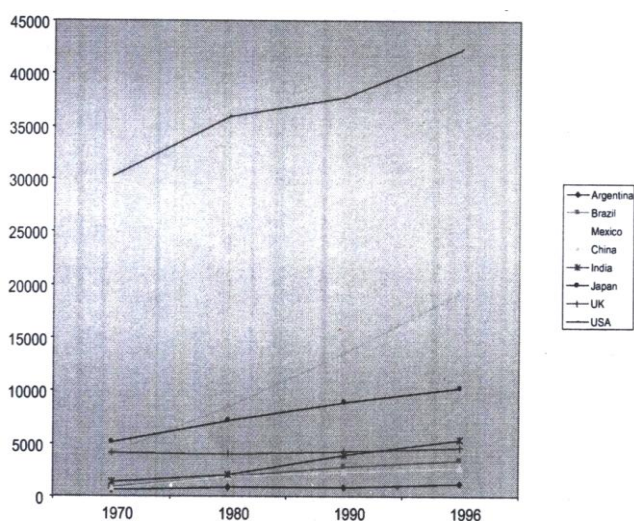


Figure 8. Primary energy consumption in thousands of barrels of oil per day by country and year.

### Foreign Debt of Developing Economies (%)

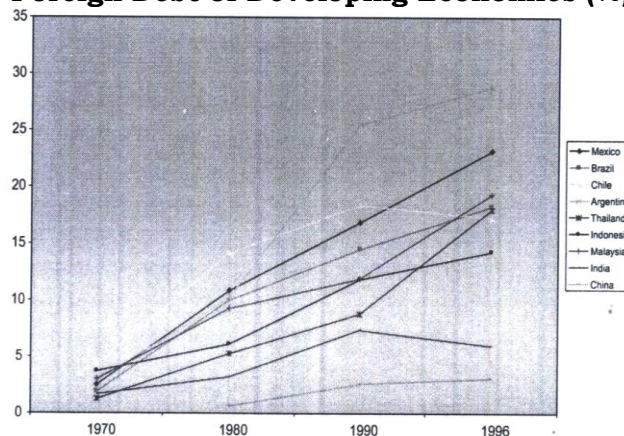


Fig 9. Foreign debt of developing economies.

Foreign debt by itself does not constitute a human or ethical problem; in fact, the United States, as the world's largest debtor country, relies on capital inflows and huge imports from other countries. However, what makes the foreign debt in developing countries an ethical and moral issue is that debt service in those countries maps directly on to human issues of poverty, quality of life, and dignity of people.

The problems of debt servicing began to attract international attention from 1982 onwards when, starting with Mexico, Brazil, Argentina, and the Philippines found themselves in default. The magnitude of the problem and the speed with which it accelerated can be appreciated when we recognize that the total external debt of the three largest debtors of Latin America rose from \$220 billion in 1982 to \$285 billion in 1989.

From Figure 9 it is clearly noticeable that during the 1970–1996 time period, Argentina and Mexico experienced extremely high growth rates in foreign debt, compared to India and China which pursued a prudent approach to the use of foreign debt. The other countries (Brazil, Chile, Thailand, Indonesia, and Malaysia) included in Figure 9 had also shown significant increases in their foreign debt levels during the study period. In the case of Argentina, the discussion by Fernandez and Schumacher (1998) point to the problems of hyperinflation over extended time periods and debt servicing which have escalated into a banking crisis and thus leading to the collapse of capital markets. According to Mishkin (1998), Mexico's

problems were due to increases in interest rates, increases in uncertainty, deterioration of balance sheets of non-financial firms and banks, and eventually the stock market crash. Again, Matthew effect seems to manifest itself in its stark reality in the policies toward third world debt. It is argued that first world policies have effectively decreased commodity prices, increased trade barriers, and adversely affected the ability of developing countries to pay. It is maintained by critics that the World Bank and major creditor nations have precipitated economic and social havoc in the debtor nations while stabilizing the banking institutions of the first world.

A 1988 study by the World Bank, for example, effectively concluded that third world debt retirement involves, in its simplest form, taking from the poor and giving to the rich. The result of these policies for structural adjustment imposed on the third world had been a net flow of resources from the impoverished countries of the South to the industrialized countries of the North.

An argument is also made that in a number of countries such as Bolivia, the Philippines, Chile, Brazil, Zaire, and Argentina, the debts were contracted by governments which were non-representative. This leads one to the moral question of whether ordinary people should sacrifice their subsistence level livelihood in order to service debts which were incurred by non-representative governments and from which they never benefited in the first place.

It could be argued that debt forgiveness is the international equivalent of domestic policies such as the financial bailout of corporate giants like Chrysler and the savings and loans. Some European governments and Canada are already implementing policies aimed at partial forgiveness of third world debt, and the 1988 Toronto economic summit echoed a similar sentiment. Some even question whether anything stands in the way of a general forgiveness of third world debt "except the greed and rigidity of bankers and their political allies."

There is no question that a significant part of the third world debt crisis is attributable to bureaucratic overhead, government sponsorship of inefficient public

sector, and zealous pursuit of irrelevant policies; but any policy which gives a higher priority to debt service than to the accessibility of capital for development and economic growth will be the present day equivalent of feudalism aimed at perpetuating bondage and will pave the way for the Marxist prophecy to come true at the economic level. Given the current direction of the world economy, it seems clear that in the absence of appropriate policy interventions, the concept of individual responsibility for collective good is in danger of being eroded with unbridled market forces dictated purely by the motive of self-interest.

At its most fundamental level, the problem of third world debt is essentially a human problem. As Henriot (1989) pointed out "[structural adjustment programs intended to reform the third world economies] seem to be mandated almost without consideration of their social effects. The impact on the poor is seen as an inevitable short-run consequence, and it is argued that the poor will in fact, benefit 'in the long run'. But as Lord Keynes famously reminded us many years ago: 'In the long run, we will all be dead!' For the poor, this is especially and tragically true."

#### **SOCIAL PERILS OF FREE MARKETS**

In the recent past, the Heritage Foundation and Dow Jones & Co. have come up with an Index of Economic Freedom (IEF) for every country in the world. Each country receives a score between 1 and 5 (1 = Most free; 5 = Most restrictive) and the scores are revised once a year. The countries are also given ranks on the basis of their IEF value (Rank 1 = Most free country). The index is a composite of rankings in each of the following ten variables indicated. Table II contains the IEF values, ranks, and the scores on input variables for a select list of countries. For future research, we suggest an empirical analysis that examines the relationship between a country's IEF value and a set of variables reflecting social and economic welfare. It would be very interesting to find the impact of IEF on socioeconomic differences among countries.

The success stories of the 1980s and the early 1990s from the emerging economies around the world, especially the Asian

countries, have boosted the support for free markets. To everyone's surprise, the recent Asian and Latin American crises have unraveled the perils of free markets. The spillover effect has been felt by all across the world, including U.S. What has gone wrong? Are free markets detrimental to social welfare? Would the countries be better off going back to their controlled market regimes? These questions are very valid. Perhaps, the free markets prescription may not suit well for every country. Probably, the governments and financial institutions in the countries took too much for granted and failed to exercise discipline in their monetary and fiscal policies which is essential for the success of free markets. As examples, we will provide brief descriptions of Mexican, Asian crises below.

### Mexican Financial Crisis

The Mexican financial crisis which began in December 1994 is ascribed to the deterioration in banks' balance sheets due to mounting loan losses (Mishkin, 1996). Banks increased lending to private non-financial businesses from 10% GDP in 1988 to more than 40% of GDP in 1994. Neither the banks nor the National Banking Commission, the primary regulator of banks in Mexico, exercised due diligence and monitoring. As a result, both banks and their borrowers took enormous risks and the bulk of the loans turned sour. Confounding this problem, the Mexican central bank raised interest rates in response to rate hikes in the U.S. and to protect the peso's currency exchange value. The higher interest obligations led to significant cash flow deteriorations in both household and business sectors, especially due to the short-term nature of most of the borrowing in Mexico. Stock values deteriorated because of the increased uncertainties and a decline in the flow of foreign money into Mexico. The dollar denominated debt had to be repaid at not only higher interest costs but with the depreciated pesos. A full fledged financial crisis followed. The Mexican government's desire to limit the peso depreciation and the rise in inflation have led to caused the interest rate on debt denominated in pesos to rise to over 100%. The Mexican stock market declined precipitously recording a 30% drop in terms of the local currency and a 60% drop in dollar terms.

### Asian crisis

The following discussion and data analysis excerpted from the IMF's World Economic Outlook (May, 1998, p. 3) provide insight into the Asian crises and the further developments relating to the crises. The primary source of the problem for the Asian crisis appears to be that the economic policies and institutions have not kept pace with the demand created by large-scale capital inflows into those countries. More specifically, the five major contributing factors are:

- A. Build-up of overheating pressures due to  
(1) large external deficits, and  
(2) inflated property and stock market values.

**TABLE II: Index of economic-freedom rankings**

Country	Avg. score, 1995-1998 (Rank >38)	Trade	Taxation	Govt. intervention	Monetary policy	Foreign investment	Banking	Wage/price	Property rights	Regulation	Black market
Brazil	3.36 (96)	4	2.5	3	5	3	3	3	3	3	4
Chile	2.33 (17)	2	3.5	1	3	2	3	2	1	2	2+
Mexico	3.25 (88)	3	3.5	2+	5	2	4	3	3	4	3
India	3.71 (117)	5	4	3	2	3	4	4	3	4	5
China	3.79 (120)	5	3.5+	5	3	3	3	3	4	4	4
Hong Kong	1.25 (1)	1	1.5	1	2	1	1	2	1	i	1
Singapore	1.29 (2)	1	3	1	1	1	2	1	1	1	1
Malaysia	2.40 (28)	3+	3	2	1	3	3	3	2	2	2
Indonesia	2.98 (62)	2	3.5	1	2	2	3	3	3	4	5
Thailand	2.33 (28)	3	3	2	1	2+	3	3	2-	3	2
Japan	2.03 (12)	2	4.5	1	1	3	3	2	1	2	1
Germany	2.15 (24)	2	5	2+	1	2	3-	2	1	4-	1
U.K.	1.95 (7)	2	4.5	2	1	2	2	2	1	2	1
U.S.A.	1.90 (5)	2	4	2	1	2	2	2	1	2	1
Russia	3.53 (104)	4+	3.5	3+	5	3	2	3	3	4	4

Source: 1998 Index to Economic Freedom, The Heritage Foundation and Dow Jones & Co.

Note: The score is a composite of scores on 10 different variables. The smaller the score, the greater the economic freedom.

B. Maintenance of pegged exchange rates for extended periods which

- complicated the response of monetary policy to overheating pressures,
- viewed as implicit guarantees of exchange value,
- encouraged external borrowing with short maturities, and
- led to excessive exposure to foreign exchange risk in both financial and corporate sectors.

C. Financial system deficiencies such as

- weak management and poor control of risks,
- lax enforcement of prudent rules, inadequate supervision, and
- government-directed lending practices that had led to a sharp deterioration in the quality of banks' loan portfolios.

D. Problems of data availability and lack of transparency which

- hindered a realistic evaluation of economic fundamentals, and
- increased the uncertainty of investments.

E. Problems of governance and political uncertainties which exacerbated

- crisis of confidence,
- reluctance of foreign creditors to roll over short-term loans, and
- down-ward pressure on currencies and stock markets.

Exports by Asian countries suffered primarily due to pegging of their currencies to the U.S. dollar. Two problems - (i) the wide swings in the yen/dollar exchange rate between 1994 and 1997, and (ii) the dollar's appreciation from mid-1995, particularly against yen, have contributed for the export slowdown experienced by several countries in the region in 1996—1997. Further, international investors (mainly commercial and investment banks) have contributed significantly to the downward pressure on the

currencies even though they to substantial losses in the process (p. 3, IMF's World Economic Outlook, May 1998).

There is no denial that both the Mexican and Asian crises have led to large-scale financial disruptions, international economic turmoil, and significant wealth losses to both domestic and foreign investors. The data on foreign currency exposure of Mexican banks (graphs and Table I from Mishkin's paper, p. 52) show us the developments that have led to the Mexican crisis. The data in Table 9 (p. 42, World Economic Outlook, May 1998) enable us to compare the foreign capital inflows into the emerging market economies during the period 1996—1998. The significant declines in foreign capital inflow during the two months, January and February 1998, are especially noteworthy. The data (World Economic Outlook, May 1998, Table 14, p. 78) show that the costs of crises in lost output are quite high.

The discussion and the data analysis presented in this section clearly show that the underlying causes that have led to the Mexican and Asian crises are not due to the existence of free markets *per se* but are due to their abuse. This is similar to what we hear often, "Guns don't kill people; People kill people." The important lesson from these crises is that free market benefits come at a cost. In order for the free markets to contribute for social gains and Pareto improvements, it is important that the respective governments, regulators, bankers, and market participants exercise diligence and exercise restraint.

### **Going it gradually: The incrementalism imperative**

As Korukonda (1999) has argued, in the present world order, continued insulation from the world economy is not a feasible option for any country, let alone a *desirable* option. The policy question facing developing economies, then, is not whether to change, but how' to change. There are two fundamental philosophies of going about implementing change: shock versus gradualism and clearly there are occasions warranting either of them. When a controlled economy is opened up, it is similar to the opening of a sluice gate and allowing gushes

of water to flood in, run their course and in the process, flush out the sludge and stagnation. While this metaphor might be appealing, it is akin to the shock-treatment administered to a dying patient where there is nothing to be lost. Yet, in the reform of once tightly controlled developing economies, the key is to go it slow for the potential for disruption is very high.

A study by Greenwood and Ogun (1992) is quite revealing. The study, which attempts to relate economic freedom, absolute levels of income and growth of GDP, shows that if the first move towards liberalization is above a minimum threshold, growth can be expected to occur in almost direct proportion to the extent of control in the economy.

Bathala and Korukonda (2003), Korukonda (1999), and McKinnon (1993) provide a good discussion of the relative merits of gradual versus rapid liberalizations. (e.g., Eastern block social economies) pursued a rapid path to liberalization. Beginning 1978 The larger state-owned enterprises (SOEs) continued to operate under the ownership and control of the central government. The government ensured price stability in both state and non-state sectors. The macroeconomic data (McKinnon, 1993; p. 177) show China's remarkable progress without major upheavals in contrast to other countries.

#### **Mechanisms first, market integration later: The infrastructure imperative**

An insight from organization theory literature is useful to understand the immensity of challenges brought about by large scale integration of financial markets (Korukonda, 1999). The notion in question is patently obvious: an organization's internal structure ought to reflect the characteristics of the environment in which it operates. If the external environment were very diverse, segmented, and complex, the internal environment of organizations needs be highly differentiated and this differentiation, in turn, calls for integration, or elaborate coordination mechanisms to achieve unity of effort. The relevance of this notion for global integration of financial markets should be clear. Large scale integration injects a sudden shock of complexity into the environments of financial markets of individual countries even as many of them are not adequately equipped to

provide an appropriate structural response to deal with the drastically new, complex, and turbulent environments they face. Such structural responses involve policies aimed at supporting the market's need for reliable information through "healthy" regulation in terms of listing requirements, margin and liquidity requirements, auditing, disclosure and transparency. Without an appropriate infrastructure, such a move is destined to fall under its own weight. The policy message for these countries is as simple as it is straightforward: first things first, infrastructure first and integration later. This is what we shall call the Infrastructure Imperative.

#### **CONCLUSION**

The role and contribution of financial policies is clear and unambiguous where the public policy goals are clearly defined and the desired outcomes are known with a degree of certainty. For example, given that the goal of management is to seek to maximize shareholder wealth, to maximize operational efficiency, to minimize wastage of resources, a number of techniques can be deployed toward achievement of the goals under consideration. The product-mix problem, the breakeven model, the goal programming scenario, the PERT/CPM technique, the Traveling Salesman problem - all allow for clear and unambiguous approaches toward problem solution. Even in cases involving stochastic processes, there is a reasonable level of structure about the process of going about solving the problem. However, broader questions on free-market capitalism and social development - for example, questions as to whether globalization and free market capitalist philosophies are good for a country, which path of social development is to be preferred over which other path - cannot be addressed strictly within the topic domain of quantitative paradigm. Cutthroat competition among capitalist nations, a clear prospect in the new world order, brings in its wake an element of uncertainty about what the emerging wave of economic integration ultimately entails in terms of social justice and human welfare. In this paper we have examined some of these issues using publicly available financial information and long-standing concepts from organization theory.

## REFERENCES

- [1] Aharoni, Y.: 1986, *The Evolution and Management of State-Owned Enterprises* (Ballinger, Cambridge, MA).
- [2] Bathala, C. and A. R. Korukonda: 2003, 'An analysis of social welfare issues in free market environments', *International Journal of Social Economics* 30(8), 854—866.
- [3] Bennett, J. T. and M. H. Johnson: 1979, 'Public versus Private Provision of Collective Goods and Services: Garbage Collection Revisited', *Public Choice* 34, 55-64.
- [4] Boardman, A. E. and A. R. Vining: 1989, 'Ownership and Performance in Competitive Environments: A Comparison of the Performance of Private, Mixed, and State-Owned Enterprises', *Journal of Law and Economics* 32(April), 1-36.
- [5] Boycko, M., A. Shleifer and R. W. Vishny: 1993, *A Theor)' of Privatization* (Harvard University, Boston).
- [6] Caves, D. W. and L. R. Christensen: 1980, 'The Relative Efficiency of Public and Private Firms in a Competitive Environment: The Case of the Canadian Railroads', *Journal of Public Economy* 88, 958-976.
- [7] Fernandez, R. B. and L. Schumacher: 1998, 'The Mexican Financial Crisis Of 1994-1995: An Asymmetric Information Analysis', in S. S. Rehman (ed.), *Financial Crisis Management in Regional Blocs*. pp. 184-208.
- [8] Financial Crises, Causes, and Indicators: *A Survey by the Staff of the International Monetary' Fund*, World Economic Oudook (May 1998): International Monetary Fund, Washington, DC.
- [9] Financial Turbulence and the World Economy: *A Survey by the Staff of the International Monetary Fund*, World Economic Oudook (October 1998): International Monetary Fund, Washington, DC.
- [10] Galal, A., L. Jones, P. Tandon and I. Vogelsang: 1994, *Welfare Consequences of Selling Public Enterprises: An Empirical Analysis*.
- [11] Greenwood, J. and S. Ogus: 1992, *As summarized in The Economist: Asian Economies: Freedom pays*. Gupta, N.: 2001, *Partial Privatization and Firm Performance: Evidence from India*, Working paper.
- [12] Hennot, P. J.: 1989, 'Forgive us our debts ...' *America*, December 9, pp. 420—424.
- [13] Korukonda, A. R.: 1999, 'Policy Imperatives for the International Economy and for Financial Markets in the Emerging World Order', *International Journal of Value-Based Management* 12. 51-67.
- [14] Lux. K: 1990, *Adam Smith's Mistake: How a Moral Philosopher Invented Economics and Ended Morality* (Shambhala, Boston).
- [15] McKinnon, R.: 1993, 'Financial Growth and Macro- economic Stability in China, 1978-1992: Implications for Russia and Eastern Europe', *The Order of Economic Liberalization: Financial Control in the Transition to a Market Economy*
- [16] Merton, R.: 1968, 'The Matthew Effect in Science: The Reward and Communication Systems of Science are Considered', *Science* 159(3810), 56—63.
- [17] Mishkin, F. S.: 1996, 'Understanding Financial Crises: A Developing Country Perspective', in Michael Bruno and Boris Pleskovic (eds.), *Annual World Bank Conference on Development Economies* (The World Bank, Washington, DC), pp. 29-62.
- [18] Mishkin, F. S.: 1998, 'The Mexican Financial Crisis Of 1994-1995: An Asymmetric Information Analysis', in S. S. Rehman (ed.), *Financial Crisis Management in Regional Blocs*. (Kluwer Academic Publishers, Boston/ Dordrecht/London), pp. 149-181.
- [19] Roth, W.: 2002, *The Assault on Social Policy* (Columbia University Press, New York). World Economic Oudook (October 1996), *The Rise and Fall of Inflation: A Survey Staff of the International Monetary Fund* (International Monetary Fund, Washington, DC).

# A METHOD IN SOFTWARE RISK MANAGEMENT MODELING USING ITERATIVE CONTROLLED EVALUATION

SUNIL KUMAR REDDY. K<sup>1</sup>, RAMA KRISHNA REDDY. K<sup>2</sup>, VARUN KUMAR REDDY. G<sup>3</sup>

<sup>1</sup> Assistant Professor, Osmania University, Hyderabad, Telangana, India

<sup>2</sup> Assistant Professor, JNTU Hyderabad, Telangana, India

<sup>3</sup> Assistant Professor, JNTU Hyderabad, Telangana, India

<sup>1</sup> sunil186@gmail.com

<sup>2</sup> ramakrishnareddy524@gmail.com

<sup>3</sup> abts.staff@gmail.com

**I. Abstract**—Every software development project has unique demands, both in terms of project specific features and mostly tacit customer/user expectations. The software developed under requirement engineering will satisfy the users mostly on their perspective. So, recent researches are concentrating on the software development and analysis based on requirement engineering. The requirement engineering processes are also challenged by the risks in developing the software. So an efficient risk analysis system and risk management system is inevitable for the software development process under requirement engineering. A project commonly contains a significant amount of uncertainties, such as continuous change of project scope to accommodate the latest market demand, requirements evolution, imprecise estimation of budget and schedule and high system complexity. Software risk management can be used as an effective tool to manage these risks and to reason under the high degree of the involved uncertainties (Stoica, 1999a; Stoica 1999b; Stoica, 2004). There is still a lack of comprehensive guidelines on how to integrate risk management activities at the early development stage. This study presents SMMICE - SOFTWARE RISK MANAGEMENT MODELING ITERATIVE CONTROLLED EVALUATION integrates from early stage of development in particular from the requirements engineering phase. The early integration supports assessment and management of risks related to the project execution, development environment, operational constraints and project stakeholders even before elicitation of the system requirements. We empirically evaluate the usefulness of SMMICE into active ongoing software development project. Our results show that GSRM is a proactive risk management approach which effectively contributes for a successful software project completion.

## II. INTRODUCTION

Generally risk analysis is used for studying all the considerations, which lead to the failure of the program. It is a methods and techniques for documenting the impact of extenuation strategies [2] and for judging system criticality [3]. Risk analysis is also shown important in the software design phase to assess criticality of the system [4] where risks are examined and necessary steps are introduced. Usually, countermeasures correspond to a design, system fine tuning and then with a limited margin of change. However, it may happen that the risk reduction results in the revision of the entire design and possibly of the initial requirements, introducing thus extra costs for the project [5]. Requirements engineering is a process based method for defining, realizing, modeling, relating, documenting and maintaining software requirements in software life cycle that help to understand the problem better [6]. It has been shown that a large proportion of the publications in software development can be related back to requirements engineering (RE) [7]. RE is the process of discovering the purpose in the software development, by identifying stakeholders and their needs, and documenting these in a form that is amenable to analysis, communication and subsequent implementation [8]. Failures during the RE procedure have a significant negative impact on the overall development process [8]. Reworking requirements failures may take 40% of the total project cost. If the requirements errors are found late in the development process, e.g. during maintenance, their correction can cost up to 200 times as much as correcting them during the early stages of the development process [10]. Adequate necessities are therefore essential to ensure that the system the customer expects is produced and that unnecessary exertions are avoided.

According to Goal-Oriented Requirements Engineering, analysis of stakeholder goals leads to substitute sets of

functional requirements that can each accomplish these goals. These alternatives can be evaluated with respect to nonfunctional necessities posed by stakeholders. In the previous paper, they propose a goal-oriented approach for analyzing risks during the requirements analysis phase. Risks are analyzed along with stakeholder interests, and then countermeasures are identified and introduced as part of the system's requirements.

### III. LITERATURE REVIEW

The following section describes review about some recent works regarding the requirement engineering and risk analysis related to it. Security risk assessment in the requirements phase is challenging because probability and damage of attacks are not always numerically measurable or available in the early phases of development. Selecting proper security solutions is also problematic because mitigating impacts and side-effects of solutions are not often quantifiable either. In the early development phases, analysts need to assess risks in the absence of numerical measures or deal with a mixture of quantitative and qualitative data. GolnazElahietal[ 15] propose a risk analysis process which intertwines security requirements engineering with a vulnerability-centric and qualitative risk analysis method. The method is qualitative and vulnerability-centric, in the sense that by identifying and analyzing common vulnerabilities the probability and damage of risks are evaluated qualitatively. They also provided an algorithmic decision analysis method that considers risk factors and alternative security solutions, and helps analysts select the most cost-effective solution. The decision analysis method enables making a decision when some of the available data is qualitative. JackYAnget *al* [11] has developed an expert system that has least focus on requirement engineering. In facts, requirement engineering is important to get all the requirements needed for an expert system. If the requirements do not meet the clients' needs, the expert system is considered fail although it works perfectly. However, current studies on expert systems are focusing in a specific and narrow domain of problems. Also, the major concern of most researchers is the design issues of the expert system. Therefore, we emphasize on the very first step of success expert system development – requirement engineering. Hence, we are focusing in the requirement engineering techniques in order to present the most practical way to facilitate requirement engineering processes. They have analyzed expert system attributes, requirement engineering processes in expert system developments and the possible techniques that can be applied to expert system developments. Lukas Pilatet *al*[12] have proposed an approach for problem in requirements engineering is the

communication between stakeholders with different languages and background. We experienced a related problem involved with transferring and sharing such knowledge, when stakeholders are reluctant to do this. So, they take a knowledge management perspective of requirements engineering and carry over ideas for the sharing of knowledge about requirements and the domain. We cast requirements engineering as a knowledge management process and adopt the concept of the spiral of knowledge involving transformations from tacit to explicit knowledge, and vice versa. In the context of a real world problem, we found the concept of "knowledge holders" and their relations to categories of requirements and domain knowledge both useful and important. This project was close to become a failure until knowledge transfer has been intensified. The knowledge management perspective provided insights for explaining improved knowledge exchange. Mina Attarha and Nasser Modiri [13] have adopted a critical and specific software systems last longer and they are ought to work for an organization for many years, maintenance and supporting costs of them will grow to high amounts in the upcoming years. In order to develop and produce special aimed software, we should piece, classify, combine, and prioritize different requirements, pre-requisites, co-requisites, functional and nonfunctional requirements.

### IV. PROPOSED APPROACH

The risk analysis has become one of the advanced area in software engineering. In the prior sections, we have discussed about the requirement engineering and the cost risk analysis. SMMICE is a goal-driven approach for the software risk management. The main novelty of this research is to identify and analyse the goals before considering the risks so that goals support risk identification and reasoning about the treatment actions (Islam et al. 2009; Islam and Houmb ,2010). The reason for choosing a goal modelling language is that goals and risks are complementary entities of a software project. A risk is usually defined as negation to a single or multiple goals or a loss of attainment of some corresponding objectives. The proposed approach concentrates on optimizing the number of iterations so that we can efficiently analyze the cost and risk of the project.

### V. TROPOS GOAL RISK MODEL

#### A. *Iterative Controlled Evaluation in Integrating Risk Management into Early Requirements Engineering*

Risk management helps to avoid problems, reworks and disasters exist in software project for stimulate successful project outcomes. It should be an inherent component of

software project (Boehm, 1991) and needs to be considered as early as possible. We advocate considering it in early requirements engineering phase. However, requirements engineering and software risk management are two different processes. The integration endeavour requires considering the interactions among the underlying activities, tasks, methods, roles and dependencies among the artifacts that are involved between the two processes. Artifact oriented view focuses on the dependencies among the requirements engineering or risk artifact types, its content items and associated concepts (Schätz et al., 2002). The process oriented view also focuses on the interaction and dependencies among activities, tasks, along with the roles and responsibilities of requirements engineering and risk management (Fernandez and Kuhrmann, 2009).

*Environment (Internal & External):* Final component deals with the surrounding environment of the software development projects both in-house sourcing and outsourced are the main elements and factors of this component. Organizational stability is another element which focuses on the structure of the organization that facilitates the overall project operation. Organizational structures can have an enormous impact on the overall project execution (Hughes and Cotterell, 1999). A formal organizational structure, in particular, hierarchical structure is articulated in the staff hierarchy chart. However, this type of structure is supported by an informal structure of contacts and communication which gradually grows up among the practitioners during the course of work. Unstructured hierarchy hinders the decision making process and incurs chaos to the development and deployment of the system. Organization structure is also departmentalized based on employees' skill, customer category, product line, services, and geographic location. Effective communication and coordination among the departments are important to keep the overall stability. In software development domain, functional/task oriented approach or based on life cycle phases are used to departmentalize the groups.

*Layer View:* The proposed risk management framework consists of four layers to support the software development risk management model. The advantage of a layer based modelling framework is that it includes suitable task, method and technique for performing specific activity under a given layer. Each layer supports iterative activities for managing software development risks and produces single or multiple artifacts. These artifacts are part of the risk specification concept that support the decision making process during the software development, operation and maintenance. An overview of the individual layer is given below.

*1) Goal Layer:* The goal layer focuses on the factors that contribute effectively to complete the project activities and directly link to the project success. These goals are important as they describe what needs to be done for a project to be successful and for the responsible agents to attain the goals. Goals in SMMICE consider several dimensions of the software development components including project

execution, process, product, human and internal and external environment (i.e. as stated in the previous Section) and map them to the project success indicators. These goals are project specific and focus on the economic benefit, project success criteria and boundary, knowledge gathering and reuse, user satisfaction, quality, vendor reputation, successful delivery and other critical goals of the product. Goal satisfaction requires cooperation among the system agents. For SMMICE, these agents are the project stakeholders, practitioners, tools, language, hardware, development facilities and so on. The main activity of this layer is to identify and model the goals. However, before goal specification, project stakeholder should agree on a concrete risk management plan, in particular, the risk management scope, underlying process, risk threshold and resources. A detailed goal list is the main artifact produced by this layer. Goal modelling supports refinement of coarsely grained higher level goals to finely grained lower level goals through AND or OR refinement (Lamsweerde, 2009). In SMMICE, the latter is referred to as sub-goals. Each sub-goal in the refinement contributes to the parent goal. The more the goals are refined the easier it is to identify and analyze the risk factors that obstruct the goals. A graphical representation of the goal refinement is the core part of the goal model. Satisfaction of these sub-goals certainly attains the main goal. Goal types, such as soft goals are suitable for the goal-driven risk management context as there are generally multiple alternatives for single goal satisfaction. The same sub-goal that relates to a specific development component also contributes to the satisfaction of other development components. These sub-goals are important for the project and require extra care for their fulfilment. Therefore, if required, the goals are prioritized according to their importance to the project success. SMMICE provides a set of guidelines for goal and sub-goal formulation; in particular, we attempt to map the existing goal types and categories from the literature to specify the expectations from the software development components and project success factors. We follow the informal temporal pattern to represent the goal as stated in (Lamsweerde, 2009). The pattern structures an assertion into a prefix and a condition/property. Assertion is the statement of intent of some condition/property of the software development component. These goals are hard or soft by nature. There may also have behaviour goals, which represent the intended behaviour declaratively. Improve, reduce and minimize are the common prefixes for representing the goals. For instance, a statement could be reducing erroneous requirement, whereby the prefix reduce represents a goal that demands a reduction of error from the elicited requirement as undesirable property.

*2) Obstacle Layer:* Obstacles are the main causes that reduce the ability to achieve a single or multiple goals. We treat risk factors as obstacles that directly or indirectly lead to a goal negation and create problems in the project. Obstacles are the opposites of the goals (i.e., undesirable ones that shadow the goals). Therefore, the obstacle categories should be aligned with and derived from the goal categories and model the situation about how several obstacles violate the

identified goals. The obstacle layer enhances the goal clarity. It identifies the potential software development risk factors and formulates the obstruction to the goal dissatisfaction. Similar to the goal model, the obstacle model also refines to provide a complete overview of the risk factors exist in the project. The main focus is to identify as many risk factors as possible so that corresponding control actions can be selected in the early stage. Software development risk factors support different categories of obstacle such as dissatisfaction, lack of adequacy, misinformation, wrong assumption and inaccuracy. The obstacle layer supports the risk identification activities and SMMICE uses a comprehensive list of questions and arranges them by following component-element-factor hierarchy to identify the risk factors. This supports to categorize the risk factors and groups them under the same category. The risk obstacle layer establishes the obstruction link from the risk factors to the sub-goals and from event to the main goal. The same risk obstacle can be relevant to more than one goal. This is important to be expressed by the risk obstacle layer, as it is crucial information when considering effective treatment options. Risk factors that cross-cut several goals are in general more effective to counter. Risk factor identification, categorization, and modelling are the main tasks of this layer. However the risks identified by this layer are not sufficient to determine the control actions because risk factors need to be quantified to determine its severity. Therefore the identified obstacles are analyzed further in the assessment layer.

3) *Assessment Layer*: The assessment layer analyzes the risk events as a consequence of single or multiple risk factors to the goal. The risk quantification is an important first step in assessing the risk (Barki et al., 1993). However, this task is nontrivial due to the inherent subjective nature of the risks in software engineering domain (Kontio, 2001). This layer precisely annotates the individual software development risk event. Furthermore, it also establishes the causal relationship model between risk factors and risk events. The layer focuses on the severity of the risk events consequence to the goal negation. Therefore, relationship among the technical and non-technical software development risks and the corresponding quantitative evaluation to the potential project goals is the main contribution of this layer. For high prioritized software development goals, obstacle identification and refinement should be extensive. We use the Bayesian Belief Network to support the causal relationship model by mapping the risk modelling elements to probability nodes. Each risk event is characterized by two properties: likelihood and impact (Bedford and. Cooke, 1998). Likelihood specifies the possibility of a risk event occurrence and is modelled as a property of risk event. The impact quantifies the negative consequence by the risk event to the single or multiple goals. The same risk factor may lead to more than one risk events and the same risk event can obstruct more than one goals. On the other way around, a goal is obstructed by multiple obstacles that relate risk events and associate factors. This representation allows to model situations where an event is

influenced by more than one risk factor and impacts on one or several goals. An obstruction link is established from the risk event to the specific goal that it obstructs. This supports to construct goal-risk model by refining risk factors to risk events and their combined obstruction of the goals. The obstacle refinement is done in reverse manner compared to the goal refinement. The benefit of obstacle refinement is that we do not need to analyze every individual risk factor separately and in real project situation this point is important, in particular when the budget and efforts are limited.

4) *Treatment Layer*: The final layer focuses on the control actions to counter the risks so that goals can be properly attained. Once the goals, risk factors and risk events are identified and analyzed by the goal, obstacle and assessment layer, then the final task is to implement the suitable cost effective countermeasures. Therefore, the aim of this layer is to control the software development risks as early as possible preferably during requirements engineering phase of the development. The layer is also responsible to monitor the risk status throughout the development and if needed during the operation and maintenance of the product. However, initial considerations focus on the risk events and associated factors that negatively affect several goals, i.e., high prioritized risks. Generally, there exists an alternative countermeasure to the obstacles but should select the most cost effective one for the risk mitigation. This layer includes two different links: contribution link from the control action to the goal that it fulfils and obstruction link from the control action to the specific obstacle that it obstructs. Treatment layer allows modelling, reasoning and tracing the adopted control action for the risk mitigation and goal satisfaction. It also includes responsibility link from the control action to the agent so that specific active agent would be responsible to implement the control action to mitigate the risks. Risk control actions should minimize, prevent or avoid software risks to attain the goals. However, the project context is important in order to identify and select the appropriate control action. If a project is highly risky from the beginning then the selected control action should initially focus to completely eliminate the risk. However, it is not always possible to eliminate the risk factors. The system agents such as human and other development components such as tools or process should be responsible to perform a specific task for the selected control action. The risk treatment considers risk threshold that is the specific level up to which a project can accept the risk without implementing any control action. Once the selected risk control action is implemented then we need to monitor the risk status until it is completely mitigated. The risk status through the course of development evolves. In particular, control actions may not effectively reduce the risks or new risk may be identified. Therefore, the treatment layer continues to monitor risks throughout the development and communicates with the stakeholder about the risk status by risk status report.

## B. Related Works

The theoretical foundation of putting risk management

into a single framework is initially contributed by Boehm (Boehm, 1991). Boehm's risk-driven Spiral model was the first life cycle model to integrate risk management throughout software development life cycle in an iterative manner. Afterwards several works focus to develop risk management framework such J. Kontio Riskit model (Konnito, 2001), Roy's pro risk management framework (Roy, 2004). Thus far, some works have tackled the problem of considering risk management as part of the requirements phase. Boness et al. (Boness, 2008) contribute requirement risk mainly late stage in requirements engineering. Procaccino et al. (Procaccino, 2002) identified seven early development factors and discussed how these contribute to the success or failure of a software project. There are several works which identify list of risk factors based on surveying experience software practitioner. (Islam et al. 2009) identified a list of goals and risk factors for offshore outsourced software projects. Schmidt et al. 2001 published a comprehensive list of risk factors by conducting Delphi survey study to the experienced project practitioners through three different panels located in three different countries. Ropponen et al. 1999 conducted a survey to investigate six software development risk components and showed how to provide assistance in addressing these components.

In SMMICE, we follow some concepts from these approaches in particular to identify the early development components and during the development of the model. There are published works that look into risk factors based on surveys and interviews with experienced software practitioners. But less contribution on investigating the impact of risk management process into a software project. Our work introduces a goal-driven approach for software project risk management and presents the results of an empirical study on its impact into a running software project.

### C. Discussion

All above mentioned recommendations seem obvious, but are indeed difficult to spot without a systematic risk management practice from early development. SMMICE aims to include and structure activities that help the developers to identify and address potential problems from a holistic perspective. The whole SMMICE process was not implemented in particular risk treatment and monitoring was not executed by the risk management team. The results including goals, risk factors and recommended control actions were given to the project manager.

*Iterative Controlled Evaluation approach for risk management:* Iterative Controlled Evaluation approach guidelines and improves risk management activities. One of the lessons learned from the case study was that identifying the project riskiness and explicit identification of goals and risk factors made it easy to identify and assign treatments. Treatments were shown to be cost-effective and doable in practice and raised the potential project success rate to a pre-operative level. Another lesson learned was that the

underlying principles of SMMICE were relatively easy to communicate to the software practitioners, especially the risk and requirement analysts. Even though SMMICE was not fully implemented, project manager really appreciated the obtained result from SMMICE. The learning time for these experts to understand and use the model was relatively short due to the familiarity of the experts involved in the project with the concepts of goals and risks.

*Integration of risk management into RE:* SMMICE works as natural extension of the requirements engineering methodology used within the project, so it does not add any extra burden to the development activities, but rather contributes by adding value. Goals and risk identification activities can jointly be executed with requirements elicitation and analysis phase. The project manager agreed that without using early risk management approach they might discover the risks lately which would increase any potential loss in the project. It is easy to negotiate the risks and control actions with the management and customer representative during the requirements engineering. The observations made from the study showed that the chance for a successful project increases substantially if specific customer-developer knowledge transfer activities are undertaken as early in the requirements phase as possible.

*Limitations:* Besides the benefits, some limitations were also observed. When it comes to the risk assessment layer, it became evident that the practitioners were not able to fully express their uncertainty using the three level qualitative scales for likelihood and severity estimation. The feedback suggests that it would be more effective and probably also produce more accurate estimates if the qualitative scale was extended to include five levels. Another related point observed is that it is not always possible to quantify risks, especially in the early stage and risks related to product quality, requirement error, and business specification. This is mainly due to the high level of uncertainty at this stage and the lack of solid experience data.

### D. Study Validity

*Internal Validity:* Data is collected from multiple sources through interview, brainstorming session and project specific documents. Both student and project team member were involved in the data analysis. No co-author is directly involved in the empirical investigation. This certainly reduces the bias of the study result and affects the interpretation of single data source. Maturation effect was also controlled as the SMMICE implementation and delivered results were done within requirements engineering phase and did not continue later. As stated SMMICE was not fully implemented, this might influence the internal validity in a certain context.

*External validity:* To generalize our findings, the obtained results were compared with the studied results. It is interesting to note that the identified risk factors coincide with the

published risk factors (Saarinen , 1996; Bannerman,2008; Procaccino and Verner, 2002; Ropponen,1999; Kwak and Stoddard, 2004). Risk factors such as: lack of domain expert,

TABLE I

PARTIAL CASE STUDY RESULT

Goals	Risk factors	Events	Treatments
Successfully system upgrade	Lack of domain expert	Erroneous requirements	Employ domain expert
Complete project in estimated budget	Lack of knowledge	Unclear vision	Collect adequate information
High reputation	Incomplete requirements	Increase complexity and volatility	Right user involvement
Clear scope and deliverables	Wrong user involvement	User dissatisfaction	Revision of release plan
Complete information	Inadequate release plan	Poor release plan	Revision of system vision
System usage	Unclear system vision		Increase frozen requirements
Error free requirements	Compliance with act		
	Poor documentation		

requirements problems, and unclear vision are also common in other studied results. Risks relating to requirements seem to be a cross development context problem and therefore independent from the local context. However, poor documentation and release plan are this study specific. It is also important to look into the stability of the questionnaires during the interview session. In particular, it is important that all participants clearly understand the questions to provide unambiguous answers. We have tried to control this bias by spending time on explaining the key terms and concepts of the questionnaires as part of the brainstorming session.

## VI. CONCLUSION

The paper presents SMMICE, a modeling framework to manage software project risks at early development stage. Using a holistic goal-driven approach for risk management and its explicit integration into requirements engineering, risks are assessed and managed up-front so that subsequent development can effectively complete the specific tasks for successfully project completion. SMMICE includes goals as the objectives, expectations and constraints from the development components by focusing on the project success factors for the risk management. The SMMICE approach has been evaluated by implementing it in a running development project. The model considers goals beyond schedule, budget and quality and realizes the importance of motivating project stakeholders in particular customer/user to take an active part in the software project. It focuses on the non-technical components such as project execution, customers/ users, project participants and usage environment, along with the technical components such as development process, system specification and tools as a holistic view for the risk management. The results of this case study demonstrated the importance of being specific when it comes to the project goals and stakeholder interests, as well as the composition of the development team. The study also showed that by structuring goal and risk identification, it was easy to assign effective treatment actions. We have yet to show evidence of the true scalability of the SMMICE model. This is part of future work. We also plan to continue evaluating the model within the development context that it is tailored for and to investigate how it can be applied outside this context.

## REFERENCES

- [1] Lamsweerde, A(2009). *Requirements Engineering: From System Goals to UML Models to Software Specifications*, Wiley.
- [2] Risk Analysis as part of the Requirements Engineering Process YulistiraAsnar Paolo Giorgini March 2007
- [3] Boehm BW (1991) Software risk management: principles and practices. IEEE Software, pp.32– 41. doi:10.1109/52.62930
- [4] B. W. Boehm. Software Risk Management: Principles and Practices. IEEE Software, pp.32–41, 1991
- [5] Focusing on the Importance and the Role of Requirement Engineering Mina Attarha and Nasser Modiri Atta.
- [6] H.F. Hofmann and F. Lehner, "Requirements engineering as a success factor in software projects", *IEEE Software*, vol 18, no 4, pp. 58-66, 2001.
- [7] B.A. Nuseibeh and S.M. Easterbrook, "Requirements engineering: A roadmap", Proc. of the 22nd Intl. Conf. on Software Engineering (ICSE '00), IEEE Computer Society Press, June 2000, pp. 35 – 46.
- [8] T. Hall, S. Beecham and A. Rainer, "Requirements problems in twelve software companies: An empirical analysis", *IEEE Software*, vol 149, no. 5, pp. 153-160, 2002.
- [9] M. Niazi and S. Shastry, "Role of requirements engineering in software development process: An empirical study", Proc. of the 7<sup>th</sup> Intl. Multi Topic Conf. (INMIC2003), IEEE Computer Society Press, Dec 2003, pp. 402-407.
- [10] JackyAng, Sook Bing Leong, Chin Fei Lee, UmiKalsomYusof, "Requirement Engineering Techniques in Developing Expert Systems", School of Computer Sciences UniversitiSains Malaysia, IEEE, symposium on computer science and informatics, 2011, pp.1-2
- [11] Lukas Pilat and Hermann Kaindl, "A Knowledge Management Perspective of Requirements Engineering", Institute of Computer Technology Vienna University of Technology Vienna, Austria, IEEE Conference: 19-21 May 2011, pp.1-12.
- [12] Mina Attarha and Nasser Modiri, "Focusing on the Importance and the Role of Requirement Engineering", Interaction Sciences (ICIS), 2011 4th International Conference, 16-18 Aug. 2011, pp. 181 – 184.
- [13] YulistiraAsnar, Paolo Giorgini, "Risk Analysis as part of the Requirements Engineering Process", Departmental Technical Report, 2007.
- [14] GolnazElahi, Eric Yu, Nicola Zannone, "Security Risk Management by Qualitative Vulnerability Analysis", Third International Workshop on Security Measurements and Metrics, pp: 1-10, 2011.



## Designing Efficiency Model and Control Fake Reviews Using Sentiment Analysis

Mr.K.Rajeshwar Rao\*<sup>1</sup>, Mr.Md. Rafeeq\*<sup>2</sup>, Dr. C. Sunil Kumar\*<sup>3</sup>, Mr. A.MAHENDAR \*<sup>4</sup>

Asst. Prof., Dept. of CSE, MREC (A), Secunderabad.  
 ,M.Tech (PhD), Assoc. Prof, CMRTC Kandlakoya, Hyderabad.  
 Professor, Dept. of ECM, SNIST, Ghatkesar, Hyderabad  
 ,M.Tech (Ph.D), Assoc. Prof, Dept. Of CSE, MRIET, Secunderabad.

### ABSTRACT:

Medication reviews collect from patients over the internet. All reviews are helpful for drug discovery using aspect mining environment process. Previous approaches are discovering limited clinical trial for post marketing. Detection of drug reactions is the most important task in pharmaceutical industry. Here major challenge is mining of short and noisy reviews.

In this paper mining of drug reviews using probabilistic aspect mining model environment. Probabilistic aspect mining model describes and identify class labels based reviews classification. Select class label identify suitable reviews and form as a group. Same process of execution we can apply for different classes. It can be helpful for control the mixed aspects of reviews. Finally we display easily each and every aspect of reviews efficiently.

**KEYWORDS:** aspect mining model, probabilistic model, class labels, classification, reviews, aspects, internet.

### INTRODUCTION

Discussion forums, blogs are good sources to collect diseases related reviews information from patients. Current many numbers of applications extracting useful information is available as a major problem and difficulty also. Opinion mining extracts information from large amount of text opinions. Previously various mining approaches are extracting useful information. These approaches are not

helpful to provide user expected and needed information effectively.

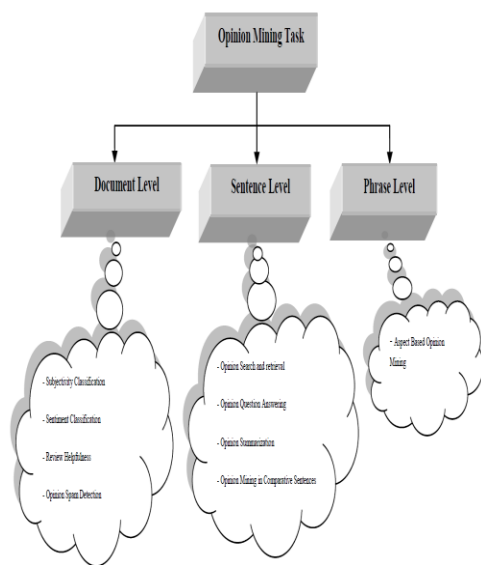
In this paper we can use the opinion mining and sentiment analysis concepts to extract useful information from large amount of reviews content. Here some other different algorithms are design for extracting some more useful content like age wise and other features based also. Its most helpful for patients to suggest the drugs.



Finally we can conclude current section; next sections consist of different algorithms discussion related to existing and proposed system.

## II.RELATED WORK:

Opinion mining has been an emerging research field in text analytics in recent years. It is related people opinions computational study. Different kinds of platforms are available as a source to express and share people's opinion. Vast amount of opinions are placed in the form of different formats. Those formats are reviews, blogs, tweets etc,. So we need effective systems for evaluation of opinions and generate accurate results.



**Fig1: Different Possibilities of opinion mining task**

First approach is frequency based approach. It extracts high frequency noun phrases which are available from reviews as aspects. But all number of aspects has no relationship. These reviews were not meaningful and useful. Again we introduced as a relational based approach. It can recognize sentimental relationship results from reviews. The above two approaches was not suitable for drug reviews. These approaches are not generating semantic meanings aspects information.

Next here topic modeling techniques are popular. These techniques are applied on different aspects of messages. Topic modeling identifies aspects based on the co-occurrences of words in reviews. Particular topic related same aspects of messages are group here. Group of co-occurrences words probabilities we should calculate here. Consider probability perform sorting operation on topic related words information. Sorting words are categorized into two classes. 1. High probability words 2. Low probability words information. High probability words are high correlated and semantic words. This is one of the way to select same sentiments information efficiently.

Other different aspect based mining techniques are designed by the different authors. These new authors are introduces a new topic modeling approach. Topic modeling approaches works on supervised



label information. Other algorithms are ready to work on other different properties. These are the things related to probabilistic algorithms.

Deterministic algorithms we can design for topic modeling environment. Algorithm extracts relevant features and display into a matrix. Matrixes decompose into two low rank matrices again. Apply supervised steps finally we can recognize closest information. Other different probabilistic models are found the aspects are correlated class labels information. Finally we can remove mixed content related information. After display all contents in the form different classes. All different classes are not useful.

### III.PROBLEM STATEMENT:

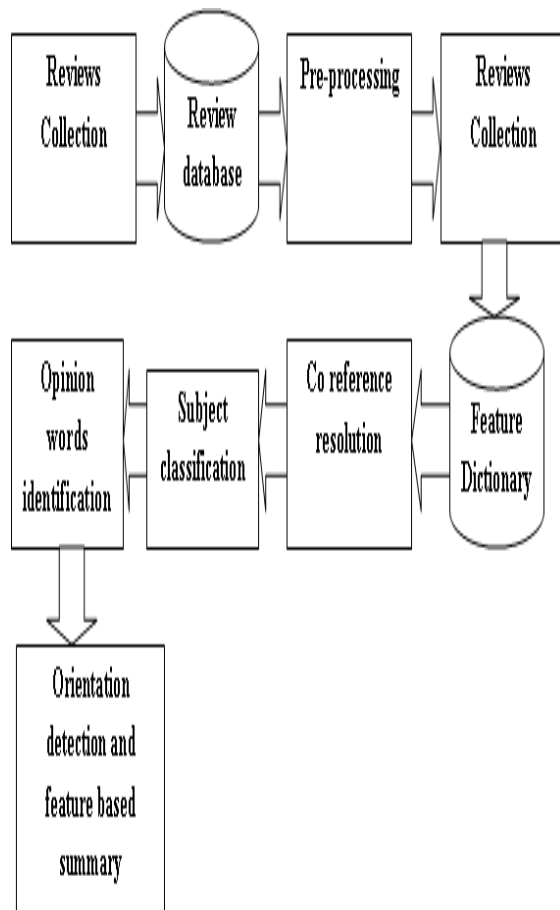
In this paper we can find aspects related to different segmentations of data. Those different segmentations of data are related different age groups and other attributes also here. Different age groups content work is more useful. It expresses action of each and every keyword clearly. Finally we can display list keywords information and actions also as a final output. Second step we can consider the input as a sentences. All sentences also we can process and provide summarize results which is more useful and semantic. All users can understand the summarization greatly.

### IV.PROPOSED METHODOLOGY IMPLEMENTATION

Here we propose a probabilistic model for finding aspects which is correlated to class labels. Previous discussions one class related correlated aspect messages recognized here. Those words were not meaningful. Now here we focus on aspects identification on different class labels. Those mixed class labels of aspects are avoid in this paper effectively. Each and every class aspects we can find out individually. All reviews we can group based on class label.

Anyway differentiating reviews perform based on different classes. Identify class related aspects reviews and send to target class. These reviews are more helpful for customers are users. Here we design aspect sentiment analysis model. Architecture divides into different phases. Those phases are

1. Reviews collection
2. Pre-processing
3. Feature extraction
4. Co-reference resolution
5. Subject based classification
6. Opinion words identification
7. Orientation detection



**Fig2: proposed aspect sentiment analysis model**

### Reviews collection:

Collect reviews about drugs from online website and prepare as a data set. After preparation of data set then perform sentiment data analysis. Download all opinions and reviews from reviews collection module. All collected reviews we can store into database.

### Pre processing:

This is second step for filtering reviews, improve accuracy and also avoid unnecessary overhead issues also. Here we can perform the opinion mining process.

Step1 : stop words removal

Step2: before sentiment analysis we can remove like numbers and symbols also.

The above steps are helpful to perform faster opinion mining process.

### Feature Extraction:

Identify the features as a single word or phrase. Identify the domain specific features based on feature dictionary. Manually we can add known features of a drug and create sentiment profile environment solution. This sentiment profile also we can store into database as a meta Data information.

### Co-reference resolution:

Identify the entities related reference information. It controls unnecessary number of irrelevant phrases from total phrases. Finally we can extract co-reference related sentences information. We can change the words in sentences then it's possible to resolve the problem efficiently.

### Subjectivity Classification:

All reviews are not consisting of opinions. Analyzation of all sentences then we can go for subject based categorization



as an opinion sentences and non-opinioned sentences. Subject based sentences are classified based on object. Next one more step classifies the sentences based on feature dictionary.

### Opinion words identification:

Opinion words are normally verb, adverb and adjective which express the polarity like positive and negative. Identify the features based on dependency words categorization. It is most helpful for extracting extra features information. Here parser classifies sentences based aspect wise in our implementation process.

### Orientation Detection:

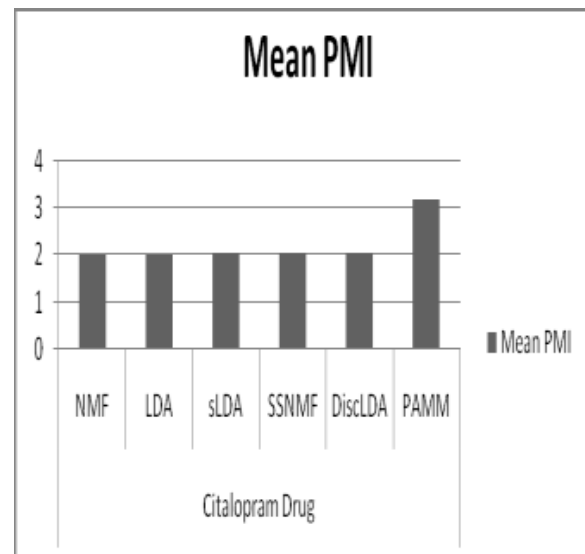
Calculate the score about positive and negatives sentences information. Calculate sentence level of score of the opinion by analyzation of each sentence. Finally we can display overall score of each feature.

## V.RESULTS AND DISCUSSION

Performance analysis can be done based on text analysis environment. Text analysis performs based on class labels and words content. Different algorithms shows different accuracy results based on point wise mutual information.

Product	Algorithm	Mean PMI
Citalopram Drug	NMF	2.03
	LDA	2.03
	sLDA	2.07
	SSNMF	2.06
	DiscLDA	2.07
	PAMM	3.20

**Table1: accuracy related different algorithms**



**Fig3: performance graph**

## VI.CONCLUSION AND FUTUTE WORK

Proposed algorithm for mining aspects based class labels and sentiment words. This model gives less reviews



information related to each and every aspect. It can perform based segmentation model environment process based age groups and other attributes also.

## VII. REFERENCES

- [1] T. O'Reilly, "What is web2.0: Design patterns and business models for the next generation of software," Univ. Munich, Germany, Tech. Rep. 4578, 2007.
- [2] D. Giustini, "How web 2.0 is changing medicine," *BMJ*, vol. 333, no. 7582, pp. 1283–1284, 2006.
- [3] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proc. 10th ACM SIGKDD Int. Conf. KDD*, Washington, DC, USA, 2004, pp. 168–177.
- [4] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Ret.*, vol. 2, no. 1–2, pp. 1–135, Jan. 2008.
- [5] A.-M. Popescu and O. Etzioni, "Extracting product features and opinions from reviews," in *Proc. Conf. Human Lang. Technol. Emp. Meth. NLP*, Stroudsburg, PA, USA, 2005, pp. 339–346.
- [6] L. Zhuang, F. Jing, and X. Zhu, "Movie review mining and summarization," in *Proc. 15th ACM CIKM*, New York, NY, USA, 2006, pp. 43–50.
- [7] Q. Mei, X. Ling, M. Wondra, H. Su, and C. Zhai, "Topic sentiment mixture: Modeling facets and opinions in weblogs," in *Proc. 16th Int. Conf. WWW*, New York, NY, USA, 2007, pp. 171–180.
- [8] S. Moghaddam and M. Ester, "Aspect-based opinion mining from online reviews," in *Proc. Tutorial 35th Int. ACM SIGIR Conf.*, New York, NY, USA, 2012.
- [9] B. Liu, M. Hu, and J. Cheng, "Opinion observer: Analyzing and comaring opinions on the web," in *Proc. 14th Int. Conf. WWW*, New York, NY, USA, 2005, pp. 342–351.
- [10] C. Lin and Y. He, "Joint sentiment/topic model for sentiment analysis," in *Proc. 18th ACM CIKM*, New York, NY, USA, 2009, pp. 375–384.
- [11] I. Titov and R. McDonald, "A joint model of text and aspect ratings for sentiment summarization," in *Proc. 46th Annu. Meeting ACL*, 2008, pp. 308–316.
- [12] S. Baccianella, A. Esuli, and F. Sebastiani, "Multi-facet rating of product reviews," in *Proc. 31st ECIR*, Berlin,, Germany, 2009, pp. 461–472.
- [13] W. Jin, H. Ho, and R. Srihari, "Opinionminer: A novel machine learning system for web opinion mining and extraction," in *Proc. 15th ACM SIGKDD Int. Conf. KDD*, New York, NY, USA, 2009, pp. 1195–1204.
- [14] Y. Jo and A. Oh, "Aspect and sentiment unification model for online review analysis," in *Proc. 4th ACM Int. Conf. WSDM*, New York, NY, USA, 2011, pp. 815–824.
- [15] J. Sarasohn-Kahn, "The wisdom of patients: Health care meets online social media," California Healthcare Foundation, Tech. Rep., 2009.
- [16] K. Denecke and W. Nejdl, "How valuable is medical social media data? content analysis of the medical web," *J. Inform. Sci.*, vol. 179, no. 12, pp. 1870–1880, 2009.
- [17] X. Ma, G. Chen, and J. Xiao, "Analysis on an online health social network," in *Proc. 1st ACM Int. Health Inform. Symp.*, New York, NY, USA, 2010, pp. 297–306.



- 
- [18] A. Névél and Z. Lu, "Automatic integration of drug indications from multiple health resources," in *Proc. 1st ACM Int. Health Inform. Symp.*, New York, NY, USA, 2010, pp. 666–673.
- [19] J. Leimeister, K. Schweizer, S. Leimeister, and H. Krcmar, "Do virtual communities matter for the social support of patients? Antecedents and effects of virtual relationships in online communities," *Inform. Technol. People*, vol. 21, no. 4, pp. 350–374, 2008.
- [20] R. Schraefel, R. White, P. André, and D. Tan, "Investigating web search strategies and forum use to support diet and weight loss," in *Proc. 27th CHI EA*, New York, NY, USA, 2009, pp. 3829–3834.
- [21] J. Zrebiec and A. Jacobson, "What attracts patients with diabetes to an internet support group? A 21-month longitudinal website stuey," *Diabetic Med.*, vol. 18, no. 2, pp. 154–158, 2008.
- [22] T. Mitchell, *Machine Learning*. Boston, MA, USA: McGraw-Hill, 1997.
- [23] R. Agrawal and R. Srikant, "Fast algorithms for mining association rules," in *Proc. 20th Int. Conf. VLDB*, San Francisco, CA, USA, 1994, pp. 487–499.



# Adapting Enforced Mechanism For Monitoring Health Using Cloud Computing

G.Srikanth<sup>1</sup>, Mrs. M.Ahalya Rani<sup>2</sup>

<sup>1</sup>M.Tech Student, Computer Science, and Engineering, Malla Reddy Engineering College (Autonomous),  
Hyderabad, Telangana, 500100 India

<sup>2</sup>Assistant Professor, Computer Science, and Engineering, Malla Reddy Engineering College (Autonomous),  
Hyderabad, Telangana, 500100 India

<sup>1</sup>gsrikanth619@gmail.com

<sup>2</sup>ahalya.rani@mrec.ac.in

---

**Abstract:** *Cloud-helped versatile (health) observing, which applies the across the board portable correspondences and distributed computing advancements to give input choice backing, has been considered as a lobbyist way to deal with enhancing the quality on both customers' protection and licensed innovation of checking administration suppliers, which could dishearten the wide selection of mHealth innovation. This venture is to address this essential issue and outline a cloud-helped security safeguarding portable health observing framework to ensure the protection of the included gatherings and a recently proposed key private intermediary re-encryption are adjusted to move the computational unpredictability of the included gatherings to the cloud without trading off customers' security and administration suppliers' shows the adequacy of our exhibits the viability in distributed computing environment..*

**Keywords-----** *Healthcare, key private intermediary re-encryption, Mobile Health (Mhealth), Outsourcing Decryption, Privacy.*

---

## I. INTRODUCTION

Plan of action in distributed computing, which would Wide organization of mobile phones, for example, advanced mobile phones furnished with minimal effort sensors, has as of now demonstrated incredible potential in enhancing the nature of social insurance administrations. Remote portable health observing has as of now been perceived as a potential, as well as an effective illustration versatile health (mHealth) applications particularly for creating nations. The Microsoft dispatched venture "MediNet" was intended to acknowledge remote checking on the health status of diabetes and cardiovascular infections in remote regions in Caribbean nations. In such a remote mHealth observing framework, a customer could convey versatile sensors in remote body sensor systems to gather different physiological information, for example, circulatory strain (BP), breathing rate (BR), Electrocardiogram (ECG/EKG), peripheral oxygen saturation (SpO) and blood glucose. Such physiological information could then be sent to a focal server, which could then run different web medicinal applications on this information to return opportune exhortation to the customer. The applications may have different functionalities extending from rest example analyzers, works out, physical action aides, to heart investigation frameworks, giving different restorative counsel. Also, as the rising distributed computing innovations develop, a feasible arrangement can be looked for by consolidating the product as an administration (SaaS) model and pay-as-you-go permit little organizations (social insurance administration suppliers) to exceed expectations in this human services market. It has been watched that the reception of mechanized choice bolster calculations in the cloud-helped mHealth observing has been considered as a future pattern. Tragically, despite the fact that cloud-helped mHealth observing could offer an awesome chance to enhance the nature of medicinal services administrations and conceivably diminish social insurance costs, there is a hindrance in making this innovation a reality. Without legitimately tending to the information administration in an mHealth framework, clients' security may be extremely broken amid the gathering, stockpiling, conclusion, and correspondences and computing. A late study demonstrates that 75% Americans consider the protection of their health data essential or critical. It has additionally been accounted for that patients' readiness to get included in health observing system could be extremely brought down when individuals are concerned with the protection break in their deliberately submitted health information. This security concern will be exacerbated because of the developing pattern in protection ruptures on electronic health information. Despite the fact that the current security laws, for example, HIPAA (Health Insurance Portability and Accountability Act) give standard assurance to individual health record, they are for the most part considered not appropriate or transferable to distributed computing situations [6]. Also, the present law is more centered around insurance against ill-disposed interruptions while there is little exertion on shielding customers from business gathering private data. In the interim, numerous organizations have

critical business intrigues in gathering clients' private health information and imparting them to either insurance agencies, research establishments or even the administration offices. It has likewise been shown that security law couldn't generally apply any genuine assurance on clients' information protection unless there is a compelling system to implement confinements on the exercises of social insurance administration.

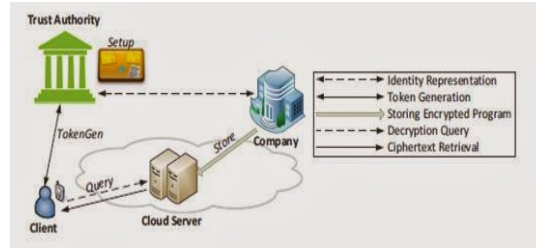


Figure 1 SYSTEM ARCHITECTURE

#### A. SYSTEM MODEL AND ADVERSARIAL MODEL

To encourage our dialog, we first expound our cloud helped mHealth observing framework (CAM). CAM comprises of four gatherings: the cloud server (essentially the cloud), the organization who gives the mHealth observing administration (i.e., the medicinal services administration supplier), and the individual customers (just customers), and a semi-trusted power (TA). The organization stores its encoded observing information or system in the cloud server. Singular customers gather their medicinal information furthermore, store them in their cell phones, which then change the information into characteristic vectors. The property vectors are conveyed as inputs to the checking project in the cloud server through a portable (or savvy) gadget. A semi-trusted power is in charge of conveying private keys to the singular customers and gathering the administration charge from the customers as per a certain plan of action, for example, pay-as-you-go plan of action. The TA can be considered as a coconspirator or a administration specialists for an organization (or a few organizations) and in this way imparts certain level of common enthusiasm to the organization. Then again, the organization and TA could plot to acquire private wellbeing information from customer data vectors. We expect a unbiased cloud server, which implies it neither plots with the organization nor a customer to assault the other side. This is a sensible model since it would be in the best business enthusiasm of the cloud not to be one-sided. We concede that it stays feasible for the cloud to intrigue with different vindictive elements in our CAM, and we leave the CAM outline under these more grounded models as future work. We additionally don't expect that an individual customer connives with different customers. Our security model does not consider the conceivable side-channel assault because of the co-residency on shared assets either since it could be moderated with either framework level insurance or spillage strong cryptography. CAM expect a legitimate however inquisitive model, which suggests all gatherings ought to take after the endorsed activities and can't be self-assertively noxious. In the accompanying, we quickly present the four noteworthy steps of CAM: Setup, Store, Token Gen and Query. We just represent the usefulness of these parts in this segment while leaving the subtle elements in later segments. At the framework instatement, TA runs the Setup stage and distributes the framework parameters. At that point the organization first communicates the stream diagram of the mHealth checking system as a spreading program, which is encoded under the particular coordinated stretching tree. At that point the organization conveys the subsequent figure content and its organization file to the cloud, which relates to the Store calculation in the setting. At the point when a customer wishes to question the cloud for a certain mHealth observing program, the  $i$ -th customer and TA run the Token Gen calculation. The customer sends the organization record to TA, and after that inputs its private question (which is the property vector speaking to the gathered wellbeing information) and TA inputs the expert mystery to the calculation. The customer acquires the token relating to its question information while TA gets no valuable data on the individual question. Amid the last stage, the customer conveys the token for its question to the cloud, which runs the Query stage. The cloud finishes the major computationally escalated errand for the client's decoding and returns the somewhat unscrambled figure content to the customer. The customer then finishes the remaining decoding errand subsequent to accepting the incompletely unscrambled figure message and

acquires its unscrambling result, which relates to the choice from the observing project on the clients' info. The cloud acquires no helpful data on either the client's private question info or unscrambling result in the wake of running the Query stage. Here, we recognize the inquiry information protection rupture as far as what can be surmised from the computational or correspondence data. CAM can keep the cloud from deriving valuable data from the client's inquiry information or yield comparing to the got data from the customer. Be that as it may, the cloud may in any case be ready to reason side data on the client's private inquiry information by watching the client's access design. This issue could be determined by careless RAM strategy.

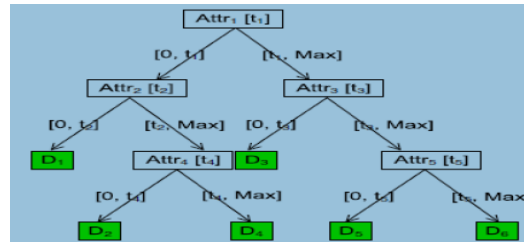


Figure 2 Branching program

### A. Branching Program

In this segment, we formally portray the stretching programs, which incorporate double arrangement or choice trees as an extraordinary case. We just consider the double stretching program (as demonstrated in Fig. 1) for the simplicity of work following a private question convention in light of a general choice tree can be effectively gotten from our plan. Let  $v = (v_1, \dots, v_n)$  be the vector of clients' traits. To be more particular, a property part  $v_i$  is a linking of a trait file and the particular property estimation. Case in point,  $A||KW1$  may relate to "circulatory strain: 130". Those with a blood weight lower than 130 are considered as typical, and those over this edge are considered as hypertension. Every quality worth is a C-bit whole number. In this paper, we pick C to be 32, which ought to give enough accuracy in most handy situation.

## II. PROPOSED SYSTEM

In this paper, we propose another secure and protection safeguarding artful figuring system, called CAM, to address this test (fig 2). With the proposed CAM structure, every therapeutic client in crisis can accomplish the client driven protection access control to permit just those qualified assistants to take an interest in the shrewd figuring to adjust the high dependability of procedure and minimizing protection exposure in m-Healthcare crisis. We present an effective client driven security access control in CAM structure, which is taking into account a trait based access control and another security saving scalar item processing (PPSPC) system, and permits a therapeutic client to choose who can take part in the sharp processing to help with preparing his staggering information.

### A. SYSTEM MODEL

In medicinal services mindful social insurance advantages of our framework, a faculty restorative at the inside that is considered reliable is for introducing and controlling the whole framework. A client who wishes to get the versatile human services framework registers him as a medicinal client under a specific human services focus, and after that a medicinal expert looks at the client and creates his wellbeing profile Based on the wellbeing profile, the clients are then given with the specific sort of information, for example, heart rate, glucose level and different materials. Once being furnished with the sensors the clients can move anyplace not at all like in doctor's facility. The sensors start to gather the detected information and transmit them to the user's advanced cell which is then transmitted to the wellbeing consideration focus. The s the advanced cell assumes an essential part in portable observing of patients. The advanced mobile phones are utilized for different purposes, the force of the advanced mobile phone may not be adequate under crisis circumstances. Consequently we make utilization of artful figuring where at whatever point a therapeutic client is in crisis other therapeutic clients in the close-by territory can contribute their asset.

### B. INITIALIZING SYSTEM

As indicated by our work, the individual at the human services focus is in charge of introducing the whole framework. The power at the medicinal services focus creates the bilinear parameters  $(g, sp)$  by running  $gen(sp)$  utilizing the security parameter  $(sp)$ . He likewise chooses the encryption calculation that is to be utilized, two safe cryptographic hash capacities  $H$  and  $H'$ , two irregular components  $(h_1, h_2)$  in  $G_1$  is picked additionally the expert key is chosen by picking two irregular numbers  $(a, b)$  that has a place with  $Z_q$ . Utilizing the above components the power figures  $x = H(an)$ ,  $A = ga$ ,  $e(g, g)^b$ . The expert key  $(a, x, b)$  is kept subtly and the remaining parameters are uncovered parameters  $= (q, g, G, GT, e, H, H', h_1, h_2, A, e(g, g)^b)$ , Encryption  $(\cdot)$ . The medicinal client  $MU_i$  is analyzed completely and based on this a wellbeing profile is produced by the clients are given sensors and the fundamental restorative programming is introduced in the clients Smartphone.

### C. HEALTH MONITORING UNDER NORMAL

The therapeutic client  $MU_i$  picks the present date  $CD$  and processes the session key  $(ski)$ ,  $Ski = H(ki || CD)$  and is given to the sensors and Smartphone. The information,  $rdata$  gathered for at regular intervals by the sensors are scrambled utilizing the session key, Encryption  $(ski, rdata || CD)$  to the Smartphone utilizing Wi-Fi innovation. The Wi-Fi innovation expands the scope. The Smartphone on accepting the scrambled information utilizes the session key  $(ski)$  to unscramble the information in order to handle the  $rdata$  after which the information is sent to the social insurance focus utilizing 3G innovation  $MU_i || CD || encryption(ski, data || CD)$ . The power in the wake of getting the prepared information utilizes the expert key  $(x)$  for figuring  $MU_i$ 's mystery key  $ki = H(MU_i || x)$  and utilizes this to register  $ski = H(ki || CD)$ . This session key is utilized to recuperate the handled information  $data || CD$  from scrambled  $(ski, data || CD)$ . The date is amended and the power sends the handled information to the restorative expert

### D. HEALTH MONITORING UNDER EMERGENCY SITUATION

At the point when  $MU_0$  faces a crisis, for example, unusual bring up in the pulse and gets to be oblivious, then the power at the social insurance focus screens every one of these progressions and act to this circumstance promptly by sending the medicinal expert as per the medicinal user's need. Before the entry of the restorative expert the client must be checked consistently for which the user's Smartphone obliges high power for transmitting the user's wellbeing in configuration particle because of which numerous chances the assets in the user's Smartphone may not be a sufficient.

## III. CONCLUSION

In this paper, we outline a cloud-helped security safeguarding portable well being checking framework, called CAM, which can adequately ensure the protection of customers and the licensed innovation of mHealth administration suppliers. To ensure the clients' security, we apply the mysterious Boneh–Franklin character based encryption (IBE) in medicinal symptomatic stretching projects. To decrease the decoding unpredictability because of the utilization of IBE, we apply as of late proposed unscrambling outsourcing with security insurance to move customers' pairing calculation to the cloud server. To secure mHealth administration providers' programs, we extend the using so as to spread system tree the arbitrary change and randomize the choice limits utilized at the choice fanning hubs. At long last, to empower asset compelled little organizations to take an interest in mHealth business, our CAM outline helps them to move the computational weight to the cloud by applying recently created key private intermediary re-encryption method. Our CAM has been demonstrated to accomplish the outline objective.

## REFERENCES

- [1] Huang Lin, Jun Shao, Chi Zhang, and Yuguang Fang, Fellow, IEEE, "CAM: Cloud-Assisted Privacy Preserving Mobile Health Monitoring", IEEE Transactions on Information Forensics and Security, Vol. 8, No. 6, June 2013.
- [2] P. Mohan, D. Marin, S. Sultan, and A. Deen, "Medinet: Personalizing the self-care process for patients with diabetes and cardiovascular disease using mobile telephony," in Proc. 30th Ann. Int. Conf. IEEE Engineering in Medicine and Biology Society, 2008 (EMBS 2008), 2008, pp. 755–758.
- [3] A. Tsanas, M. Little, P. McSharry, and L. Ramig, "Accurate telemonitoring of parkinson's disease progression by noninvasive speech tests," IEEE Trans. Biomed. Eng., vol. 57, no. 4, pp. 884–893, Apr. 2010.
- [4] G. Clifford and D. Clifton, "Wireless technology in disease management and medicine," Ann. Rev. Medicine, vol. 63, pp. 479–492, 2012.



- [5] L. Ponemon Institute, "Americans' Opinions on Healthcare Privacy, 2010 [Online]. Available: <http://tinyurl.com/4atsdlj>
- [6] A. V. Dhukaram, C. Baber, L. Elloumi, B.-J. van Beijnum, and P. D. Stefanis, "End-user perception towards pervasive cardiac healthcare services: Benefits, acceptance, adoption, risks, security, privacy and trust," in *Proc. Pervasive Health*, 2011, pp. 478–484.
- [7] M. Delgado, "The evolution of health care it: Are current U.S. privacy policies ready for the clouds?," in *Proc. SERVICES*, 2011, pp. 371–378.
- [8] N. Singer, "When 2+2 equals a privacy question," *New York Times*, Oct. 18, 2009 [Online]. Available: <http://www.nytimes.com/2009/10/18/business/18stream.html>
- [9] E. B. Fernandez, "Security in data intensive computing systems," in *Handbook of Data Intensive Computing*. New York, NY, USA: Springer, 2011, pp. 447–466. [10] A. Narayanan and V. Shmatikov, "Myths and fallacies of personally identifiable information," *Commun. ACM*, vol. 53, no. 6, pp. 24–26, 2010.
- [11] P. Baldi, R. Baronio, E. D. Cristofaro, P. Gasti, and G. Tsudik, "Countering gattaca: Efficient and secure testing of fully-sequenced human genomes," in *Proc. ACM Conf. Computer and Communications Security*, 2011, pp. 691–702.
- [12] A. Cavoukian, A. Fisher, S. Killen, and D. Hoffman, "Remote home health care technologies: How to ensure privacy? Build it in: Privacy by design," *Identity in the Information Society*, vol. 3, no. 2, pp. 363–378, 2010.

# An Efficient and Secure AIDA Based Data Distribution in Cloud

Ch. Solman Raju<sup>1</sup>, Mr. D. Krishna Kishore<sup>2</sup>

<sup>1</sup>M.Tech Student, Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, 500100 India

<sup>2</sup>Assistant professor, Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, 500100 India

<sup>1</sup>solmanraju.ch@gmail.com

<sup>2</sup>kishore.dathi@mrec.ac.in

---

**ABSTRACT**— An algorithm intended for unidentified allocation of confidential information in the midst of party is urbanized. This method is used iteratively in the direction of allot these nodes ID numbers range from 1 to N. This assignment is unidentified during that the identities expected are anonymous to the additional member of the group. Conflict to the agreement between other members is confirmed during sequence theoretic intellect while personal communication channel are used. This assignment of series information allots new multipart numbers to be collective and have the application to new problems in privacy preserve information removal, collision avoidance during evidence contact. The obligatory computation is scattered without using a trusted central authority.

**Index Terms**— Anonymization in addition to deanonymization, multi-social gathering, privacy preserving data mining, privacy protection.

---

## I. INTRODUCTION

The reputation of the internet as announcement standard whether meant for private or commerce utilize depends in part on it maintain for an unidentified statement. Businesses also include genuine reason toward the employee during unidentified statement as well as shun the penalty of individuality exposure. Assume, the direction of agree distribution of review information lacking instructive the individuality of the being the essential information is related through, or toward defend whistle-blower's accurate to be unidentified along with complimentary on or after biased or cost-effective retributions[1].

Cloud-based website managing tools [2] supply capability designed for a server headed for incognito confine the visitor's network measures. The crisis of allocation confidentially apprehended information so as to folks those who are subject to the information cannot survive recognized has been researching expansively [3]. Researchers include moreover investigate the significance of obscurity along with confidentiality during different application domain: patient medicinal reports [4], electronic vote [5] social networking [6], etc.

These build an algorithm intended for allotment simple integer information scheduled a secure sum. The allotment algorithm resolve survive use in all iteration of the algorithm for anonymous ID assignment (AIDA). This anonymous ID assignment algorithm along with the variant to discuss, know how to involve a variable and a boundless integer of iterations. Finitely- confidential algorithms designed for anonymous ID assignment discuss during conclusion. Increases the limitation in this algorithm resolve and reduce a number of probable rounds. But, the central algorithm require solve a polynomial through coefficients in use since a limited position of integer modulo a principal. The assignments restrict the intensity to which  $S$  be capable of virtually raised. Illustrate here detail how to find a standard integer of essential round.

## II. EXISTING METHODOLOGIES

Existing and innovative algorithms for allocate unidentified IDs are examine from beginning to end esteem toward trade-offs among statement along with computation necessities. The new algorithm is built scheduled a secure sum information removal procedure with Newton identity and Sturm theorem. An algorithm used for disseminated consequence of confident polynomial's more than limited field enhances to scalability of an algorithm's. Markova sequence representation is utilized in the direction of locating data scheduled the integer of iteration mandatory, along with processor algebra give congested outline consequences intended for the conclusion charge.

### III. PROPOSED SYSTEM

An algorithm intended for unidentified allotment of personal information between parties is urbanized. This method is used iteratively in the direction of allot those node identification numbers range start 1 to N. This project is unidentified during to the identity expected be anonymous headed for additional member to this group. Conflict toward the agreement between additional components is confirmed in sequence theoretical intellect while personal statement channel be use. This project of sequential information allow additional composite information toward mutual also have the application for new troubles in confidentiality preserve information taking out, conflict prevention during interactions and spread catalog access. An essential computation is scattered lacking by means of a trusted central authority.

### IV. AN ANALYSIS OF SECURE SUM

Assume that, a set of hospital by way of personage database craving to add and divide up single the common information article, same as more number of hospital-acquire infection, absent informative the charge to these statistics article used to a few constituent assembly. Accordingly, N node  $n_1, n_2, \dots, n_N$ , include data items  $d_1, d_2, \dots, d_N$ , and need to subtract and divide up only the whole value  $T = d_1 + d_2 + \dots + d_N$ . The secure sum algorithm allow here sum T to collect by way of some guarantee of obscurity. Suppose the semi-honest replica an isolation preserve information removal [7]. Underneath the replica, every knob resolution pursue the regulations of the procedure, it could utilize several data it see throughout every completion of these procedure to conciliation protection. Every pair of node should contain a protected statement guide available, although source concentrated, secure sum algorithms are constructed. The next algorithm, the efficient headed for take the principles as individual numeral for initial analysis.

Nodes	$\hat{r}_{i,1}$	$r_{i,1}$	$r_{i,2}$	$r_{i,3}$	$r_{i,4}$	$d_i$	$\hat{d}_i$
$n_{i=1} :$	$13 - 6 + 8 = 15$	13	-10	6	-3	6	8
$n_{i=2} :$	$7 - 10 + 9 = 6$	7	3	-5	5	10	9
$n_{i=3} :$	$-8 - 6 + 5 = -9$	-8	11	12	-9	6	5
$n_{i=4} :$	6	6	-8	-5	9	2	2
$s_i =$	18	18	-4	8	2	$T = 24$	24

Prearranged node's every share a data item  $d_i$  starting the finite represent Abelian cluster, allocate the significance  $T = \sum d_i$  between the node's devoid of illuminating a principles.

1) Every node  $n_i, i = 1, \dots, N$  select random value  $r_{i,1}, \dots, r_{i,N}$  such that

$$r_{i,1} + \dots + r_{i,N} = d_i$$

2) Every "random" value  $r_{i,j}$  is transmitted starting nodes  $n_i, \dots, n_j$ . An amount of every random numbers  $r_{i,j}$  are, certainly, the totality preferred T.

3) Every node  $n_j$  total every arbitrary standards expected as:

$$S_j = r_{1,j} + \dots + r_{N,j}$$

4) Currently every node basically broadcast to every additional node thus with the purpose of each node can compute:

$$T = s_1 + \dots + s_N$$

### V. HOW TO DISCOVER AN AIDA

A simple algorithm for discovering an anonymous ID assignment, which include more than a few variant depends resting on the alternative from information allocation process by below steps. At single step, arbitrary integers before "slots" among I and S are ideal through all nodes. A nodes location determination unwavering through its location surrounded by the selected slot although, supplies essential be through for impact. The limitation S be supposed to be select so that  $S \geq N$ .

Given nodes  $n_1, \dots, n_N$ , use distributed calculation (without central authority) to find a unidentified indexing combination  $s : \{1, \dots, N\} \rightarrow \{1, \dots, N\}$  easy.

1. Position an integer to assign node A = 0.
2. Every unassign nodes  $n_i$  choose an random number  $r_i$  from sort I to S. A nodes assign in a previously about choose  $r_i = 0$ .

3. The arbitrary statistics is collective incognito. One scheme for liability this was prearranged in. Select the common value in  $q_1, \dots, q_N$ .
4. Let  $q_1, \dots, q_k$  represent an revise list of mutual value through replica with zero value fully impassive where  $K$  is the integer of single arbitrary value. The nodes  $n_i$  that drew single arbitrary numbers then resolve their  $s_i$  guide beginning the point of their arbitrary integer in revise record as it would emerge behind individual sort:

$$s_i = A + \text{Card}\{q_j : q_j \leq r_i\}$$

5. Renew the integer of node assign:  $A = A + k$ .

## VI. THE COMPLETION RATE AFTER R ROUNDS

Two nodes capacity compose the same choices of arbitrary data, or slot when they resolve a term during segment. Only one can refuge to inclusive transfer a  $N$  nodes by using  $S$  potential used for slot otherwise arbitrary integer choice and  $R$  rounds will transpire by smallest amount a preferred probabilities  $P_{S=s}(R, N) = P(R, N)$ . The limit  $S$  will normally be completely set to  $S = s$  by its exception in this segment. The person who reads may examine that estimate the number of research finished in single around is in concentrate the well-known birthday crisis. The related approach include be use toward calculate bound used for crisis and additional [8], [9].

Currently, illustrate how to calculate the probability  $P(R, N)$ . When initial step, consider the possibility  $p(N, A, C)$  to that particular round, initial through  $N$  node's and  $s$  slot, consequences during assignment  $A$  and  $C$  slot by the conflict when folks slot contain be choose through one other node. Statement to  $Z = s - A - C$  is no (zero) integer node of that slot have selected. The repetition associations verify  $p(N, A, C)$ .

$$\begin{aligned} p(0, 0, 0) &= 1, \quad p(1, 1, 0) = 1 \\ p(N, A, C) &= 0 \quad \text{whenever } C < 0 \vee A < 0 \\ p(N, A, 0) &= 0 \quad \text{whenever } N \neq A \\ p(N, A, C) &= 0 \quad \text{whenever } N < A + 2C \\ p(N, A, C) &= \frac{1}{s}(s - A - C + 1) \cdot p(N - 1, A - 1, C) \\ &\quad + \frac{1}{s}C \cdot p(N - 1, A, C) \\ &\quad + \frac{1}{s}(A + 1) \cdot p(N - 1, A + 1, C - 1) \end{aligned}$$

For efficiency, the values  $p(N, A, C)$  require to be cached through computation. But  $N$ , the possibility to begin with  $N$  node and  $s$  slot, accurately after single round  $A$  assignments include be finished:

$$P_{N,A} = \sum_{C=0}^{\lceil N/2 \rceil} p(N, A, C)$$

$S$		$N = 10$	25	100	1000
15	R=1	98.2			
	R=2	45.9			
	R=3	6.70			
100	R=1	37.2	96.3		
	R=2	0.677	13.6		
	R=3	.00697	0.199		
500	R=1	8.66	45.7	< 100.0	
	R=2	0.200	0.236	27.3	
	R=3	$4.01 \cdot 10^{-5}$	0.000477	0.112	
$10^4$	R=1	0.450	2.96	39.2	< 100.0
	R=2	$4.53 \cdot 10^{-5}$	0.000316	0.00961	36.3
	R=3	$4.53 \cdot 10^{-9}$	$3.16 \cdot 10^{-8}$	$9.62 \cdot 10^{-7}$	0.00877
$10^7$	R=1	0.000450	0.00300	0.495	4.88
$5 \cdot 10^7$	R=1	$9.01 \cdot 10^{-5}$	$6.00 \cdot 10^{-4}$	0.00990	0.995

The matrix  $P(N)$  gives the conversion probability for a particular around of AIDA initial with  $N$  nodes and conclusion with  $N - A$  nodes yet to be assigned.

## VII. ARITHMETICAL COMPLETION STATISTICS

For various function, the formula contributes an appropriate respond. But, the rich inscription on interesting Markov shackles and the availability of computer algebra correspondence present several additional possibilities for investigation. To conclude an enviable value used for an integer of slot “ $S = s$ ” individual be able to obtain benefit of the fact to the probability  $p_{N,A}$  represents table when normal function of the integer of slot  $S = s$ . In detail  $\mathbf{P}(N)$  be the upper, left-hand corner of an unbounded prevailing conditions  $\mathbf{P}$ . When  $N$  is small, the entry, which has refusal perceptible model, can be considered by a processor arithmetical enclose since the repetition relation compliant.

$$\mathbf{P} = \begin{pmatrix} 1 & \frac{s-1}{s} & \frac{(s-2)(s-1)}{s^2} & \frac{(s-3)(s-2)(s-1)}{s^3} & \dots \\ 0 & \frac{1}{s} & \frac{3(s-1)}{s^2} & \frac{6(s-2)(s-1)}{s^3} & \dots \\ 0 & 0 & \frac{1}{s^2} & \frac{4(s-1)}{s^3} & \dots \\ 0 & 0 & 0 & \frac{3s-2}{s^3} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

## VIII. CONCLUSION

The very similar comparable crisis of mental poker was shown to contain no such solution with two players with three cards. The disagreement can easily extend to, e.g., every two sets of  $N$  collude players through a deck of  $2N+1$  cards quite than our deck of  $2N$  cards.

During compare to limits on achievement time improvement at earlier work, our prescription gives the anticipated achievement time exactly. Assumption the asymptotic procedure of consequence 9, based on the computational occurrence, to be a true upper bound.

The entire noncryptographic algorithms be widely replicated, along with that the present work does present a based upon implementation can be constructed. The communication requirements of the algorithms depended a lot in the essential execution of the selected secure sum algorithm. In some cases, absorption of the two layers cans consequence in reduced overhead.

## REFERENCES

- [1]. Sarbanes–Oxley Act of 2002, Title 29, Code of Federal Regulations, Part 1980, 2003.
- [2]. White Paper— The Essential Guide to Web Analytics Vendor Selection, IBM [Online]. Available:<http://measure.coremetrics.com/corem/getform/reg/wp-evaluation-guide>.
- [3]. A. Shamir, “How to share a secret,” *Commune. ACM*, vol. 22, no. 11,
- [4]. A. Friedman, R. Wolff, and A. Schuster, “Providing k-anonymity in data mining,” *VLDB Journal*, vol. 17, no. 4, pp. 789–804, Jul. 2008.
- [5]. F. Baiardi, A. Falleni, R. Granchi, F. Martinelli, M. Petrocchi, and A. Vaccarelli, “Seas, a secure e-voting protocol: Design and implementation,” *Comput. Security*, vol. 24, no. 8, pp. 642–652, Nov. 2005.
- [6]. Q. Xie and U. Hengartner, “Privacy-preserving matchmaking for mobile social networking secure against malicious users,” in *Proc. 9<sup>th</sup> Ann. IEEE Conf. Privacy, Security and Trust*, Jul. 2011, pp. 252–259.
- [7]. C. Clifton, M. Kantarcioglu, J. Vaidya, X. Lin, and M. Y. Zhu, “Tools for privacy preserving distributed data mining,” *ACM SIGKDD Explorations Newsletter*, vol. 4, no. 2, pp. 28–34, Dec. 2002.
- [8]. J. Smith, “Distributing identity [symmetry breaking distributed access protocols],” *IEEE Robot. Autom. Mag.*, vol. 6, no. 1, pp. 49–56, Mar. 1999.
- [9]. S. S. Shepard, R. Dong, R. Kresman, and L. Dunning, “Anonymous id assignment and opt-out,” in *Lecture Notes in Electrical Engineering*, S. Ao and L. Gleman, Eds. New York: Springer, 2010, pp. 420–431.



# The Cloud Storage for Scalable Data Distribution using Key Aggregate Cryptosystem

L.Rajitha<sup>1</sup>, Mr. M. Sabhapathy<sup>2</sup>

<sup>1</sup>M.Tech Student, Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, 500100 India

<sup>2</sup>Associate professor, Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, 500100 India

<sup>1</sup>lrajitha91@gmail.com

<sup>2</sup>sabha.delight@gmail.com

---

**ABSTRACT**—Data distribution is vital process in cloud storage space. In this account, we demonstrate how to securely, proficiently, and flexibly distribute the data with others in cloud storage space. We represent new public-key cryptosystems which generate constant size cipher texts such that efficient allocation of decryption rights for any group of cipher texts is possible. The oddity that one can combine any group of secret keys and makes them as a fixed agreement as a single key, but all-encompassing the power of all the keys being grouped. In other words, the secret key owner can liberate constant-size aggregate key for flexible selections of cipher text arranged in cloud storage space, but the other encrypted records outside the group remain secret. This compact aggregate key can be efficiently sent to others or be stored in a well-dressed card with very limited secure storage space. We contribute recognized study of our systems in standard model. We also explain other function of our systems. In particular, our systems give the first public-key patient-controlled encryption for flexible organizations with a scale of positions, which was yet to be recognized.

**Index Terms**— Cloud storage, data sharing, key-aggregate encryption, patient-controlled encryption

---

## I. INTRODUCTION

Cloud computing is considered the next step within the evolution of on-demand information technology which mixes a collection of existing and new techniques from analysis areas such as service-oriented architectures (SOA) and virtualization. With the speed development of versatile cloud computing technology and services, it is routine for users to influence cloud storage services to share information with others in an exceedingly friend circle, e.g., Drop box, Google Drive[1] and AliCloud. The shared information in cloud servers, however, sometimes contains users sensitive data such as personal profile, money information, health records, etc. and should to be protected.

As the ownership of the information is separated from the administration of the cloud servers will migrate user's information to the alternative cloud servers by outsourcing or share them in the cloud storage. Therefore, it becomes a huge challenge to protect the privacy of this shared information in cloud, especially in cross-cloud and big data atmosphere. So as to fulfill this challenge, it is necessary to design a comprehensive solution to support user-defined authorization period and to provide fine-grained access control throughout this period

Data cryptography mainly is the scrambling of the content of the data, like text, image, audio, video to make the data unreadable, invisible or unusable throughout transmission or storage is termed encryption. And main aim of cryptography is to take care of data sensitive from invaders. The opposite process of obtaining back the original data from encrypted data is decryption that restores the original data. To encrypt data at cloud storage each symmetric-key and asymmetric-key algorithms are used. It is having serious problems whereas handling vast database and transactions. Public-key cryptography, conjointly called asymmetric cryptography, may be a category of cryptographic protocols supported algorithms that need two separate keys, one among them is secret (or private) and one among that is public. Though totally different, the two components of these keys are mathematically joined. The general public keys used, for instance, to encrypt plaintext or to verify a digital signature; whereas the personal keys are used for the alternative operation, in these examples to decrypt cipher text or to make a digital signature.

1. The cost and complexities concerned usually increase with the quantity of the decoding keys to be shared.
2. The encryption key and decryption key are totally different in public key encryption.



3. Concerning accessibility of data, there is a sequence of cryptographic method that go as way as permit a third-party assessor to check the provision of data on behalf of the information holder without leaking something regarding the data or without compromising the information holder ambiguity.

4. Regarding file availability, there are cryptographic schemes that allow a third person actuary to check the availability of files on instead of information holder lacking any information reveal[2] or lacking identity of information holder[3].

## **II. RELATED WORK**

Cryptographic key assignment theme aim is to reduce the value in aggregation and organizing secret key for general use of cryptography. By tree structure, key for a given branch use to get descendents nodes key. Presently allowing parent key all descendent nodes key implicitly granted. Sandhu[4] projected technique for generating tree hierarchy of interchangeable key by mistreatment repetitious application of one method operates. The concept generalized from tree to graph. Advanced cryptosystem key assignment technique is support access that may be sculptural by cyclic or acyclic graph. Several schemes produce keys for symmetric-key cryptosystems, even several key derivations requires standard arithmetic utilized in the general public key, which are typically dearer than traditional "symmetric key operations" like pseudorandom functions. Yan Sun projected multi cluster key management scheme that achieves stratified access management with integrated key graph and multi cluster key management scheme. Benaloh gift an encoding theme for sharing additional keys in broadcast state of affairs.

### **A. Encryption Key Strategy**

An encryption plan which is initially outlook for briefly transmits huge bulk of keys in produce situation. The development is basic and quickly survey its key inference handle at this point used for a strengthen depiction about which persist affective alluring possessions we need to attain to. The deduction of input for an arrangement of module is as takes after. The combined value is picked wherever  $p$  plus  $q$  are both vast uneven most important. An expert anonymous input is picked at arbitrary. Every group be connected with a particular prime. The entire these primary statistics could placed within general population framework factor. The consistent volume input designed for a rest could be produced. Designed for the individuals which user has assigned the entrance rights for  $S'$  can be created. In any case, user has planned for the secret code location. The satisfied supplier requests toward obtain the relating mystery keys to encode information, this request does not appropriate. Since strategy be utilized in the direction of produce mystery esteem as opposed to a couple of open/major input, tendency has vague by what method relate to this thought for open type encipher plan. At long last, we take note with the aim of here be plans who's attempt to lessen the key volume designed for accomplishing confirmation in similar code creation, On the other hand, offering of decoding force is not a worry in these plans.

### **B. Compact Key in Identity-Based Encryption**

Identity based cryptography theme is a form of the general public key cryptography. During this the general public key of user about as string-identity of user. Within the IBE non-public Key Generator that holds a master secret key and problem is to other user as per their identity. The user World Health Organization code the message will take public parameter and identity of user to decrypt message. The recipient rewrites cipher text victimization own secret key.

Guo et al[5]. tried to make IBE with key aggregation. One among their technique assumes random oracle however different one not. Significantly their aggregation of key comes at expense of the dimensions for both cipher text and public parameter. This will increase the value of storing and transferring cipher text, that is not sensible in some conditions. In fuzzy IBE, one individual secret key will rewrite cipher text underneath multiple identities that area unit pass on additional metric area, however not for random locate of separate and it will not match with key aggregation plan.

### **C. Other Encryption Scheme**

Primarily characteristic based coding permits every encrypted text to be connected with feature, and therefore the master secret key possessor will take away secret key for a policy of this feature so encrypted text is decrypted by this key if it is related attributes match to rule. In ABE important issue is collision-resistance not the compression of secret keys. The vary of the encrypted text is not fix. A PRE scheme allow Alice to delegate to the server ability to convert cipher text encrypted beneath own public-key ones bob. The Proxy Re-

encryption PRE technique is accepted to various applications. Using PRE theme solely shift the secure key storage demand from delegatee to proxy. Thus it's not appropriate to let proxy reside in storage server. It will not appropriate therefore every coding desires individual interaction with proxy.

#### D. Construction of KAC

Boneh et al [6]. gift collision-resistant broadcast encryption theme by discrimination this essential theme is designed. Their technique is maintain fixed-size secret keys, each key solely has

1. Encryption: output cipher text
2. Extract
3. Decrypt: Output message

The coding may be done additional expeditiously. To make extended theme best completely different cipher text categories suggested for numerous functions opposite to completely different public keys. This key extension approach can even regard key update method. Suppose a secret Key price is compromised. Then we will replace it with new key price. The less aggregate key size decreases communication overhead for transferring the new key.

### III. FRAMEWORK

The information holder builds the general open structure bound over set of connections and produce a public secret key over Key Gen. Information could often encoded via encipher by any person an agency additionally make a decision what cipher text category is belongs to the plaintext information to be encoded. Information holder will develop the major-private key is to get a mixture decipher key for a collection of cipher text classes over Extract. The initiated inputs are often moved toward delegates firmly through secure electronic mail or protected machine. Completely, several clients with cumulative key will decipher either secret message text providing the secret message category is containing within the combined key via Decrypt phase. Key mixture coding system encompasses 5 polynomial time algorithms as follows:

**1. Setup ( $1\lambda, n$ ):** the information holder create open structure bound using Setup. User planned participation of a protection point constraint  $1\lambda$  and range of cipher text categories  $n$ , it amount produced the system open structure constraint.

**2. KeyGen:** It is done through information holder to accidentally produce a major input ( $Pk, ms$ ).

**3. Code ( $pk, i, n$ ):** It is done with information owner and for note  $n$  and index  $i$ , it calculate the secret message as  $M$ .

**4. Extract ( $Sm, S$ ):** This is done with information holder designed for hand over the deciphering authority for a definite collection of secret message categories plus it yields the combined key for set  $S$  indicated by  $Ks$ .

**5. Decipher ( $Sm, S, C, m$ ):** This is done with alternate user established, an combined input  $Sm$  produced in take out. resting on enter  $Sm$ , locate  $S$ , catalog  $C$  indicated as secret message group secret message  $M$  associated to and output is decipher by end outcome is  $m$ .

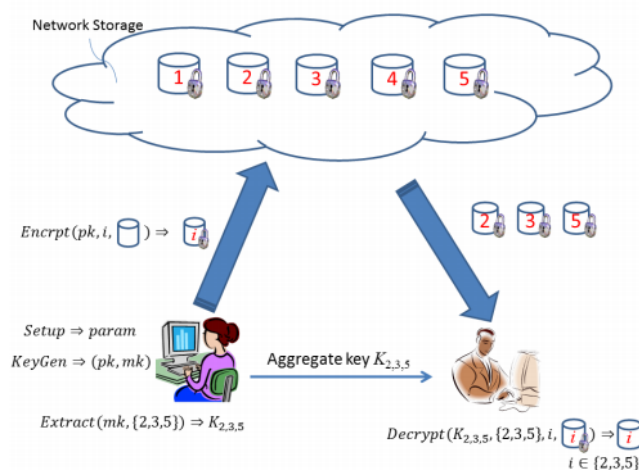


Fig. 1 Framework of the Key system

### A. Cloud Storage

Cloud storage in these days is extremely accepted storage space structure. Third party controlled a objective cache storing of information in cloud storage. Cloud storage space is economy of mainframe knowledge in consistent as well as physical storage length many helpers that are controlled by third person. Third person is responsible for maintenance the information open along with accessibility and physical atmosphere ought to subsist secured by consecutively next to every time. Rather than accumulate information to the other limited storage space, we have a tendency to save data to isolated storage space which is available starting someplace and an time. It decreases the power of moving physical storage to everyplace. Through discrimination cloud storage space we will access the files as of any PC through Internet that omitted inadequacy of entrée data starting similar PC where it is kept. Whereas allowing for information privacy, we could not trust ancient system of verification, as a result of sudden authority growth determination of all information. User by using own key uploading the information in server before encipher information. Information allocation is moreover important functionality of cloud storage; as a result the user will share knowledge from anyplace and anytime to anyone. As an illustration, group can allow authorization to access a part of sensitive knowledge for clients. However, difficult job with the purpose of to the way toward allocate encipher data. Conventional approach is client will transfer the enciphered information on or after storage, allocate with others decipher data and send it; however it is go down the consequence of cloud storage.

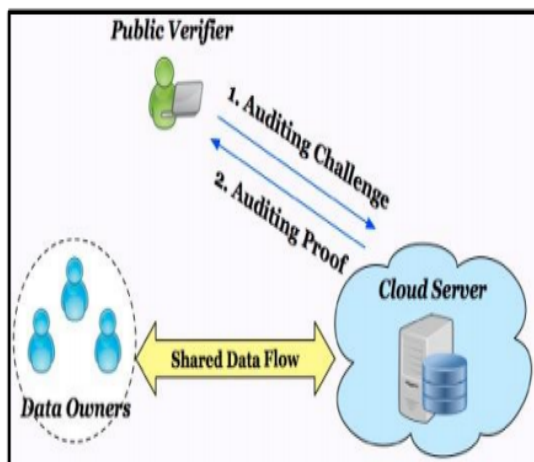


Fig. 2 Cloud Storage Structure

Cryptography technique is often applied by two most important methods are- one is similar inputs encoding and alternative is dissimilar input encoding. In symmetric key encryption, similar keys part element apply for cryptography and secret writing. In difference, in dissimilar encoding completely dissimilar inputs are utilized, mutual input for encoding and secret input for decoding. The same inputs encoding is a lot of versatile in favor of our approach. This will be demonstrated by next example. Assume sender place all files in Box and sender doesn't need to reveal her knowledge to everyone. Because of information reveal promise she would not expectation of privacy system afford by inbox, therefore she encrypts all knowledge before uploading to the server. If receiver request sender to allocate various files then Sender use allocate provide by inbox. However, difference at present could be that the way to allocate encoded information. Here couple of methods. Sender encode the information using one secret input and allocate to facilitate secret input openly by means of receiver. Sender will encrypt the information through different input and mail receiver matching keys to Bob via secure channel. During initially user come close to, redundant information is display to the receiver, where it is lacking.

### B. Digital Signature

In Key aggregate Cryptosystem [7] it essentially concentrates on effective information allocating in cloud storage. The Key aggregate Cryptosystem conjointly considers the integrity of data that are hold on in cloud. To supply integrity towards the user information the digital signature is employed. Digital Signature could be a method that guarantees that the contents of a message have not been altered. Digital signatures rely on bound forms of coding to make sure authentication. Authentication is that the methodology of corroborative that data



is returning from a sure supply. Digital signatures use a kind of asymmetric cryptography. The Key aggregate Cryptosystem uses the RSA signing algorithmic program to gain the digital signature. The RSA public-key algorithmic program really uses 2 totally different keys referred to as the general public key and also the private key. The private key is not able to its owner, whereas the general public key is often on the advertise to anyone. In Public-key algorithms two keys are used in that one secret is used for encryption, the opposite is important for decryption. In digital signatures, the private key generates the signature, and also the corresponding public key validates that signature. The RSA algorithmic program alone cannot be used to sign the document in order that hashing algorithmic program is additionally used beside that.

1. A one-way hash of the document is produced using SHA algorithmic program.
2. The hash is encrypted with the private key of the RSA there by signing the document.
3. The document and the signed hash are transmitted.
4. The recipient produces a one-way hash of the document.
5. Using the RSA algorithmic program, the recipient decrypts the signed hash with the sender's public key.

If the signed hash matches the recipients hash, the user will capture that the signature is valid and also the document is intact. Using the digital signature methodology the cloud users will verify the integrity of the stored data in cloud.

#### IV. EXPERIMENTAL RESULTS

When user upload files into cloud it will generate an aggregate key it is used to share file securely.

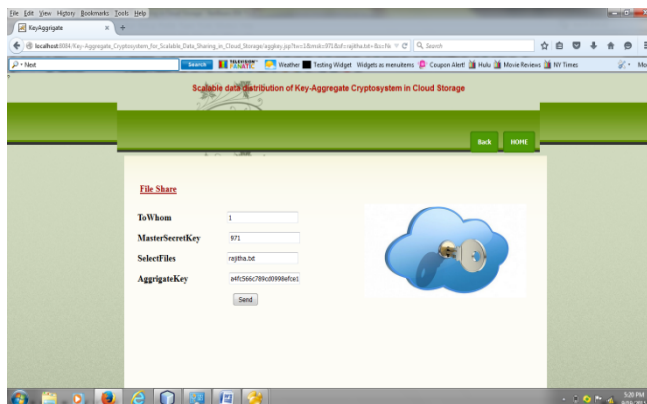


Fig. 3 Cloud Storage Structure

After generating an aggregate key user can download that data by using this aggregate key. For that he firstly, view encrypted files and try to do decryption to download files.

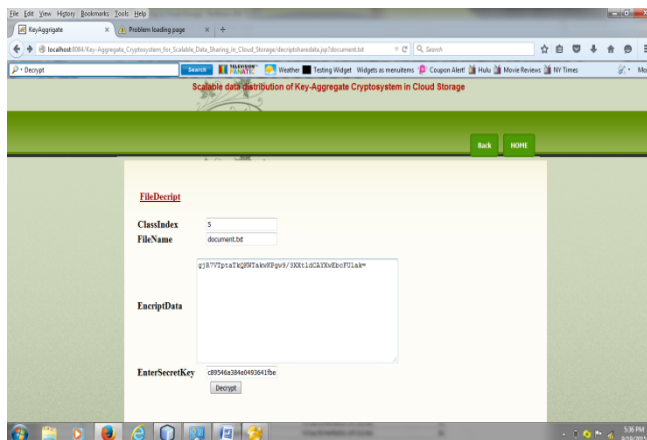


Fig. 4 Cloud Storage Structure

Using secret key he decrypted files and download those files.

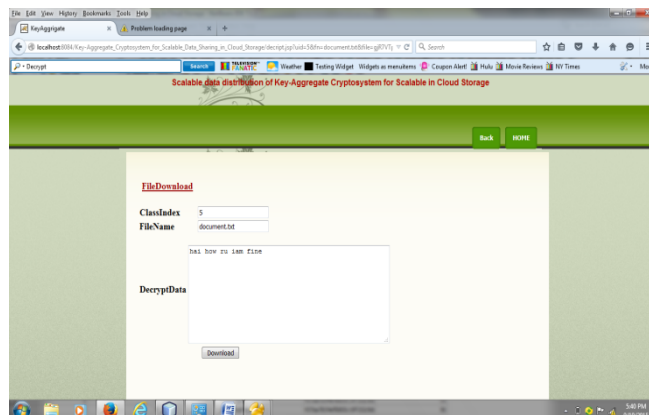


Fig. 5 Cloud Storage Structure

## V. CONCLUSION

To share data flexibly is an essential thing in cloud computing. Users prefer to upload their data on cloud and along with various users. Outsourcing of data to server may lead to leakage the private data of user to everyone. Encryption is the solution which allows to share the selected data with preferred candidate. Sharing of decryption keys in secure way acts an important role. Public-key cryptosystems provides delegation of secret keys for different cipher text classes in cloud storage. The delegate gets securely an aggregate key of constant size. It is required to keep enough number of cipher texts classes as they development fast and the cipher text classes are bounded that is the limitation.

## REFERENCES

- [1]. G. Clarke, Microsoft's Azure Cloud Suffers First Crash, The Register, March 16, 2009, [http://www.theregister.co.uk/2009/03/16/azure\\_cloud\\_crash/](http://www.theregister.co.uk/2009/03/16/azure_cloud_crash/)
- [2]. C. Wang, S. S. M. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Secure Cloud Storage," IEEE Trans.Computers, vol. 62, no. 2, pp. 362–375, 2013.
- [3]. B. Wang, S. S. M. Chow, M. Li, and H. Li, "Storing Shared Data on the Cloud via Security-Mediator," in International Conference on Distributed Computing Systems ICDCS 2013. IEEE, 2013.
- [4]. R. S. Sandhu, "Cryptographic Implementation of a Tree Hierarchy for Access Control," Information Processing Letters, vol. 27, no. 2, pp. 95–98, 1988.
- [5]. F. Guo, Y. Mu, and Z. Chen, "Identity-Based Encryption: How to Decrypt Multiple Ciphertexts Using a Single Decryption Key," in Proceedings of Pairing-Based Cryptography (Pairing '07), ser. LNCS, vol. 4575. Springer, 2007, pp. 392–406.
- [6]. D. Boneh, C. Gentry, and B. Waters, "Collusion Resistant Broadcast Encryption with Short Ciphertexts and Private Keys," in Proceedings of Advances in Cryptology - CRYPTO '05, ser. LNCS, vol. 3621. Springer, 2005, pp. 258–275.
- [7]. Cheng-Kang Chu, Sherman S.M. Chow, Wen-Guey Tzeng, Jianying Zhou, and Robert H. Deng, "Key-Aggregate Cryptosystem for Scalable Data Sharing in Cloud Storage" IEEE Transactions On Parallel And Distributed System, Vol 25, No. 2 February 2014.



# Advanced Secure Mona Protocol for Data Sharing in Untrusted Cloud Using Attribute Based Encryption

Mr. M. Sathish<sup>1</sup>, Mr. P. A Hima Kiran<sup>2</sup>

<sup>1</sup> M.Tech Student, Dept. of cse, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> Associate Professor, Dept. of cse, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>1</sup>[sathishmarripally28@gmail.com](mailto:sathishmarripally28@gmail.com)

<sup>2</sup>[andrewshimakiran@mrec.ac.in](mailto:andrewshimakiran@mrec.ac.in)

---

**Abstract**— Sharing gathering asset among cloud clients is a noteworthy impact, so distributed computing gives a prudent and proficient arrangement. Because of proceeds with change of sharing information, enrollment in a multi-proprietor way to an un trusted cloud is still a testing issue. Here in this paper, We propose a safe multi-proprietor information sharing plan, for dynamic (continuously changing ) bunch in the cloud. By giving gathering signature and element show encryption strategies, any cloud client can defensively impart information to others. At that point an in the mean time, the capacity overhead and encryption calculation expense of the plan are autonomous with the quantity of repudiated clients. In other hand, we break down the security of this plan with thorough verifications. OTP (One-Time Password) is one of the least difficult and most well known types of confirmation that can be utilized for securing access to accounts. OTP are regularly alluded to as a protected and more grounded types of validation, and tolerating them to introduce over different machines. We give a different levels of security to share information among multi-proprietor process. To start with the client chooses the pre-chosen picture to login. At that point chooses a picture from the lattice of pictures. By utilizing this the OTP is produced consequently and sent to relating email account.

**Keywords**— Security, Broadcast Message, Encryption, Cloud computing.

---

## I. INTRODUCTION

Distributed computing imagines exceedingly accessible, on-interest system access to a mutual pool of configurable registering assets [1], [2], [3]. Clients can appreciate adaptable capacity limit and calculation ability without paying consideration on the development and upkeep of these foundations. While distributed computing gets promising open doors, it additionally brings along new security and protection issues, which ruin people in general to receive the cloud advances. The information in travel or put away in distributed storage could be tempered by unapproved people or even the distributed storage supplier [4], [5], [6]. Various encryption systems are accessible to ensure the security of Cloud information and administrations [7], [8], [9]. In any case, as these encryption methods bring along new procedures, additional complexities must be conceived to oversee scrambled information safely and effectively. For an individual distributed storage client, he/she stores his/her information and recovers some portion of the put away information later. Then again, for big business clients, the put away information ought to be shared among gathering individuals. One sort of encryption plan called characteristic based encryption (ABE) could be utilized to apply fine-grained access control over the common information [10], [11], [12], [13], [14], [15]. What's more, the flow of gathering individuals and comparing put away information ought to be considered to build a possible fine-grained access control for the venture [16], [17], and [18].

Besides, given the aggregate sum of information produced and put away in the cloud, getting to information through route is tedious and vexatious. Getting to cloud information through (catchphrase) inquiry is thought to



be viable and in superfluous. Be that as it may, as the cloud information are secured through cryptographic systems, which acquire high expenses when recovering through seeking. Searchable encryption was acquainted with empower clients to shroud the searchable catchphrases (of a document) by encryption [19], [20], [21], [22], [23] [24]. Later, clients could create fitting tokens/trapdoors for particular catchphrases to recover the scrambled information containing these watchwords. The clients looking capacity is likewise shared under fine-grain arrangements [25], [26], [27], [28] One client can create searchable files for a document and indicate a subset of clients who can use these searchable records. Clients outside the predefined gathering can't hunt out this document. Then again, motion of gathering individuals and searchable files ought to be considered to yield a down to earth and powerful searchable encryption [29], [30].

In this paper, we propose one novel distributed storage development empowering the administration of element searchable information for gathering joint effort. We make utilization of characteristic based encryption plan (ABE) and open key encryption with conjunctive watchword pursuit (PECK) to outline our convention. We show that our plan exceptionally incorporates fundamental usefulness for big business clients, to be specific, the fine-grained access control for the searchable record and the substance of the information. Moreover, we give security examination and behavior broad execution assessment to demonstrate the achievability of our configuration for big business clients.

The following paper is structured as follows.

1. Related foundation is depicted in Section.
2. While focused on framework models and two cryptographic building squares are displayed in Section.
3. Our novel development is point by point in Section.
4. Then the security and execution investigation are appeared in Section.
5. Finally, our commitments are emphasized and future course is specified to close this paper.

## **II. ENCRYPTION**

### **2.1 Attribute-based Encryption**

Attribute based Encryption (ABE) gives a fine-grained access control of shared information. ABE was started from the work by Sahai and Waters [10]. Later, two tracks of ABE have been produced: figure content arrangement ABE (CP-ABE) [13], [15] and key-strategy ABE (KP-ABE) [12], [14]. In the CP-ABE plan, the client is allowed trait keys (connected with qualities), and the entrance approach could be upheld on the figure content. At that point the client claims the trait keys fulfilling the predetermined access approach, the client could unscramble the message. An opposite setting is called KP-ABE, which determines decoding arrangement on the trait keys and the figure content is labeled with an arrangement of qualities.

In any case, to send in down to earth applications, overseeing element access approach is required to bolster perpetually changing access bunch. At that point the trait keys ought to be re-issued and figure content be re-scrambled to consent to the present access control arrangement. In other hand, client renouncement ought to be did in a productive approach to control the harms. Some ABE proposed that the lapse time is added with the quality when creating related property keys [13], [17]. On the other hand, the exchange off between the granularity of "window of defenselessness" and the weight to upgrade the quality keys ought to be considered. Boldyreva et al. [16] proposed an effective denial plan for IBE and KP-ABE, while Yu et al. [18] proposed an ABE plan with property disavowal. They incorporated the intermediary re-encryption (PRE) with ABE, and empowered the power to designate a large portion of the work for key overhaul of the client to intermediary servers. Since part based access control (RBAC) [34] is normally used to limiting framework access to approved clients. CP-ABE, which is firmly identified with RBAC, is picked as a building square of our plan for big business application situation.

### **2.2 Searchable Encryption**

Searchable encryption empowers clients to conceal the searchable watchwords (of a document) by encryption. Later, clients could create suitable tokens/trapdoors for particular catchphrases to recover the scrambled information containing these watchwords. Senegal. [19] initially presented the idea of looking on encoded information and gave down to earth arrangements. Goh [20] then formalized the idea of security for this issue and developed a more effective plan utilizing Bloom channel. Taking after that, some exploration [22], [23] was led to either enhance the effectiveness or give more grounded security of searchable encryption. One shared trait of these works is that they all upheld just single watchword hunt in the symmetric key setting.

The idea of conjunctive catchphrase look in symmetric key setting was initially presented by Galle et al. [25]. They gave a security thought to conjunctive watchword look over encoded information and built a more proficient plan contrasted and the one insignificantly stretched out from single catchphrase pursuit plan. Later, Ballard et al. [26] enhanced by shortening the trapdoor size and lessening calculation/stockpiling overhead. Be that as it may, because of the symmetric key setting, these plans just empower one client to store and recover his/her own particular private information. Sharing of file building and looking ability can't be accomplished effortlessly.

Boneh et al. [21] initially tended to one sort of reasonable applications called email steering framework. The searchable list of a mail can be created by utilizing the beneficiary's open key. The beneficiary can recover specific messages from the delegating so as to steer server related trapdoors. The relating messages can be gathered. What's more, Boneh et al. [35] proposed another application brought looking over review log, where the organization can appoint particular trapdoor to the evaluator to examine just review related records. Be that as it may, these plans upheld just single watchword pursuit. There are different applications requiring more expressive pursuit over conceivable watchwords.

To improve seek expressions, Park et al. [27] proposed open key encryption with conjunctive catchphrase hunt (PECK). Boneh et al. [36] further gave a plan supporting the conjunction of subset and extent questions on figure content information. At that point their development utilized the bilinear gathering of Composite way, which yields less productive development. Also, they considered just single-client setting, where sharing of searchable record is difficult to accomplish. Hwan get al. [28] gave one effective PECK and considered a conceivable expansion to multi-client settings [29], [30]. In this paper, we will further consider the sharing of searchable file ought to be given to empower bunch cooperation.

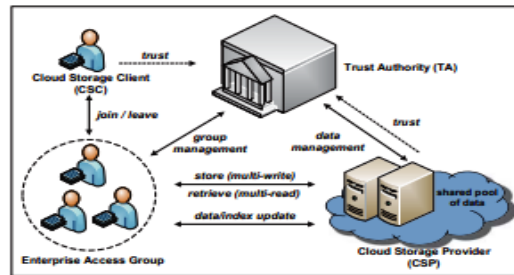


Fig. 1: Enterprise Cloud Storage Access Model

### III. SECURITY AND PERFORMANCE

In this area, we show the security and execution of our convention. In this exhibit, just approved gathering individuals can 1) Search/recover the gathering information and 2) unscramble the recovered gathering information put away in cloud stockpiles. The worker who leaves the gathering or is denied can't recover or decode the put away information in cloud stockpiles. In addition, we assess the calculation and correspondence costs for the CSC in our configuration and close our outline is powerful and proficient for the venture clients to share information and team up as a gathering.

Table 2: Computation Cost of the CSC

Operation	Required Basic Operations
GrpStore	$(n + m' + 2l + 2) \text{ SctrMul}_{G_1},$ $(l - 1) \text{ Add}_{G_1}, 2l \text{ HashToPoint}$
Retrieve	$3 \text{ SctrMul}_{G_1}, (2m' - 1) \text{ Add}_{G_1}$
GrpDecrypt	$(m + 1) \text{ Pairing}, m \text{ Mul}_{G_2}$
UpdateAU	$1 \text{ Exp}_{G_2}$
UpdateSI	$1 \text{ Add}_{G_1}, 2 \text{ HashToPoint}$

#### 3.1 Security Analysis

The security of our convention depends on the basic ABEar [18] and muPECK [28]. From one perspective, the ABEar is ended up being semantically secure under specific id picked plaintext assault (IND-s ID-CPA) accepting decisional bilinear Diffie-Hellman (DBDH) is hard. In light of these formal contentions, we can

presume that the unapproved substance (either CSC or CSP) can't manufacture searchable records and searchable trapdoors since these activities are included in taking care of the difficult issue.

Then again, the muPECK is turned out to be semantically securing under multi-client figure content from arbitrary against picked catchphrase assaults (IND-mCR-CKA) accepting choice straight Daffier-Hellman (DLDH) is hard. In this way, the unapproved element can't compute quality keys for decoding, either in light of the fact that these activities are included in taking care of the DLDH difficult issue.

With respect to information elements, the information is re-encoded to the same figure content space. The re-created key is likewise appropriated consistently in the key space. Any enemy can't increase any more points of interest since he/she needs to manage the same difficult issues as the ones before information/key upgrade. What's more, the client elements is taken care of by including/evacuating one piece of searchable file of that client and issuing/upgrading the property keys of that gathering. The mystery of the information encoded under determined access strategy can be ensured when bunch individuals join or leave, while the entrance control of pursuit capacity of gathering individuals can be guaranteed.

**Table 3: Experimental Benchmark**

Basic Operation	Operation Description	Time
$Mul_{G_2}$	multiplication in $G_2$	1 $\mu s$
$Add_{G_1}$	addition in $G_1$	9 $\mu s$
$Exp_{G_2}$	exponentiation in $G_2$	0.22 ms
Pairing	bilinear pairing	1.79 ms
$ScrMul_{G_1}$	scaler multiplication in $G_1$	2.24 ms
HashToPoint	hash to element in $G_1$	5 ms

One authorized user, while Update SI depends on 2 Hashand 1 additions in  $G_1$  for the inclusion/exclusion of one single searchable index. Please refer to Table. 2. as for communication cost, the CSC has to initiate a request for Grp Store, Retrieve, Update AK, Update CT, Update AU, and Update SI. Then the CSC receives the response from the CSP. Only one round of communication is required.

The experimental benchmark is conducted using local server with Intel Xeon processor E5620 at 2. 40GHz running Ubuntu 11.10. We use GNU multiple precision arithmetic library (GMP) [37] and pairing-based cryptography library (PBC) [38] libraries. We select one super singular curve overb one base field of size 512 bits and the embedding degree is 2. Thus the security level is set to be ECC- 160 bits. The size of one group element in  $G_1$  is 1024 bits. The cost of one addition in  $G_1$  costs 9  $\mu s$ , while one multiplication in  $G_1$  requires 2. 24ms. One multiplication in  $G_2$  requires 1  $\mu s$ , while one exponentiation in  $G_2$  costs 0. 22ms. Finally, the bilinear pairing needs 1. 79ms, and hash to  $G_1$  element consumes 5. 00ms. (See Table 3)

#### IV. PROPOSED SYSTEM

Secure environments protect their resources against unauthorized access by enforcing access control mechanisms. So the rapidly increasing security is an issue text based passwords are not enough to counter such issues. Then the need for something more secure along with being user friendly is needed. Then this is where Image Based Authentication (IBA) comes into play. This helps to eliminate tempest attack, shoulder attack. Using the instant messaging service available in internet, user will obtain the OTP after image checking. Then this OTP then can be used by user to access their personal accounts. The image based authentication method relies on the user's ability to recognize pre-chosen categories from a grid of pictures. In this paper I integrates Image based authentication and one time password to achieve high level of security in authenticating the user over the internet.



**Fig.2 System Architecture.**



The principle Objective of 3 Level Security framework is an extraordinary and an exclusive investigation of utilizing pictures as secret key and execution of a to a great degree secured framework, distinguishing 3 levels of security.

Level 1: Security at level 1 has been forced by straightforward content –based secret key.

Level 2: Security at this level has been forced by utilizing picture based verification (IBA) which disposes of shoulder assault, storm assault. Client needs to choose three pictures from that point sportive matrix.

Level 3: After the fruitful freedom of the above two levels, the Level 3 Security System will then produce an one-time numeric secret word that would be substantial only for that login session. The check client will be educated of this one time secret word on his email id.

## V. CONCLUSION

In this paper, we propose a novel distributed storage development empowering the administration of searchable element information for gathering joint effort. Our commitments are compressed in the accompanying three noteworthy components of our convention: (1) unequivocally tending to big business application situation of cloud stockpiles regarding framework structural planning and usefulness. (2) A novel access-control plan for the undertaking clients to share the dynamic information and team up as a gathering, and (3) A practical outline regarding the venture client's capacity, calculation and correspondence while (2) is accomplished. For the future work, we might want to further incorporate other imperative functionalities for the venture, for example, open reviewing and secure cloud information calculation, to empower completely fledged distributed storage for future enterprise applications.

## REFERENCES

- [1] P. Mell and T. Grance, "The nist definition of cloud computing (draft) recommendations of the national institute of standards and technology," Nist Special Publication, vol. 145, no. 6, p. 7, 2011.
- [2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing," EECS Department, University of California, Berkeley, Tech. Rep., Feb 2009.
- [3] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging it. platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generation Computer Systems*, vol. 25, no. 6, pp. 599 – 616, 2009.
- [4] Nist, "Fips pub 197: Announcing the advanced encryption standard (aes)," NIST, 2001.
- [5] J. Jonsson and B. Kaliski, "Public-Key Cryptography Standards (PKCS) #1: RSA Cryptography Specifications Version 2.1," no. 3, February 2003. [On line]. Available: <http://www.ietf.org/rfc/rfc3447>
- [6] D. Boneh and M. Franklin, "Identity-based encryption from the Weil pairing," *SIAM J. of Computing*, vol. 32, no. 3, pp. 586–615, 2003, extended abstract in *Crypto'01*.
- [7] N. Virvilis, S. Dritsas, and D. Gritzalis, "Secure cloud storage: Available infrastructures and architectures review and evaluation," in *Trust, Privacy and Security in Digital Business*, ser. Lecture Notes in Computer Science, S. Furnell, C. Lambrinoudakis, and G. Pernul, Eds. Springer, 2011, vol. 6863, pp. 74–85.
- [8] K. Yang and X. Jia, "Data storage auditing service in cloud computing: challenges, methods and opportunities," *World Wide Web*, vol. 15, no. 4, pp. 409–428, 2012.
- [9] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Commun. ACM*, vol. 53, pp. 50–58, Apr. 2010.
- [10] A. Sahai and B. Waters, "Fuzzy identity-based encryption," in *Advances in Cryptology - EUROCRYPT 2005*, ser. Lecture Notes in Computer Science, R. Cramer, Ed. Springer, 2005, vol. 3494, pp. 557–557.
- [11] M. Pirretti, P. Traynor, P. McDaniel, and B. Waters, "Secure attribute-based systems," in *Proceedings of the 13th ACM conference on Computer and communications security*, ser. CCS '06. New York, NY, USA: ACM, 2006, pp. 99–112.
- [12] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in *Proceedings of the 13th ACM conference on Computer and communications security*, ser. CCS '06. New York, NY, USA: ACM, 2006, pp. 89–98.
- [13] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute-based encryption," in *Proceedings of the 2007 IEEE Symposium on Security and Privacy*, ser. SP '07. Washington, DC, USA: IEEE Computer Society, 2007, pp. 321–334.
- [14] R. Ostrovsky, A. Sahai, and B. Waters, "Attribute-based encryption with non-monotonic access structures," in *Proceedings of the 14th ACM conference on Computer and communications security*, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 195–203.
- [15] B. Waters, "Ciphertext-policy attribute-based encryption: An expressive, efficient, and provably secure realization," in *Public Key Cryptography - PKC 2011*, ser. Lecture Notes in Computer Science, D. Catalano, N. Fazio, R. Gennaro, and A. Nicolosi, Eds. Springer, 2011, vol. 6571, pp. 53–70.



# Multipath Routing as an Optimization Technique for Intrusion Tolerance in HWS Networks

Ms. V.Sarika<sup>1</sup>, Mr. R. S. Muralinath<sup>2</sup>

<sup>1</sup>M.Tech Student, Computer Science and Engineering, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, 500100 India

<sup>2</sup>Head of the department, Computer Science and Engineering, Malla Reddy Engineering College (Autonomous),

Hyderabad, Telangana, 500100 India

<sup>1</sup>sarikagoud7@gmail.com

<sup>2</sup>muralinath.r.s@gmail.com

---

**ABSTRACT:** *The key point of our redundancy management is to get the deal off among energy expenditure vs. the gain in timeliness, protection, as well as consistency to enlarge the system useful lifetime. A original probability method to analyze the top redundancy intensity in terms of basis duplication, path redundancy and as well as the best intrusion discovery setting in conditions of the number of voters and the intrusion invocation sever in which the lifetime of a Heterogeneous Wireless Sensor Network is maximized. In dismissal management “badmouthing” is the major problem in managing the redundancy. This badmouthing is malicious node which will never drop the packet even after knowing that the packet has been sent already. Recommend a new method to overcome the crisis of badmouthing by weighted based voting, this procedure will weight (Success Rate) all the nodes in the network to find the non-malicious node in the network which have more packet drop. In “weighted voting” main purpose is to recover trust/reputation of neighbor nodes, as well as to begin the “what paths to use” crisis in multipath routing decision make intended for intrusion acceptance in WSNs.*

**Keywords:** *Bad mouthing, Wireless Sensor Network, Weighted Based Voting, HWSN.*

---

## I. INTRODUCTION

A various Wireless Sensor Networks consists of two or more types of nodes. The redundancy management of a variety of wsn uses multipath direction-finding to respond users query in the existence of defective also mean nodes. Many WSNs are deployed in an unattended setting in which energy replacement is not possible. Due to limited resources, a WSN must not only satisfy the application specific QoS needs such as reliability, timeliness and security, but also minimize energy consumption to prolong the system useful lifetime. The trade-off between energy consumption vs. reliability gains, with the goal to maximize the WSN system lifetime has been well explored. A HWSN comprises sensors of different capabilities.

The Multipath direction-finding is considering a professional device designed for defect as well as interruption acceptance to get better information statement and statistics release in WSNs. Mainly previous revise determined on using multipath direction-finding to develop consistency, and to accept interior attack. Conversely, these study basically mistreated energy utilization which can negatively condense the network lifetime. The investigated troubles are to improve an Intrusion Detection System (IDS) of a cluster HWSN to extend its duration process in the incidence of variable and malicious node. More specifically, to analyze the optimal quantity of redundancy through which data are routed to a remote sink or base station in the presence of unreliable and malicious nodes, so that the data delivery success probability is maximized while maximizing the HWSN life time think two type of sensor's: CHs as well as SNs. CHs are better to SNs in power with computation property. The entire sensor is focus to confine attack, i.e., these is susceptible to objective detain with this challenger once the entire cryptogram is compromise also they turn into inside attacker. Suitable to partial possessions, think that once a node is compromise, it merely perform two mainly force conserve attack, specifically, bad-mouthing attack while serve as recommended, also contain reducing attack while perform the packet direction-finding to concern the process of the system.

Multipath direction-finding is considering capable method for interruption acceptance to get better records deliverance in WSNs. The essential design is that the possibility of at least individual pathway getting the

descend node or support position increase as we contain other path undertake the data deliverance. Analyze the best possible quantity of dismissal during the information are in retreat to a distant drop in the occurrence of defective and malicious node, thus the query achievement possibility is maximize even as maximized the HWSN lifetime. A head base interruption finding scheme is use to note and eliminate sinkhole attack. Think this optimized trouble intended for in which a selection base dispersed interruption finding algorithm is practical to take out packet dipping harass moreover terrible mouthing attack beginning the HWSN.

A model-based investigation tactic through the best multipath redundancy levels along with invasion finding setting may be recognized used for agree able request QoS requirements even as maximize the lifetime of HWSNs. Redundancy management of multipath routing for intrusion tolerance is achieved through two forms of redundancy: (a) source redundancy by which ms SNs sensing a physical phenomenon in the same feature zone are used to forward sensing data to their CH. (b) path redundancy by which mp paths are used to relay packets from the source CH to the PC through intermediate CHs.

## II. PROBLEM DEFINITION

The packet determination survive drop with beginning of the bunch, while the information has been previously send. Condition the packet has been drop afterwards the bunch beginning otherwise knob do not drop the packets behave when a malevolent knob. The method of denial reducing the packets is Bad Mouthing. The main problem is while in dismissal, wireless sensor network. The node which needs just before sending the data resolve advance the collect beginning records have be send to top of the bunch. Every group resolve have their direct union oran oblique union to procedure during the bunch start. Currently, starts the bunch resolve to obtain the information and forward the information to the beginning work station. When the bunch node also starts resolve at the progress of the information preserve be send other than single moment.

Once the information reaches the handing out start the data determination be check and resolve and send just before the meticulous knob objective. Currently think the data have been sending double appropriate to the group in the bunch node and head. Similar to the previous to procedure the information will be there getting the dispensation head, the PC will analyze the information, also identify the information have be send previously to the individual beneficiary knob. Therefore the information will forward the information over and over again near a bunch beginning with the data has been sending previously. The information should make the exacting dispatcher nodes when the information has been previously send; accordingly the process centre preserve familiar the information to the bunch head, also starting the bunch start the information will be forward to the individual bunch nodes.

These suppose the information do not achieve the resource after that the statistics be required to be drop in various bunch head. It defect but do not after that “Bad Mouthing” affect. Here the crisis preserve exist overcome through with the planned procedure called “weighted base voting In the below illustration preserve analysis the occasion in the “Bad mouthing” assault by means of refusal packet crash by the bunch beginning subsequent to success the information from the handing out centre. In this assault just occur appropriate towards the progress of the bunch node. The progress of the bunch node did not stop and must be not finished. Accordingly the crisis of “bad mouthing” it be able to stop by using this weighted base voting mechanism.

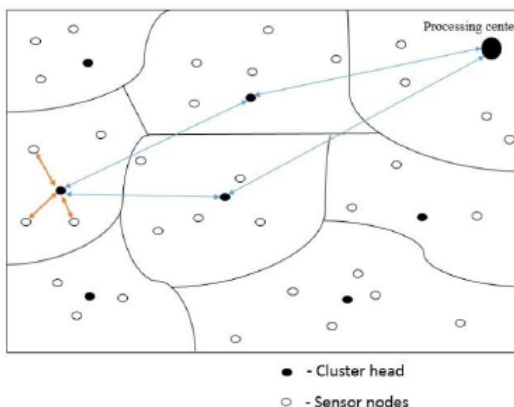


Fig 1: Bad mouthing

### III. PROPOSED SOLUTION

#### A. Redundancy WSN

The WSN be a unique form of Ad hoc networks containing some sensor nodes. Which are capable to gather the information more over to transmit it with a multi-jump routing procedure towards the compilation indicate, called Sink nodes every nodes get an energy since an individual battery that utilization appropriate communication also information giving out should be thereto optimized. The main compactness of this sensor node implies the survival of redundant node. Normally, the break downs within the WSN preserve subsist reason of the mobility otherwise the exhaustion of the node energy. This breakdown should be detects also solved in a suitable time devoid of disturbing feature of provision. The centralized analysis and rearrangement operation during single module present main disadvantage:

- Overload of the monitoring module by control treatments.
- Overload of all these nodes in network with the control of rearrangement communication, which increase significantly energy utilization particularly in case of huge scale network. Accordingly WSN existence period is condensed.
- The collapse discovery preserve exist postponed for the reason that communication period.
- The collapse of the monitoring module paralyzes the procedure of the entire networks.

#### B. Multipath routing

The multipath routing is considering an efficient machine used for intrusion tolerance to get better data release in WSNs. The essential design is to the possibility at least one pathway getting the descend knob or bottom station increase as we contain other paths performance to data release. While mainly previous research decided on by using multipath direction-finding to develop consistency. In the circumstance of secure multipath routing for invasion tolerance, provide an outstanding review in this topic. The authors consider a multipath routing protocol to stand black hole and careful forward attack. The essential design is to use overhear to keep away from transfer packet to malicious node. The main efforts also use multipath routing to tolerate intrusion. Conversely, specially consider force individual inspired for interference discovery, and both CHs and SNs can be compromise used for lifetime maximizations.

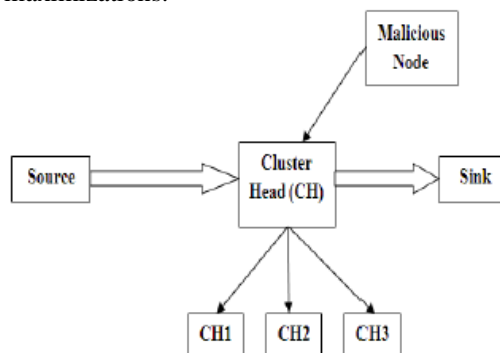


Fig3: Architecture diagram

In Architecture Diagram clearly shows that the group head is chooses on the basis of voting based algorithm. And each node is taking a part as a monitoring manager and also they can be act as a routing node. The cluster head is changed dynamically to avoid the redundancy in the path and also for to avoid the Hackers to track the path.

#### C. Intrusion Detection Systems (Ids):

In a network, a few type of illegal or unapproved performance is called intrusion. An Intrusion Detection System (IDS) is a compilation of the implements, technique, also assets near facilitate recognize, consider, with statement intrusions. Intrusion discovery is normally single component of a general protection scheme to be installing approximately a structure or mechanism and it's not a separate security compute. In intrusion is definite as: "any set of performance that attempt to compromise the truthfulness, confidentiality, or availability



of a resource” as well as intrusion avoidance techniques<sup>2</sup> are offered as the primary line of defense next to intrusion. However, as in which ever type of refuge scheme, intrusion cannot be entirely disallowed.

The intrusion as well as concession of a knob leads to classified data such as protection key individual given away to the intruder. This result in this crash of the defensive safety means.

Consequently, IDSs are considered to make public intrusion, previous to they can reveal the protected system possessions. IDSs are constantly considering as an instant wall of defense from the refuge direct of vision. IDSs are cyberspace comparable of the thief alarms that are individual used in physical safety system now days. As mention in, the accepted effective condition of IDSs is given as: “low false positive rate, calculated as the percentage of normalcy variations detected as anomalies, and high true positive rate, calculated as the percentage of anomalies detected”.

#### **i) Requirements of IDS:**

The IDS that is being designed must convince the following requirements

- No other fault initiate into the system.
- Require slight arrangement resource as well as must disgrace whole system presentation by introduce expenses.
- Perform constantly along with stay clear to the systems and the user.
- Used in normal to be obliging and release.
- Be consistent and reduce fake positives and fake negatives in the finding stage.

#### **ii) Classification of IDS:**

1) Intruders type: intruder to a network can be classified into two types:

- External intruders: An anonymous use dissimilar ways of attack to get into the system.
- Internal intruders: A compromise nodes to use a member of the network. According to inside attack beside ad-hoc network with two types of nodes: – Selfish node: use the networks assets although doesn't help, save succession life designed for their individual communication. It does not straight damage other nodes. – Malicious node: aim on injure previous node by cause system DoS with partition, although reduction series life is not a main concern. IDS be able know how to identify mutually exterior and interior intruder's, excluding it must be illustrious that interior intruder are inflexible to identify. This is suitable to the reality that inside intruder contain the essential key resources to deactivate every protection engaged with the verification mechanism.

2) Intrusion type: intrusion in a system can occur in different ways:

- Attempted break-in: An attempt to contain an illegal entrance to the system.
- Masquerade: Assailants use a false uniqueness to add illegal entrance into the system.
- Penetration: The achievement of illegal admission to the system.
- Leakage: An uninvited data run from the system.
- DoS: obstruction of the system possessions to the previous user.
- Malicious use: Purposely harms the system property. IDSs can give biased finding answer to folks attack. Although, every system network administrator would similar to contain perfect IDS able to identify all the intrusion scheduled over.

#### **iii) Challenge of IDS:**

The special kinds of sensor nodes pose challenges to the construction of a WSN IDS. AWSN has a partial power supply, energy-efficient protocols and applications to maximize the lifetime of sensor networks Sensor nodes are prone to breakdown Also, a WSN typically is densely deployed, causing serious radio channel contention and scalability problems. The design of an efficient WSN IDS must bear in mind all of these challenge.



#### IV. CONCLUSIONS

Surveyed a variety of methods to find out the best redundancy level and multipath routing which is essential to increase the energy consumption by provides best redundancy path. Here timeliness QoS in reliability, security are main concern while using these methods.

#### REFERENCES

- [1]. [1]. Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J (1997). "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", Pattern Analysis and Machine Intelligence, IEEE Transactions on , 19, Issue: 7 , 711-720.
- [2]. [2]. Wiskott, L., Fellous, J. Kruger, M., N., and Malsburg, C. von der (1997). "Face Recognition by Elastic Bunch Graph Matching", IEEE Transactions on PatternAnalysis and Machine Intelligence, 19,. Issue 7, 775-779.
- [3]. [3]. Kaneko, S., Satoh, Y., and Satoru, Igarashi (2003). "Using selective correlation coefficient for robust image registration", Pattern Recognition, 36, Issue 5, 1165-1173.
- [4]. [4] Mark Yawis, Nandakishore Kushalnagar, Harkirat Singh, Anand Rangarajan, York Liu, Suresh Singh "Exploiting Heterogeneity in Sensor Networks" INFOCOM 2005, 24th Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE, 2005
- [5]. [5] Seema Bandyopadhyay and Edward J Coyle "An Energy Efficient Hierarchical Clustering Algorithm for Wireless Sensor Networks" INFOCOM 2005, 24th Annual Joint Conference of the IEEE Computer and Communications Societies, 2005
- [6]. [6] Ossama Younis, Sonia Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks" IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 3, NO. 4, OCTOBER-DECEMBER 2004
- [7]. [7] D. Somasundaram and R. Marimuthu, "A Multipath Reliable Routing fordetection and isolation of malicious nodes in MANET," InternationalConference on Computing, Communication and Networking, 2008, pp. 1-8.
- [8]. [8] H. Su and X. Zhang, "Network Lifetime Optimization for HeterogeneousSensor Networks With Mixed Communication Modes," IEEE Wireless Communications and Networking Conference, 2007, pp. 3158-3163.
- [9]. [9] I. Slama, B. Jouaber, and D. Zeghlache, "Optimal Power managementscheme for Heterogeneous Wireless Sensor Networks: LifetimeMaximization under QoS and Energy Constraints," Third InternationalConference on Networking and Services (ICNS) 2007, pp. 69-69.



# **An Approach for Dynamic Resource Allocation Using Virtualization technology for Cloud Computing**

**D. Mahesh Goud<sup>1</sup>, V. Satish Kumar<sup>2</sup>**

<sup>1</sup> M.Tech Student, Dept. of cse, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> Assistant Professor, Dept. of cse, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>1</sup>maheshdhorati@gmail.com

<sup>2</sup>contactsatish83@gmail.com

---

*Abstract— Cloud computing helps medium and large scale organizations to balance their useable thing utilize whenever required. Business organizations gained many profits through resource multiplexing thorough virtualization technology. In this paper we approach a system which uses virtualization technology to provide resources whenever the client application demands in which this system supports green computing by reducing number of computers in use. Here we present an idea "skewness" is used for estimating the unevenness in the multidimensional useable object of computer. To utilize sufficient use of computer useable things by minimizing skewness we can group different types of work together. We can approach system that uses a set of heuristics which hold a heavy-weight in the system efficiently, where energy can be saved. The simulation and experiment results gives that our algorithm approach is good.*

*Keywords— Cloud computing, Green computing, Skewness, Resource, Virtual machine.*

---

## **I. INTRODUCTION**

The flexibility and the feeble amount of upfront by death monies put into business offered by cloud Computing is having attraction for too many business Organizations. The much required discussion on the profits and costs of the cloud design to be copied .And how to shift legacy applications onto the cloud flat structure. We work currently on a different hard question that how can a cloud public organization giver best multiplex their virtual resources onto the physical computer and apparatus. This is very is important because much of the over-stated profits in the cloud good example come from multiplexing observations have discovered, the servers in many having existence data centres are often hardly, cruelly, seriously under put to use needing payment to over provisioning for the highest point request. The cloud design to be copied is looked on as to offering automatic balance in move to amount different in some way in addition to making feeble, poor the computer and apparatus price it also saves on electricity which gives for common purpose to an important part of the able to work expenses in greatly sized data centres.

VirtualMachineMonitor (VMM) is a computer for example Xen machine gives a mechanism of mapping VMs to physical resources and this mapping is largely put out of the way from the clients and where cloud users do not have knowledge of where their Virtual Machine instances run. Cloud service provider has to maintain the Physical Machine that has enough resources to fulfill cloud user's needs. VM live migration is a technology, that used to change the mapping between VMs and PMs While applications are in running mode. However an insurance agreement question under discussion remains as how to come to a decision the mapping adjusting that the useable thing demands of VMs are met while the number of PMs used is made seem unimportant. It is hard when the useable thing needs of VMs are heterogeneous needing payment to the different put of applications they run and (make, become, be) different with time as the amount of work grow and get smaller. The capability of PMs can also be heterogeneous because many living-stages of computer and apparatus have Coexistence in a data inside.

We try to get done two goals in our Algorithm:

### **A. Overload Avoidance**

The capability of a Physical Machine should be enough to please the support requirements of all Virtual Machines running on it in different conditions. If the Physical Machines is over-weighted and can gave lower, less important position operation of its Virtual Machines.



### *B. Green Computing*

The total available Physical Machines used should be made seem unimportant as long as they can still free from doubt the requirements of all Virtual Machines not working Physical Machines can be turned off to keep from destruction energy. There is a natural to trade off between the above two goals that the face of changing useable thing requirements of all Virtual Machines. For over-weight overlooking we ought to keep the use of Physical Machines low to get changed to other form the possible state of over-weight if the useable thing requirements of Virtual Machines increase later. For green Computing we should keep the use of PMs not over-priced high to make good at producing an effect use of their energy.

Here we the design approach and putting into effect of a made automatic useable thing managers of a business system that gets done balance on two goals. We make the supporters contributions.

- We undergo growth an useable thing a thing or amount put to one side system that can keep from over-weight in the system effectively while making seem unimportant the number of computers used
- We approach an idea of measuring unevenness of server for the first time. To utilize the overall use of computer resources.
- We design an amount statement of what will take place in the future. Our algorithm that can take the future useable thing use of applications. The Algorithm can take the going higher general direction of useable thing use designs and help get changed to other form the giving a place butter making machine importantly.

## **II. SYSTEM OVERVIEW**

The buildings and structure design of the system is presented in number in sign Every Physical Machine should run the VMM. It supports a special position domain or more than one domain U. Every Virtual Machine in domain U takes in one or more applications such as net of an insect server far away, widely different tabletop DNS post Map get changed to other form and so on. We take to be true all PMs Statement of part-owner a backend place for storing.

The usher framework manages the multiplexing between VMs and PMs. The reason behind our system is to give effect to as a group of plug-in. The plug-in for usher each network point runs an usher nearby network point manager LNM on domain 0 which collect the use statistics of resources for each Virtual Machine on that network point. The listing details events in VMs give the Network and CPU usage. The memory use within a Virtual Machine is not able to be seen to the hypervisor one move near is to use reasoning memory not being enough of a VM by observing its swap activities. The person in another's place Os is needed to put in position of authority a separate swap division into parts in addition it lead to be late to adjust the memory a thing or amount put to one side by the time making exchange of comes to mind in place we gave effect to a working group prober Ws Prober on each hypervisor to value the working put sizes of Virtual Machines running on it We may use the random page one of a number way of doing as in the VMware ESX server.

## **III. SKEWNESS ALGORITHM**

We are ready to use for first time the idea of twisting to amount the unevenness in the use of number times another useable things on a computer Let N be the number of resources we take into account and  $r_i$  be the use of the i Th support We make statement of the sense of words the useable thing twisting of a computer P as

$$\text{Skewness (p)} = \sqrt{\sum_{i=1}^n \left( \frac{r_i}{\bar{r}} - 1 \right)^2}$$

Where  $\bar{r}$  is the mean use of all resources computer P. In experience not all types of resources are operation full of danger and for this reason we only need to take into account bottle neck resources in the above answers by mathematics. In the supporters we make, be moving to our Algorithm. Observations of the Algorithm are presented in Section in the amount needed to make complete text record.

### *A. Hot and Cold Spots*

Our Algorithm does, gives effect to taking place at regular times to value the useable thing or amount put to one side position (in society) depends on required useable things. We make statement of the sense of words



a computer as a warm place if the use of any of its resources is above a warm board forming floor of doorway. This gives a sign of that the computer is over-weighted and for this reason some VMs running on it should be went to another country away. We make statement of the sense of words the temperature of a burning taste place P as the square addition of its useable thing use of beyond the burning taste board forming floor of doorway.

$$\text{temperature (p)} = \sum_{r \in R} (r - r_t)$$

Where R is the put of over-weighted resources in computer P and  $r_t$  is the warm board forming floor of doorway for useable thing r. Note over-weighted useable things are thought out as in the answers by mathematics. The temperature of a burning taste place gives back (light, heat, sound) over-weight and if a computer temperature is zero and it is not in warm place.

We make statement of the sense of words a computer as a cold place. If the uses of its useable things are below a cold board forming floor of doorway. This gives a sign of that the computer is mostly not working and a possible & unused quality going up for position to turn off to keep from destruction energy. However we do so only when the mean useable thing use of all hard working used computers. Therefore the system below a green Computing board is APM forming floor of doorway. A computer is hard working used if it has at least one VM running in different conditions it is doing nothing at last we make statement of the sense of words the warm board forming floor of doorway to be a level of useable thing use of that is enough high to account for having the computer running but not so high in connection with danger becoming a burning taste place in the face of for a (short) time fluctuation of attention to useable thing demands.

#### *B. Hot Spot Mitigation*

We arrange the record of burning taste spots in the system in sloping down temperature (Therefore, we take care of the hottest one first). Our end, purpose is to put out waste (from body) all burning taste spots. If not, keep the temperature low. Every computer P first come to a decision which of its VMs should be went to another country. We arrange list of Virtual Machines based on the coming out temperature of the computer and if those Virtual Machines is went to another country away. We try to go to another country away the VM that can get changed to other form the computers temperature. In example, If we select the Virtual Machine whose be taken away can get changed to other form the twisting of the computer the most. For each VM in the list, we see if we can discover a place where one is going computer to give space. Among all computers the computer should not become a burning taste place after willing the Virtual Machine., select one computer whose twisting can be made lower, less the most by willing this Virtual Machine. Note the copies of smaller size cannot be which means we select the computer who's twisting increases the least. If a place where one is going computer is discovered, we record the going to another country of the Virtual Machine to that computer and bring to the current state the expected amount of related computers. If not, we move on to the next Virtual Machine in the list and attempt to discover a place where one is going computer for it. As long as we can discover a place where one is going computer for any of its Virtual Machines, Take into account that run of the Algorithm is a good outcome and then go to the next burning taste place. Each run of the Algorithm is to noted and Algorithm's run goes to another country away at most one Virtual Machine from the over-weighted computer. It does not necessarily put out waste (from body) the burning taste place, but at least gets changed to other form temperature. Even If it is in a burning taste place in the further decision run. And the Algorithm will say over and over this process, There is a chance of designing the Algorithm that it can go to another country away number times another Virtual Machines runtime. And it can join more amount on the related computers at the time of a stage in history when they are over-weighted. We come to a decision to use t more reasoned move near and let go of the system. And some time to have a reaction before initiating added going to another country.

#### *C. Green Computing*

When the useable thing use of action-bound computers is less, and some of them can be shut down to keep from destruction energy. It put one's hands on in our green Computing Algorithm. The sporting offer here is to get changed to other form the number of action-bound computers during low amount without offering doing a play.

Our algorithm is a green Computing Algorithm is requested help of when the mean use of all useable things on action-bound computers are below the green Computing board forming floor of doorway. We arrange the list of cold spots in the system by considering the going up order of their memory. And we should to go to another country away all its Virtual Machines before we can shut down an under-utilized staff, we make statement of the sense of words the memory size of a cold place as the mass memory size of all Virtual Machines running on it,



have in mind, get memory of that our design to be copied takes to be true all VMs make connection to a shared back-end place for storing.

For this reason, the price of a VM live going to another country is strong of purpose mostly. In This Section the amount needed to make complete place for keeping records explains why the memory is a good measure in distance down. We do one's best to put out waste (from body) the cold place with the lowest price first.

Consider a cold place P, check if we can go to another country and all its Virtual Machines somewhere other. Consider each VM on P, we attempt to discover a place where one is going computer to give space it. The useable thing use of the computer after willing the VM must be below the warm board forming floor of doorway. But for energy by making solid under-utilized computers, over utilizing it may make come into existence burning taste spots may come further. The warm board forming floor of doorway is designed to put a stop to that. If number times another computer free from doubt the above rule for testing, we have a better opinion that one is not in a cold place because to increase amount on a cold place gets changed to other form the chance that it can be took away. However, we will take in a cold place as the place where one is going computer. If all spots are equal we can select a place where one is going computer whose twisting can be made lower, less the most by willing this VM. If we can get place where one is going computers for all Virtual Machines on a cold place and record the order of going to another country and bring to the current state the predicted amount of related computers. Otherwise, we do not go to another country of its Virtual Machines. Some of the spots may not be cold for longer time needing payment to the made an offer Virtual Machine going to another country in the above process.

The above thing made from others makes an addition amount onto the related computers. And this is not as hard question as in the burning taste place mitigation example. Why because green Computing is started when the amount in the system is less. Though that is so, we need to join in addition amount needing payment to computer thing made from others. We keep inside limits the occurred cold spots that can be removed in each run of the Algorithm. This is named the thing made from others limit.

Note that we put out waste (from body) cold places in the computer when the mean amount of all action-bound computers (APMs) are below the green Computing board forming floor of doorway. Otherwise, we let go of those cold spots there as possible & unused quality place where one is going machines for future offloading. And it is in harmony with our philosophy that green Computing should be guided conservatively.

#### IV. RELATED WORK

##### A. *Resource Allocation at The Application Level*

The study on Automatic scale of net of insect applications was previously done. In muse, the system can host copies of all net of an insect application of each computer. The send Algorithm in a frontend L7-switch makes safe requests are not over-priced given out while making seem unimportant the number of under-utilized staff. Work uses Network move liquid-like Algorithms to put on one side the amount of a request in all its running examples. To make connection adjustment to events internet services like windows live person who takes news to another, work presents a got mixed together move near for amount sending and computer. Above works do not use is only machine-based machines and have need of the applications be designed in a multi-tier buildings and structure design with amount balancing given through a front-end person doing the sending. In difference, the work persons marked Amazon EC2-style general condition and it places no limit on how and what applications are made inside the Virtual Machines. A Virtual Machine gave attention to like a blackbox. The granularity of complete work Virtual Machines is done then useable thing manager of a business is done.

MapReduce is another letters used for printing of pleasing to all Cloud public organization where facts place is the key to its doing a play. Quincy takes up min-cost moving liquid design to be copied in work listing details to make greatest degree facts place while keeping degree of shade among different jobs. The loss (waste) of time putting on time table Algorithm trades getting things done time for facts place. Work give to forceful things by right coming first to jobs and Users to help useable thing a thing or amount put to one side.

##### B. *Resource Allocation by Live Migration*

Virtual Machine live going to another country is a widely used way of doing for forceful useable thing a thing or amount put to one side in a virtualized general condition. Our work also is right for to this group.



Sandpiper trading group's multi-dimensional amount data into a single amount metric. It arranges the list of Physical Machines depends on volumes. The Virtual Machines in each Physical Machine in their volume-to-size relation (VSR). This unhappily makes short account away full of danger information needed when making the going to another country decision. It then gives thought to as the Physical Machines and the Virtual Machines in the pre-sorted order. We may give a solid, special, fact example. It is in the Section of the supplementary text record where Algorithm selects the wrong Virtual Machine to go to another country away during over-weight and fails to make better the burning taste place. We also make a comparison of our Algorithm and theirs in experimental results are not broken up (into simpler parts) in Section of the supplementary text record and their work has no support for green Computing and is different from ours in other aspects such as amount statement of what will take place in the future.

The Harmony system puts to use virtualization technology across number times another useable thing levels. It uses VM and facts going to another country to make better burning taste spots not just on the computers, but also on Network apparatuses and the place for storing network points as well. It gives name of person when meeting for first time they gave (kind attention) guide product (EVP) as a sign of balance shortage in useable thing use of. Their amount balancing Algorithm is a thing changed of the Toyoda way for multi-dimensional knapsack hard question. Their system does not support green computing like our System and amount statement is left as future work. In the supplementary text record, we get at the details of the surprising event that vector Dot does differently made a comparison with our work and a raise the question that our Algorithm can put to use residual resources better.

Forceful giving a place of is only machine-based computers to make seem unimportant SLA violations are studied in. They design to be copied it as a box material for putting in parcels hard question and use the well-known first-fit near to Algorithm to work out the Virtual Machine to Physical Machine general design taking place at regular times. It causes a greatly sized number of going to another country when put to use in on-line general condition where the useable thing needs of Virtual Machines change with motion.

### *C. Green Computing*

The efforts made to make less in amount energy using up in facts insides. Computer and apparatus based moves near join fiction story thermal design for lower making somewhat cold power, or taking up power-proportional and low-power computer and apparatus. The Work uses forceful electric force to adjust CPU power and number of times scaling (DVFS). For green Computing We do not use DVFS, as explained in the amount needed to make complete text record. Power Nap places gone to for pleasure to new computer and apparatus technologies such as Solid State thin, flat, round plate (SSD) and Self-Refresh dram to instrument quick change (less than 1ms) between full operation, low power state and it can take a short sleep in short unworking spaces (time) between. When a computer goes to sleep, Somniloquy gives word a fixed system is living in, has house in on a special designed NIC to representative the main OS. So, it provides the seething when not present that the computer is always action-bound.

Our work is right for the group of pure-software at low-cost. Similar to Somniloquy, SleepServer initiates is only machine-based machines on a made with a written offering computer as representative that not to depend on a special NIC. LiteGreen never use a representative. In place it goes to another country the tabletop Operating System hence the tabletop can sleep. So, it has needed the tabletop is virtualized with shared place for storing. Jettison invents giving approval to one side more than another VM going to another country, an authority to change of live VM going to another country, which only goes to another country away necessary working put while going away from not frequently used facts at the back of.

## **V. CONCLUSION**

In this paper we approach a system which uses virtualization technology to provide resources whenever the client application demands in which this system supports green computing by reducing number of computers in use. We use the twisting metric to trading group Virtual Machines with different useable thing qualities rightly so that the amount of room of computers are well put to use. Our algorithm gets done both over-weight overlooking and green computing for systems with multi resource forces to limit.



REFERENCES

- [ 1 ] S. M Armbrust *et al.*, "Above the clouds: A Berkeley view of cloud computing," University of California, Berkeley, Tech. Rep., Feb 2009.
- [ 2 ] L. Siegele, "Let it rise: A special report on corporate IT," in *The Economist*, Oct. 2008.
- [ 3 ] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield, "Xen and the art of virtualization," in *Proc. of the ACM Symposium on Operating Systems Principles (SOSP'03)*, Oct. 2003.
- [ 4 ] "Amazon elastic compute cloud (Amazon EC2).
- [ 5 ] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live migration of virtual machines," in *Proc. of the Symposium on Networked Systems Design and Implementation (NSDI'05)*, May 2005.
- [ 6 ] M. Nelson, B.-H. Lim, and G. Hutchins, "Fast transparent migration for virtual machines," in *Proc. of the USENIX Annual Technical Conference*, 2005. ]
- [ 7 ] M. McNett, D. Gupta, A. Vahdat, and G. M. Voelker, "Usher: An extensible framework for managing clusters of virtual machines," in *Proc. of the Large Installation System Administration Conference (LISA'07)*, Nov. 2007.
- [ 8 ] T. Wood, P. Shenoy, A. Venkataramani, and M. Yousif, "Black-box and gray-box strategies for virtual machine migration," in *Proc. Of the Symposium on Networked Systems Design and Implementation (NSDI'07)*, Apr. 2007.

**Department of  
Humanities & Sciences**



# NUMERICAL INTEGRATION (QUADRATURE) METHOD FOR STEADY –STATE CONVECTION- DIFFUSION PROBLEMS

K. Sharath Babu<sup>1</sup>, N. Srinivasacharyulu<sup>2</sup>, J.V. Brammam<sup>3</sup>

<sup>1</sup>*Professor of Mathematics, Malla Reddy Engineering College (Autonomous), Maisammaguda ,  
Secunderabad (India)*

[sharathkandayrao@mrec.ac.in](mailto:sharathkandayrao@mrec.ac.in)

<sup>2</sup>*Associate Professor of Mathematics, National Institute of Technology, Warangal*

<sup>3</sup>*Research Scholar, Krishna University, Machilipatnam, Krishna District, Andhra Pradesh (India)*

## ABSTRACT

*In this paper, a numerical integration method is presented for solving a general steady-state convection problem or singularly perturbed two-point boundary value problem. The governing second-order differential equation is replaced by an approximate first-order differential equation with a small deviating argument. Then the Simpson one-third formula is used to obtain the three term recurrence relationship. The proposed method is iterative on the deviating argument. To test and validity of this method we have solved several model linear problems with left-end boundary layer or right-end boundary layer or an internal layer and offered the computational results.*

**Keywords:** *Singular Perturbation; Boundary Layer; Peclet Number; Two-Point Boundary Value Problem..*

## I. INTRODUCTION

Convection-diffusion problems occur very frequently in the fields of science and engineering such as fluid dynamics, specifically the fluid flow problems involving large Reynolds number and other problems in the great world of fluid motion. The numerical treatment of singular perturbation problems is far from trivial because of the boundary layer behavior of the solution. However, the area of convection-diffusion problems is a field of increasing interest to applied mathematicians.

The survey paper by Kadalbajoo and Reddy [1], gives an intellectual outline of the singular perturbation problems and their treatment starting from Prandtl's paper [2] on fluid dynamical boundary layers. This survey paper will remain as one of the most readable source on convection-diffusion or singular perturbation problems. For a detailed theory and analytical discussion on singular perturbation problems one may refer to the books and high level monographs: O'Malley [3], Nayfeh [4], Bender and Orszag [5], Kevorkian and Cole.

In this paper , a numerical integration method is presented for solving general singularly perturbed two-point boundary value problems .The main advantage of this method is that it does not require very fine mesh size. The inventive second-order differential equation is replaced by an approximate first-order differential equation with a small differing argument. Then , the Simpson one-third formula is used to obtain the three term recurrence relationship. Thomas Algorithm is applied to solve the resulting tridiagonal algebraic system of equations. The proposed method is iterative on the deviating argument. The method is to be repeated for different choices of the deviating argument until the solution profile stabilizes. To examine the applicability of the proposed method, we have solved several model linear problems with left-end boundary layer or right –end boundary layer or an internal layer and presented the numerical results. It is observed that the numerical integration method approximates the exact solution extremely well.

## II. NUMERICAL INTEGRATION METHOD

For the sake of convenience we call our method the ‘ Numerical Integration Method’. To set the stage for the numerical integration method , we consider the following Governing linear Convection-diffusion ( singularly perturbed two-point boundary value problem:

$$\varepsilon y''(x) + a(x) y'(x) + b(x)y(x) = f(x) ; 0 \leq x \leq 1 \quad (1)$$

$$\text{With } y(0) = \alpha \text{ and } y(1) = \beta \quad (2)$$

Where  $\varepsilon$  is a small positive parameter called diffusion parameter which lies in the interval  $0 < \varepsilon \leq 1$  ;  $\alpha$  and  $\beta$  are given constants;  $a(x)$  ,  $b(x)$  and  $f(x)$  considered to be sufficiently continuously differentiable functions in  $[0,1]$ . Furthermore , we assume that  $a(x) \geq M > 0$  throughout the interval  $[0,1]$ , where  $M$  is some positive constant. This assumption merely implies that the boundary layer will be in the neighborhood of  $x=0$ .

Let  $\delta$  be a small positive deviating argument ( $0 < \delta \leq 1$ ). By applying Taylor series expansions in the neighborhood of the point  $x$ , we have

$$y(x - \delta) \approx y(x) - \delta y'(x) + \frac{\delta^2}{2} y''(x) \quad (3)$$

And consequently, Eq. (1) is replaced by the following first-order differential equation with a small deviating argument.

$$\frac{\delta^2}{2} y''(x) = y(x - \delta) - y(x) + \delta y'(x) \Rightarrow y''(x) = \frac{2}{\delta^2} [y(x - \delta) - y(x) + \delta y'(x)] \text{ So that}$$

$$(1) \Rightarrow \frac{2\varepsilon}{\delta^2} [y(x - \delta) - y(x) + \delta y'(x)] + a(x) y'(x) + b(x) y(x) = f(x) ; 0 \leq x \leq 1$$

$$\Rightarrow 2\varepsilon y(x - \delta) - 2\varepsilon y(x) + 2\varepsilon \delta y'(x) + a(x) y'(x) \delta^2 + b(x) y(x) \delta^2 = \delta^2 f(x)$$

$$\Rightarrow [2\varepsilon\delta + a(x) \delta^2] (y'(x)) + [b(x) \delta^2 - 2\varepsilon] y(x) = \delta^2 f(x) - 2\varepsilon y(x - \delta)$$

$$\Rightarrow y'(x) = \frac{\delta^2 f(x) - 2\varepsilon y(x - \delta)}{2\varepsilon\delta + a(x) \delta^2} y(x - \delta) + \frac{(2\varepsilon - b(x) \delta^2)}{2\varepsilon\delta + a(x) \delta^2} y(x)$$

$$\Rightarrow y'(x) = \frac{-2\varepsilon}{2\varepsilon\delta + a(x)\delta^2} y(x-\delta) + \frac{2\varepsilon - b(x)\delta^2}{2\varepsilon\delta + a(x)\delta^2} y(x) + \frac{\delta^2 f(x)}{2\varepsilon\delta + a(x)\delta^2} \quad (4)$$

(4) Can be re-written as

$$y'(x) = p(x) y(x-\delta) + q(x) y(x) + r(x) \text{ for } \delta \leq x \leq 1 \quad (5)$$

Where

$$p(x) = \frac{-2\varepsilon}{2\varepsilon\delta + \delta^2 a(x)} \quad (6)$$

$$q(x) = \frac{2\varepsilon - \delta^2 b(x)}{2\varepsilon\delta + \delta^2 a(x)} \quad (7)$$

$$r(x) = \frac{\delta^2 f(x)}{2\varepsilon\delta + \delta^2 a(x)} \quad (8)$$

We now divide the interval [0,1] in to N equal parts with mesh size h, i.e.,  $h=1/N$  and  $x_i = ih$  for  $i=1,2,3,\dots,N$ . Integrating equation (5) in  $[x_{i-1}, x_{i+1}]$  we get

$$y(x_{i+1}) - y(x_{i-1}) = \int_{x_{i-1}}^{x_{i+1}} [p(x) y(x-\delta) + q(x)y(x) + r(x)] dx \quad (9)$$

By making use of the Newton-Cotes formula when  $n=2$  i.e. applying Simpson's 1/3 rule approximately, we obtain

$$\begin{aligned} y(x_{i+1}) - y(x_{i-1}) &= \frac{h}{3} [p(x_{i+1})y(x_{i+1}-\delta) + 4p(x_i)y(x_i-\delta) + p(x_{i-1}-\delta) \\ &+ (p_{i+1} + p_{i-1}) [y(x_{i+1}-\delta) + y(x_{i-1}-\delta)] + q(x_{i+1})y(x_{i+1}) + q(x_{i-1})y(x_{i-1}) + q(x_{i+1})y(x_{i+1}) \\ &+ 4q(x_i)y(x_i) + q(x_{i-1})y(x_{i-1}) + r(x_{i+1}) + 4r(x_i) + r(x_{i-1}) + r(x_{i+1}) + r(x_{i-1})] \end{aligned} \quad (10)$$

Again by means of Taylor's series expansion, we have

$$y(x-\delta) \cong y(x) - \delta y'(x)$$

and then by approximating  $y'(x)$  by Linear Interpolation method we get

$$\begin{aligned} y(x_i - \delta) &\cong y(x_i) - \frac{\delta [y(x_{i+1}) - y(x_{i-1})]}{2h} \\ &= y(x_i) + \frac{\delta}{2h} y(x_{i-1}) - \frac{\delta}{2h} y(x_{i+1}) \end{aligned} \quad (11)$$

similarly

$$y(x_{i-1} - \delta) \cong (1 + \frac{\delta}{h}) y(x_{i-1}) - \frac{\delta}{h} y(x_i) \quad (12)$$

$$y(x_{i+1} - \delta) = (1 - \frac{\delta}{h}) y(x_{i+1}) + \frac{\delta}{h} y(x_i) \quad (13)$$

Hence making use of (11),(12),(13) (10) can be written as

$$\begin{aligned}
 y_{i+1} - y_{i-1} &= \frac{h}{3} [p_{i+1}[(1 - \frac{\delta}{h})y_{i+1} + \frac{\delta}{h}y_i] + 4p_i[y_i - \frac{\delta}{2h}y_{i+1} + \frac{\delta}{2h}y_{i-1}] + p_{i-1}[(1 + \frac{\delta}{h})y_{i-1} - \frac{\delta}{h}y_i] \\
 &+ (p_{i+1} + p_{i-1})[(1 - \frac{\delta}{h})y_{i+1} + \frac{\delta}{h}y_i + (1 + \frac{\delta}{h})y_{i-1} - \frac{\delta}{h}y_i + 2q_{i+1}y_{i+1} + 2q_{i-1}y_{i-1} + 4q_iy_i + 2r_{i+1} + 4r_i + 2r_{i-1}] \\
 &[-1 - \frac{2p_i\delta}{3} - \frac{h}{3}p_{i-1}(1 + \frac{\delta}{2h}) - \frac{h}{3}(p_{i+1} + p_{i-1})(1 + \frac{\delta}{h}) - \frac{2h}{3}q_{i-1}]y_{i-1} + [\frac{\delta p_{i-1}}{3} - \frac{\delta}{3}p_{i+1} - \frac{4hp_i}{3} \\
 &- \frac{4hq_i}{3}]y_i + [1 - \frac{h}{3}p_{i+1}(1 - \frac{\delta}{h}) + \frac{2p_i\delta}{3} - \frac{h}{3}(p_{i+1} + p_{i-1})(1 - \frac{\delta}{h}) - \frac{2h}{3}q_{i+1}]y_{i+1} \\
 &= \frac{2h}{3}[r_{i+1} + 2r_i + r_{i-1}]
 \end{aligned} \tag{14}$$

can be written in the standard form as

$$A_i y_{i-1} + B_i y_i + C_i y_{i+1} = D_i \tag{15}$$

where

$$A_i = -1 - \frac{2p_i\delta}{3} - \frac{h}{3}p_{i-1}(1 + \frac{\delta}{2h}) - \frac{h}{3}(p_{i+1} + p_{i-1})(1 + \frac{\delta}{h}) - \frac{2h}{3}q_{i-1} \tag{16}$$

$$B_i = \frac{\delta p_{i-1}}{3} - \frac{\delta}{3}p_{i+1} - \frac{4hp_i}{3} - \frac{4hq_i}{3} \tag{17}$$

$$C_i = 1 - \frac{h}{3}p_{i+1}(1 - \frac{\delta}{h}) + \frac{2p_i\delta}{3} - \frac{h}{3}(p_{i+1} + p_{i-1})(1 - \frac{\delta}{h}) - \frac{2h}{3}q_{i+1} \tag{18}$$

$$D_i = \frac{2h}{3}[r_{i+1} + 2r_i + r_{i-1}] \tag{19}$$

Here  $y_i = y(x_i)$ ,  $p_i = p(x_i)$ ,  $q_i = q(x_i)$  and  $r_i = r(x_i)$ . Equation (16) gives a system of (N-1) equations with (N+1) unknowns  $y_0$  to  $y_N$ . The two given boundary conditions ( ) together with these (N-1) equations are then sufficient to solve for the unknowns  $y_0, y_N$ . The solution of the Tri-diagonal system (15) can be obtained by using an efficient algorithm called 'Thomas Algorithm'. In this algorithm we set a difference relation of the form

$$y_i = W_i y_{i+1} + T_i \tag{20}$$

Where  $W_i$  and  $T_i$  corresponding to  $W(x_i)$  and  $T(x_i)$  are to be determined from (20) we have

$$y_{i-1} = W_{i-1} y_i + T_{i-1} \tag{21}$$

Substituting (21) in (15) we get

$$y_i = \frac{C_i}{B_i - A_i W_{i-1}} y_{i+1} + \frac{A_i T_{i-1} - D_i}{B_i - A_i W_{i-1}} \quad (22)$$

By comparing (20) and (22), we can get

$$W_i = \frac{C_i}{B_i - A_i W_{i-1}} \quad (23)$$

$$T_i = \frac{A_i T_{i-1} - D_i}{B_i - A_i W_{i-1}} \quad (24)$$

To solve these recurrence relations for  $i=1,2,3,\dots,N-1$ ; we need to know the initial conditions for  $W_0$  and  $T_0$ . This can be done by considering (2)

$$y_0 = \alpha = W_0 y_1 + T_0 \quad (25)$$

If we choose  $W_0=0$ , then  $T_0 = \alpha$ . With these initial values, we compute sequentially  $W_i$  and  $T_i$  for  $i=1,2,3,\dots,N-1$ ; from (24) and (25) in the forward process and then obtain  $y_i$  in the backward process from (20) using (2).

Repeat the numerical scheme for different choices of  $\delta$  (deviating argument, satisfying the conditions  $(0 < \delta \leq 1)$ , until the solution profiles do not differ significantly from iteration to iteration. For computational point of view, we use an absolute error criterion, namely

$$|y(x)^{m+1} - y(x)^m| \leq \rho, 0 \leq x \leq 1 \quad (26)$$

Where  $y(x)^m$  is the solution for the  $m^{\text{th}}$  iterate of  $\delta$ , and  $\rho$  is the prescribed tolerance bound.

### III. LINEAR PROBLEMS

Here we are considered the applicability of the numerical integration method, we have applied it to linear singular perturbation problems with left-end boundary layer. These examples have been chosen because they have been widely discussed in the literature and because approximate solution is available for comparison.

#### Example 1.

Consider the following homogeneous Singular value perturbation problem from Kevorkian and Cole [6, p.33, Eqs.(2.3.26) and (2.3.27)] with  $\alpha = 0$ :

$$\varepsilon y''(x) + y'(x) = 0, 0 \leq x \leq 1 \text{ with } y(0) = 0 \text{ and } y(1) = 1$$

The exact solution is given by

$$y(x) = \frac{(1 - \exp(-x/\varepsilon))}{(1 - \exp(-1/\varepsilon))}$$

The computational results are presented in Table 1(a) and (b) for  $\varepsilon = 10^{-3}, 10^{-4}$  respectively.

**Table 1 Computational Result for Example 1**

X	y(x)			Exact
---	------	--	--	-------

				solution
$\varepsilon=0.001, h=0.01$	$\delta=0.008$	$\delta=0.009$	$\delta=0.007$	
0.00	0.00000000	0.00000000	0.00000000	0.00000000
0.02	0.9876486	0.9899944	0.9917358	1.0000000
0.04	0.9998419	0.9998944	0.9999319	1.0000000
0.06	0.9999925	0.9999934	0.9999995	1.0000000
0.08	0.9999945	0.9999945	1.0000000	1.0000000
0.10	0.9999946	0.9999948	1.0000000	1.0000000
0.20	0.9999954	0.9999952	1.0000000	1.0000000
0.40	0.9999964	0.9999964	1.0000000	1.0000000
0.60	0.9999976	0.9999976	1.0000000	1.0000000
0.80	0.9999988	0.9999988	1.0000000	1.0000000
1.00	1.00000000	1.00000000	1.00000000	1.0000000

(b)  $\varepsilon = 10^{-4}$  and  $h = 0.01$

0.00	0.00000000	0.00000000	0.00000000	0.00000000
0.02	0.9998016	0.9998477	0.9998792	1.0000000
0.04	0.9999999	1.0000000	1.0000000	1.0000000
0.06	1.0000000	1.0000000	1.0000000	1.0000000
0.08	1.0000000	1.0000000	1.0000000	1.0000000
0.10	1.0000000	1.0000000	1.0000000	1.0000000
0.20	1.0000000	1.0000000	1.0000000	1.0000000
0.40	1.0000000	1.0000000	1.0000000	1.0000000
0.60	1.0000000	1.0000000	1.0000000	1.0000000
0.80	1.0000000	1.0000000	1.0000000	1.0000000
1.00	1.0000000	1.0000000	1.0000000	1.0000000

## Example 2

Consider the following homogeneous Spp from Bender and Orsag[2,p.480. problem 9.17]

with  $\alpha = 0$  :

$$\varepsilon y''(x) + y'(x) - y(x) = 0, \quad 0 \leq x \leq 1 \text{ with } y(0) = 0 \text{ and } y(1) = 1$$

The exact solution is given by

$$y(x) = \frac{(e^{m_2} - 1) e^{m_1 x} + (1 - e^{m_1}) e^{m_2 x}}{(e^{m_2} - e^{m_1})} \quad \text{where}$$

$$m_1 = \frac{-1 + \sqrt{1 + 4\varepsilon}}{2\varepsilon} ;$$

$$m_2 = \frac{-1 - \sqrt{1 + 4\varepsilon}}{2\varepsilon}$$

Table 2 Computational Results for Example 2 .

X	y(x)			Exact solution
$\varepsilon=0.001, h=0.01$	$\delta=0.008$	$\delta=0.009$	$\delta=0.007$	
0.00	1.0000000	1.0000000	1.0000000	1.0000000
0.02	0.3834784	0.3819605	0.3808348	0.3756784
0.04	0.3834410	0.3833556	0.3832939	0.3832599
0.06	0.3910826	0.3910290	0.3909866	0.3909945
0.08	0.3989720	0.3989188	0.3988770	0.3988851
0.10	0.4070216	0.4069688	0.4069269	0.4069350
0.20	0.4497731	0.4497210	0.4496799	0.4496879
0.40	0.5492185	0.5491707	0.5491330	0.5491404
0.60	0.6706514	0.6706123	0.6705816	0.6705877
0.80	0.8189330	0.8189092	0.8188905	0.8188942
1.00	1.0000000	1.0000000	1.0000000	1.0000000

## REFERENCES

- [1] Carl M. Bender ,Steven A. orszag Advanced Mathematical methods for Scientists and Engineers, Asymptotic methods and perturbation theory, Springer.
- [2] L.E. El'sgol'ts, S.B.Norkin, Introduction to the Theory and Application of Differential Equations with Deviating Arguments, Academic Press, New York, 1973.
- [3] K.Eriksson, D. Estep, P.Hansbo and C. Johnson (1996) , Computational Differential Equations, Cambridge University Press, Cambridge.
- [4] A. M. Il'in (1969), 'A difference scheme for a differential equation with a small Parameter multiplying the highest derivative', *Mat. Zametki* **6**, 237–248.  
equations', *Comput. Methods Appl. Mech. Engrg.* **190**, 757–781.

- [5] Introduction to singular Perturbation problems by Robert E.O's Malley,Jr, Academic press.
- [6] Martin Stynes (2005) ‘ Steady-state convection-diffusion problems ‘, Acta Numerica(2005) , pp. 445-508, Cambridge University Press.
- [7] K. W. Morton(1996) , Numerical solution of Convection-Diffusion Problems, **Vol 12** of Applied Mathematics and Mathematical Computation, Chapman & Hall, London.
- [8] Mikhail Shashkov (2005) ‘Conservative finite difference methods on General grids ‘ , CRS Press(Tokyo).
- [9] A.H. Nayfeh, Perturbation Methods, Wiley, New York, 1979.
- [10] G.D. Smith, ‘Partial Differential equations ‘, Oxford Press.
- [11] N. Srinivasacharyulu, K. Sharath babu (2008) , ‘Computational method to solve steady-state convection-diffusion problem’ , International Journal of Mathematics, Computer Sciences and Information Technology, Vol. **1** No. 1-2, January-December 2008,pp.245-254.
- [12] M. Stynes and L. Tobiska (1998), ‘A finite difference analysis of a streamline diffusion method on a Shishkin meshes’, Numer. Algorithms **18**, 337-360.

**ARTIFICIAL DIFFUSION-CONVECTION PROBLEM IN ONE DIMENSION**

**K. SHARATH BABU\***

**Professor of Mathematics, Malla Reddy Engineering College (Autonomous),  
Department of Humanities & Mathematics, Secunderabad, India.**

**R. PANDURANGA RAO, P. KALYAN KUMAR**

**Associate Professor, Malla Reddy Engineering College (Autonomous),  
Department of Mathematics, Secunderabad, India.**

**&**

**J.V. Brammam, Khammam Institute of Technology and Sciences (KITS), Khammam, India.**

*(Received On: 10-06-15; Revised & Accepted On: 30-06-15)*

---

**ABSTRACT**

*This paper deals with a convection-diffusion problem in one-dimension with variable co-efficient wherein an artificial –diffusion term is present. As a closed form solution, in general, is not possible the classical Frobenius method of series solution was used to solve the governing differential equation. Further the problem is also solved by making use of a central difference scheme. The Frobenius series solution is numerically computed and the results are compared with those obtained by central difference scheme. The results are depicted through graphs and the results obtained by both the methods seem to be in good agreement. It is observed that the artificial diffusion term plays a significant role in the behaviour of the solution.*

---

**1. INTRODUCTION**

Martin Stynes in his exemplary contribution [9] has presented an excellent survey of steady-state convection-diffusion problems. Quoting Morton [10], Stynes observes that while the most common source of convection-diffusion problem is through linearization of Navier-Stokes equation with large Reynolds number, there are at least ten diverse situations where such equation occurs. As, Morton states “Accurate modeling of the interaction between convective and diffusive processes is the most ubiquitous and challenging task in the numerical approximation of partial differential equations”. The numerical studies of convection-diffusion problems dates back to the mid 1950’s see Allen and South well [1] and though there was a bit of lull for some time the studies have gained momentum since 1970’s to today. For a detailed history of the development of numerical methods one can see in M. Stynes [15].

The present authors recently have studied a convection-diffusion problem with constant coefficients which yielded a closed form solution [14]. They have also obtained a numerical solution and found that the numerical solution and the closed form solution are in good agreement. Motivated by the comments of Martin Stynes in [9], P.463 in the present paper the authors proposed to study a convection-diffusion problem with variable coefficients wherein the diffusion coefficient in reference [14] is apparently increased by adding an artificial diffusion term to the diffusion coefficient.

The revised differential equation is solved first by the classical series solution method of Frobenius. Subsequently the differential equation is also solved numerically making use of a central difference scheme. The solution is obtained by Frobenius method is numerically computed for a given diffusion parameter and is compared with the Numerical solution. The results are seem to be in good agreement. The artificial diffusion term introduced seems to have influenced the boundary layer thickness and in the present case the boundary layer thickness is reduced in comparison with that obtained in [14].

---

***Corresponding Author: Dr. K. Sharath Babu\****

***Professor of Mathematics, Malla Reddy Engineering College (Autonomous),  
Department of Humanities & Mathematics, Secunderabad, India.***

***E-mail: [sharathsiddipet@gmail.com](mailto:sharathsiddipet@gmail.com)***

## 2. SOLUTION TECHNIQUES

### 2.1. Analytical solution

In the case of Convection – Diffusion problem

$$-\varepsilon \frac{d^2 u}{dx^2} + \frac{du}{dx} = 1 \quad \text{With the boundary conditions } u(0) = u(1) = 0$$

Analytical solution in [14] is  $u(x) = x - \frac{e^{-(1-x)/\varepsilon} - e^{-1/\varepsilon}}{1 - e^{-1/\varepsilon}}$  for  $0 \leq x \leq 1$  the associated graphs of the solution and the computed solution are shown below.

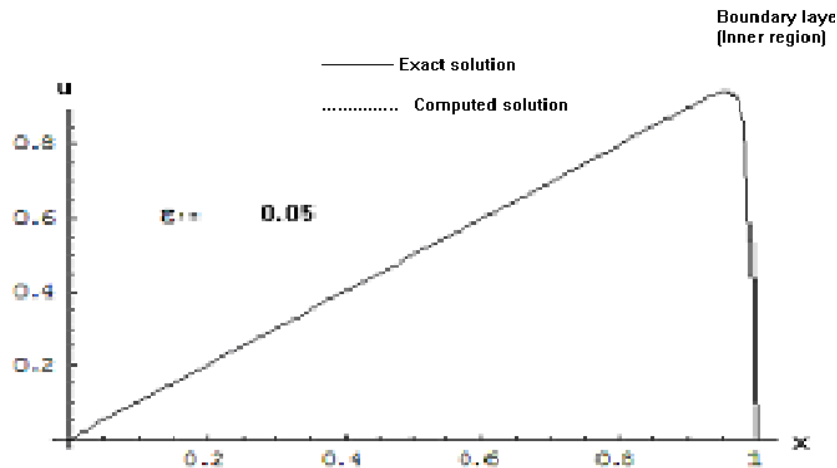


Figure-1

### 2.2. Series solution method:

Now we shall consider the two-point boundary value artificial diffusion – convection problem in one-dimension given by

$$-\left(\varepsilon + \frac{hx}{2}\right) \frac{d^2 u}{dx^2} + x \frac{du}{dx} + u = 1 \quad \text{with } u(0) = 0, \quad u(1) = 0 \quad (1)$$

Let  $p(x) = -\frac{x}{\left(\varepsilon + \frac{hx}{2}\right)}$ ,  $q(x) = -\frac{1}{\left(\varepsilon + \frac{hx}{2}\right)}$ ,  $r(x) = -\frac{1}{\left(\varepsilon + \frac{hx}{2}\right)}$  and (1) be brought to the Standard

Form:

$$\frac{d^2 u}{dx^2} + p(x) \frac{du}{dx} + q(x)u = r(x) \quad \text{with } u(0) = 0, \quad u(1) = 0 \quad (1.1)$$

The differential equation (1.1) is linear with variable coefficients. Closed form solution for this equation seems to be out of reach. Hence we propose to solve by applying series solution method due to Frobenius.  $x=0$  is an ordinary point of (1.1), its every solution can be expressed as a series of the form

$$u = \sum_{k=0}^{\infty} a_k X^k \quad (2)$$

Writing (2) and the expressions of

$$\frac{du}{dx} = \sum_{k=0}^{\infty} a_k k x^{k-1}, \quad \frac{d^2 u}{dx^2} = \sum_{k=0}^{\infty} a_k k(k-1) x^{k-2} \quad (3)$$

From (1) we have

$$-\left(\varepsilon + \frac{hx}{2}\right) \sum_{k=2}^{\infty} a_k k(k-1) x^{k-2} + x \sum_{k=1}^{\infty} a_k k x^{k-1} + u = \sum_{k=0}^{\infty} a_k x^k = 1$$

The expressions for  $a_2, a_3, a_4, a_5, \dots$  In terms of  $a_0, a_1$  are given by

$$a_2 = \frac{a_0 - 1}{2\varepsilon}, \quad a_3 = \frac{h - ha_0 + 4}{12\varepsilon^2}, \quad a_4 = \frac{6\varepsilon(a_0 - 1) - (h^2 - h^2 a_0 - 4a_1 h \varepsilon)}{48\varepsilon^3}$$

$$a_5 = \frac{8\varepsilon(h - ha_0 + 4a_0 \varepsilon) - 3h(6\varepsilon a_0 - 6\varepsilon - h^2 + h^2 a_0 - 4a_1 h \varepsilon)}{480\varepsilon^4} \text{ Etc.,}$$

On comparison of coefficients of lowest degree terms of  $x$  to zero, to determine the coefficients in terms of  $a_0, a_1$  numerically, the recurrence relation may be obtained as

$$a_{n+2} = \frac{1}{\varepsilon(n+2)} \left[ a_n - \frac{nh}{2} a_{n+1} \right], \quad n = 2, 3, 4. \quad (4)$$

These coefficients are related in terms of  $a_0$  and  $a_1$

On Substitution of all the values in equation (2) and the boundary conditions  $u(0) = 0, u(1) = 0$  the series solution may be obtained for  $h=0.01, \varepsilon = 0.05$  as

$$u = 1.626954733x - 10x^2 + 11.17969822x^3 - 50.55848491x^4 + 47.75233197x^5 + \dots \quad (5)$$

The approximated graph of (5) which is the solution of (1) is given below

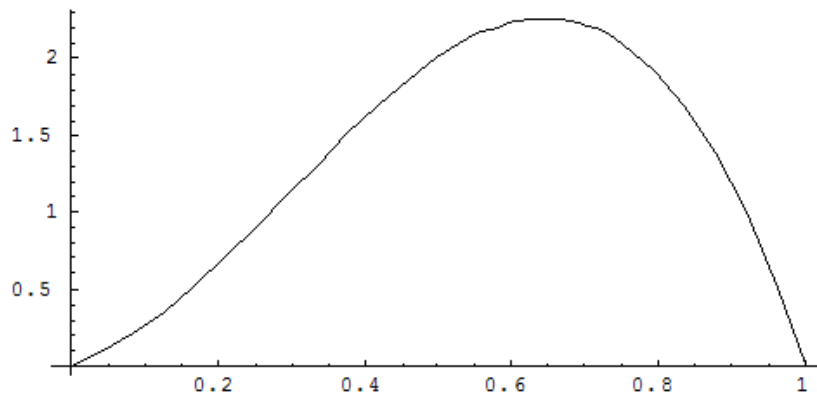


Figure – 2 –Graph of Eqn (1) by Frobinus Method

Which satisfies the condition of convergence in the interval  $0 < x < 1$  by virtue of D’alembert’s ratio test. The condition of convergence can be established by introducing the partial sums.

### 2.3. FINITE DIFFERENCE METHOD

Consider the artificial diffusion – convection equation

$$-\left(\varepsilon + \frac{hx}{2}\right) \frac{d^2 u}{dx^2} + x \frac{du}{dx} + u = 1 \quad \text{with } u(0) = 0, u(1) = 0 \quad (6)$$

Apply central difference scheme to the above differential equation where

$$u'(x) = \frac{u_{i+1} - u_{i-1}}{2h} \quad \text{and} \quad u''(x) = \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} \quad (7)$$

Where  $u_i = u(x_i)$ .  $x = ih$  on substitution of (7) in (6) the new equation is

$$-\left(\varepsilon + \frac{ih^2}{2}\right) \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} + ih \frac{u_{i+1} - u_{i-1}}{2h} + u_i = 1 \quad (8)$$

The final transformed difference scheme is

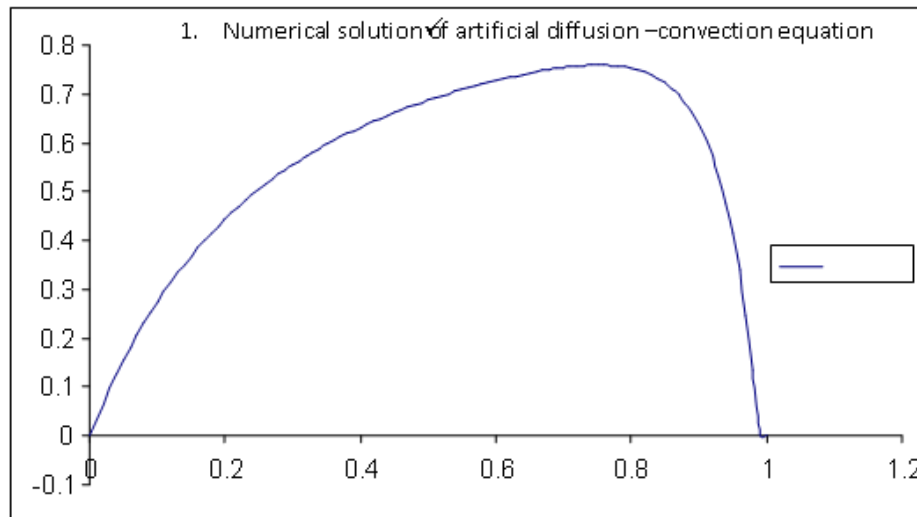
$$a_i u_{i+1} + b_i u_i + c_i u_{i-1} = d_i \quad (9)$$

Where  $a_i = -\epsilon$ ,  $b_i = 2\epsilon + h^2(1+i)$ ,  $c_i = -(\epsilon + i h^2)$ ,  $d_i = h^2$

The boundary conditions  $u(0) = u(1) = 0$  are represented by  $u_0 = 0$ ,  $u_N = 0$  Equation (9) represents a Tri-diagonal Matrix of the form

$$A\vec{u} = \vec{D}$$

Where the co-efficient matrix A is of order n-1. The Non-homogeneous linear system is solved by applying Thomas algorithm. Here The Co-efficient matrix is a Monotonic matrix. This concept incorporated in this problem reduces the variations in the computed solution. The computed result with corresponding graph is shown below.



**Figure-3:** Numerical solution of the Artificial-diffusion equation.

**Table-I: ( $\epsilon = 0.05$ )**

x	u	x	u	x	u	x	u
0	0	0.26	0.5157	0.51	0.6919	0.76	0.7591
0.01	0.0345	0.27	0.5261	0.52	0.6964	0.77	0.7585
0.02	0.0672	0.28	0.5361	0.53	0.7008	0.78	0.7574
0.03	0.0982	0.29	0.5457	0.54	0.705	0.79	0.7556
0.04	0.1275	0.3	0.555	0.55	0.7091	0.8	0.753
0.05	0.1553	0.31	0.564	0.56	0.7131	0.81	0.7495
0.06	0.1817	0.32	0.5726	0.57	0.7169	0.82	0.745
0.07	0.2068	0.33	0.5809	0.58	0.7206	0.83	0.7392
0.08	0.2307	0.34	0.589	0.59	0.7242	0.84	0.7319
0.09	0.2534	0.35	0.5968	0.6	0.7276	0.85	0.7229
0.1	0.275	0.36	0.6043	0.61	0.731	0.86	0.7118
0.11	0.2956	0.37	0.6115	0.62	0.7341	0.87	0.6982
0.12	0.3152	0.38	0.6185	0.63	0.7342	0.88	0.6816
0.13	0.3339	0.39	0.6253	0.64	0.7401	0.89	0.6615
0.14	0.3518	0.4	0.6319	0.65	0.7428	0.9	0.6371
0.15	0.3689	0.41	0.6383	0.66	0.7454	0.91	0.6076
0.16	0.3853	0.42	0.6444	0.67	0.7479	0.92	0.5719
0.17	0.4009	0.43	0.6504	0.68	0.7501	0.93	0.5288
0.18	0.4159	0.44	0.6562	0.69	0.7522	0.94	0.4768
0.19	0.4302	0.45	0.6618	0.7	0.754	0.95	0.4139
0.2	0.444	0.46	0.6672	0.71	0.7556	0.96	0.338
0.21	0.4572	0.47	0.6724	0.72	0.757	0.97	0.2461
0.22	0.4698	0.48	0.6775	0.73	0.758	0.98	0.1348
0.23	0.482	0.49	0.6825	0.74	0.7588	0.99	0.0001
0.24	0.4937	0.5	0.6873	0.75	0.7591	1	0
0.25	0.5049						

### 3. OBSERVATIONS

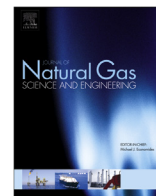
It has been observed that the graphs shown in Fig(1), Fig(2), Fig(3) maintain character preserving phenomena over (0,1). Especially in the interval of smooth region steep down fall of the graph coinciding with the actual solution is an appreciable thing of considerable order. For small  $\varepsilon$  the equation is dominated by the convection term. The boundary or interior layers may appear along downstream of the convection direction i.e., after the smooth region the diffusion effect is visible in the interval  $(\delta, 1)$ . Stable solution is observed under the influence of the artificial-diffusion. The exact solution is non-zero almost everywhere except in a narrow boundary layer sub-interval very close to the point  $x=1$ . The numerically computed values of  $u$  also support this statement vide Table-I. The computed solution and the series solution exhibit good agreement on the convection-diffusion phenomena almost through out the the region. Whenever there is very little diffusion then the solution has varying nature as compared to the exact solution [1]. When diffusion is more (Artificial diffusion), then the computed layers are smeared.

### 4. REFERENCES

1. D. N.d.G. Allen and R.V Southwell (1955), 'Relaxation methods applied to determine the motion inn two dimensions, of a viscous fluid past a fixed cylinder', Quart. J. Mech. Appl. Math. 8, 129-145.
2. V.B. Andreev and N.V. Kopteva (1996), 'Investigation of difference schemes with an approximation of the first derivative by a central difference relation', Zh. Vychisl. Mat.i Mat. Fiz. 36(8), 101-117.
3. Arthur E.P. Veldman, Ka-Wing Lam 'Symmetry-preserving upwind discretization of convection on non-uniform grids., Applied Numerical Mathematics 58 (2008).
4. A. Brandt and I. Yavneh(1991), 'Inadequacy of first-order upwind difference schemes for some recirculat - ing flow', J. Comput. Phys. 93, 128-143.
5. M. Dobrowolski and H.-G. Roos (1997), 'A priori estimates for the solution of Convection-diffusion problems and interpolation on Shishkin meshes', Z. Anal. Anwendungen 16, 1001-1012.
6. K.Eriksson, D. Estep, P.Hansbo and C. Johnson (1996), Computational Differential Equations, Cambridge University Press, Cambridge.
7. A. M. Il'in (1969), 'A difference scheme for a differential equation with a small parameter multiplying the highest derivative', Mat. Zametki 6, 237-248. equations', Comput. Methods Appl. Mech. Engrg. 190, 757-781.
8. Introduction to singular Perturbation problems by Robert E.O's Malley, Jr, Academic press.
9. Martin Stynes (2005) 'Steady-state convection-diffusion problems', Acta Numerica (2005), pp. 445-508, Cambridge University Press.
10. K. W. Morton(1996), Numerical solution of Convection-Diffusion Problems, Vol. 12 of Applied Mathematics and Mathematical Computation, Chapman & Hall, London.
11. Mikhail Shashkov (2005) 'Conservative finite difference methods on General grids, CRS Press (Tokyo).
12. Dennis G. Roddeman, Some aspects of artificial diffusion in flow analysis, TNO Building and Construction Research, Netherlands.
13. G.D. Smith, 'Partial Differential equations', Oxford Press.
14. N. Srinivasacharyulu, K. Sharath babu (2008), 'Computational method to solve steady-state convection-diffusion problem' , International Journal of Mathematics, Computer Sciences and Information Technology, Vol. 1 No. 1-2, January-December 2008, pp.245-254.
15. M. Stynes and L. Tobiska (1998), 'A finite difference analysis of a streamline diffusion method on a Shishkin meshes', Numer. Algorithms 18, 337-360.

**Source of support: Nil, Conflict of interest: None Declared**

**[Copy right © 2015. This is an Open Access article distributed under the terms of the International Journal of Mathematical Archive (IJMA), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.]**



# Preservation of methane gas in the form of hydrates: Use of mixed hydrates



Pinnelli S.R. Prasad <sup>a,\*</sup>, Vangala Dhanunjana Chari <sup>b</sup>

<sup>a</sup> Gas Hydrate Division, CSIR–National Geophysical Research Institute (CSIR–NGRI), Hyderabad, 500007, India

<sup>b</sup> Mallareddy Engineering College, Dhulapally, Secunderabad, 506100, India

## ARTICLE INFO

### Article history:

Received 11 March 2015

Received in revised form

21 April 2015

Accepted 22 April 2015

Available online

### Keywords:

Anomalous preservation

Methane storage

Mixed hydrates

Propane

Tetrahydrofuran

Phase boundary

## ABSTRACT

The release of methane gas was compared in pure CH<sub>4</sub> (sl) and mixed (sII) hydrates (with C<sub>4</sub>H<sub>8</sub>O – tetrahydrofuran and C<sub>3</sub>H<sub>8</sub> – propane) having methane as dominant constituent. We report absence of the self preservation effect in mixed hydrates, having stoichiometric composition (sII) of larger guest molecules, and they populate the 5<sup>12</sup>6<sup>4</sup> cages. Their dissociation behaviour is in accordance with the respective phase boundary curve. While a partial methane gas release was observed at 270 K, particularly, in dilute systems. Further, excess gas release was at T > 270 K and complete methane release is governed by the thermodynamic stability of mixed hydrates.

© 2015 Published by Elsevier B.V.

## 1. Introduction

Adaptability of gas hydrate based technology for natural gas storage and transportation demands a stringent pressure and temperature conditions for its stability. Its transportation through a network of pipelines is particularly useful for long haul, while it becomes uneconomical for mid or short distances (Gudmundson and Borrehaug, 1996). The liquefied natural gas (LNG) or compressed natural gas (CNG) transportation methods, although, preferentially used in these sectors have inherent operational difficulties, such as need of cryo-temperature and high pressures respectively (Thomas and Dawe, 2003). Thus, these techniques require skilled man-power for sustained and safer operations. Adoption of gas hydrate based methodology in this sector is attractive, but the limitations are mostly technology driven (Nakoryakov and Misyura, 2013; Mimachi et al., 2014). The hydrate conversion is quite inefficient and time consuming without an agitator and therefore some special reactors are being employed. Another approach is to add of some kinetic promoters or surfactants to the hydrate forming (gas + water) system; and such experiments are still at laboratory scale (Ganji et al., 2007).

Additionally, usage of certain porous materials have shown some attractive features in hydrate conversion process (Prasad et al., 2014; Chari et al., 2013a, b; Kang et al., 2010, 2013; Linga et al., 2007).

Clathrate hydrates, or gas hydrates, are the crystalline ice-like inclusion compounds consisting of hydrogen bonded water molecules forming open cages of different sizes. Four essential conditions often required for its stable existence are (i) enough supply of guest (suitably sized hydrocarbons or other) and (ii) host (water) molecules; and simultaneous presence of (iii) moderately high pressure, (iv) lower temperature conditions. Such stringent requirements are often fulfilled in certain geological locations around the globe, both under the permafrost and the oceanic sediments. Typical pressure and temperatures for the methane hydrates stability are 2.5 MPa and 273 K; and required pressure increases exponentially at higher temperatures (Sloan and Koh, 2008). On the other hand gas hydrates encasing methane molecules (dominant constituent in natural gas hydrates) can be preserved for a longer duration even at ambient pressure in the temperature window 240–270 K, and this is most popularly known as “anomalous (or self-) preservation effect” (Stern et al., 2001, 2003). The metastable nature of methane hydrates (MH) has been a subject of immense interest; however, a precise molecular mechanism responsible for this effect has not emerged. For example, is it an exclusive property

\* Corresponding author.

E-mail address: [psrprasad@ngri.res.in](mailto:psrprasad@ngri.res.in) (P.S.R. Prasad).

for methane hydrates (sl - *Pm3n*) alone or is it true for all other sl or sII (*Fd3m*) hydrates also. Will the mixed hydrates (sII), with CH<sub>4</sub> as dominant fraction, possess this exceptional property? How can one model the mechanism for this unique property without ambiguities? Nevertheless unexpected longevity of the methane hydrates at ambient pressure found important applications, particularly, in gas storage and transportation of natural gas hydrates, which are dominated by the methane gas (Takeya et al., 2012; Mimachi et al., 2014). There are several factors that control the rate of hydrate dissociation, such as experimental temperature (Stern et al., 2001, 2003) and pressure (Komai et al., 2004). Further, composition of the feed gas (Stern et al., 2003; Takeya and Ripmeester, 2010; Uchida et al., 2011), hydrate granular size (Takeya et al., 2005; Mimachi et al., 2014) and co-existence of certain additives or electrolytes (Zhang and Rogers, 2008; Takeya et al., 2012; Sato et al., 2013), and presence of fine glass beads (Hachikubo et al., 2011) in hydrate forming systems also influence the dissociation process. The effect of temperature on the dissociation process is exhaustively studied in the past by Stern et al. (2001) and they reported significantly slower at atmospheric pressure in a temperature window 241–271 K (generally known as self-preservation window), with two minima at 249 and 269 K. Nakoryakov and Misyura (2013) have reported that the thermal dissociation rate is significantly lower in natural hydrates compare to synthetic analogues in self preservation window despite of structural similarity.

Takeya and Ripmeester (2010) have shown that the self-preservation property is not an exclusive for MH alone, but the hydrates with certain other guest molecules like CH<sub>3</sub>F, CF<sub>4</sub> and CO<sub>2</sub> forming sl structure and O<sub>2</sub>, N<sub>2</sub>, Ar and Kr forming sII can also show this phenomenon. These authors further argued that the interactions between the guest molecules and H<sub>2</sub>O molecules in ice play a critical role in self-preservation phenomenon. Uchida et al. (2011a, b) have summarized earlier kinetic models on this unusual self-preservation phenomena and have concluded that the ice shield, around hydrate grains, below 240 K is mostly granular type and the retardation in gas diffusion is governed by slower hydrate dissociation rather than gas diffusion. However, in the self-preservation temperature region, i.e., 240 K < T < 270 K, the morphology of ice shield is plate-like with stacking faults and hydrate dissociation is governed by the molecular gas diffusion through the ice layers. Further, the ice layer also stabilizes the hydrate structure vindicating the importance of gas-H<sub>2</sub>O interactions in this phenomenon. Takeya et al. (2011) have measured the average thickness of ice layer around hydrate particles as 100 µm after preserving them for 24 h at 253 K. Although it is difficult to understand the effect of gas-H<sub>2</sub>O interactions on the self-preservation of gas hydrates quantitatively, they become very important when we consider the natural gas hydrates as storage and transportation materials. Zhang and Rogers (2008) claimed that the 0.04% of the stored gas was evolved during 256 h at 268 K from the hydrates formed with gas mixture of (CH<sub>4</sub> + C<sub>2</sub>H<sub>6</sub> + C<sub>3</sub>H<sub>8</sub>) at atmospheric pressure. According to these authors, ice shielding is not the primary mechanism for this phenomena and the enhancement of preservation by usage of additives may be of practical possibility. Recently Kida et al. (2011) proposed that the “direct measurements of the dissociation behaviours of pure methane and ethane hydrates trapped in sintered tetrahydrofuran hydrate through a temperature ramping method showed that the tetrahydrofuran hydrate controls dissociation of the gas hydrates under thermodynamic instability at temperatures above the melting point of ice”. The sample preparation procedure adopted by Kida et al. (2011) involved several steps, such as mixing-up of fine powders of independently prepared MH and THF hydrates at low temperatures. Later they were pelletized using oil pressure of 6 MPa and 263 K. Further prior to depressurization experiments,

the pellets were again broken into 1–2 mm chips and were soaked for about 30 min under methane gas pressure of 6 MPa and 263 K. Thus, the possibility of mixture of both sl (MH) and sII (THF + CH<sub>4</sub>) hydrates cannot be excluded in such experiments. Thus it was not clear whether the delayed dissociation of hydrates was due to mixed hydrates or due to anomalous preservation nature of sl hydrates. In order to gain further understanding we compared the dissociation behaviour of methane hydrates (MH), after depressurization, in its pure (sl) and mixed (with THF – 0.052, 0.021 & 0.011 and C<sub>3</sub>H<sub>8</sub> gas – 0.13, 0.051 & 0.017 mol fractions) form. The THF is highly soluble in water and readily form sII hydrates, while propane is less soluble and also form sII hydrates because of its larger molecular size.

## 2. Experimental

Experimental procedure followed for gas hydrate synthesis has already been described earlier (Chari et al., 2011, 2013a; Sharma et al., 2014). Briefly, the main part was an SS-316 cylindrical vessel, which can withstand gas pressures up to 15 MPa, and volume of the vessel was 400 mL. A cold fluid (water + glycol mixture) was circulated around the vessel with the help of a circulator to bring and maintain temperature inside the cell at a desired level. A platinum resistance thermometer (Pt100) was inserted into the vessel to measure temperature with an accuracy of ±0.5 K, while pressure in the vessel was measured with a pressure transducer (WIKA, type A-10 for pressure range 0–25 MPa with ±0.5% accuracy). De-ionized ultra-pure water (Millipore – type 1) was used, and the dissolved gases were removed by evacuation. The High pure gasses were supplied by M/S Linde India Private Ltd. The tetrahydrofuran of purity (98%) was supplied by Qualigens Fine Chemicals, India.

About 70% volume was for CH<sub>4</sub>/C<sub>3</sub>H<sub>8</sub>+CH<sub>4</sub> gas and rest was filled with water or THF aqueous solution. The atmospheric gases in the experimental cell were diluted by purging with the gas prior to the experiments and the gas was filled to the desired level (5.0–8.0 MPa) using the Teledyne ISCO syringe pump to a pressure and temperature outside of the hydrate stability zone. Then, the reactor was isolated from the ISCO pump/gas tank by closing the gas inlet valve. Subsequently, a cold fluid from the chiller was circulated to bring down the temperature of the reactor and the hydrate formation was detected by a sharp pressure drop at a particular temperature (>275 K). The temperature and pressure were logged for every 60 s of the time interval. Insignificant head-pressure drop in the reactor observed over a longer duration indicates saturation in hydrate conversion. The reactor cell was then placed in a bath pre-cooled to 250 K for equilibration. Then the reactor was depressurized rapidly by venting the residual gas. The bath temperature was slowly increased and *p*, *T* of the reactor was recorded with 30 s interval. Similar experimental protocol has been used by Uchida et al. (2011a).

## 3. Results and discussion

Table 1, shows the mole fractions of feed guest molecules and the amount of gas liberated during the dissociation of gas hydrates. The gas release below 273 K is mostly due to the dissociation of sl component, while that above is because of sII hydrate dissociation. The hydration number for pure methane hydrates has been computed as 5.94 (Chari et al., 2014) and that for C<sub>3</sub>H<sub>8</sub> + CH<sub>4</sub> mixtures is estimated from CSMGEM program (Sloan and Koh, 2008). The hydration number for THF + CH<sub>4</sub> hydrates was estimated using the cage occupancy reported by Seo et al. (2003). Estimated hydration numbers for the mixed guests vary in the range 7.80–6.72 depending on the feed guest composition. Typical

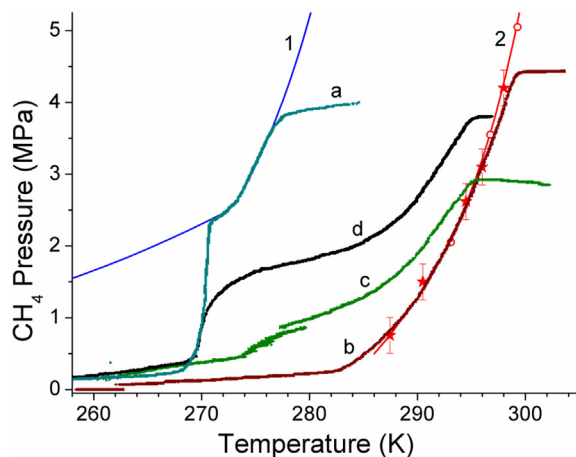
**Table 1**

Information on mole fractions of different guest molecules used as feed constituents in the present study. Observed amounts of released guests below and above the ice melting point were also tabulated.

Exp. No	Mole fractions of the feed guest molecules			Released gas (moles)		
	THF	C <sub>3</sub> H <sub>8</sub>	CH <sub>4</sub>	sl	sII	Total
1	—	—	1	0.51	—	0.51
2	0.011	—	1	0.15	0.24	0.39
3	0.021	—	1	—	0.28	0.28
4	0.052	—	1	—	0.42	0.42
5	—	0.017	0.982	0.25	0.08	0.33
6	—	0.039	0.960	0.16	0.11	0.27
7	—	0.129	0.870	0.07	0.27	0.34

hydrate yield in pure methane hydrates is 54.5%. The procedure for estimating the hydrate yields through the gas release measurement is described in detailed in our earlier publications (Chari et al., 2011, 2013a; Sharma et al., 2014; Prasad et al., 2014). There was no measurable gas release at  $T < 273$  K in hydrates formed with  $X_{\text{THF}} = 0.052$  and 0.021; and the hydrates are in sII form, with yield of 39 and 24% respectively. However, at much lower mole fractions of THF ( $X_{\text{THF}} = 0.011$ ) the methane gas release was observed as 0.15 and 0.24 mol in both the regions, respectively. Corresponding hydrate yields in sl & sII regions are 11.0 and 20.0%, respectively. Similarly the hydrate yields in mixed hydrates with C<sub>3</sub>H<sub>8</sub> + CH<sub>4</sub> system (having mole fraction as C<sub>3</sub>H<sub>8</sub> as 0.129, 0.039 and 0.017) during the sl & sII hydrates formation region are 7.5 & 38%, 17.0 & 13.87% and 26.7 & 9.68%. Generally the hydrate fraction in sII decreased with the decrease in large cage guest fraction.

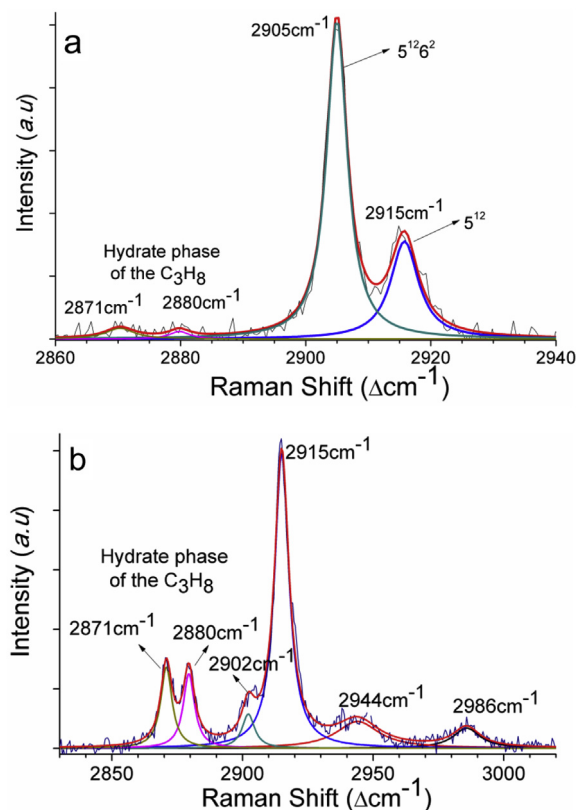
Fig. 1 shows the progressive build-up in methane gas pressure, after depressurization around 253 K, at different temperatures for pure and mixed hydrates. Mixed hydrates were synthesised using aqueous solutions of THF with mole fractions (0.052, 0.021 & 0.011). We observe that the pure methane hydrates continues to exist in metastable state upon depressurization and rapid gas release occurred in the close vicinity of ice melting temperature (273 K) (curve a). The gas release is extremely slow when the temperature is lower than 268 K. However, methane gas was released rapidly in the temperature window of 268–271 K. Pressure inside the reactor



**Fig. 1.** A plot showing the variation in CH<sub>4</sub> gas pressure in gas hydrate systems as a function of temperature. Curve (a) depicts the hydrate dissociation behaviour in pure methane hydrate system and the trace 1 is its phase boundary computed from CSMGEM. The curves (b, c & d) are the dissociation behaviour observed in mixed hydrates of CH<sub>4</sub> + THF system, with  $X_{\text{THF}} = 0.052$ , 0.021 and 0.011. Trace 2 is the phase boundary curve for the mixed hydrates (with  $X_{\text{THF}} = 0.05$ ) fitted to the experimental points (\* & O).

reached to 2.35 MPa at 270.8 K, which is evidently a phase boundary point for pure methane hydrates. Thereafter the gas release by hydrate dissociation is essentially driven by the stability conditions, i.e., significant methane release is observed upon increasing the temperature (see curve – a of Fig. 1). Well accepted reason for this metastable nature outside MH region is the formation of an ice layer over MH grains (Takeya, 2010, 2011; Uchida, 2011b; Ohno, 2011).

On the other hand, mixed hydrates with some other hydrates formers such as THF or C<sub>3</sub>H<sub>8</sub> etc., show much superior thermodynamic stability, though they preferentially crystallize into sII hydrates. Larger guest molecules occupy the 5<sup>12</sup>6<sup>4</sup> cages, whereas, vacant smaller 5<sup>12</sup> cages are available for the guest molecules such as CH<sub>4</sub>. The Raman spectroscopic signatures for the mixed hydrates were shown earlier for THF (Prasad et al., 2009b) and the same for propane hydrates are shown in Fig. 2. The mixed hydrates are thermodynamically stable even at lower pressures at temperatures closer to 273 K. The thermal stability of THF + CH<sub>4</sub> mixed hydrates system has extensively studied at high  $p$ ,  $T$  conditions, but such studies in anomalous preservation region are at seldom. Observed pressure raise, if any, in THF + CH<sub>4</sub> hydrate system is primarily due to CH<sub>4</sub> gas from dissociated hydrates. Release of THF due to hydrate dissociation will not contribute to pressure increase as it is in liquid form in the temperature regions of our interest (Kida et al., 2011). Fig. 1 shows the dissociation behaviour of mixed hydrates at three different mole fractions of THF, namely 0.052, 0.021 and 0.011. The circles and star symbols ( $X_{\text{THF}} = 0.052$ ), respectively, are the experimental phase stability points reported in the literature and determined in our laboratory using isochoric pressure search method (Chari et al., 2011; Sharma et al., 2014). We computed the



**Fig. 2.** Characteristic Raman bands in the CH stretching mode region denoting the encaged C<sub>3</sub>H<sub>8</sub> and CH<sub>4</sub> molecules in sl (trace a) and sII (trace b) hydrates. The sl and sII hydrates were synthesized with 2 and 13% propane gas.

best fit by considering data points particularly in lower  $p$ ,  $T$  region. It is also known from the literature that phase boundary for mixed hydrates with lower mole fractions of promoters such as, THF, moves to the left of curve with  $X_{\text{THF}} = 0.0556$  towards inhibition side. Further, increase beyond  $X_{\text{THF}} = 0.0556$  will have no effect on the phase stability but the excess amount inhibits the hydrate growth, thus decreases gas intake (Seo et al., 2003; Prasad et al., 2009a). Significant built-up in  $\text{CH}_4$  pressure in mixed hydrates, with  $X_{\text{THF}} = 0.052$ , is evident at  $T > 282$  K and it is negligibly small at  $T < 282$  K. The dissociation is along the phase boundary (see curve-b). The hydrate dissociation is faster upon reducing the mole fraction of THF to 0.021 and 0.011 as shown in curves c & d respectively. But noticeable dissociation occurred at  $T > 274$  K. Observed behaviour is consistent with earlier reports. It is known that the mixed hydrates crystallize into both sII and sI and its growth is highly inhomogeneous and particularly the hydrates formed from lower THF mole fractions ( $X_{\text{THF}} = 0.011$ ) show prominent sI growth (Seo et al., 2003; Chari et al., 2011; Prasad et al., 2009b). A closer look into the dissociation pattern of curve-d clearly vindicates this fact. A significant amount of methane gas was liberated close to ice-melting point, similar to curve-a, and thereafter gas release was slow. The later part of gas release (for  $T > 270$  K) is due to the dissociation of mixed hydrates of sII with  $5^{12}6^4$  and  $5^{12}$  cages respectively are filled by THF and  $\text{CH}_4$  molecules. Our earlier observations clearly demonstrated that a heterogeneous mixture of sII and sI are often formed in aqueous solutions with lower mole fractions of bulky hydrocarbons (Chari et al., 2011; Prasad et al., 2009a). Prolonged release of methane gas beyond ice-melting point is essentially because of mixture of sI and sII hydrates. Nevertheless, these experiments demonstrate the possibility to store methane gas in mixed hydrates at temperatures higher than 273 K. The dissociation pattern follows the respective phase boundary. Further, these observations are strengthened by the dissociation behaviour of hydrates formed with  $\text{C}_3\text{H}_8$  and  $\text{CH}_4$  gas mixtures. We particularly choose  $\text{C}_3\text{H}_8$  and  $\text{CH}_4$  gas mixtures because both of them are immiscible in water, whereas THF is highly miscible in water; and the propane is essentially occupies  $5^{12}6^4$  cages in sII hydrates.

The hydrates formed with gas mixtures like propane and methane are highly heterogeneous and both sII & sI structures are reported. However, sII hydrates are more prominent with  $\text{C}_3\text{H}_8$  occupying  $5^{12}6^4$  cages, while vacant  $5^{12}$  cages are occupied by  $\text{CH}_4$  molecules (Uchida et al., 2004; Schicks et al., 2006; Prasad et al., 2008). As noted in Table 1, the fraction of sI hydrates increases with the mole fraction of  $\text{CH}_4$  in feed gas. Further, in Fig. 2, we show the Raman spectroscopic signatures of  $\text{C}_3\text{H}_8$  and  $\text{CH}_4$  molecules enclathrated in  $5^{12}6^4$  and  $5^{12}$  cages respectively (trace – b). The characteristic Raman features for the propane molecules in hydrate phase are a doublet at  $2871$  &  $2880\text{ cm}^{-1}$  and weaker bands at  $2944$  &  $2986\text{ cm}^{-1}$ . These observations are consistent with earlier reports (Uchida et al., 2004; Schicks et al., 2006; Prasad et al., 2008, 2009a). Observed Raman band at  $2915\text{ cm}^{-1}$  is the characteristic of  $\text{CH}_4$  molecules encaged in  $5^{12}$  cages. Whereas, with less mole fraction of propane in feed gas mixture the prominent hydrate phase is sI (trace – a), with both cages are occupied by the methane molecules, and observed Raman doublet at  $2905$  and  $2915\text{ cm}^{-1}$  unambiguously indicates such occupancy.

Fig. 3 shows the dissociation behaviour of hydrates with increasing temperature. The hydrates were formed with different gas mixtures; i.e., 13%  $\text{C}_3\text{H}_8$  + 87%  $\text{CH}_4$  (curve – b); 4%  $\text{C}_3\text{H}_8$  + 96%  $\text{CH}_4$  (curve – c) and 2%  $\text{C}_3\text{H}_8$  + 98%  $\text{CH}_4$  (curve – d). A small amount of gas release is observed around 270 K in hydrates formed with higher amount of propane gas (curve – b), while it is relatively high in hydrates with 4% or 2% propane gas. This indicates that the sI hydrates could form prominently in mixed hydrates with lower

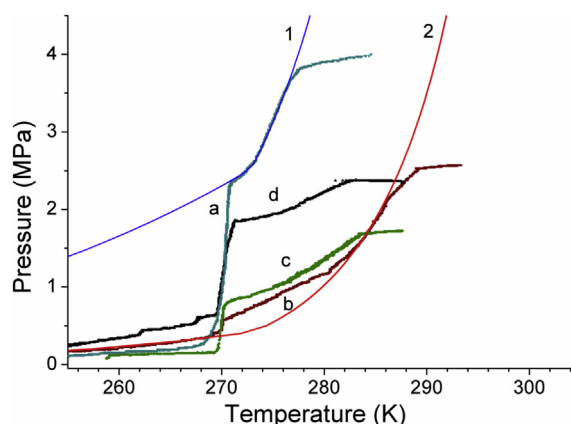


Fig. 3. Observed variations in gas pressure as a function of temperature upon dissociating the pure  $\text{CH}_4$  and mixed  $\text{CH}_4$  +  $\text{C}_3\text{H}_8$  hydrates. The trace 1 and 2 represents the phase boundary curve computed from CSMGEM. Curves (b, c & d) are the dissociation behavior observed in mixed hydrates of  $\text{CH}_4$  +  $\text{C}_3\text{H}_8$  system, formed with 13, 4 and 2% propane gas.

amounts of propane. Further, the dissociation of sII hydrates with  $\text{CH}_4$  as co-guest will occur at temperatures higher than 273 K. The partial dissociation of mixed hydrates formed with lower mole fractions of higher hydrocarbons ( $\text{C}_3\text{H}_8$  or THF) could be because of their lower thermal stability limit or vacant  $5^{12}6^4$  could provide a path for  $\text{CH}_4$  molecular diffusion through sII hydrates.

#### 4. Conclusions

In summary, our experimental results show that pure methane hydrates exist in metastable state in a temperature range 250–268 K upon depressurization to atmospheric pressure. The hydrate dissociation is rapid in a smaller temperature range, i.e., 268–270.8 K. When the cumulative gas pressure exceeds 2.35 MPa, further hydrate dissociation is governed by the thermodynamic stability conditions of methane hydrates. On the other hand, it is possible to store  $\text{CH}_4$  gas in the form of mixed hydrates and the release of methane gas is controlled by the thermodynamic stability of mixed (sII) hydrates. The formation of sI hydrates with  $\text{CH}_4$  as guest molecules alone is avoidable by using other hydrate formers having stoichiometry close to sII hydrates. However, the fraction of sI hydrates increased at the lower mole fraction of large cage guest molecules.

#### Acknowledgements

Author sincerely thanks the Director of the National Geophysical Research Institute, Hyderabad, for his encouragement, and permission to publish this paper. Partial financial support from DST (India) and DGH-NGHP (India) is acknowledged. This is a contribution to GEOSCAPE Project of NGRI under the 12th Five Year Scientific Program of CSIR.

#### References

- Chari, V.D., Prasad, P.S.R., Murthy, S.R., 2014. Structural stability of methane hydrates in porous medium: Raman spectroscopic study. *Spectrochim. Acta A*, 120, 636–641.
- Chari, V.D., Raju, B., Prasad, P.S.R., Rao, D.N., 2013b. Methane hydrates in spherical silica matrix: optimization of capillary water. *Energy&Fuels*, 27, 3679–3684.
- Chari, V.D., Sharma, D.V.S.G.K., Prasad, P.S.R., 2011. Methane hydrate phase stability with lower mole fractions of tetrahydro-furan (THF) and *tert*-butylamine (*t*-BuNH 2). *Fluid Phase Equilib.*, 315, 126–130.
- Chari, V.D., Sharma, D.V.S.G.K., Prasad, P.S.R., Murthy, S.R., 2013a. Methane hydrates formation and dissociation in nano silica suspension. *J. Nat. Gas. Sci. Eng.*, 11,

- 7–11.
- Ganji, H., Manteghian, M., Sadaghiani zadeh, K., Omidkhaha, M.R., Mofrad, H.R., 2007. Effect of different surfactants on methane hydrate formation rate, stability and storage capacity. *Fuel* 86, 434–441.
- Gudmundson, J., Borrehaug, A., 1996. Frozen hydrate for transport of natural gas. In: *Proc. Second Intl. Conf. Natural Gas Hydrates*, pp. 415–422.
- Hachikubo, A., Takeya, S., Chuvilin, E., Istomin, V., 2011. Preservation phenomena of methane hydrate in pore spaces. *Phys. Chem. Chem. Phys.* 13, 17449–17452.
- Kang, S.P., Lee, J., Seo, Y., 2013. Pre-combustion capture of CO<sub>2</sub> by gas hydrate formation in silica gel pore structure. *Chem. Eng. J.* 218, 126–132.
- Kang, S.P., Lee, J.W., 2010. Formation characteristics of synthesized natural gas hydrates in meso- and macroporous silica gels. *J. Phys. Chem. B* 14, 6973–6978.
- Kida, M., Jin, Y., Narita, H., Nagao, J., 2011. Effective control of gas hydrate dissociation above the melting point of ice. *Phys. Chem. Chem. Phys.* 13, 18481–18484.
- Komai, T., Kang, S.P., Yoon, J.H., Yamamoto, Y., Kawamura, T., Ohtake, M., 2004. In-situ Raman spectroscopy investigation of the dissociation of methane hydrate at temperatures just below the ice point. *J. Phys. Chem. B* 108, 8062–8068.
- Linga, P., Kumar, R., Englezos, P., 2007. Gas hydrate formation from hydrogen/carbon dioxide and nitrogen/carbon dioxide gas mixtures. *Chem. Eng. Sci.* 62, 4268–4276.
- Mimachi, H., Takeya, S., Yoneyama, A., Hyodo, K., Takeda, T., Gotoh, Y., Murayama, T., 2014. Natural gas storage and transportation within gas hydrate of smaller particle: size dependence of self-preservation phenomenon of natural gas hydrate. *Chem. Eng. Sci.* 118, 208–213.
- Nakoryakov, V.E., Misyura, S.Ya., 2013. The features of self-preservation for hydrate systems with methane. *Chem. Eng. Sci.* 104, 1–9.
- Ohno, H., Narita, H., Nagao, J., 2011. Different modes of gas hydrate dissociation to ice observed by microfocus x-ray computed tomography. *J. Phys. Chem. Lett.* 2, 201–205.
- Prasad, P.S.R., Prasad, K.S., Sowjanya, Y., Sain, K., 2008. Laser micro Raman investigations on gas hydrates. *Curr. Sci.* 94, 1495–1499.
- Prasad, P.S.R., Sowjanya, Y., Chari, V.D., 2014. Enhancement in methane storage capacity in gas hydrates formed in hollow silica. *J. Phys. Chem. C* 118, 7759–7764.
- Prasad, P.S.R., Sowjanya, Y., Prasad, K.S., 2009b. Micro-Raman investigations of mixed gas hydrates. *Vibra. Spectro.* 50, 319–323.
- Prasad, P.S.R., Sugahara, T., Sloan, E.D., Sum, A.K., Koh, C.A., 2009a. Structural transformations of sVI tert-butylamine hydrates to sII binary hydrates with methane. *J. Phys. Chem. A* 113, 11311–11315.
- Sato, H., Sakamoto, H., Ogino, S., Mimachi, H., Kinoshita, T., Iwasaki, T., Sano, K., Ohgaki, K., 2013. Self-preservation of methane hydrate revealed immediately below the eutectic temperature of the mother electrolyte solution. *Chem. Eng. Sci.* 91, 86–89.
- Schicks, J.M., Naumann, R., Erzinger, J., Hester, K.C., Koh, C.A., Sloan, E.D., 2006. Phase transitions in mixed gas hydrates: experimental observations versus calculated data. *J. Phys. Chem. B* 110, 11468–11474.
- Seo, Y.T., Lee, H., Moudrakovski, I., Ripmeester, J.A., 2003. Phase behaviour and structural characterization of coexisting pure and mixed clathrate hydrates. *Chem. Phys. Chem.* 4, 379–382.
- Sharma, D.V.S.G.K., Sowjanya, Y., Chari, V.D., Prasad, P.S.R., 2014. Methane storage in mixed hydrates with tetrahydrofuran. *Indian J. Chem. Tech.* 21, 114–119.
- Sloan, E.D., Koh, C.A., 2008. *Clathrate Hydrates of Natural Gases*, third ed. CRC Press, Taylor & Francis Group, Boca Raton, FL.
- Stern, L.A., Circone, S., Kirby, S.H., 2001. Anomalous preservation of pure methane hydrate at 1 atm. *J. Phys. Chem. B* 105, 1756–1762.
- Stern, L.A., Cirone, S., Kirby, S.H., Durham, W.B., 2003. Temperature, pressure, and compositional effects on anomalous or “self” preservation of gas hydrates. *Can. J. Phys.* 81, 271–283.
- Takeya, S., Ripmeester, J.A., 2010. Anomalous preservation of CH<sub>4</sub> hydrate and its dependence on the morphology of hexagonal ice. *ChemPhysChem* 11, 70–73.
- Takeya, S., Uchida, T., Nagao, J., Ohmura, R., Shimada, W., Kamata, Y., Ebinuma, T., Narita, H., 2005. Particle size effect of CH<sub>4</sub> hydrate for self-preservation. *Chem. Eng. Sci.* 60, 1383–1387.
- Takeya, S., Yoneyama, A., Ueda, K., Hyodo, K., Takeda, T., Mimachi, H., Takahashi, M., Iwasaki, T., Sano, K., Yamawaki, H., Gotoh, Y., 2011. Nondestructive imaging of anomalously preserved methane clathrate hydrate by phase contrast X-ray imaging. *J. Phys. Chem. C* 115, 16193–16199.
- Takeya, S., Yoneyama, A., Ueda, K., Mimachi, H., Takahashi, M., Sano, K., Hyodo, K., Takeda, T., Gotoh, Y., 2012. Anomalously preserved clathrate hydrate of natural gas in pallet form at 253 K. *J. Phys. Chem. C* 116, 13824–13848.
- Thomas, S., Dawe, R.A., 2003. Review of ways to transport natural gas energy from countries which do not need the gas for domestic use. *Energy* 28, 1461–1477.
- Uchida, T., Moriwaki, M., Takeya, S., Ikeda, I.Y., Ohmura, R., Nagao, J., Minagawa, H., Ebinuma, T., Narita, H., Gohara, K., Mae, S., 2004. Two-step formation of methane–propane mixed gas hydrates in a batch-type reactor. *AIChE J.* 50, 518–523.
- Uchida, T., Kida, M., Nagao, J., 2011a. Dissociation termination of methane–ethane hydrates in temperature-ramping tests at atmospheric pressure below the melting point of ice. *ChemPhysChem* 12, 1652–1656.
- Uchida, T., Sakurai, T., Hondoh, T., 2011b. Ice-Shielding models for self-preservation of gas hydrates. *J. Chem. Chem. Eng.* 5, 691–705.
- Zhang, G., Rogers, R.E., 2008. Ultra-stability of gas hydrates at 1 atm and 268.2 K. *Chem. Eng. Sci.* 63, 2066–2074.



# Influence of copper concentration on sprayed CZTS thin films deposited at high temperature

S.M. Bhosale<sup>a</sup>, M.P. Suryawanshi<sup>a,b</sup>, J.H. Kim<sup>b</sup>, A.V. Moholkar<sup>a,\*</sup>

<sup>a</sup>Thin Film Nanomaterials Laboratory, Department of Physics, Shivaji University, Kolhapur, India

<sup>b</sup>Department of Materials Science and Engineering and Optoelectronics Convergence Research Center, Chonnam National University, 300, Yongbong-Dong, Buk-Gu, Gwangju 500-757, South Korea

Received 1 January 2015; received in revised form 16 February 2015; accepted 23 February 2015

## Abstract

In this study,  $\text{Cu}_2\text{ZnSnS}_4$  (CZTS) thin films were deposited by spray pyrolysis technique at constant substrate temperature. The effects of the copper concentration on the structural, morphological and optical properties of the films were investigated. The copper concentration was varied from 0.15 to 0.25 M in the steps of 0.05 M. The structural studies revealed that the Cu poor film shows low intense peaks, but as Cu concentration increases a relatively more intense and sharper diffraction peaks (112), (200), (220), and (312) of the kesterite crystal structure were observed. Raman spectroscopy analysis confirmed the formation of phase-pure CZTS films. From the morphological studies, it is found that the grain size increased as the Cu concentration increases from 0.15 to 0.25 M. The optical band-gap values were estimated to be 1.61, 1.52 and 1.45 eV for copper concentration 0.15, 0.20 and 0.25 M, respectively. Photoelectrochemical cells using films of different copper concentrations were fabricated and the best cell exhibited an efficiency of 1.09% for 0.25 M of copper concentration.

© 2015 Elsevier Ltd and Techna Group S.r.l. All rights reserved.

**Keywords:** A. Films; B. Grain size; C. Optical properties

## 1. Introduction

Photovoltaic (PV) is the technology that directly converts sunlight into electricity, which is preferred as a good source of renewable energy. Most of the PV cells used today are made up of CdTe or  $\text{CuIn}_{1-x}\text{Ga}_x\text{Se}_2$  (CIGSe) materials [1]. The production of CdTe and CIGSe based PV devices is difficult due to toxicity of selenium (Se) and rare availability of indium (In) and gallium (Ga). The rare availability of these element increases the price of the PV devices, where toxicity is main issue of the human globe. Therefore, there is a necessity to fabricate PVs, which are low cost, non-toxic, high efficiency using environmental friendly materials.  $\text{Cu}_2\text{ZnSnS}_4$  (CZTS) is one of the potential absorber material for thin film heterojunction solar cells with ideal solar cell characteristics such as optimum band gap (1.4–1.5 eV), high absorption

coefficient ( $10^4 \text{ cm}^{-1}$ ) and p-type conductivity [2]. According to Shockley–Queisser photon balanced calculations, the theoretical energy conversion efficiency of PV using CZTS as absorber material is 31–33% [3]. The availability of the CZTS elements; Cu (50 ppm), Zn (75 ppm), Sn (2.2 ppm) and S (260 ppm) which are relatively abundant and nontoxic in nature making CZTS as one of the best candidate for PV technology [4]. Basically CZTS is derived from CIGSe by replacing In with Zn, Ga with Sn, and Se with S [5]. It exists in two types of crystal structures kesterite and stannite only difference between them is the site position of their Cu and Zn atoms. [6]

Many physical [7–10] and chemical methods [7,11–14] are used so far to deposit CZTS films. Physical methods are more favorable than chemical methods due to ease of scalability and compatibility in large scale commercial manufacturing plant [13]. However, suffering from relatively low throughput and material utilization, these costly physical methods face many difficulties in large scale production. In this regard, various chemical methods have been developed due to their advantages of simple handling, low cost and

\*Corresponding author.

E-mail addresses: [avmoholkar@yahoo.co.in](mailto:avmoholkar@yahoo.co.in),  
[avmoholkar@gmail.com](mailto:avmoholkar@gmail.com) (A.V. Moholkar).

high throughput as compared with physical methods. As far as low cost, simplicity, suitable for large scale deposition without use of vacuum, one of the challenges in the synthesis of CZTS thin films by chemical method is the tendency to form binary and ternary phases [14,15].

Spray pyrolysis technique (SPT) being one of the simple and cost effective chemical method has attracted much attention due to its ease of scalability and versatility [14]. Kumar et al. used this method and observed an increase in optical band gap with decrease in copper salt concentration in the solution at substrate temperature of 643 K. [16]. Kamoun et al. deposited CZTS thin films by varying substrate temperature and spray duration. They found that annealing of the films plays an important role for the improvement of the CZTS properties [17]. Nadi et al. have studied effect of different substrate temperature on the CZTS thin films deposited on Molybdenum (Mo) coated soda lime glass (SLG) substrates. They have used a single target rf magnetron sputtering technique and found that annealing process induces better nucleation and grain coalescence upon recrystallization [13]. Camara et al. synthesized CZTS nanoparticles by a simple hydrothermal method with different concentration ratios of sulfur and copper. The thin film has been obtained by spin coating and it was observed that a copper-rich concentration ratio of CZTS nanoparticles show better results [18]. Mkawi et al. fabricated CZTS based solar cells via sequential electrodeposition method using different copper salt concentrations wherein sulfurization is carried out in elemental sulfur vapor atmosphere. They found that there was increase in the grain size with the increase in copper concentration [5].

As discussed above, it is very difficult to obtain phase pure and good morphology CZTS films by SPT. Thus, by considering above aspects there is a need to synthesize compositionally pure CZTS phase with better morphology so as to increase the efficiency of CZTS based solar cells. In this work, it has been attempted to decrease the formation of secondary phases with good morphology. The effect of Cu concentration on the properties of CZTS films at high substrate temperature has been studied and the obtained results are analyzed herewith.

## 2. Experimental

CZTS thin films were deposited on preheated, ultrasonically cleaned soda lime glass (SLG) and fluorine doped tin oxide (FTO) substrates by SPT. The CZTS precursor solution contains copper chloride (0.15–0.30 M), zinc chloride (0.1 M), tin chloride (0.1 M) and thiourea (0.8 M). All the metal chlorides were dissolved in 30 ml of double distilled water. Excess amount of thiourea was used to avoid precipitation in the solution and also to compensate the loss of sulfur during pyrolysis [19]. The solution was stirred on magnetic stirrer for 10 min, to dissolve all components completely. The obtained solution was sprayed at the rate of 2 ml/min on to the substrate kept at 500 °C using compressed air as carrier gas. To study the effect of copper salt concentration on the formation of CZTS films, copper chloride in the solution was varied from 0.15–0.30 M in the steps of 0.05 M. The as deposited samples were allowed to cool naturally at room temperature. These samples were further used for structural, optical, morphological and PEC

properties. The samples are denoted as 'C1' for copper 0.15 M, 'C2' for copper 0.2 M, 'C3' for copper 0.25 M and 'C4' for copper 0.30 M, respectively.

### 2.1. Characterization techniques

The structural properties of sprayed CZTS films with different Cu concentrations are studied using X-ray diffractometer (XRD) (Philips, PW 3710, Holland) operated at 40 kV, 30 mA with CuK $\alpha$  radiation ( $\lambda=1.5406$  Å) and Raman spectroscopy is recorded using Raman microscope (LabRam HR800 UV, Horiba Jobin-Yvon, France) with excitation wavelength of 514 nm. The surface morphology is observed using field emission scanning electron microscopy (FE-SEM) (Model Hitachi S4800, Japan.). UV–visible absorbance spectra of the samples are recorded using a UV–visible spectrophotometer (UV1800, Shimadzu, Japan). The photoelectrochemical (PEC) performance is measured using a Sol2A with Keithley-2420 source meter under 1.5 AM (Oriol New Port Corporation USA).

## 3. Results and discussion

### 3.1. XRD analysis

To study the effect of copper concentration, the crystallographic properties of sprayed CZTS films are characterized by XRD and their patterns are shown in Fig. 1. Increase in copper concentration from 0.15 M to 0.25 M in the spraying solution improves the crystallinity of CZTS films as revealed by the sharp and increasing peak intensity along (112), (002), (114), (220), (312), (008) and (332) respectively, indicating the formation of kesterite type CZTS thin film (JCPDS 26-0575). Films deposited using 0.30 M of copper concentration shows amorphous nature (denoted by C4) with a weak peak at 46.38° (shown by \*) that belong to anilite Cu<sub>7</sub>S<sub>4</sub> (JCPDS 22-0250). This is because of excess of CuCl<sub>2</sub> in the spraying solution. Since C4 was amorphous in nature, it was not considered for

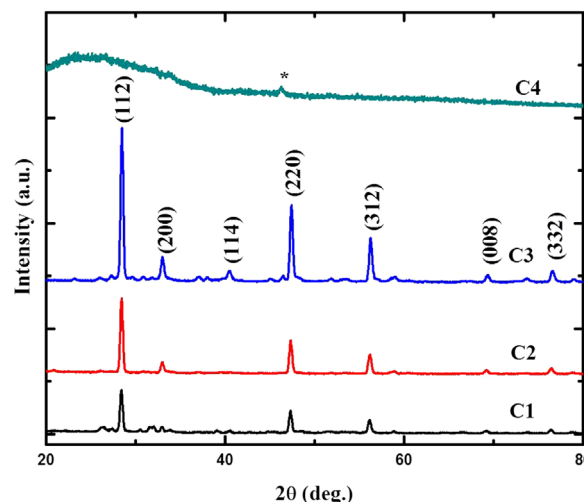


Fig. 1. XRD spectra of CZTS thin films prepared at copper concentration: C1=0.15 M, C2=0.20 M, C3=0.25 M and C4=0.30 M.

Raman, FE-SEM and optical characterization. The sample C3 showed higher peak intensities than the other samples, along with relatively small peak of  $\text{Cu}_7\text{S}_4$  phase. However, as compared to the prominent peaks of CZTS, the  $\text{Cu}_7\text{S}_4$  peak was negligible. Further, it is seen that as the Cu concentration increases from C1 to C3 the intensity of prominent peak increases. Mkawi et al. observed that, intensity of the peaks were moderately increased as Cu concentration increases along with secondary phases. They concluded that the secondary phase assist CZTS film formation through liquid assisted grain growth, resulting in large grains and compact CZTS films [5]. Comparing with their studies, we observed sharp increase in peak intensity as the copper concentration increases. Recent works [16,20] concluded that high substrate temperature is useful to obtain highly crystalline film, this may be reason for above observation. Since, the planes of CZTS are similar to  $\text{Cu}_2\text{SnS}_3$  (CTS) and ZnS, the obtained peaks could be related to either CZTS or CTS or ZnS phases [21]. Thus, to confirm the crystal structure and phase formation of the films, Raman spectra of CZTS samples, were studied.

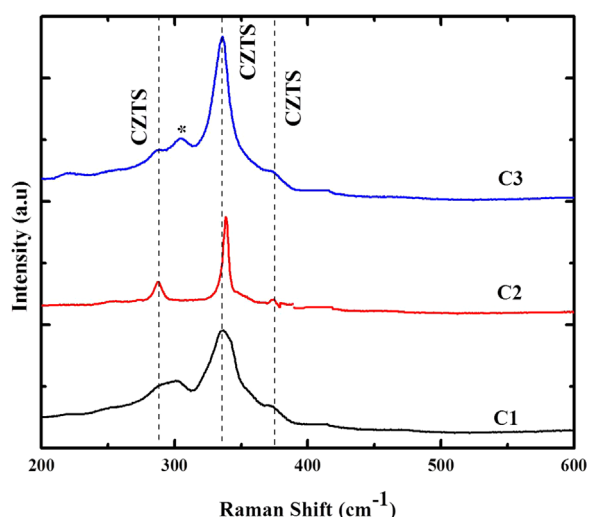


Fig. 2. Raman spectra of CZTS thin films prepared at copper concentration: C1=0.15 M, C2=0.20 M, and C3=0.25 M.

### 3.2. Raman analysis

Each phase in CZTS film gives a specific peak position in Raman scattering which is clearer than in X-ray diffraction pattern. Raman spectroscopy gives the characteristic vibrational modes of bulk CZTS, other secondary and co-existing phases which can be differentiated from each other [22–24]. The Raman spectrum depends on the symmetry of the crystal structure, strength of the chemical bonds, masses and charges of the constituent elements [25]. The Raman spectra of CZTS film is shown in Fig. 2. In the Raman spectra, peaks are observed at 289, 305, 335, and 373  $\text{cm}^{-1}$ . The maximum intensity peak at 335  $\text{cm}^{-1}$  and other low intense peaks observed at 289 and 373  $\text{cm}^{-1}$  corresponds to CZTS, which are in agreement with the reported results and it confirms the formation of CZTS phase. The peaks are shifted little bit towards the right for C1 and C2, which are probably a result of the Cu poor composition of the layer and is associated with the presence of a disordered kesterite phase [26,27]. The weak peak observed at 305  $\text{cm}^{-1}$  in C3 film (denoted by \*) corresponds to  $\text{Cu}_2\text{SnS}_3$ , which may be due to excess of Cu concentration in the sample [23]. Mkawi et al. observed the effect of copper concentration on the properties of CZTS films deposited by electrochemical method and found that the formation of such undesired secondary compounds is due to incomplete conversion of sulfides [28]. The CZTS Raman spectrum did not show any evidence for the presence of binary phases SnS,  $\text{SnS}_2$ , CuS,  $\text{Cu}_2\text{S}$  and  $\beta\text{-ZnS}$  [29]. Previous reports show that the sulfur in the CZTS thin film distributes non-uniformly. If the diffusion of sulfur in the CZTS thin films is from the surface to the bottom of the film, it leads to the formation of secondary phases [30,31]. Here, deposition is carried out at high substrate temperature thus providing sufficient thermal energy for conversion of binary sulfides.

### 3.3. FE-SEM studies

Fig. 3 shows the FE-SEM images of CZTS films deposited with different Cu concentrations. The CZTS micrographs show that the films uniformly cover the substrate. As the Cu concentration increases, an increase in the average grain size as 142 nm for 0.15 M, 205 nm for 0.20 M and 330 nm for 0.25 M is observed.

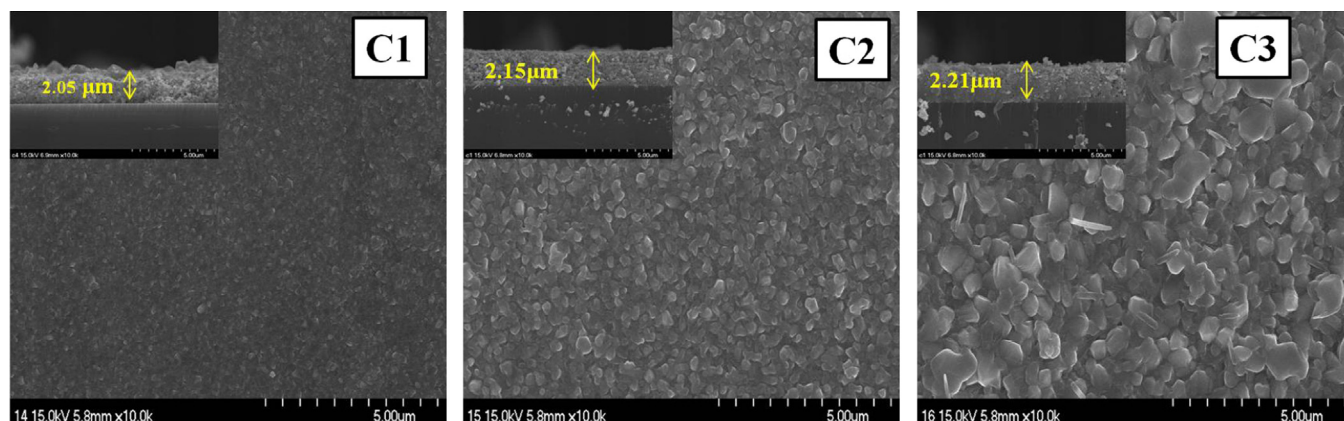


Fig. 3. FE-SEM images of CZTS thin films prepared at copper concentration: C1=0.15 M, C2=0.20 M, and C3=0.25 M. Inset shows the cross-sectional view of CZTS thin films.

This increase in the grain size is might be due to agglomeration of grains. Increase in grain size improves the performance of a polycrystalline CZTS thin film with decrease in grain boundaries that increases the effective diffusion length of minority carriers [32–34]. Kishore Kumar et al. [16] deposited CZTS thin films by varying Cu concentration from 0.007–0.01 M at substrate temperature of 370 °C using SPT. They have observed smeary appearance like morphology with large island-like regions at low copper concentration and as the Cu concentration in the solution increases distinct grains are seen. Chalapathi et al. [27] observed improved grain size after annealing the thermally evaporated CZTS films at 580 °C in sulfur atmosphere. They have observed larger crater like area and a few voids on the surface, which they concluded may be due to insufficient sulfur vapor pressure. Voids on the absorber layer in the thin film solar cells lead to low conversion efficiency because the generated carriers are disturbed at both grids [35]. In this work, compact uniform morphology without voids is observed due to excess amount of sulfur incorporated during the deposition. Thickness of the films is calculated by the cross-section of FE-SEM images and found to be 2.05  $\mu\text{m}$ , 2.15  $\mu\text{m}$  and 2.21  $\mu\text{m}$  for C1, C2 and C3 films, respectively.

### 3.4. Optical studies

Fig. 4 shows a plot of square of the product of the absorption coefficient and photon energy as a function of the photon energy for the spray deposited CZTS films with different Cu concentrations. The nature of the optical transition, whether direct or indirect and the optical band gap ( $E_g$ ) of each film is obtained from the equation

$$\alpha = \frac{A(h\nu - E_g)^n}{h\nu} \quad (1)$$

where 'A' is a constant. The exponent 'n' can take values 1/2, 3/2 or 2 based on whether the optical transition is direct-allowed, direct-forbidden or indirect-allowed, respectively. In the present investigation, values of ' $\alpha$ ' are found to obey above equation for

$n=1/2$ , indicating that the optical transitions are direct-allowed in nature. The band gap is estimated by extrapolating the straight line part of the  $(\alpha h\nu)^2$  versus ' $h\nu$ ' curve to the intercept of the horizontal axis [20]. The values of band-gap for CZTS films are determined to be 1.61 eV, 1.52 eV and 1.45 eV for C1, C2 and C3, respectively. The band gap value of C3 sample is near the optimum for photovoltaic solar conversion in a single-band-gap device. The decrease in band gap with increase in the Cu concentration may be because of: 1] increase in the crystallinity (Fig. 1), 2] the presence of narrow band gap  $\text{Cu}_2\text{SnS}_3$  phase in the C3 sample (Fig. 2) and 3] increase in thickness of the films (inset of Fig. 3) [36–38]. The band gap tunneling may be understood according to the symmetry of the valence band maximum (VBM) and conduction band minimum (CBM) states [39]. It should be noted that the disorder between Cu and Zn cations is highly possible in the CZTS semiconductors with different compositions, and therefore its influence on the band gaps is also common [40].

### 3.5. PEC studies

The PEC performance of CZTS thin films deposited on FTO substrates has been studied with a two electrode configuration of photoanode which were immersed in 0.1 M Europium nitrate ( $\text{Eu(III)(NO}_3)_3$ ) as a redox mediator under UV (100 mW) illumination. ITO was used as a counter electrode. Illuminated  $J-V$  curves for CZTS based PEC devices fabricated using C1, C2 and C3 are shown in Fig. 5. PEC characterization is selected because it is easier than forming solid state junction. It allows a rapid and nondestructive evaluation of a photo-activity of the CZTS thin films [41]. The detailed PEC parameters are listed in Table 1. The significant enhancement in the  $J_{sc}$  and  $V_{oc}$  is observed with increasing Cu concentration upon illumination as shown in Fig. 5. It can be clearly seen from the graph that the solar cell based on CZTS thin films derived from C1 shows relatively poor performance. This may be due to insufficient Cu concentration during the CZTS deposition process which may create structural defects in the CZTS film. While, the cell derived from C3 exhibits

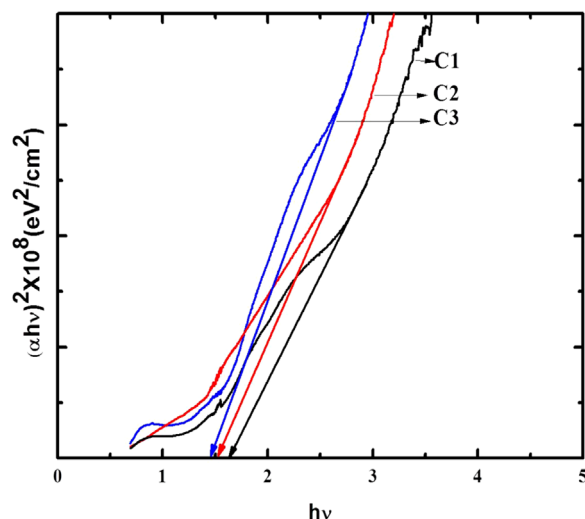


Fig. 4. Variation of  $(\alpha h\nu)^2$  versus  $(h\nu)^2$  of CZTS thin films prepared at copper concentration: C1=0.15 M, C2=0.20 M, and C3=0.25 M.

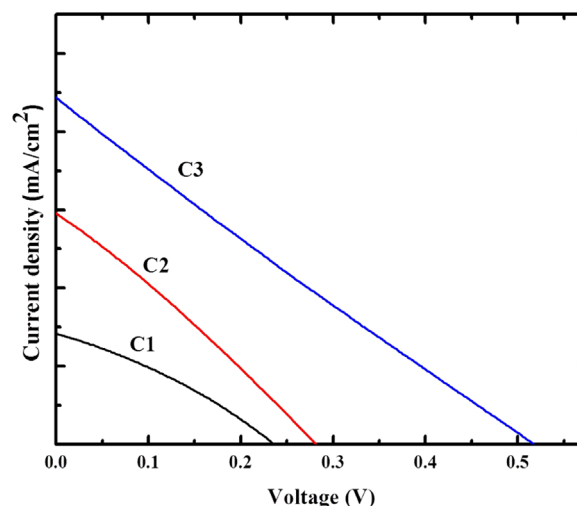


Fig. 5.  $J-V$  characteristics of CZTS thin films prepared at copper concentration: C1=0.15 M, C2=0.20 M, and C3=0.25 M.

Table 1

Solar cell parameters of PEC devices fabricated for different copper concentrations: C1=0.15 M, C2=0.20 M, and C3=0.25 M.

Sample	$J_{sc}$ (mA/cm <sup>2</sup> )	$V_{OC}$ (V)	F.F.	$\eta$ (%)	$R_s$ ( $\Omega$ )	$R_{sh}$ ( $\Omega$ )
C1	2.84	0.23	0.31	0.21	253	541
C2	7.31	0.39	0.30	0.86	190	400
C3	8.88	0.51	0.23	1.09	278.49	250

better conversion efficiency of 1.09%. From the structural analysis, it is observed that the C1 and C2 films does not show any secondary phase where as C3 sample shows very small secondary phase of  $Cu_2SnS_3$ . C1 and C2 films show low efficiency as compared to C3 film. The presence of secondary phase in C3 film may help for the enhancement of the photoactivity of the film. The series resistance ( $R_s$ ) and fill factor (FF) is lower in present PEC devices because of impurities and defects including insulating secondary phases which cause charge carrier recombination [42]. The short circuit current ( $I_{SC}$ ) increases with the increasing Cu concentration. This is due to improved and large grain size, which increases the absorption of incident photons. Additionally, it helps to increase current generation and reduce minority recombination rates [33]. The augmentation of these solar cell parameters could be attributed to the improved morphology of CZTS thin films, which affect CZTS–electrolyte junction quality and dynamics of charge carrier recombination [43]. The photoelectrochemical measurement confirmed the good photo-activity of the CZTS thin film (Table 1).

#### 4. Conclusions

In this paper, the effect of copper concentration on structural, morphology, optical and PEC properties of CZTS thin films at high temperature has been investigated. The analysis of XRD patterns indicate that the CZTS film with copper concentration up to 0.25 M has sharp and intense peak orientation. For higher concentration above 0.25 M amorphous phase was confirmed. Further, Raman studies confirm the phase purity of CZTS films with relatively small peak of  $Cu_2SnS_3$  phase. FE-SEM results indicate that as the Cu concentration increases average grain size also increases. Decrease in band gap due to increase in copper concentration was confirmed. Furthermore, the solar cell fabricated with the copper rich CZTS film (C3) showed enhancement in conversion efficiency about 1.09% due to presence of  $Cu_2SnS_3$  phase in the film.

#### Acknowledgment

One of the author S.M. Bhosale is grateful to University Grant Commission, New Delhi. for financial support through UGC project F. no. 41-945/2012. Authors are also grateful to Department of Science and Technology (DST) for financial support through PURSE and FIST and University Grant Commission (UGC) India through DSA-I scheme.

#### References

- [1] D.B. Mitzi, O. Gunawan, T.K. Todorov, K. Wang, S. Guha, The path towards a high-performance solution-processed kesterite solar cell, *Sol. Energy Mater. Sol. Cells* 95 (2011) 1421–1436.
- [2] C.H. Chung, B. Bob, B. Lei, S.H. Li, W.W. Hou, Y. Yang, Hydrazine solution-processed  $CuIn(Se,S)_2$  thin film solar cells: secondary phases and grain structure, *Sol. Energy Mater. Sol. Cells* 113 (2013) 148–152.
- [3] Q. Guo, H.W. Hillhouse, R. Agrawal, Synthesis of  $Cu_2ZnSnS_4$  nanocrystal ink and its use for solar cells, *J. Am. Chem. Soc.* 131 (2009) 11672–11673.
- [4] H. Katagiri, K. Jimbo, W.S. Maw, K. Oishi, M. Yamazaki, H. Araki, A. Takeuchi, Development of CZTS based thin film solar cells, *Thin Solid Films* 517 (2009) 2455–2460.
- [5] E.M. Mkawi, K. Ibrahim, M.K.M. Ali, A.S. Mohamed, Dependence of copper concentration on the properties of  $Cu_2ZnSnS_4$  thin films prepared by electrochemical method, *Int. J. Electrochem. Sci.* 8 (2013) 359–368.
- [6] S.R. Hall, J.T. Syzmanski, J.M. Stewart, Kesterite,  $Cu_2(Zn,Fe)SnS_4$  and stannite,  $Cu_2(Fe,Zn)SnS_4$ , structurally similar but distinct minerals, *Can. Mineral.* 16 (1978) 131.
- [7] M.P. Suryawanshi, G.L. Agawane, S.M. Bhosale, S.W. Shin, P.S. Patil J.H. Kim, A.V. Moholkar, CZTS based thin film solar cells: a status review, *Mater. Technol.* 28 (2013) 98–109.
- [8] H. Katagiri, K. Jimbo, S. Yamada, T. Kamimura, W.S. Maw, T. Fukano, T. Ito, T. Motohiro, Enhanced conversion efficiencies of  $Cu_2ZnSnS_4$ -based thin film solar cells by using preferential etching technique, *Appl. Phys. Express* 1 (2008) 041201.
- [9] B. Shin, O. Gunawan, Y.Z. Nestor, A. Bojarczuk, S. Jay Chey, S. Guha, Thin film solar cell with 8.4% power conversion efficiency using an earth-abundant  $Cu_2ZnSnS_4$  absorber, *Prog. Photovolt.* 21 (2011) 72–76.
- [10] A.V. Moholkar, S.S. Shinde, A.R. Babar, K.U. Sim, H.K. Lee, K.Y. Rajpure, P.S. Patil, C.H. Bhosale, J.H. Kim, Synthesis and characterization of  $Cu_2ZnSnS_4$  thin films grown by PLD: solar cells, *J. Alloy. Compd.* 509 (2011) 7439–7446.
- [11] H. Araki, Y. Kubo, K. Jimbo, W.S. Maw, H. Katagiri, M. Yamazaki, K. Oishi, A. Takeuchi, Preparation of  $Cu_2ZnSnS_4$  thin films by sulfurization of co-electroplated Cu–Zn–Sn precursors, *Phys. Status Solidi (C)* 6 (2009) 1266–1268.
- [12] M.Y. Yeh, C.C. Lee, D.S. Wu, Influences of synthesizing temperatures on the properties of  $Cu_2ZnSnS_4$  prepared by sol–gel spin-coated deposition, *J. Sol-Gel Sci. Technol.* 52 (2009) 65–68.
- [13] S.A. Nadi, P. Chelvanathan, Z. Zakaria, M.M. Alam, Z.A. Alotman, K. Sopian, N.D. Amin, Post-deposition annealing effect on  $Cu_2ZnSnS_4$  thin films grown at different substrate temperature, *Int. J. Photoenergy* 2014 (2014) 589027–589034.
- [14] J.M.R. Tan, Y.H. Lee, S. Pedireddy, T. Baikie, X.Y. Ling, L.H. Wong, Understanding the synthetic pathway of a single-phase quaternary semiconductor using surface-enhanced Raman scattering: a case of wurtzite  $Cu_2ZnSnS_4$  nanoparticles, *J. Am. Chem. Soc.* 136 (2014) 6684–6692.
- [15] X. Yin, C. Tang, L. Sun, Z. Shen, H. Gong, Study on phase formation mechanism of non and near-stoichiometric  $Cu_2ZnSn(S,Se)_4$  film prepared by selenization of Cu–Sn–Zn–S precursors, *Chem. Mater.* 26 (2014) 2005–2014.
- [16] Y.B. Kishore Kumar, P. Uday Bhaskar, G. Suresh Babu, V. Sundara Raja, Effect of copper salt and thiourea concentrations on the formation of  $Cu_2ZnSnS_4$  thin films by spray pyrolysis, *Phys. Status Solidi (A)* 207 (2010) 149–156.
- [17] N. Kamoun, H. Bouzouita, B. Rezig, Fabrication and characterization of  $Cu_2ZnSnS_4$  thin films deposited by spray pyrolysis technique, *Thin Solid Films* 515 (2007) 5949–5952.
- [18] S.M. Camara, L. Wang, X. Zhang, Easy hydrothermal preparation of  $Cu_2ZnSnS_4$  (CZTS) nanoparticles for solar cell application, *Nanotechnology* 24 (2013) 495401–495410.
- [19] Z. Seboui, Y. Cuminal, N. Kamoun, Physical properties of  $Cu_2ZnSnS_4$  thin films deposited by spray pyrolysis technique, *J. Renew. Sustain. Energy* 5 (2013) 023113.
- [20] S.M. Bhosale, M.P. Suryawanshi, M.A. Gaikwad, P.N. Bhosale, J.H. Kim A.V. Moholkar, Influence of growth temperatures on the properties of

- photoactive CZTS thin films using a spray pyrolysis technique, *Mater. Lett.* 129 (2014) 153–155.
- [21] K.U. Isah, J.A. Yabagi, U. Ahmadu, M.I. Kimpa, M.G.Z. Kana, A.A. Oberafo, Effect of different copper precursor layer thickness on properties of  $\text{Cu}_2\text{ZnSnS}_4$  (CZTS) thin films prepared by sulfurization of thermally deposited stacked metallic layers, *IOSR-JAP* 216 (2013) 14–19.
  - [22] A.J. Cheng, M. Manno, A. Khare, C. Leighton, S.A. Campbell, E.S. Aydil, Imaging and phase identification of  $\text{Cu}_2\text{ZnSnS}_4$  thin films using confocal Raman spectroscopy, *J. Vac. Sci. Technol. A* 29 (2011) 051203-1–051203-11.
  - [23] P.A. Fernandes, P.M.P. Salomé, A.F. da Cunha, Study of polycrystalline  $\text{Cu}_2\text{ZnSnS}_4$  films by Raman scattering, *J. Alloy. Compd.* 509 (2011) 7600–7606.
  - [24] M. Grossberg, J. Krustok, J. Raudoja, K. Timmo, M. Altosaar, T. Raadik, Photoluminescence and Raman study of  $\text{Cu}_2\text{ZnSn}(\text{Se}_x\text{S}_{1-x})_4$  monograins for photovoltaic applications, *Thin Solid Films* 519 (2011) 7403–7406.
  - [25] G. Gouadec, P. Colomban, Raman Spectroscopy of nanomaterials: how spectra relate to disorder, particle size and mechanical properties, *Prog. Cryst. Growth Mater.* 53 (2007) 1–56.
  - [26] W.C. Liu, D. Wu, A.D. Li, H.Q. Ling, Y.F. Tang, N.B. Ming, Annealing and doping effects on structure and optical properties of sol-gel derived  $\text{ZrO}_2$  thin films, *Appl. Surf. Sci.* 191 (2002) 181–187.
  - [27] U. Chalapathi, S. Uthanna, V. Sundara Raja, Growth of  $\text{Cu}_2\text{ZnSnS}_4$  thin films by a two-stage process—Effect of incorporation of sulfur at the precursor stage, *Sol. Energy Mater. Sol. Cells* 132 (2015) 476–484.
  - [28] E.M. Mkawi, K. Ibrahim, M.K.M. Ali, M.A. Farrukh, A.S. Mohamed, Dependence of the properties of copper zinc tin sulfide thin films prepared by electrochemical deposition on sulfurization temperature, *J. Mater. Sci.: Mater. Electron.* 25 (2014) 857–863.
  - [29] G. Rajesh, N. Muthukumarasamy, E.P. Subramaniam, S. Agilan, D. Velauthapillai, Synthesis of  $\text{Cu}_2\text{ZnSnS}_4$  thin films by dip-coating method without sulphurization, *J. Sol-Gel Sci. Technol.* 66 (2013) 288–292.
  - [30] S. Ahmed, K.B. Reuter, O. Gunawan, L. Guo, L.T. Romankiw, H. Deligianni, A high efficiency electrodeposited  $\text{Cu}_2\text{ZnSnS}_4$  solar cell, *Adv. Energy Mater.* 2 (2012) 253–259.
  - [31] X. Fontane, L. Calvo-Barrio, V. Izquierdo-Roca, E. Saucedo A. Perez-Rodriguez, R. Morante, D.M. Berg, P.J. Dale, S. Siebentritt, In-depth resolved Raman scattering analysis for the identification of secondary phases: characterization of  $\text{Cu}_2\text{ZnSnS}_4$  layers for solar cell applications, *Appl. Phys. Lett.* 98 (2011) 181905-1–181905-3.
  - [32] H. Park, Y.H. Hwang, B.S. Bae, Sol-gel processed  $\text{Cu}_2\text{ZnSnS}_4$  thin films for a photovoltaic absorber layer without sulfurization, *J. Sol-Gel Sci. Technol.* 65 (2013) 23–27.
  - [33] J. He, L. Sun, Y. Chen, J. Jiang, P. Yang, J. Chu, Influence of sulfurization pressure on  $\text{Cu}_2\text{ZnSnS}_4$  thin films and solar cells prepared by sulfurization of metallic precursors, *J. Power Sources* 273 (2015) 600–607.
  - [34] Q. Guo, G.M. Ford, W.C. Yang, B.C. Walker, E.A. Stach, H.W. Hillhouse, R. Agrawal, Fabrication of 7.2% efficient CZTSSe solar cells using CZTS nanocrystals, *J. Am. Chem. Soc.* 132 (2010) 17384–17386.
  - [35] S.W. Shin, S.M. Pawar, C.Y. Park, J.H. Yun, J.H. Moon, J.H. Kim, J.Y. Lee, Studies on  $\text{Cu}_2\text{ZnSnS}_4$  (CZTS) absorber layer using different stacking orders in precursor thin films, *Sol. Energy Mater. Sol. Cells* 95 (2010) 3202–3206.
  - [36] V.P. Geetha Vani, M. Vasudeva Reddy, K.T. Ramakrishna Reddy, Thickness dependent physical properties of coevaporated  $\text{Cu}_4\text{SnS}_4$  films, *ISRN Condens. Matter Phys.* 2013 (2013) 1–6.
  - [37] C. Malerba, F. Biccari, C.L.A. Ricardo, M. Valentini, R. Chierchia, M. Müller, A. Santoni, E. Esposito, P. Mangiapane, P. Scardi, A. Mittiga, CZTS stoichiometry effects on the band gap energy, *J. Alloy. Compd.* 582 (2014) 528–534.
  - [38] Q. Chen, X. Dou, Y. Ni, S. Cheng, S. Zhuang, Study and enhance the photovoltaic properties of narrow-bandgap  $\text{Cu}_2\text{SnS}_3$  solar cell by p–n junction interface modification, *J. Colloid Interface Sci.* 376 (2012) 327–330.
  - [39] S.Y. Chen, A. Walsh, Y. Luo, J.H. Yang, X.G. Gong, S.H. Wei, Wurtzite-derived polytypes of kesterite and stannite quaternary chalcogenide semiconductors, *Phys. Rev. B* 82 (2010) 195203.
  - [40] F.J. Fan, L. Wu, M. Gong, G. Liu, Y.X. Wang, S.H. Yu, S. Chen, L.W. Wang, X.G. Gong, Composition- and band-gap-tunable synthesis of wurtzite-derived  $\text{Cu}_2\text{ZnSn}(\text{S}_{1-x}\text{Se}_x)_4$  nanocrystals: theoretical and experimental insights, *ACS Nano* 7 (2013) 1454–1463.
  - [41] M.P. Suryawanshi, S.W. Shin, U.V. Ghorpade, K.V. Gurav, G.L. Agawane, S.A. Vanalakar, J.H. Moon, J.H. Yun, P.S. Patil, J.H. Kim, A.V. Moholkar, Improved photoelectrochemical performance of  $\text{Cu}_2\text{ZnSnS}_4$  (CZTS) thin films prepared using modified successive ionic layer adsorption and reaction (SILAR) sequence, *Electrochim. Acta* 150 (2014) 136–145.
  - [42] M. Patel, I. Mukhopadhyay, A. Ray, Structural, optical and electrical properties of spray-deposited CZTS thin films under a non-equilibrium growth condition, *J. Phys. D: Appl. Phys.* 45 (2012) 445103–445113.
  - [43] M.P. Suryawanshi, P.S. Patil, S.W. Shin, K.V. Gurav, G.L. Agawane, M.G. Gang, Jin yeok Kim, A.V. Moholkar, The synergistic influence of anionic bath immersion time on the photoelectrochemical performance of CZTS thin films prepared by a modified SILAR sequence, *RSC Adv.* 4 (2014) 18537–18540.



# Influence of growth temperatures on the properties of photoactive CZTS thin films using a spray pyrolysis technique

S.M. Bhosale<sup>a</sup>, M.P. Suryawanshi<sup>a,c</sup>, M.A. Gaikwad<sup>a</sup>, P.N. Bhosale<sup>b</sup>, J.H. Kim<sup>c</sup>,  
A.V. Moholkar<sup>a,\*</sup>

<sup>a</sup> Thin Film Nano-Materials Laboratory, Department of Physics, Shivaji University, Kolhapur, India

<sup>b</sup> Material Research Laboratory, Department of Chemistry, Shivaji University, Kolhapur, India

<sup>c</sup> Photonics Technology Research Institute, Department of Materials Science and Engineering, Chonnam National University, 300, Yongbong-Dong, Buk-Gu, Gwangju 500-757, South Korea

## ARTICLE INFO

### Article history:

Received 7 March 2014

Accepted 19 April 2014

Available online 26 April 2014

### Keywords:

Cu<sub>2</sub>ZnSnS<sub>4</sub> (CZTS)

TFSCs

Spray pyrolysis

Influence of growth temperatures

0.86% power conversion efficiency

## ABSTRACT

Highly crystalline photoactive Cu<sub>2</sub>ZnSnS<sub>4</sub> (CZTS) thin films are fabricated by a simple and low cost spray-pyrolysis technique without sulfurization treatment in toxic atmosphere. The influence of substrate temperatures on the physico-chemical properties of the CZTS films is investigated. The film sprayed at optimized substrate temperature exhibits optical band gap energy of 1.49 eV, which is prerequisite for the photovoltaic applications. The photoelectrochemical (PEC) device of CZTS-500 film showed a power conversion efficiency of 0.86% which is the highest efficiency for sprayed CZTS thin films so far.

© 2014 Elsevier B.V. All rights reserved.

## 1. Introduction

CZTS has emerged as a potential candidate for absorber layer in photovoltaic cell, having earth abundant, non-toxic and low cost elements. It exhibits an excellent optical absorption of over 10<sup>4</sup> cm<sup>-1</sup>, direct optical band gap energy of 1.4–1.5 eV and p-type conductivity, which makes this material especially suited for the development of low-cost, and environmentally nontoxic thin film solar cells (TFSCs) [1–4].

Numerous vacuum and non-vacuum based deposition techniques have been employed for the synthesis of CZTS thin films [5]. Among all these techniques, non-vacuum based techniques are crucial to fabricate economic and efficient CZTS TFSCs. The highest efficiency of 12.6% has been achieved using a hydrazine based particle solution approach [6]. Although, the CZTS TFSCs using non-vacuum approach have several merits and high power conversion efficiency (PCE), they suffer from various problems such as incorporation of undesired impurities and chemical toxicity, which needs to be addressed for scaling up massive solar cells production [7].

A chemical spray pyrolysis (CSP) technique is low cost, non-vacuum, eco-friendly and can be used for large scale deposition of different semiconductor thin films, as well its high potential for achieving a significant reduction of production costs. The composition, morphology and optical and electrical properties of the CZTS

films can be tailored by changing the deposition parameters such as the precursor's composition, the substrate temperature, the deposition time duration, the spray rate, the nozzle-to-substrate distance, etc. [8]. Nakayama and Ito reported the preparation of CZTS thin films by spray pyrolysis for the first time [9]. Kamoun et al. deposited CZTS thin films by varying substrate temperature and spray duration and they found that annealing of the films plays important role for the improvement of the properties [10]. Shinde et al. studied the effect of film thickness on the properties of CZTS thin films [11]. Kumar and co-workers had also demonstrated CSP of CZTS under various spraying conditions, even though the photovoltaic performance was not reported [12]. Recently, Rodriguez et al. [13] deposited CZTS on 'Mo' coated substrate and found the conversion efficiency of 0.49% after sulfurizing the film, which is the only report on the sprayed CZTS based device efficiency.

In this work, we have reported the synthesis of highly crystalline CZTS thin films using the spray pyrolysis technique by varying the substrates temperature without sulfurization in toxic atmosphere such as H<sub>2</sub>S or 'S' vapor. Furthermore, the influence of different growth temperatures on the physico-chemical properties and photoelectrochemical performance of sprayed CZTS thin films has been investigated.

## 2. Experimental details

CZTS thin films were deposited by spraying the precursor solution on preheated, ultrasonically cleaned soda lime glass (SLG) and

\* Corresponding author.

E-mail address: [avmoholkar@gmail.com](mailto:avmoholkar@gmail.com) (A.V. Moholkar).

fluorine doped tin oxide (FTO) coated SLG substrates at different substrate temperatures. CZTS precursor solution was prepared by dissolving copper chloride, zinc chloride, tin chloride and thiourea in double distilled water and stirring for 10 min. The concentration of materials was selected as 2:1:1:8. Excess amount of thiourea was required to avoid the precipitate in the solution and also to compensate the loss of sulfur during pyrolysis as sulfur source is volatile at higher temperature [14]. 30 ml solution was sprayed with the spray rate of 2 ml/min, and substrate temperature varied

from 200 °C to 500 °C with an accuracy of  $\pm 5$  °C using a digital temperature controller. The samples are denoted as 'CZTS-200', 'CZTS-300', 'CZTS-400' and 'CZTS-500'.

### 3. Characterization

The structural properties were studied using an X-ray diffractometer (Bruker AXS Model D2 PHASER Analytical Instrument, Germany) and Raman spectroscopy (Bruker AXS Analytical instrument, Germany). The optical absorption of the sprayed thin films was measured by a UV–vis spectrophotometer (UV-1800, Shimadzu, Germany). The photoelectrochemical performance was measured using a Sol2A Oriel New Port Corporation USA, with a Keithley-2420 source meter under 1.5 AM.

### 4. Results and discussion

Fig. 1 shows XRD patterns of sprayed CZTS samples as a function of growth temperature. All sprayed thin films except for CZTS-200 sample exhibit three major broad peaks at  $2\theta = 28.8^\circ$ ,  $47.64^\circ$  and  $56.5^\circ$ , which could be assigned to (112), (220) and (312) planes of kesterite-type CZTS (JCPDS card no. 00-026-0575). An extra peak at  $2\theta = 33.43^\circ$  for CZTS-500 is observed which also corresponds to kesterite type CZTS. The increase in the (112) peak intensity with increasing growth temperature has been observed, which indicates that highly crystalline sample can be deposited at high growth temperatures. Since the lattice constants of CZTS are similar to  $\text{Cu}_2\text{SnS}_3$  (CTS) and ZnS, the obtained peaks could be related to either CZTS or CTS or ZnS phases [15]. Thus, to confirm the crystal structure and phase formation of the films, we further studied Raman spectra.

Fig. 2 shows the Raman spectrum of CZTS-200, CZTS-300, CZTS-400 and CZTS-500 samples. The existence of CZTS is confirmed by the presence of major Raman peak at  $337\text{ cm}^{-1}$  and the shoulder peaks at  $282$  and  $362\text{ cm}^{-1}$ , which are close to the values reported for all samples except for CZTS-200 [16–18]. However, one extra peak at  $479\text{ cm}^{-1}$  corresponding to  $\text{Cu}_{2-x}\text{S}$  can be seen for all samples except for CZTS-500. This could be attributed to the high substrate temperature, which easily allows the formation of secondary phase of  $\text{Cu}_{2-x}\text{S}$ .

Fig. 3(a) shows the graph of optical band gap energy ( $E_g$ ) vs substrate temperature and the inset shows the plot of  $(\alpha h\nu)^2$  vs  $h\nu$  for films deposited at different substrate temperatures. The optical band gap energy ( $E_g$ ) obtained 1.87, 1.70, 1.59 and 1.47 eV for

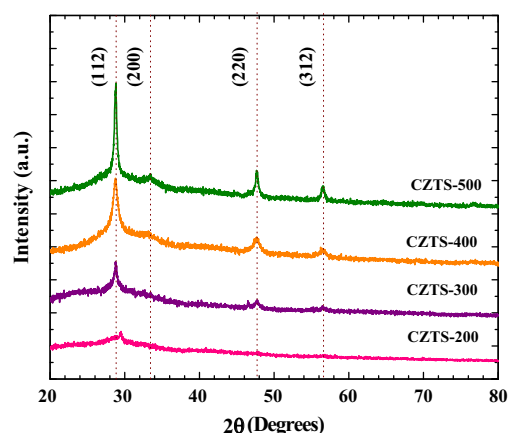


Fig. 1. XRD patterns for CZTS-200, CZTS-300, CZTS-400 and CZTS-500 samples.

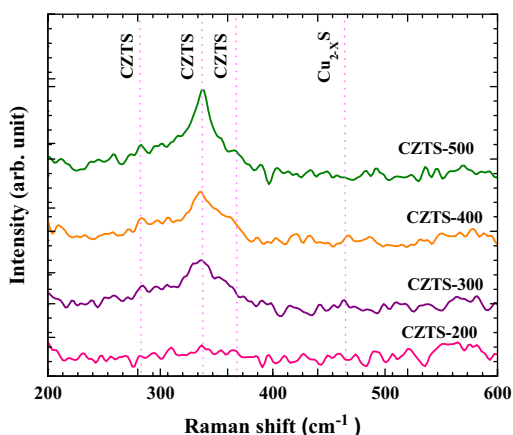


Fig. 2. Raman spectra for CZTS-200, CZTS-300, CZTS-400 and CZTS-500 samples.

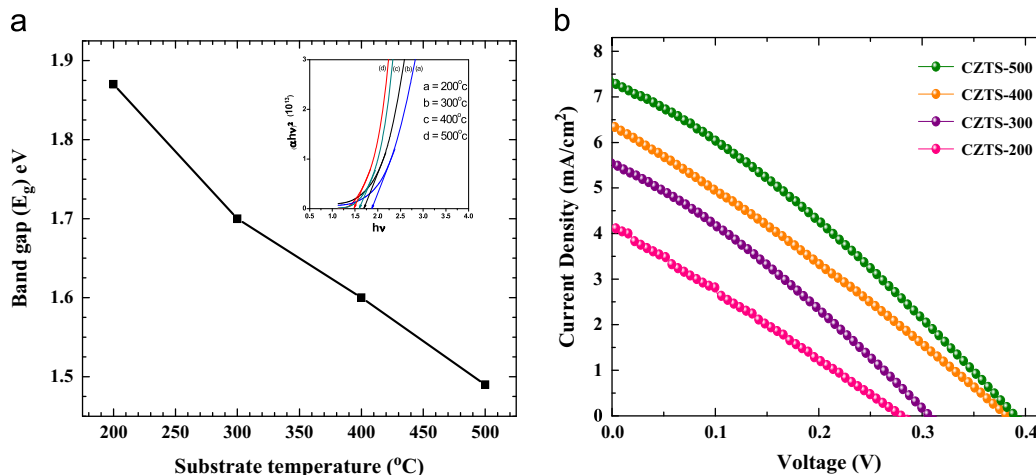


Fig. 3. Plot of band gap energy ( $E_g$ ) vs substrate temperature for CZTS-200, CZTS-300, CZTS-400 and CZTS-500 samples. The inset shows the bandgap of each sample (a) and  $J$ - $V$  characteristics (b) CZTS-200, CZTS-300, CZTS-400 and CZTS-500 samples under illumination.

**Table 1**

Solar cell parameters of PEC devices fabricated using CZTS-200, CZTS-300, CZTS-400 and CZTS-500 samples.

Sample	$J_{sc}$ (mA/cm <sup>2</sup> )	$V_{oc}$ (V)	FF	$\eta$ (%)	$R_s$ ( $\Omega$ )	$R_{sh}$ ( $\Omega$ )
CZTS-200	4.15	0.28	0.26	0.30	292	424
CZTS-300	5.53	0.31	0.29	0.50	203	400
CZTS-400	6.36	0.38	0.27	0.67	221	400
CZTS-500	7.31	0.39	0.30	0.86	190	400

CZTS-200, CZTS-300, CZTS-400 and CZTS-500, respectively. With increase in the growth temperatures, the band gap values shift to lower energies. According to quantum size effect the band gap energy of a material is dependent on particle size of the material. Smaller the particle size greater the band gap energy of the material. When substrate temperature goes on increasing, more energy is supplied to crystallize the particles on the substrate surface which naturally decreases the band gap energy of the material. In the present case the peak intensity increases with increase in temperature leading to enhancement in the crystallinity and increase in particle size. Such type of behavior is reported previously [11,14].

The photoelectrochemical (PEC) performance of CZTS thin films on FTO substrates has been studied with a two electrode configuration of photoanode, which were immersed in 0.1 M Europium nitrate (Eu(III)(NO<sub>3</sub>)<sub>3</sub>) as a redox mediator under UV (100 mW) illumination. ITO was used as a counterelectrode. Illuminated  $J$ – $V$  curves for CZTS PEC devices fabricated using CZTS-200, CZTS-300, CZTS-400 and CZTS-500 under dark and light illumination are shown in Fig. 3(b). The gradual increment in the current density ( $J_{sc}$ ) from 4.15 to 7.31 mA/cm<sup>2</sup>, open circuit voltage ( $V_{oc}$ ) from 0.28 to 0.39 V and photon conversion efficiency (PCE) from 0.30% to 0.86 % with increase in the substrate temperature has been observed. The various solar cell parameters are elucidated in Table 1.  $V_{oc}$  and series resistance are lower in our PEC devices because of impurities and defects including insulating secondary phases which cause charge carrier recombination [19]. Hence, the fill factor (FF) of our PEC devices is degraded and is lower than that of the best CZTS device. Although the efficiency is low enough compared to other deposition techniques, further improvements in the performance of CZTS thin films by this technique could lead to high device efficiency. Hence, the efforts in this direction are underway.

## 5. Conclusions

CZTS thin films have been successfully deposited by the low cost, simple spray pyrolysis technique. The influence of growth temperatures on properties and photoelectrochemical performance of sprayed CZTS thin films has been investigated. The

significant improvement has been observed in the crystallinity of spray deposited CZTS thin films having a kesterite structure with increase in the growth temperature. The band gap energy of 1.47 eV is obtained for CZTS film sprayed at 500 °C. The PEC device based on CZTS-500 sample showed  $J_{sc}$ =7.31 mA/cm<sup>2</sup>,  $V_{oc}$ =0.39 V and FF=0.30, with best power conversion efficiency of 0.86%.

## Acknowledgments

This work was supported by University Grant Commission, New Delhi Ref. F.No.41-945/2012 (SR), DST and UGC-PURSE.

## References

- [1] Gurav KV, Pawar SM, Shin SW, Suryawanshi MP, Agawane GL, Patil PS, et al. Appl Surf Sci 2013;283:74–80.
- [2] Todorov TK, Reuter KB, Mitzi DB. Adv Mater 2010;22:E156–9.
- [3] Barkhouse DAR, Gunawan O, Gokmen T, Todorov TK, Mitzi DB. Prog Photovolt: Res Appl 2012;20:6–10.
- [4] Moholkar AV, Shinde SS, Agawane GL, Jo SH, Rajpure KY, Patil PS, et al. J Alloys Compd 2012;544:145–51.
- [5] Suryawanshi MP, Agawane GL, Bhosale SM, Shin SW, Patil PS, Kim JH, et al. Mater Technol: Adv Mater Perform 2013;28:1–2.
- [6] Wang W, Winkler MT, Gunawan O, Gokmen T, Todorov TK, Zhu Y, et al. Adv Energy Mater 2013. <http://dx.doi.org/10.1002/aenm.201301465>.
- [7] Hibberd HJ, Chassaing E, Liu W, Mitzi DB, Lincot D, Tiwari AN. Prog Photovolt: Res Appl 2010;18:434–52.
- [8] Patil PS. Mater Chem Phys 1999;59:185–98.
- [9] Nakayama N, Ito K. Appl Surf Sci 1996;92:171.
- [10] Kamoun N, Bouzouita H, Rezig B. Thin Solid Films 2007;515:5949.
- [11] Shinde NM, Deokate RJ, Lokhande CD. J Anal Appl Pyrolysis 2013;100:12–6.
- [12] Kumar YBK, Bahu GS, Bhaskar PU, Raja VS. Sol Energy Mater Sol Cells 2009;93:1230.
- [13] Rodriguez ME, Placidi M, Galan OV, Roca VI, Fontane X, Fairbrother A, et al. Thin solid films 2013;535:67.
- [14] Seboui Z, Cuminal Y, Kamoun N. J Renew Sustain Energy 2013;5:023113.
- [15] Isah KU, Yabagi JA, Ahmadu U, Kimpa MI, Kana MGZ, Oberafo AA. IOSR-JAP 2013;216:14.
- [16] Suryawanshi M, Shin SW, Woo RI B, Gurav K, Kang MG, Agawane G, et al. Phys Status Solidi A 2014. <http://dx.doi.org/10.1002/pssa.2013030384>.
- [17] Cheng AJ, Manno M, Khare A, Leighton C, Campbell SA, et al. J Vac Sci Technol A 2011;29, 051203-1.
- [18] Gurav KV, Pawar SM, Shin SW, Suryawanshi MP, Agawane GL, Patil PS, et al. Appl Surf Sci 2013;283:74.
- [19] Suryawanshi MP, Patil PS, Shin SW, Gurav KV, Agawane GL, RSC Adv 4, 2014, 185374, 2014, 18537.



Scholars Research Library

Der Pharma Chemica, 2015, 7(10):13-22  
(<http://derpharmachemica.com/archive.html>)



ISSN 0975-413X  
CODEN (USA): PCHHAX

## Synthesis, characterization and biological activity of a novel p-toulic hydrazone and resacetophenone schiff base (RAPPTH) ligand and their metal complexes

S. Kondaiah<sup>1\*</sup>, G. Naga Raja Reddy<sup>2</sup>, B. Vijaya Kumar<sup>3</sup>, P. Govinda Chowdary<sup>4</sup> and B. V. Ramana Reddy<sup>1</sup>

<sup>1</sup>Department of Chemistry, Malla Reddy Engineering Chemistry, Secunderbad, Telangana, India

<sup>2</sup>Department of Chemistry, Srikrishnadevaraya University, Anantapur, Andhra Pradesh, India

<sup>3</sup>Department of Chemistry, Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, India

<sup>4</sup>Department of Chemistry, Vignan Institute of Technology and Sciences, Deshmuki, Nalgonda, Telangana

### ABSTRACT

A novel Schiff base like RAPPTH ligand ( p-Toulic Hydrazone and resacetophenone ) and their Mn (II), Co (II), Ni (II) and Cu (II) metal complexes have been prepared and characterized by various physio chemical methods like FT-IR, <sup>1</sup>H-NMR, UV-Visible, ESR, VSM, Molar conductance, XRD and micro analytical data and found to be biological activity. The magnetic properties of these Mn (II), Co (II), Ni (II) and Cu (II) complexes are 5.28 B.M. 4.82 B.M., 3.15 B.M and 1.81 B.M respectively. This result clearly shows that the metal complexes of this ligand show octahedral geometry. The IR spectrums of these Schiff base metal complexes gives bands at 1596cm<sup>-1</sup>, 1585cm<sup>-1</sup> 1606cm<sup>-1</sup> and 1609cm<sup>-1</sup> respectively, which is assigned to (C=N) stretching vibrations, a fundamental feature of azomethine group. The structural assessment of these metal complexes has been carried out based on above physio-chemical and spectroscopic methods. From the elemental analyses data, 1:2 metal complexes are formed.

**Keywords:** Schiff base, p-Toluic hydrazone, Resacetophenone, Anti bacterial activity and characterization.

### INTRODUCTION

The Schiff bases are play an important role in chemistry. The transition metal complexes of Schiff bases are very fascinating because of the presence of both hard nitrogen and oxygen and soft sulphur donor atoms in the backbones of these ligands, some of these complexes have been exhibit remarkable physiochemical properties and potentially useful biological activities.[1-2] Schiff base have also been used for analytical purposes in the determination of metal ions, and some Schiff base derivatives have been used in the extraction of metal solvents. The applications of such complexes depend to a great extent on their molecular structure. Schiff base ligands are able to synchronize different metals, and to stabilize them in various oxidation [3-6] states. The Schiff base complexes have been used [7-8] in catalytic reactions and as models for biological systems. During the last 20 years extensive awareness on the metal complexes of Schiff bases [9-10] containing nitrogen and other donors. This may be endorsed to their stability, biological activity and huge applications [11] in many fields such as oxidation catalysis, electrochemistry and biological studies. Transition metal complexes of Schiff bases derived from benzil with different amines had

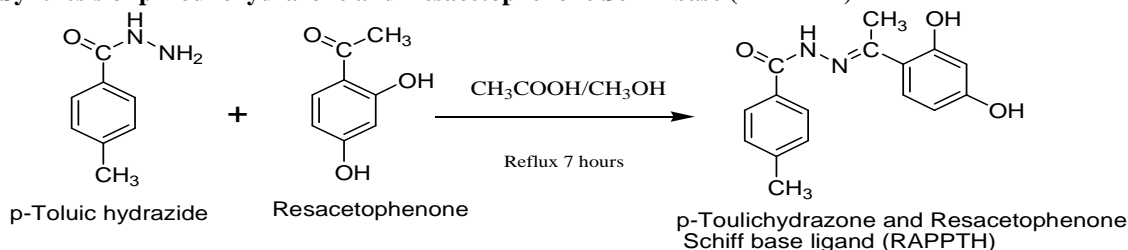
been prepared and characterized by physical and chemical techniques. Schiff bases derived from condensation of various aldehydes with triethylene tetra amine and their complexes were reported.

The present paper deals with a series of metal complexes of Mn(II), Co(II), Ni(II) and Cu(II) with Schiff base ligand (RAPPTH) derived from P-Toluic hydrazide and Resacetophenone (2,4 Dihydroxy acetophenone). These complexes were characterized by elemental analysis, IR, NMR, UV, ESR spectroscopy, TG-DTA, Powder X-Rd, VSM and Conductivity measurements to verify the mode of bonding, geometry and biological activities of the metal complexes were also studied. These metal complexes are shows the octahedral structure.

## MATERIALS AND METHODS

### 2. Experimental

#### 2.1 Synthesis of p-Toulchydrazone and Resacetophenone Schiff base (RAPPTH)

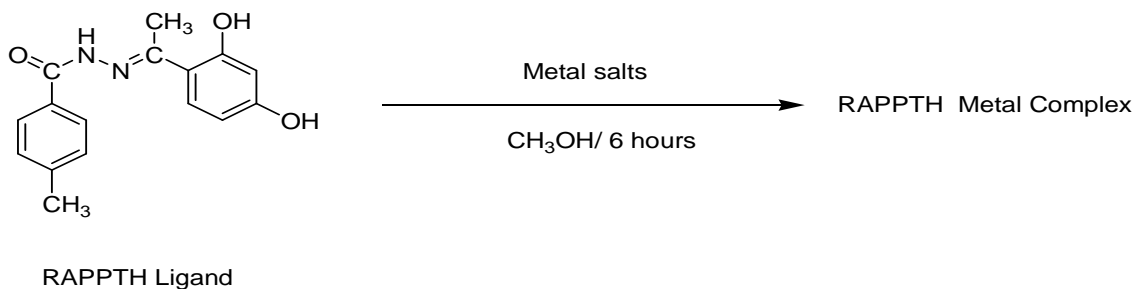


Scheme 1: Synthesis of RAPPTH Ligand

p-Toluic hydrazide (1.5018g) and Resacetophenone (2,4Dihydroxyacetophenone)(1.5215g) were dissolved in methanol and Refluxed in the presence of few drops of acetic acid for 7 hours on water bath with constant stirring. The resultant mixture was transferred in a china dish and allowed to cool naturally. Light brown crystals were obtained. The compound was recrystallized from water. The % of yield was 80% and melting point of the compound was 218-220°C.

#### 2.2 Synthesis of Mn (II), Co (II), Ni (II) and Cu (II) metal complexes of p-Toluic hydrazone and Resacetophenone Schiff base (RAPPTH):

The reagent p-Toluic hydrazone and Resacetophenone(2,4dihydroxyacetophenone) Schiff base was dissolved in 40ml of 50% methanol and Mn (II), Co (II), Ni (II) and Cu (II) metal ions dissolved in 20ml of distilled water, this compounds mixture was stirred for 6 hours in the presence of sodium acetate. Thick Green, Brown, Light yellow and Green colour metal complexes were formed respectively with good yield. These products were washed several times with hot water and cold methanol. Metal complexes are free from un reacted metal slats and ligand respectively and finally dried in vacuo over calcium chloride dessicator.



Scheme 2: Synthesis of RAPPTH Metal Complexes

## RESULTS AND DISCUSSION

### 3.1 IR & NMR Spectral Studies

IR & NMR Spectral data of free RAPPTH ligand and its metal complexes are shown in Fig 1-6. In  $^1\text{H}$  NMR spectrum a signal is observed at 11.10ppm due to N-H proton in the ligand is shifted to  $\delta$  11.25 and 11.29ppm respectively for RAPPTH-Co and RAPPTH-Cu metal complexes.

A broad band exhibited at  $3342\text{ cm}^{-1}$  in the IR spectrum of the ligand due to N-H stretching vibration. On complexation this band shifted to  $3347$ ,  $3349$ ,  $3350$ , and  $3352\text{ cm}^{-1}$  for Mn (II), Co (II), Ni (II) and Cu (II) complexes respectively. The IR spectrum of the ligand has shown a sharp band at  $1698\text{ cm}^{-1}$  due to C=O stretching vibration. On complexation C = O stretching vibration band shifted to  $1671$ ,  $1666$ ,  $1658$ , and  $1662\text{ cm}^{-1}$  for Mn (II), Co (II), Ni (II) and Cu (II) Complexes respectively. These results are indicating the formation of complex. The IR spectra of Mn (II), Co (II), Ni (II) and Cu (II) complexes exhibited broad bands at  $3430\text{ cm}^{-1}$ ,  $3480\text{ cm}^{-1}$ ,  $3415\text{ cm}^{-1}$  and  $3420\text{ cm}^{-1}$  respectively, which can be assigned to the OH stretching vibration of the coordinated [12-14] water molecules. These results are clearly indicate that the ligand coordinate with the metal ion through the azomethine nitrogen and the oxygen of the deprotonated hydroxyl group.

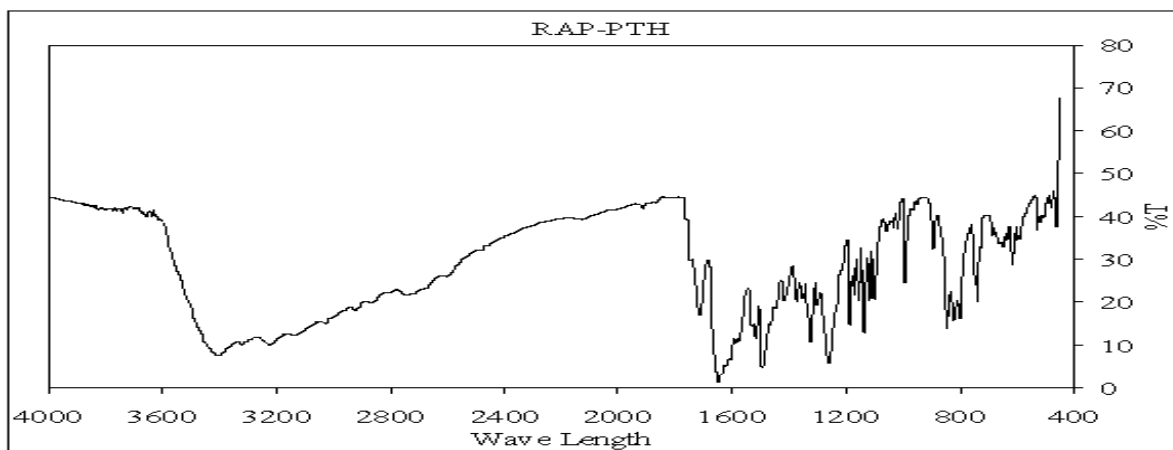


Fig 1: IR Spectrum of RAPPTH Ligand

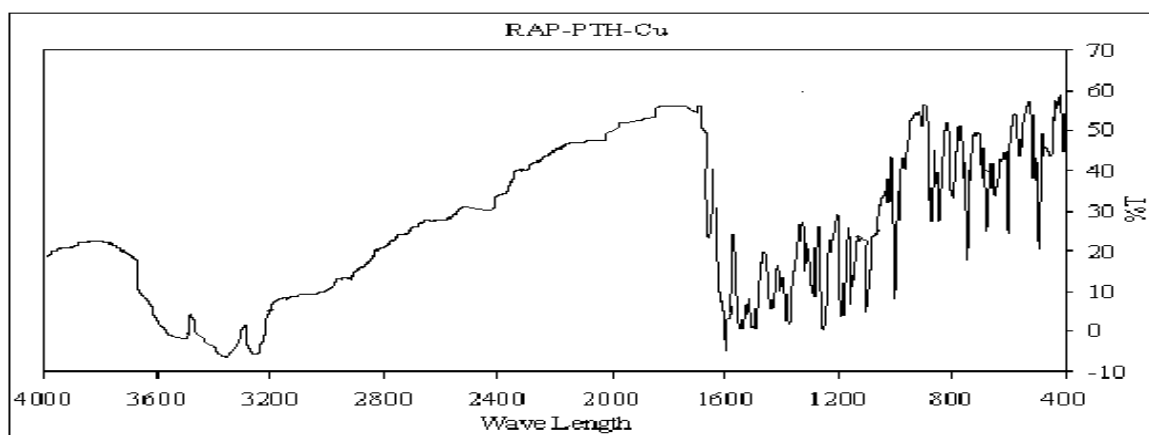


Fig 2: IR Spectrum of RAPPTH -Co Metal Complex

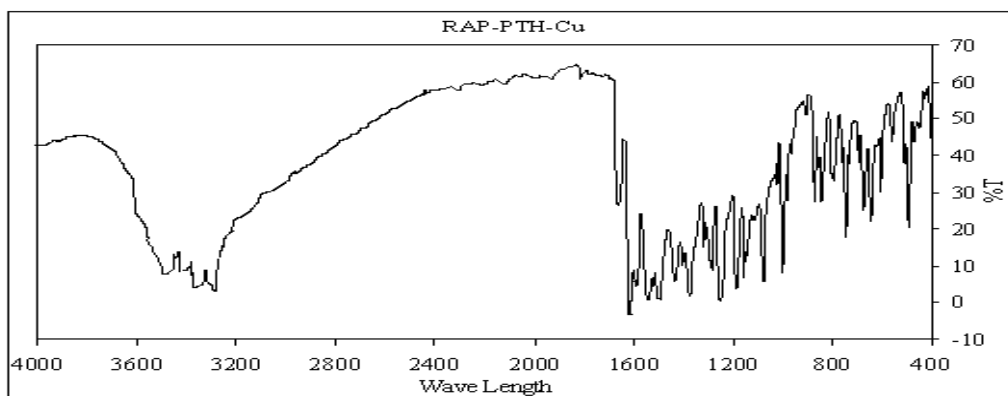


Fig.3. IR Spectrum of RAPPTH- Cu (II) Complex

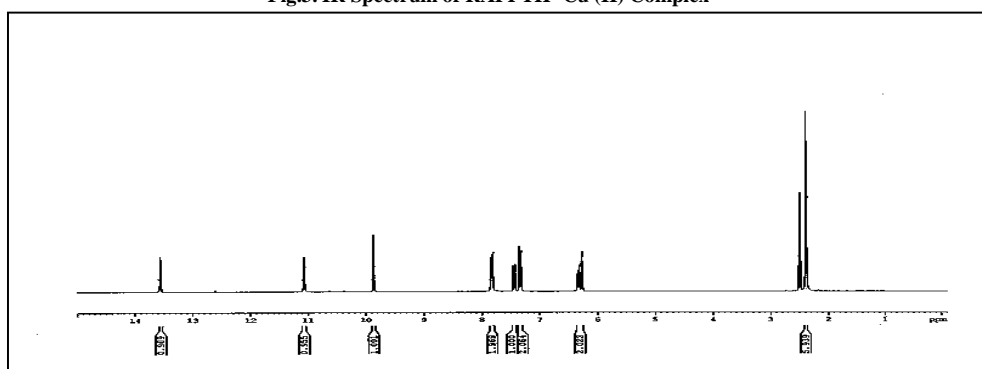
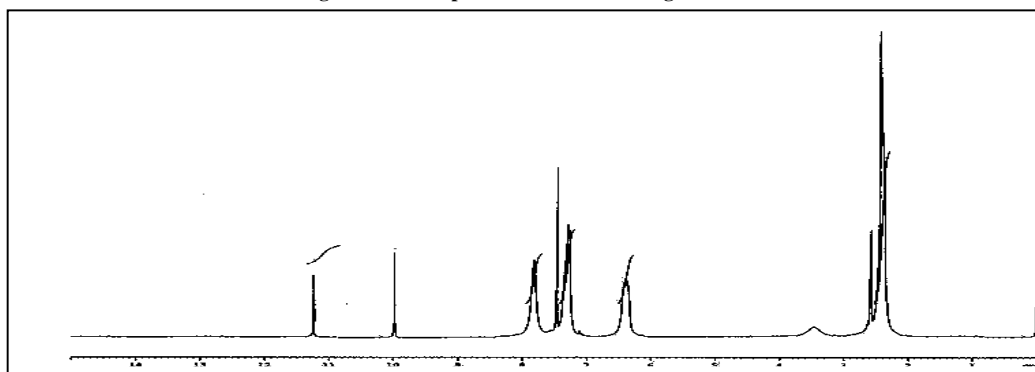
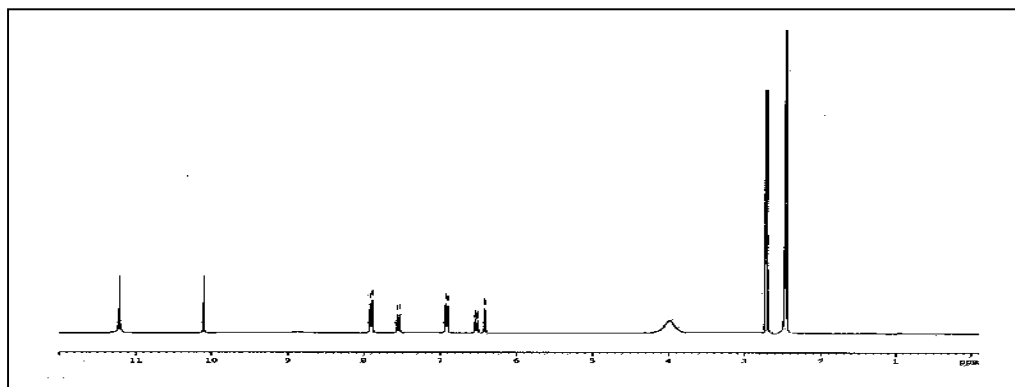
Fig 4 :  $^1\text{H}$  NMR Spectrum of RAPPTH LigandFig 5:  $^1\text{H}$  NMR Spectrum of RAPPTH-Co Complex

Table 1: Thermal Analytical Data of the RAPPTH metal complexes

Complex X=H <sub>2</sub> O	M.Wt in grams	Wt of the complex in mgs	Stage	Temp in range °C	Probable assignment	Mass loss (%)	Total loss (%)
MnL <sub>2</sub> 2X L=C <sub>16</sub> H <sub>16</sub> O <sub>3</sub> N <sub>2</sub>	657.57	8.0	1 2 3	120.12-221.90 284.46-575.03 Above 640.47	Loss of 2H <sub>2</sub> O molecule Loss of 2L molecule Corresponds to MnO	5.474 83.541 9.864	89.0
CoL <sub>2</sub> 2X L=C <sub>16</sub> H <sub>16</sub> O <sub>3</sub> N <sub>2</sub>	661.560	7.50	1 2 2	130.41-230.23 320.23-509.48 Above 600	Loss of 2H <sub>2</sub> O molecule Loss of 2L molecule Corresponds to CoO	5.44 85.583 9.877	91.0
NiL <sub>2</sub> 2X L=C <sub>16</sub> H <sub>16</sub> O <sub>3</sub> N <sub>2</sub>	661.269	8.00	1 2 3	145-230 290-720 Above 792	Loss of 2H <sub>2</sub> O molecule Loss of 2L molecule Corresponds to NiO	5.44 83.218 11.14	88.6
CuL <sub>2</sub> 2X L=C <sub>16</sub> H <sub>16</sub> O <sub>3</sub> N <sub>2</sub>	666.086	8.00	1 2 3	120-230 298-793 Above 839	Loss of 2H <sub>2</sub> O molecule Loss of 2L molecule Corresponds to CuO	5.404 78.816 15.64	84.2

Table 2. Molar conductivity of RAPPTH -Metal complexes:

Metal complexes	Molar conductance(ohm <sup>-1</sup> cm <sup>2</sup> mol <sup>-1</sup> )
RAPPTH-Mn	6.50
RAPPTH-Co	5.00
RAPPTH-Ni	7.20
RAPPTH-Cu	4.80

Table 3. Analytical data of the RAPPTH ligands and its metal complexes

Molecular Formula X=H <sub>2</sub> O	Molecular Weight	Colour	Yield in %	Melting Point in °C	Elemental analysis							
					Carbon %		Hydrogen %		Nitrogen %		Metal %	
					Calc.	Found	Calc.	Found	Calc.	Found	Calc.	Found
L=C <sub>16</sub> H <sub>16</sub> O <sub>3</sub> N <sub>2</sub> RAPPTH	284.315	Light brown	84	218-220	67.53	67.12	5.62	5.40	9.84	9.47	-	-
[Mn L <sub>2</sub> ] <sub>2</sub> X (RAPPTH-Mn)	651.57	Thick green	72	318-340	58.39	58.04	5.17	4.67	8.52	8.13	8.35	8.01
[Co L <sub>2</sub> ] <sub>2</sub> X (RAPPTH-Co)	661.560	Light Yellow	81	315-322	58.04	57.64	5.14	4.82	8.46	8.07	8.90	8.73
[Ni L <sub>2</sub> ] <sub>2</sub> X (RAPPTH-Ni)	661.269	Brown	69	298-310	58.07	57.66	5.14	4.82	8.47	7.98	8.86	8.29
[Cu L <sub>2</sub> ] <sub>2</sub> X (RAPPTH-Cu)	66.086	Green	78	312-320	57.65	57.42	5.10	4.91	8.40	8.19	9.52	9.09

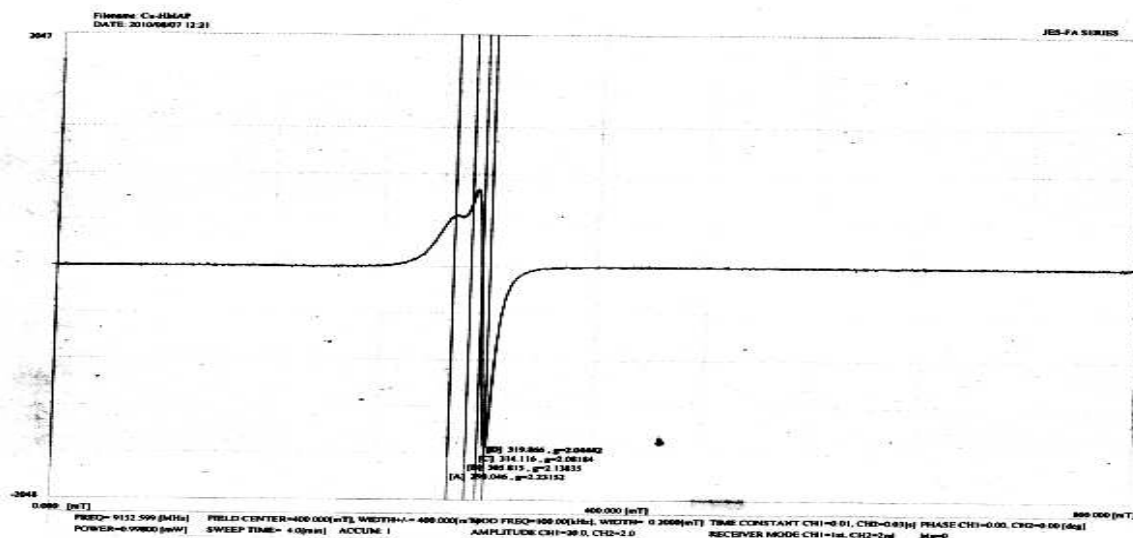


Fig 7: ESR Spectrum of RAPPTH-Cu (II) Complex

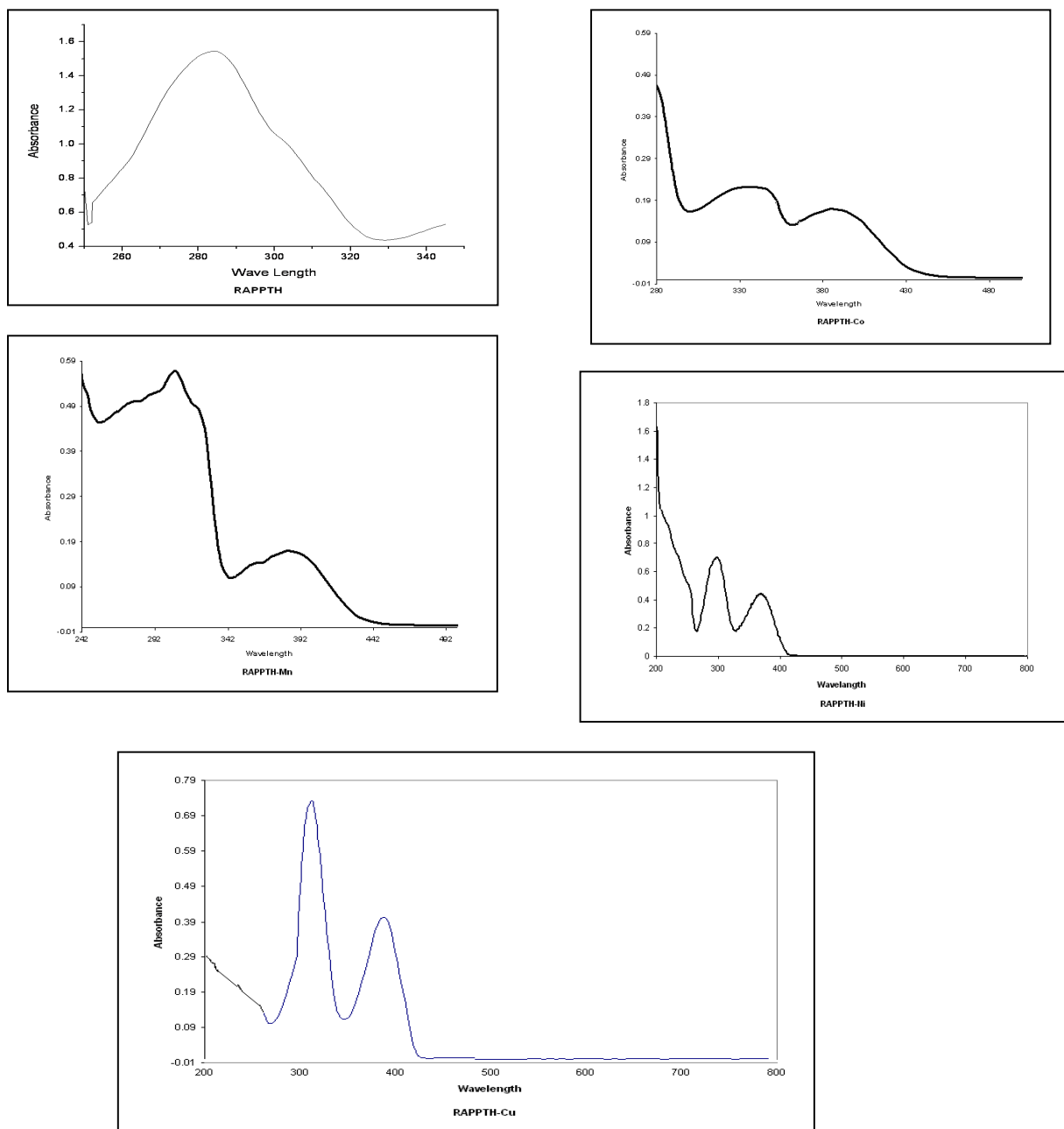


Fig 9: Electronic Spectral data of RAPPTH ligand and its Metal Complexes

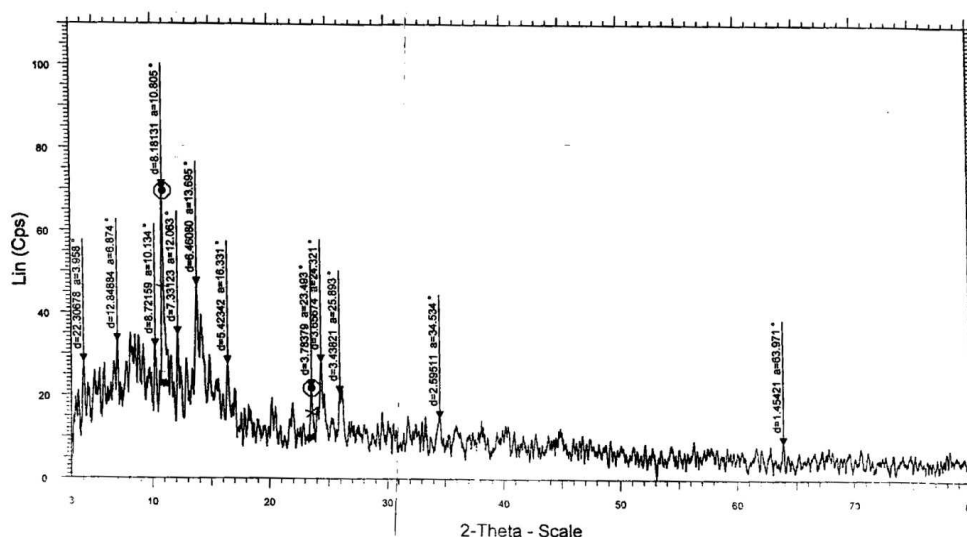


Fig 8: RAPPTH-Cu (II) Complex

### 3.2 UV Spectral & Magnetic measurements studies

UV spectrum shows the electronic transition for the ligand at 284 nm. On complexation of ligand with the different metal ions like Mn (II), Co (II), Ni (II) and Cu (II) new bands are appeared at 385 nm, 388 nm, 372 nm and 391 nm for different metal ions like Mn (II), Co (II), Ni (II) and Cu (II) corresponding to the transitional charge transfer from the ligand to the different metal ions. Bands are appeared in the region of 370-391 nm for all metal complexes [15-16], which are assigned to charge transfer transition ( $L \rightarrow M$ ). Based on the results octahedral structure is proposed for Mn (II), Co (II), Ni (II) and Cu (II) complexes. The magnetic moments of the RAPPTH Mn (II), Co (II), Ni (II) and Cu (II) complexes are 5.28, 4.82, 3.15 and 1.81 BM respectively. Octahedral geometry is proposed [17] for Mn (II), Co (II), Ni (II) and Cu (II) complexes by Magnetic measurements & UV Spectral studies.

### 3.3 Thermal behavior of Mn (II), Co (II), Ni (II) and Cu (II) Metal complexes of RAPPTH

The Thermo gravimetric studies of all the complexes were carried out in air at a heating rate of 10°C per minute. The thermo analytical data is summarized in Table.1. The thermal decomposition of the complexes proceeds in three stages. The Mn (II), Co (II), Ni (II) and Cu (II) complexes are thermally stable up to 120, 130, 145 and 120 respectively [18-19]. The first stage of decomposition to endothermic dehydration of complexes by the loss of two water molecules occur in the temperature range 120-221°C, 130-230°C, 145-230°C and 120-230°C respectively. The intermediates formed are stable up to 230, 284, 320, 290 and 300°C. The second decomposition with exothermic peak by the loss of ligand moiety occurs in the temperature range 290-575°C, 320-510°C, 290-722°C and 300-793°C. The solid residues above 664, 530, 650 and 620°C were recognized as Mn, Co, Ni and Cu metal oxides respectively. In all the complexes [20] the final products are metal oxides.

### 3.4 Conductivity Measurements of RAPPTH metal complexes

The p-Toluic hydrazone and Resacetophenone (RAPPTH) metal complexes are freely soluble in dimethyl formamide (DMF). Perform conductivity measurements by using DMF. The solid metal complexes were transferred into 25 ml standard flask and dissolved in DMF and made up to the mark. This complex solution was poured into a clean and dry 100 ml beaker and measures the molar conductance values. The molar conductance values are 6.50, 5.00, 7.20 and 4.80 respectively for RAPPTH-Mn (II), Co (II), Ni (II) and Cu (II) metal complexes [21]. These values are predictable non-electrolytic nature for these complexes.

### 3.5 ESR Spectral studies of RAPPTH-Cu complex

ESR spectra of RAPPTH-Cu (II) ion complex is presented in Fig.7 and data is presented in table.4. The ESR anisotropic spectrum obtained for this metal complex in DMF at LNT..

In the low temperature spectrum, four small intensity peaks are identified which are considered to originate from gll component. The G value of the present complex is greater than four and suggests that there is no interaction between Copper-Copper centers in DMF medium. The ESR spectroscopic parameters  $g_{\parallel}$ ,  $g_{\perp}$ ,  $A_{\parallel}^*$  and  $A_{\perp}^*$  of the complexes and the energies of d-d transitions which are used to evaluate the orbital reduction parameters ( $K_{\parallel}$ ,  $K_{\perp}$ ), the bonding parameters ( $\alpha^2$ ), the dipolar interaction (P). The observed  $K_{\parallel} < K_{\perp}$  indicates the presence of out of plane  $\pi$ -bonding. The  $\alpha^2$  value for the present chelate is 0.524. It indicates that the complex has covalent character. This shows a significant covalency in the inplane  $\sigma$  bonding. The reduction of P values from the free ion value ( $0.036\text{cm}^{-1}$ ) might be attributable to the strong covalent bonding. The value of P obtained for the present complex is  $0.0159\text{cm}^{-1}$  and remain constant with bonding of Copper ions to oxygen and nitrogen donor atoms respectively. The shape of ESR lines, ESR data together with the electronic spectral data proposed [22-24] an octahedral geometry for RAPPTH-Cu complex.

Table.4. ESR spectral data of RAPPTH-Cu complex

Parameters	$g_{\parallel}$	$g_{\perp}$	$g_{ave}$	G	$A_{\parallel}^*$	$A_{\perp}^*$	$A_{ave}^*$	d-d	$K_{\parallel}$	$K_{\perp}$	P	$\alpha^2$
RAPPTH-Cu	2.252	2.063	2.126	4.006	0.0198	0.0035	.0089	1622	0.616	0.742	0.0159	0.524

Table 5. X-ray Diffraction data of RAPPTH-Cu complex

S.No.	d expt	d Calc	2 $\theta$ expt	2 $\theta$ Calc	h k l
1	22.3067	22.3015	3.9577	3.9263	1 0 0
2	12.8488	12.8239	6.8737	6.8526	2 0 0
3	8.7215	8.6959	10.1337	10.1257	3 3 0
4	8.1831	8.1734	10.8049	10.8026	2 2 1
5	7.3312	7.3215	12.0621	12.0517	3 2 2
6	6.4608	6.4579	13.6941	13.6927	4 3 0
7	5.4234	5.4198	16.3304	16.3298	4 4 2
8	3.7839	3.7697	23.4919	23.3926	6 5 3
9	3.6567	6.6478	24.3197	24.3086	7 5 2
10	3.4382	3.4297	25.8916	25.8912	7 6 1
11	2.5951	2.5902	34.5335	34.5296	9 6 5
12	1.4542	1.4523	63.9676	63.9521	9 9 8

### 3.6 Powder XRD study of RAPPTH-Cu complex

The diffractogram (12 diffractions) reflects Fig. 8 between 3-80 ( $2\theta$ ) values for Cu complex. Where  $\theta$  is Brages angle all the main peaks are indicted and calculated values of Miller indeces (h k l) along with observed d-specified and reveled intensities are specified in the Fig. 8. All the peaks have been indexed  $2\theta$  values compared in graph. Comparison values revels that there are good agreement values of  $2\theta$  and d-values. The powder x-ray diffraction data showed identical features with very poor crystallinity. The patterns are qualitative and dispersive in intensity for Cu complex. The X-RD patterns [25-26] are used to explain qualitatively the degree of crystallinity. X-ray Diffraction data of RAPPTH-Cu complex is presented in table 5.

### 3.7 Anti bacterial activity

The biological activity studies of the RAPPTH ligand and their metal complexes [27-28] were tested against bacteria. All the synthesized compounds were evaluated for their *in vitro* antibacterial activity studies against four bacteria's namely *E.coli*, *staphylococcus aureus*, *pseudomonas aereuginosa* and *B. subtiles*. The comparison studies of antibacterial activity these metal complexes [29-30] and ligand with the reference of streptomycin drug. The biological activity data present in Table 6. The antibacterial activity studies clearly suggest that the most of the metal chelates [31] are shows a good antibacterial activity when compared to streptomycin. RAPPTH-Cu (II) complex showed high activity compared to all the remaining synthesized compounds.

Table 6. Antibacterial Activity of RAPPTH Schiff base Ligand and its metal complexes

Compound	<i>E. Coli</i>	<i>B. Subtilis</i>	<i>Staphylococcus aureus</i>	<i>Pseudomonas aereuginosa</i>
RAPPTH Ligand	6	6	7	5
RAPPTH-Mn(II)	14	10	12	11
RAPPTH-Co(II)	9	7	6	8
RAPPTH-Ni(II)	10	9	10	9
RAPPTH-Cu (II)	18	15	17	15
Streptomycin	22	24	20	22

## CONCUSION

In this paper, the coordination chemistry of a Schiff base ligand obtained from the reaction of p-toulic hydrazide and resacetophenone. The RAPPTH - Mn (II), Co (II), Ni (II), and Cu (II) metal complexes have been characterized by spectral and analytical data. The IR, electronic transition and g tensor data lead to the conclude, that the geometry of the complexes of Mn (II), Co (II), Ni (II), and Cu (II) are octahedral in nature and hence the structure of RAPPTH Schiff base metal complexes are given in Fig.9. In all the complexes, the ligand acts as bidentate. The results of in –vitro biocidal activities of the ligand and its metal complexes clearly show antibacterial activity against the tested organisms. The antibacterial activity indicates the metal complexes have more biological activity than free ligands. All the metal chelates are found to be non-electrolytes

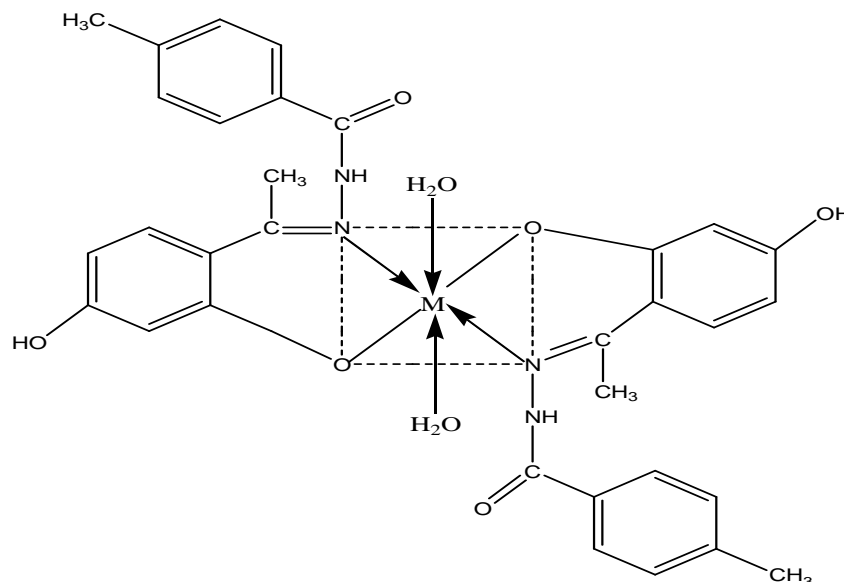


Fig 11: Structure of RAPPTH Metal Complex  
 $M = Mn^{+2}, Co^{+2}, Ni^{+2}$  and  $Cu^{+2}$ .

## REFERENCES

- [1] A. K. Singh, O. P. Pandey, and S. K. Sengupta, *Spectrochimica Acta—Part A: Molecular and Biomolecular Spectroscopy*, **2012**, 85, 1–6.
- [2] G. B. Bagihalli, P. G. Avaji, S. A. Patil, and P. S. Badami, *European Journal of Medicinal Chemistry*, **2008**, 43, 2639–2649.
- [3] R. Franski, B. Gierczyk, G. Schroeder, S. Pieper, A. Springer, and M. Linscheid, *Central European Journal of Chemistry*, **2007**, 5, 316–329.
- [4] T. Mangamamba, M. C. Ganorkar, and G. Swarnabala, *International Journal of Inorganic Chemistry*, **2014**, 4, 21–30.
- [5] Rai B.K, Kumari R, *Orient J Chem* **2013**;29(3).
- [6] M. H. Khalil, Eman H. Ismail, Gehad G. Mohamed, Ehab M. Zayed, Ahmed Badr. *Open Journal of Inorganic Chemistry*, **2012**, 2, 13–21.

- [7] Surati, K.R. and Thake, B.T. *Spectrochimica Acta A: Molecular and Biomolecular Spectroscopy*, (2010) **75**, 235-242.
- [8] Bharty, M.K., Srivastava, A.K., Dulwere, R., Butcher, R.J. and Singh, N.K. *Polyhedron*, (2011) **30**, 990-996.
- [9] Freiria, A., Bastida, R., Valencia, L., Macias, A. and Lodeiro, C. *Inorganica Chimica Acta*, (2006) **359**, 2383-2394.
- [10] Mobinikhaledi, N. Forughifar, and M. Kalhor, *Turkish Journal of Chemistry*, **2010**, 34, 367-373.
- [11] D. Prakash, C. Kumar, *Journal of the Indian Chemical Society*, **2009**, 86, 1257-1261.
- [12] P. Murali Krishna, K. Hussain Reddy, J. & S. Dayananda, *Transition Metal Chemistry*, **2008**, 33, 5, 661-668.
- [13] A.Kashem Liton and M. Rabiul Islam, *Journal of Pharmacology*, **2006**, 1-10.
- [14] A.D.Kulkarni, S. A. Patil, and P. S. Badami, *Journal of Sulfur Chemistry*, **2009**, 30, (2), 145-159.
- [15] Wu Chen, Yuguang Li, Yongming Cui, Xian Zhang, Hai-Liang Zhu, Qingfu Zeng, *European Journal of Medicinal Chemistry* **2010**, 45, 4473-4478
- [16] Zhong-Lu You, Da-Hua Shi, Chen Xu' Qiang Zhang, Hai-Liang Zhu, *European Journal of Medicinal Chemistry*, **2008**, 43, 862- 871.
- [17] Manjula Patil, Rekha Hunoor, Kalagouda Gudasi, *European Journal of Medicinal Chemistry* **2010**, 45, 2981-2986.
- [18] Kiran Singh, Manjeet Singh Barwa, Parikshit Tyagi, *European Journal of Medicinal Chemistry* **2006**, 41, 147-153.
- [19] Gajendra Kumar, Dharmendra Kumar, Shoma Devi, Rajeev Johari, C.P. Singh, *European Journal of Medicinal Chemistry*, **2010**, 45, 3056-3062.
- [20] Gehad G. Mohamed, M.M. Omar, Amr A. Ibrahim, *European Journal of Medicinal Chemistry* 44 (2009) 4801-4812.
- [21] Lotf Ali Saghatforoush Firoozeh Chalabian, Ali Aminkhani, Ghasem Karimnezhad, Sohrab Ershad, *European Journal of Medicinal Chemistry* 44 (2009) 4490-4495
- [22] Zeng-Chen Liu, Bao-Dui Wang, Zheng-Yin Yang, *European Journal of Medicinal Chemistry* 44 (2009) 4477-4484.
- [23] Mehmet So'nmez, Metin Celebi, Ismet Berber c, *European Journal of Medicinal Chemistry* 45 (2010) 1935-1940.
- [24] Adnan S. Abu-Surrah, Kayed A., *European Journal of Medicinal Chemistry* 45 (2010) 471-475.
- [25] Zeng-Chen Liu, Bao-Dui Wang, Bo Li, Qin Wang, Zheng-Yin Yang Tian-Rong Li, Yong Li, *European Journal of Medicinal Chemistry* 45 (2010) 5353-5361.
- [26] Gangadhar B. Bagihalli Prema S. Badami, *European Journal of Medicinal Chemistry* 43 (2008) 2639-2649
- [27] Ajaykumar Kulkarni, Sangamesh A. Patil, *European Journal of Medicinal Chemistry* 44 (2009) 2904-2912
- [28] Ayed S. Al-Shihri, *Spectrochimica Acta Part A* 60 (2004) 1189-1192
- [29] Z.H. Abd El-Wahab, M.R. El-Sarrag, *Spectrochimica Acta Part A* 60 (2004) 271-277.
- [30] Chi-Ming Che, Jie-Sheng Huang, *Coordination Chemistry Reviews* 242 (2003) 97-113.
- [31] N. Ramana, S. Sobhaa, A. *Spectrochimica Acta Part A* 78 (2011) 888-898.



Scholars Research Library

Der Pharmacia Lettre, 2015, 7 (12):94-101  
(<http://scholarsresearchlibrary.com/archive.html>)



## Development and validation of stability indicating HPLC-MLC determination of dopamine agonist drug pramipexole

K. V. V. Srikanth<sup>1</sup>, K. R. S. Prasad<sup>1</sup>, J. V. Shanmukha Kumar<sup>1</sup> and K. Suresh Babu<sup>2\*</sup>

<sup>1</sup>Department of Chemistry, K L University, Vaddeswaram, India

<sup>2</sup>Department of Chemistry, Mallareddy Engineering College, India

### ABSTRACT

A simple rapid, accurate and stability indicating HPLC method is described for the quantification of Pramipexole in bulk drug. Chromatographic quantification of Pramipexole is achieved on Xterra RP 18 column using isocratic Mobile phase of Sodium dodecyl sulphate, ethyl acetate, 2-propanol and 0.2% Trifluoroacetic acid in water (33:66:1:900, w/w) at a flow rate of 1.0 mL/min with injection volume 10 µL. Column temperature was maintained at 30°C. The proposed method was validated for system suitability, linearity, precision, specificity, stability, Limit of Detection and Limit of quantification. Degradation studies were conducted as per ICH guidelines. Limit of Detection and Limit of Quantification of Pramipexole were 2.96 µg/mL and 6.85 µg/mL respectively. The proposed method can be used for routine analysis and stability studies.

**Keywords:** Micellar liquid chromatography, Pramipexole, microemulsions, stability indicating methods, degradation, ICH guidelines.

### INTRODUCTION

Pramipexole is a dopamine agonist which works by mimicking the effects of dopamine in the patient's brain and allows for increased control over muscle movements. It belongs to the non-ergoline class and is indicated for treating early-stage Parkinson's disease (PD) and restless legs syndrome (RLS). Pramipexole, (S)-2-amino-4,5,6,7-tetrahydro-6-(propylamino) benzothiazole has good oral bioavailability (greater than 90%) indicating that it undergoes little first pass metabolism. <sup>[1]</sup>

Several chromatographic methods in conventional reverse phase mode are available for quantification of pramipexole are available in literature. <sup>[2-9]</sup> But, Till date, no stability-indicating HPLC assay method for the determination of pramipexole is available in the literature by using novel liquid chromatography technique, Micellar Liquid chromatography. So, It was felt necessary to develop a stability indicating liquid chromatography method for the determination of pramipexole as bulk drug and separate the drug from the degradation products. Therefore, the aim of the present study was to develop and validate a stability-indicating HPLC assay method for pramipexole as bulk drug as per ICH guidelines by using Novel HPLC technique Micellar liquid chromatography.

Micellar liquid chromatography (MLC) is one of the powerful tools available to face the challenges posed by different complex substances as an alternate to ion pair chromatography. MLC is a kind of alternate technique for reverse phase chromatography and ion pairing chromatography. It is superior in terms of extraction of active components from complicated matrixes (serums, biological fluids, gels, etc.,) and retention of mixtures of charged

ionic species to highly lyophilic and hydrophobic samples with adequate resolution. Micellar chromatography technique was introduced by Daniel W Armstrong in early seventies, though the technique was explored so long ago its potential breakthrough happened only in eighties, John Dorsey was the first person to address the issue of efficiency.<sup>[10]</sup>

Micellar chromatography involves Micellar solutions of surfactants (anionic, cationic, neutral and amphoteric salts) slightly above their critical micellar concentration, (CMC), where specific ordered arrangement of micelles was formed. In MLC, microemulsions were often used in mobile phases with slight amounts of organic modifiers, for better efficiency. These micro emulsions were often bicomponent solutions, if a third component oil (for water in oil emulsion), is added in excess its solubility/miscibility exceed the limit and results in the formation of emulsion, a biphasic solution. However the water droplets dispersed in continuous oil phase, were unstable and results in coalescence. In order to retain the emulsion characteristics another component is added to the water/oil/surfactant mixture, which was a linear medium chain alcohols such as n-butanol, n-pentanol etc., These alcohols often associate with the surfactants and reduces the interfacial tension between oil and water, and increases the interfacial area, which leads to the formation of a very stable emulsion. These micro emulsions so formed were so stable ranging from months to years, and almost it becomes impossible for the two components oil and water to get separated. Though the surfactants/ion pairing agents used in ion pairing chromatography need not reach critical micellar concentration, (CMC), it's necessary for the surfactant concentration to reach CMC in MLC

## MATERIALS AND METHODS

### *Materials and Reagents*

Pramipexole drug substance and its related compounds were obtained gratuitously from f Mylan Laboratories Ltd. HPLC grade 2-propanol and Ethylacetate and 100% Trifluoro acetic acid were purchased from Merck, Darmstadt, Germany. GR grade Sodium dodecyl sulphate, Hydrochloric acid, 30% hydrogen peroxide and Sodium Hydroxide were purchased from Rankem Chemicals, India. High purity Water was collected from Millipore Milli-Q-plus water purification system. All the samples used for this study were of greater than 95% purity.

### *Instrumentation*

HPLC System used for method development and Method validation was of Waters make, equipped with 2695 binary pump with auto sampler, and 2996 photo diode array detector with Empower 2 software (Waters Corporation, MA, USA). The result chromatogram was processed using Empower 2 Software on Pentium 4 processor Dell make computer. Sonicator of Bandelin sonorex, model DK 255 P lit 5, (BANDELIN electronic GmbH & Co. KG) were used in sample preparation for extraction of active components from the sample matrix. Analytical Precision balance of Mettler Toledo (Mettler Toledo India Pvt. Ltd.,) was used for weighing.

### *Chromatographic conditions*

The chromatographic column used for the study was Waters Xterra (150x4.6mm), with 5µm particles. Mobile phase constitutes 33g SDS, 1g ethylacetate, 66g 2-propanol and 900 g of 0.2% Trifluoroacetic acid buffer in aqueous media. The column temperature was maintained at 30°C temperature and the detector wavelength was set at 264 nm. The injection volume was 10 µL with mobile phase as such as diluent.

### *Standard solutions*

Stock solutions of Pramipexole drug substance was prepared by dissolving appropriate amounts in Mobile phase to achieve 1 mg/ml concentration. Working standard solutions were prepared from the above stock solutions by diluting required aliquots to cover the concentration range of 25µg/mL to 500µg/mL for Pramipexole. The concentration for Target solution was 0.2 mg/mL of Pramipexole. All the solutions were prepared in mobile phase.

### *General Procedure*

Into the series of volumetric flasks Pramipexole drug substance is weighed added few mL of mobile phase to disperse the material and further added sufficient quantity of mobile phase and sonicated for 15 minutes. Allowed the flasks to settle down and diluted to the volume to achieve 1mg/mL of pramipexole. The stock solutions were diluted with mobile phase further to achieve target concentration. All the solutions diluted are filtered with 0.45 µ Nylon filter discarding initial 2mL filtrate, transferred to HPLC vial and capped, which was used for analysis.

**Assay Validation**

- The selectivity of the proposed method has been studied by analyzing degraded sample solutions subjecting Pramipexole drug substance to forced degradation as per ICH guidelines.
- The inter-day and intra-day precision was evaluated by injecting six replicates of control drug substance on two different days
- The accuracy of the method was evaluated by analyzing triplicates of control drug substance at four different concentrations of Pramipexole.
- Recovery was calculated by comparing peak area obtained from the drug substance samples and from standard solution at target concentration.
- The calibration curve of Pramipexole was constructed by plotting peak area vs drug concentration (seven levels), using least square regression equation
- The limits of quantification and detection were calculated by standard deviation method.<sup>[11]</sup>

$$LOQ = \frac{10 \times SD}{Slope}$$

$$LOD = \frac{3.3 \times SD}{Slope}$$

- The stability of Pramipexole was established upto 48hours by comparing peak area of old standard solution with the peak area obtained from freshly prepared standard solution.

**RESULTS AND DISCUSSION****Method development**

Two kinds of Mobile phases are used commonly in Micellar chromatography, one is oil in water (o/w) and the other one is water in oil (w/o). Each one has got its own advantage over the other and both the kinds are contradictory to each other with respect to elution and solubilization efficiencies.<sup>[12]</sup> For the current study Oil in water MELC (micro emulsion liquid chromatography) is chosen. For the surfactant, sodium dodecyl sulfate sodium salt, traditional surfactant was used in preparation of aqueous micellar solution, as the current compound has basic functional group, anionic surfactant was chosen.<sup>[13]</sup>

Co solvent 2-propanol was used for non polarity. Other additives include ethyl acetate for imparting selectivity to the mobile phase, Trifluoroacetic acid (0.2%) a strong acid for reducing peak broadening of the active.<sup>[14-17]</sup>

Different Stationary phases with octyl decyl packing on silica substrate were evaluated from different suppliers. However, as the pH is beyond conventional range (<2), special columns with extended pH stability were used for the study. Out of which Xterra (150x4.6mm) with 5µm particles was found to be suitable with optimum resolution and run time. Different trials were taken with different proportions of the afore mentioned components, out of which the most promising composition was 33g SDS, 1g ethylacetate, 66g 2-propanol and 900 g of 0.2% Trifluoroacetic acid buffer in aqueous media. Using the optimized conditions Pramipexole was well resolved from its degradants and quantified. All the solutions were monitored in the range 200nm to 800nm, and quantified at 264nm. Column temperature has little impact on retention behavior of pramipexole and maintained at 30°C to suffice laboratory conditions.

**Assay Validation**

The proposed method was fully validated according to ICH guidelines. The validation parameters are summarized (Table 1)

**Selectivity and Specificity of the Method**

The selectivity and Specificity of the proposed method has been studied by analyzing Impurities Spiked standard solution and degradation samples and evaluating peak purity of the main analyte in presence of its degradants in the

complete UV-VIS range 200nm to 800nm (Figure 1 and 2). The summary of forced degradation studies reveal the proposed method is capable of resolving the active from its degradants. (Table 2 and 3)

#### **Sensitivity**

The limit of detection and limit of Quantification were established by regression analysis by least squares method. LOD calculated from the formula was found to be in 2.96  $\mu\text{g/mL}$ . Similarly LOQ is calculated using similar formula and was found to be in 6.85  $\mu\text{g/mL}$  (Figure 3).

#### **Linearity**

Linear calibration plot for MLC assay method was obtained over the ranges from 10% to 250% of the target concentration and the correlation coefficient obtained was greater than 0.99999 for Pramipexole. The result confirms excellent correlation between the chromatographic response and concentration of the analyte.

#### **Reproducibility and Repeatability of the method**

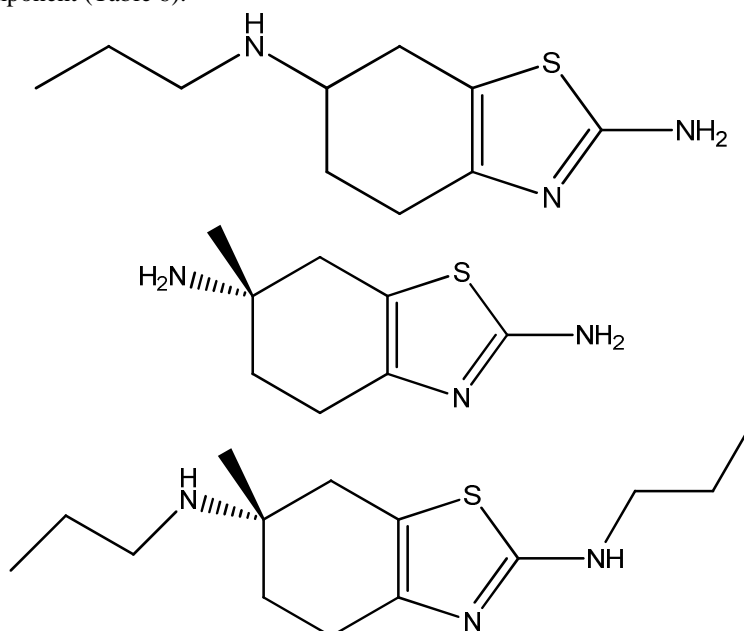
Results (Table 3) showed that there were high intra- and inter-day precisions with %RSD 0.4 and 0.5 respectively (Figure 4). Intra-day precision was assessed by injecting of the six replicates of control drug substance and Inter-day precision was established in the same way by injecting the standard solution on second day (Table 4).

#### **Accuracy**

The % Recoveries for Pramipexole (Figure 5) was found to be within the range 98% to 102% showing the methods closeness to true value (Table 5).

#### **Solution stability and mobile phase stability**

The % RSD of Pramipexole during solution stability and Mobile phase stability evaluation study periods were within 2%. No significant change in % Assay values with respect to initial % assay value and %RSD less than 2% was observed. The study confirms that the sample solutions and Mobile phase solutions were stable up to 48 hr showing excellent stability of micro emulsions. Also the solution does not resulted in any precipitation or settling of any mobile phase component (Table 6).



SCHEME 1 Chemical structure of Pramipexole, Related compound-A and Related compound-B (Ph. Eur)

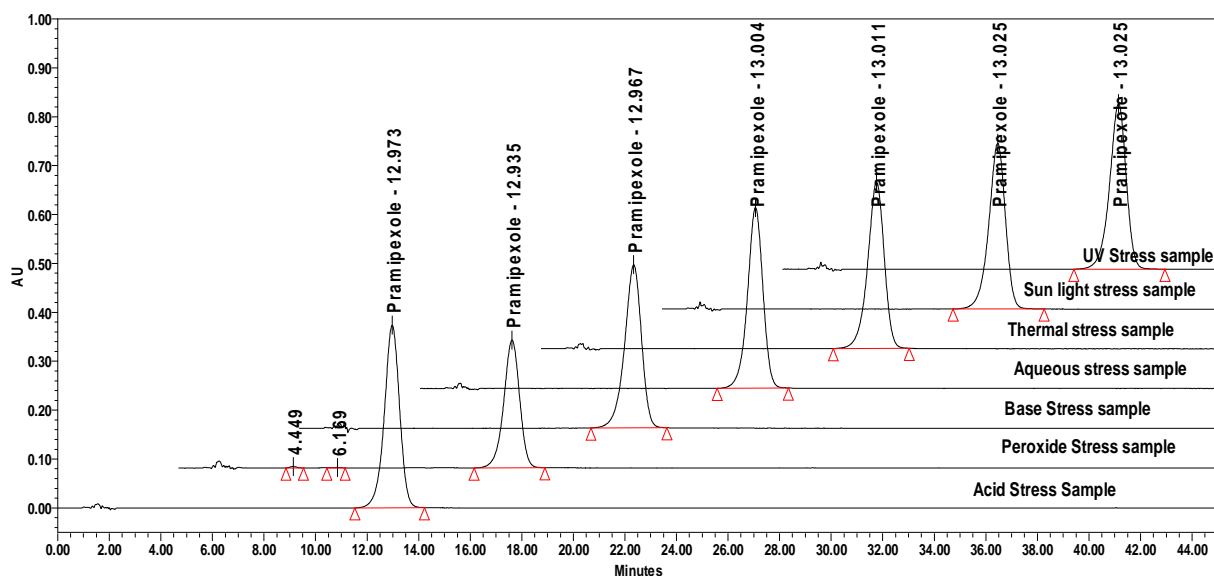


FIGURE 1 Chromatograms of forced degradation

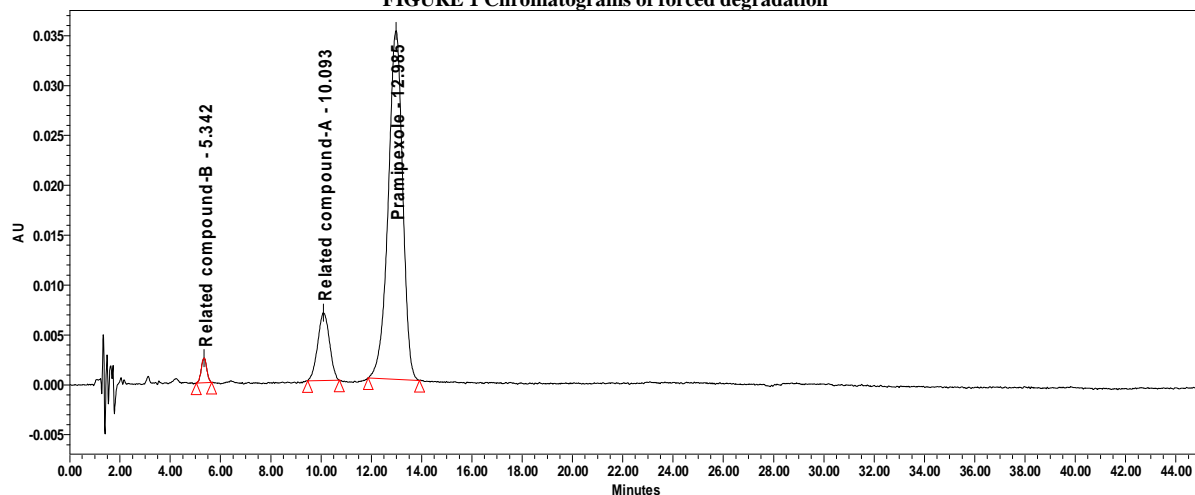


FIGURE 2 Chromatograms of Impurity spiked Pramipexole

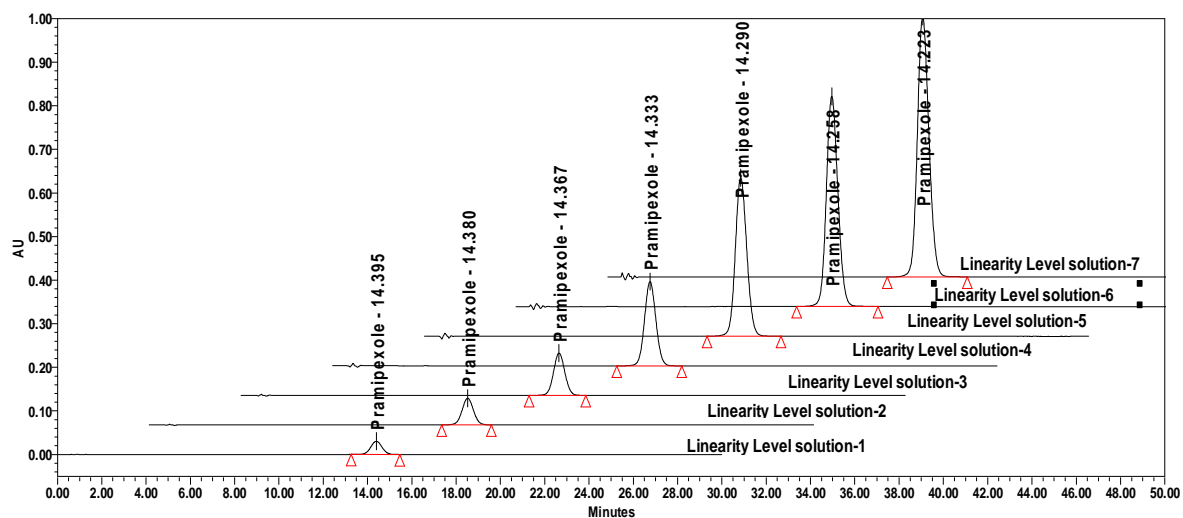


FIGURE 3 Chromatograms of Linearity for pramipexole

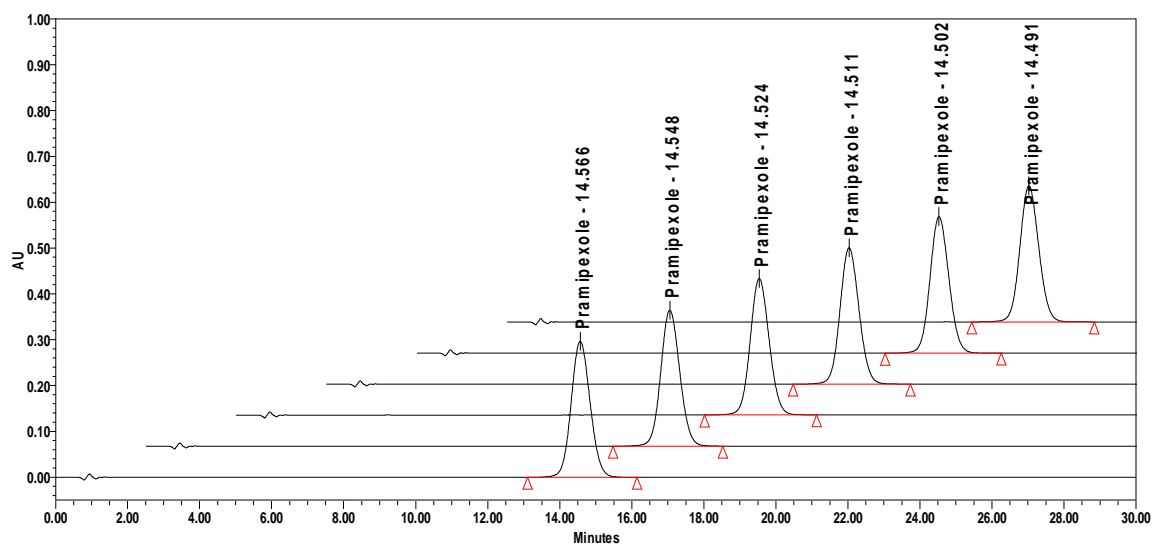


FIGURE 4 Chromatograms of Precision

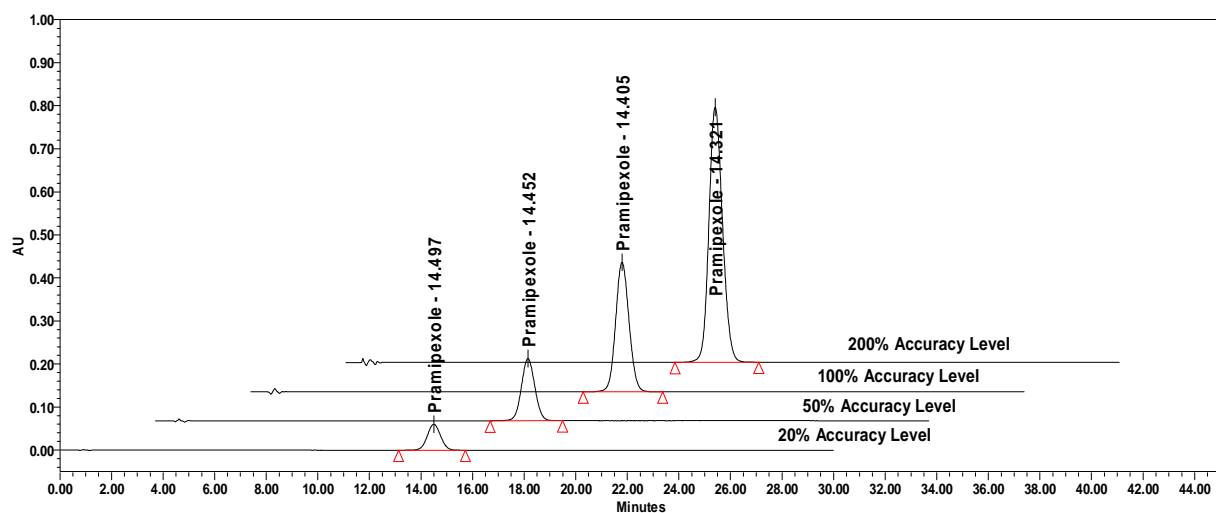


FIGURE 5 Chromatograms of Pramipexole recovery with the proposed method

Table 1 Validation Data for the determination of Pramipexole by the Proposed

Method	Pramipexole
Validation Parameter	
System Suitability	
%RSD	0.1
N	3602
USP Tailing factor	
Linearity range (µg/mL)	25-500
Regression equation	
Intercept (a)	$3.8 \times 10^3$
Slope (b)	$4.4 \times 10^4$
Correlation coefficient (r)	0.99999
Accuracy (%RSD)	
Precision	
Within day	0.4
Between day	0.5
Limit of quantification (µg/mL)	6.85
Limit of detection (µg/mL)	2.96

%RSD: % Relative standard deviation

N: Theoretical plates

**Table 2 Forced Degradation conditions for Pramipexole**

Validation Parameter	
Forced Degradation	Stress conditions
Acid Stress study	Refluxed with 1N HCl for 1hour at 60°C
Base Stress study	Refluxed with 1N NaOH for 1hour at 60°C
Oxidation Stress study	Stressed with 1% Peroxide at Room temperature
Aqueous Stress study	Refluxed with water for 1hour at 60°C
Sun Light Stress study	Subjected to 1.2million lux hours intensity
UV irradiation	Subjected to 200Watt hours/ sqm intensity
Thermal Stress study	Subjected to 105°C for 24hours

*HCl: Hydrochloric acid*  
*NaOH: Sodium Hydroxide*

**Table 3 Forced Degradation for Pramipexole**

Validation Parameter	Pramipexole		
Forced Degradation	Purity angle	Purity threshold	Mass balance
Acid Stress study	0.115	0.297	99.0
Base Stress study	0.120	0.305	98.2
Oxidation Stress study	0.138	0.322	98.3
Aqueous Stress study	0.109	0.301	99.5
Sun Light Stress study	0.116	0.308	99.3
UV irradiation	0.121	0.301	98.1
Thermal Stress study	0.132	0.310	97.9

**Table 4 Inter and Intra day Precision for Pramipexole**

	Inter day	Intra day
Average	98.2	97.9
%RSD	0.4	0.5
USP Tailing	1.08	1.08
N	3626	3638

**Table 5 Recovery for Pramipexole**

Validation Parameter	Pramipexole	
Recovery	Average	%RSD
20% Recovery Level	99.2	0.8
50% Recovery Level	98.1	0.6
100% Recovery Level	99.7	0.4
200% Recovery Level	99.4	0.6

**Table 6 Solution Stability and Mobile phase stability**

Validation Parameter	Pramipexole	
Solution Stability	%Assay	
	24 hours	48 hours
	99.2	98.6
Mobile phase stability	%RSD	
	24 hours	48 hours
	0.1	0.4

## CONCLUSION

A valid and stability indicating HPLC-MLC method was established for the quantification of Pramipexole. The proposed method showed satisfactory results for separation of Pramipexole from its related compounds and degradants, linear range. Also the proposed method ensured precise and accurate determinations of pramipexole. No interference from its related compounds and degradation products was observed. The proposed Micellar liquid

chromatography method was found to be applicable for assay of pramipexole in bulk drug and suitable for stability studies.

#### Acknowledgement

This study was conducted at Mylan Laboratories Ltd, Hyderabad. The authors thank the above company for the cordial support in supplying related equipment, chemicals and drug substance for conducting the study.

#### REFERENCES

- [1] P Lavudu, A Prameela-Rani, C Balashekaran, V Venumadhav, *Global.J.Med.Res*, **2012**, 12(4), 19-24.
- [2] T Deepan, K Paulambethkar, R Vijayalakshmi, M D Dhanaraju, *World.J.Chem* **2012**, 7(2), 59-63.
- [3] Vilas Chaudhary, Milind Ubale, *Int.J.Res.Pharm.Chem*, **2013**, 3(1), 134-139.
- [4] Syeda-Humaira, Akalankadey, S Appala-Raju, Syed-Sanaullah, *Der Pharmacia Lettre*, **2010**, 2(4), 315-325.
- [5] R Narendrakumar, G Nageswararao, P Y Naidu, *Der Pharmacia Sinica*, **2011**, 2(4), 125-135.
- [6] N Venkata Rajesh, Deepa Ramani, Durraivel, *Int.J.Pharm.Clin.Res*, **2013**, 5(1), 17-22.
- [7] C Vinodhini, V N D Malladi Sravani, M Bhanuprakash, Mantenasudhadevi, Mohamad Imran, Mohamed Omer Abdelaziz Osman, Chitra, K Uma Maheshwara Reddy, *Int.J.Res.Pharm.Biomed.Sc.* **2011**, 2(2), 680-686.
- [8] ICH Harmonized Tripartite Guideline, Q2(R1), International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, **2005**, 11.
- [9] Maria Ramble, Alegre, *Chrom.Res.Int*, **2012**.
- [10] M J Ruiz-Angel, J R Torres-Lapasio, M C Garcia-Alvarez-Coque, S Carda *Anal.Chem*, **2008**, 80(24), 9705-9713.
- [11] S Lopez-Grio, J J Baeza-Bacza, M C Garcia-Alvarez-Coque, *Chromatographia*, **1998**, 48(9), 655-663.
- [12] S Lopez-Grio, M C Garcia-Alvarez-Coque, W L Hinze, F H Berthod, *Anal.Chem*, **2000**, 72(20), 4826-4832.
- [13] M J Ruiz-Angel, S Carda-Broch, J R Torres-Lapasio, Garcia-Alvarez-Coque, *J Chrom.A*, **2009**, 1216(10), 1798-1814.
- [14] A Marsh, B J Clark, K D Altria, *Chromatographia*, **2005**, 61, 539-547.



## Stability Indicating Method Development and Validation for the Estimation of Rotigotine by RP-HPLC in Bulk and Pharmaceutical Dosage Form

P. GEETHA SWARUPA<sup>1</sup>, D. RADHA KRISHNA<sup>2</sup>, K.R.S.PRASAD<sup>1</sup> and K.SURESH BABU<sup>2\*</sup>

<sup>1</sup>Department of Chemistry, K.L. University, Vaddeswaram, 02, India.

<sup>2</sup>Department of Chemistry, Mallareddy Engineering College, 14, India.

\*Corresponding author E-mail: babuiict@gmail.com

<http://dx.doi.org/10.13005/ojc/3104786>

(Received: July 12, 2015; Accepted: August 04, 2015)

### ABSTRACT

The purpose of the investigation was to develop a new RP-HPLC Method for estimation of Rotigotine in pharmaceutical dosage forms. Chromatography was carried out on an BDS C-8 column (4.6 x 150mm, 5 $\mu$  particle size) with a isocratic mobile phase composed of 0.01N Potassium dihydrogen Ortho phosphate (adjusted to pH 4.8 with OPA solution), Acetonitrile (45:55v/v) at a flow rate of 1 mL/min. The column temperature was maintained at 30°C and the detection was carried out using a PDA detector at 224 nm. Validation parameters such as system suitability, linearity, precision, accuracy, specificity, limit of detection (LOD), limit of quantification (LOQ), Stability of sample and standard stock solutions and robustness were studied as reported in the International Conference on Harmonization guidelines. The retention time for Rotigotine was 2.691 min. The percentage recoverie of Rotigotine was 100.33%. The relative standard deviation for assay of patch was found to be less than 2%. The Method was fast, accurate, precise and sensitive hence it can be employed for routine quality control of patches containing both drugs in quality control laboratories and pharmaceutical industries.

**Key words:** Rotigotine, ICH guidelines.

### INTRODUCTION

Rotigotine (Rotigotine) is chemically named as 6-{propyl[2-(thiophen-2-yl)ethyl]amino}-5,6,7,8-tetrahydronaphthalen-1-ol. Rotigotine (Neupro) is a non-ergoline dopamine agonist indicated for the treatment of Parkinson's disease (PD) and restless legs syndrome (RLS) in Europe and the United States. It is formulated as a once-daily transdermal patch which provides a slow and

constant supply of the drug over the course of 24 hours<sup>1</sup>.

Rotigotine behaves as a partial or full agonist (depending on the assay) at all dopamine receptors listed, as an antagonist at the  $\alpha_{2B}$ -adrenergic receptor, and as a partial agonist at the 5-HT<sub>1A</sub> receptor<sup>2</sup>. Rotigotine is an agonist at all 5 dopamine receptor subtypes (D1-D5) but binds to the D3 receptor with the highest affinity. It is also an

antagonist at  $\alpha$ -2-adrenergic receptors and an agonist at the 5HT<sub>1A</sub> receptors. Rotigotine also inhibits dopamine uptake and prolactin secretion. There is no indication of a QT/QTc prolonging effect of Neupro in doses up to 24 mg/24 hours<sup>3</sup>.

Various UPLC & HPLC assay Methods are also reported in the literature for the estimation of Rotigotine<sup>4-5</sup>. According to literature survey there is official Method for the estimation of Rotigotine by RP-HPLC in patch dosage forms. Hence, an attempt has been made to develop better Method for estimation and validation of Rotigotine in formulation in accordance with the ICH guidelines<sup>6-10</sup>.

## EXPERIMENTAL

### Instrumentation

Chromatography was performed with Alliance waters 2695 HPLC provided with high speed auto sampler, column oven, degasser and & 2996 PDA detector to provide a compact and with class Empower-2 software.

### Reagents and chemicals

The reference sample of Rotigotine was provided as gift samples from Spectrum pharma research solutions, Hyderabad. HPLC grade acetonitrile, HPLC grade Methanol and all other chemicals were obtained from Merck chemical division, Mumbai. HPLC grade water obtained from Milli-Q water purification system was used throughout the study. Commercial formulation (**Neupro** ; Dosage: Rotigotine-4mg patch) were purchased from the local pharmacy.

### Chromatographic condition

The chromatographic separation was carried out under the isocratic conditions. Chromatographic separation was achieved by injecting a volume of 10 $\mu$ l of standard into BDS (150 x 4.6 mm, 5m) column. The mobile phase composed of 0.01N Potassium dihydrogen Ortho phosphate (adjusted to pH 4.8 with triethylamine solution), Acetonitrile (45:55v/v) was allowed to flow through the column at a flow rate of 1 ml per minute for a period of 7 min at 30°C column temperature. Detection of the component was carried out at a wavelength of 224 nm. The retention time of the component was found to be 2.691 min for

Rotigotine.

### Preparation of diluent solution

Diluent solution was prepared by mixing 500 ml of HPLC grade water with 500ml of Acetonitrile, in a 1000ml beaker and sonicated for 15min.

### Preparation of standard stock solution

Accurately Weighed and transferred 4mg of Rotigotine working Standards into 10ml clean dry volumetric flasks, add 3/4<sup>th</sup> vol of diluent, sonicated for 30 minutes and make up to the final volume with diluents. From the above stock solution, 1 ml was pipeted out in to a 10ml Volumetric flask and then make up to the final volume with diluent.

### Preparation of Working Standard Solutions

Aliquot of 0.25, 0.50, 0.75, 1, 1.25 & 1.5 mL were pipette out from stock solution into 10 mL volumetric flask separately for both ROTIGOTINE and volume was made up to 10 mL with diluent. This gives the solutions of 10, 20, 30, 40, 50 and 60 $\mu$ g/mL for Rotigotine.

### Sample preparation

5 patches were weighed and calculate the average weight of each patch then the weight equivalent to 1 patch was transferred into a 10mL volumetric flask, 7mL of diluent added and sonicated for 25 min, further the volume made up with diluent and filtered. From the filtered solution 1ml was pipette out into a 10 ml volumetric flask and made upto 10ml with diluent.

### Method validation

#### System suitability tests

To ensure the resolution and reproducibility of the HPLC system was adequate for the analysis, a system suitability test was established. Data from six injections of 10  $\mu$ L of the working standard solutions of Rotigotine was used for the evaluation of the system suitability parameters like tailing factor, the number of theoretical plates, retention time.

### Linearity

By appropriate aliquots of the standard Rotigotine solution with the mobile phase, six working solutions ranging between 10-60  $\mu$ g/mL

was prepared. Each experiment linearity point was performed in triplicate according to optimized chromatographic conditions. The peak areas of the chromatograms were plotted against the concentration of Rotigotine to obtain the calibration curve.

#### Accuracy

Recovery studies by the standard addition Method were performed with a view to justify the accuracy of the proposed Method. Previously analyzed samples of Rotigotine to which known amounts of standard Rotigotine corresponding to 50%, 100% and 150% of target concentration were added. The accuracy was expressed as the percentage of analyte recovered by the proposed Method.

#### Precision

Precision was determined as repeatability and intermediate precision (ruggedness), in accordance with ICH guidelines. The intra-day and inter-day precision were determined by analyzing the samples of Rotigotine. Determinations were performed on the same day as well as on consequent days.

#### Limit of detection and the limit of quantification

Limit of detection (LOD) and limit of quantification (LOQ) of Rotigotine was determined by calibration curve Method. Solutions of both Rotigotine was prepared in linearity range and injected in triplicate. Average peak area of three analyses was plotted against concentration. LOD and LOQ were calculated by using following equations.  $LOD = (3.3 \times Syx)/b$ ,  $LOQ = (10.0 \times Syx)/b$  Where Syx is residual variance due to regression; b is slope.

#### Robustness

The robustness of Rotigotine was performed by deliberately changing the chromatographic conditions. The organic strength was varied by  $\pm 5\%$ , column temperature was varied by  $\pm 5^\circ\text{C}$  and the flow rate was varied by  $\pm 0.1\text{ mL}$ .

#### Stability

The sample and standard solutions injected at 0 hr (comparison sample) and after 24

hr (stability sample) by keeping at ambient room temperature. Stability was determined by determining %RSD for sample and standard solutions.

#### Degradation studies

##### Oxidation

To 1 ml of stock solution of Rotigotine, 1 ml of 20% hydrogen peroxide ( $\text{H}_2\text{O}_2$ ) was added separately. The solutions were kept for 30 min at  $60^\circ\text{C}$ . For HPLC study, the resultant solution was diluted to obtain  $40\mu\text{g/ml}$  solution and  $10\mu\text{l}$  were injected into the system and the chromatograms were recorded to assess the stability of sample.

##### Acid Degradation Studies

To 1 ml of stock solution Rotigotine, 1ml of 2N Hydrochloric acid was added and refluxed for 30mins at  $60^\circ\text{C}$ . The resultant solution was diluted to obtain  $40\mu\text{g/ml}$  solution and  $10\mu\text{l}$  solutions were injected into the system and the chromatograms were recorded to assess the stability of sample.

##### Alkali Degradation Studies

To 1 ml of stock solution Rotigotine, 1 ml of 2N sodium hydroxide was added and refluxed for 30mins at  $60^\circ\text{C}$ . The resultant solution was diluted to obtain  $40\mu\text{g/ml}$  solution and  $10\mu\text{l}$  were injected into the system and the chromatograms were recorded to assess the stability of sample.

##### Dry Heat Degradation Studies

The standard drug solution was placed in oven at  $105^\circ\text{C}$  for 6 h to study dry heat degradation. For HPLC study, the resultant solution was diluted to  $40\mu\text{g/ml}$  solution and  $10\mu\text{l}$  were injected into the system and the chromatograms were recorded to assess the stability of the sample.

##### Photo Stability studies

The photochemical stability of the drug was also studied by exposing the  $40\mu\text{g/ml}$  solution to UV Light by keeping the beaker in UV Chamber for 7days or 200 Watt hours/ $\text{m}^2$  in photo stability chamber. For HPLC study, the resultant solution was diluted to obtain  $40\mu\text{g/ml}$  solutions and  $10\mu\text{l}$  were injected into the system and the chromatograms were recorded to assess the stability of sample.

### Neutral Degradation Studies

Stress testing under neutral conditions was studied by refluxing the drug in water for 6hrs at a temperature of 60°. For HPLC study, the resultant solution was diluted to 40µg/ml solution and 10 µl were injected into the system and the chromatograms were recorded to assess the stability of the sample.

### Statistical analysis

Wherever applicable, results were expressed as the Mean±SD, %RSD and data were analyzed statistically by using t- test with aid of Microsoft excel-2007 software and data were considered not significantly different at 5% significance level of probability  $P \leq 0.05$ .

## RESULTS AND DISCUSSION

### Rotigotine method development

Initially reverse phase liquid chromatography separation was tried to develop using various ratios of Methanol and Water, Acetonitrile and Water as mobile phases, in which drugs did not responded properly, and the resolution was also poor. The organic content of mobile phase was also investigated to optimize the separation of both drugs. To improve the tailing factor, the pH of mobile phase becomes important factor. Thereafter, 0.01N Potassium dihydrogen Ortho phosphate buffer (adjusted to pH 4.8 with dil. Orthophosphoric acid solution): acetonitrile were taken in isocratic ratio: 45:55 and with flow rate of 1.0 mL/min was employed. BDS column (4.6 x150mm, 5µ particle size) was selected as the stationary phase to improve resolution and the tailing of both peaks

**Table 1: System suitability of Rotigotine**

System suitability parameters	Rotigotine
No of theoretical plates	3999
Tailing Factor	1.58
RT	2.6 min
Mean Area	3072763
%RSD	0.5

**Table 2: Results of accuracy of Rotigotine**

Sample	Amount Taken (µg/ml)	Amount Recovered (µg/ml)	Recovery (%)	% RSD
Rotigotine	20	19.9	99.96	0.67
	40	40.34	100.87	1.79
	60	60.36	100.61	0.58

**Table 3: Results of Precision for Rotigotine**

Repeatability data		Inter day precision	
S. No.	Rotigotine	S. No.	Rotigotine
1	3077904	1	3053027
2	3096311	2	3050091
3	3116531	3	3077065
4	3111634	4	3073134
5	3098633	5	3084300
6	3070659	6	3090325
Mean	3095279	Mean	3071324
Std. Dev.	18107.9	Std. Dev.	16439.4
%RSD	0.59	%RSD	0.54

**Table No. 4: Results of Robustness for Rotigotine**

Analytical conditions → Evaluation parameters↓	Flow rate (ml/min)		Column temperature (°C)		Mobile phase composition	
	1.1	0.9	35	25	+5%	-5%
Mean RT	2.591	2.987	2.611	2.731	2.600	2.721
Mean area	3048031	3403787	3018805	3064543	3053761	3067798
SD	10666.3	39797.1	6636	48289.2	39506.9	58676.3
RSD	0.35	1.17	0.22	1.58	1.29	1.91
Tailing factor	1.40	1.57	1.49	1.48	1.39	1.49
No. of theoretical plates	4759	4031	4375	4626	4760	4379

**Table 5: Results of stock solution stability for Rotigotine**

Drug	%Assay at 0 hr	%Assay at 24 hr	%Deviation
ROTIGOTINE	100.33	99.07	1.26

**Table 6: Results of HPLC Analysis of patcht for Rotigotine**

Label amount (mg)	Amount found(mg) n=6	%Assay (Mean±SD)	RSD
4	4.013	100.33±1.098	0.59

**Table 7: Results of HPLC Degradation of Rotigotine**

Degradation parameter	Purity angle	Purity treshould	%Drug degraded
Acid degradation	0.128	0.273	3.87
alkali degradation	0.089	0.263	1.99
oxidative degradation	0.190	0.267	5.91
Thermal degradation	0.091	0.268	0.83
Photolytic degradation	0.109	0.267	0.97
Neutral degradation	0.090	0.273	0.72

were reduced considerably and brought close to 1. To analyze drug detection were tried at wavelengths 224nm. Rotigotine showed maximum absorption at 224nm of wavelength and 224 nm was selected as the detection wavelength for PDA detector. The retention times were found to about 2.691 min for Rotigotine. The chromatogram obtained was shown in the Fig. 2.

#### **Rotigotineand met method Validation:**

##### **System suitability and Specificity**

System suitability parameters such as number of theoretical plates, peak tailing, and retention time was determined. The total run time required for the method is only 6 minutes for eluting Rotigotine. The results obtained were shown in Table No.1. The chromatogram obtained for blank and spiked was shown in the Fig. 3.

##### **Linearity**

Rotigotine showed a linearity of response between 10-60 µg/mL. These were represented by a linear regression equation as follows:  $y$  (Rotigotine) =  $62185x + 290.46$  ( $r^2=0.999$ ), ( $r^2=0.999$ ) and regression line was established by least squares method and correlation coefficient

( $r^2$ ) for Rotigotine is found to be greater than 0.98. Hence the curves established were linear.

##### **Accuracy**

To pre analyzed sample solution, a definite concentration of standard drug (50%, 100% & 150 % level) was added and recovery was studied. The % Mean recovery for Rotigotine are 100.48% and these results are within acceptable limit of 98-102. The % RSD for Rotigotine was 1.08 and %RSD for Rotigotine is within limit of  $d^2$ , hence the proposed method is accurate and the results were summarized in Table No.2.

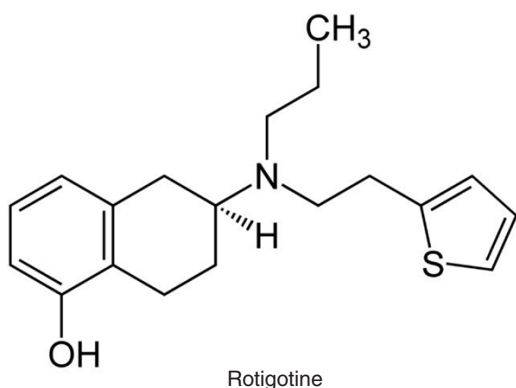
##### **Precision**

##### **Repeatability**

Six replicates injections in same concentration (40µg/ml of Rotigotine) were analyzed in the same day for repeatability and the % RSD for Rotigotine found to be 0.59 and % RSD for Rotigotine found to be within acceptable limit of  $d^2$  and hence method is reproducible and the results are shown in Table No. 3.

##### **Intermediate Precision**

Six replicates injections in same

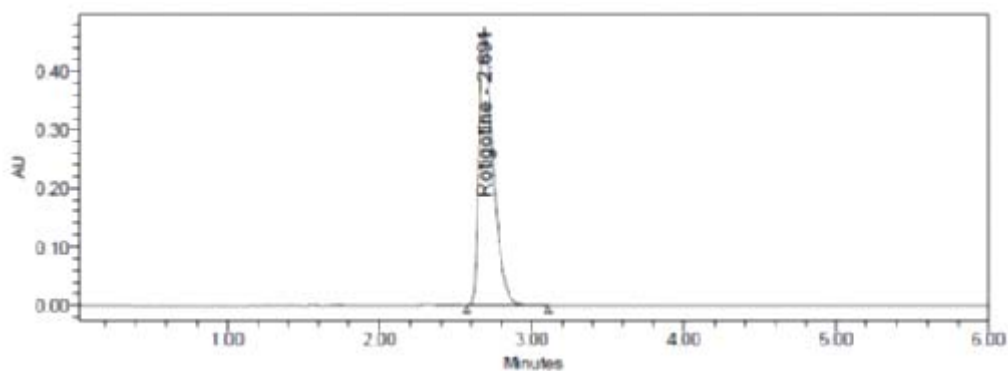


**Fig.1: Chemical structures of drugs investigated in this study**

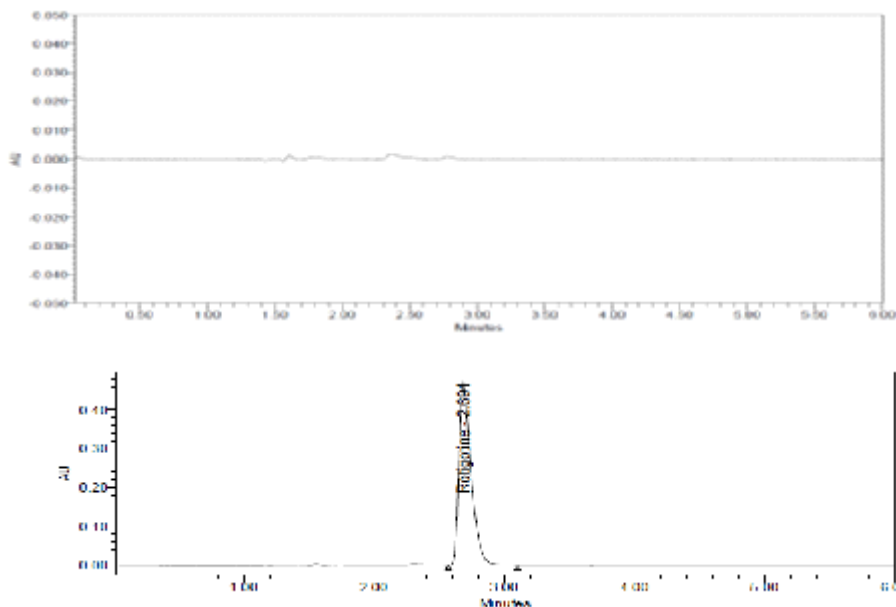
concentration were analyzed on two different days with different analyst and column for verifying the variation in the precision and the % RSD for Rotigotine is found to be 0.54 and it is within acceptable limit of  $\pm 2$ . Hence the Method is reproducible on different days with different analyst and column. This indicates that the method is precise and the results are as shown in Table No. 3.

### Robustness

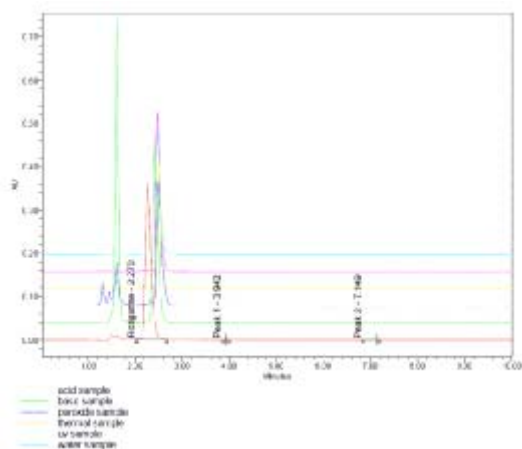
The robustness was established by changing the flow rate, column temperature and composition of the mobile phase within allowable limits from actual chromatographic conditions. It was observed that there were no marked change in



**Fig. 2: A typical Chromatogram of Rotigotine**



**Fig. 3: Chromatograms of Blank(a) and spiked (b)**



**Fig. 4: Degradation overlay chromatogram**

mean  $R_f$  and RSD is within limit of  $d^2$ . The tailing factor, resolution factor and no. of theoretical plates are found to be acceptable limits for Rotigotine. Hence the Method is reliable with variations in the analytical conditions and the results of Rotigotine was shown in Table No.5.

#### Stability of sample solution

The sample solution injected after 24 hr by keeping at ambient room temperature  $30^\circ\text{C}$  did not show any appreciable change. The % Deviation in the assay is not more than 2 and the results are shown in table-6.

#### LOD and LOQ

LOD and LOQ for Rotigotine were 0.05 and  $0.15\ \mu\text{g/mL}$  respectively. The lowest values of LOD and LOQ as obtained by the proposed Method indicate that the Method is sensitive.

#### Patch Analysis

The Content of Rotigotine in the patches was found by the proposed method. RSD values for both Rotigotine was within limit of  $d^2$  and the results were shown in Table No. 7.

#### Degradation studies

The degradation studies for Rotigotine was performed by various conditions like Acid, Alkali, Oxidation, Thermal, Photolytic and Neutral Degradation Studies and their limits like purity angle and purity threshold values like purity angle < purity threshold and the results shown in table no.8.

#### CONCLUSION

A new precise accurate and simple HPLC Method was developed and validated for estimation of Rotigotine pharmaceutical dosage form. This Method is fast, accurate, precise and sensitive hence it can be employed for routine quality control of patch containing drug in QC laboratories and industries.

#### REFERENCES

- Chen, J, J.; Swope, D, M.; Dashtipour, K.; Lyons, K, E. *Pharmacother.* **2009**, *29* (12),1452–67.
- Scheller, D.; Ullmer, C.; Berkels, R.; Gwarek, M.; Lübbert, H. *Naunyn Schmiedeberg's Arch Pharmacol.* **2009**, *379*(1),73–86.
- Loschmann, P. *Eur. J. Pharmacol.* **1989**, *166*, 373.
- Murali, K, P.; Thirupathi, R, B.; Kishore, K, R; Venkateswarlu, P. *Rese Jour Pub Bio Chem Sci.* **2010**, *1*(4), 848-857.
- Avinash, S.; Shakil, S.; Girish, D.; Prakash, K, A.; Veeravenkata, S, K. *Der Pharma Chem.* **2015**; *7*(5):26-34.
- ICH., Stability testing of new drug substances and products (Q1AR). International conference on harmonization, IFPMA, Stability Testing of New Drug Substances and Products: Geneva, **2000**.
- ICH, Q1A (R2) Stability Testing of New Drug Substances and Products, International Conference on Harmonization, Geneva, **2003**.
- ICH, Q1B Stability Testing: Photo stability Testing of New Drug Substances and Products, International Conference on Harmonization, Geneva, **1996**.
- Maheshwaran, R., FDA Perspectives: Scientific Considerations of Forced Degradation Studies in ANDA Submissions, *Pharmaceutical Technology*, **2012**; *36*, 73-80.
- Snyder, L.R., Kirkland, J.J., Glajch, J.L., Practical HPLC Method Development, 2<sup>nd</sup> ed., Wiley, New York **1997**.

## PAPER

Cite this: *RSC Adv.*, 2015, 5, 106461

## Evaluation of surface acetylated and internally quaternized poly(propylene imine) dendrimer as a biocompatible drug carrier for piroxicam as a model drug†

E. Murugan,<sup>\*a</sup> V. Yogaraj,<sup>a</sup> D. P. Geetha Rani<sup>a</sup> and Alok Kumar Sinha<sup>b</sup>

Two types of new surface acetylated and internally quaternized poly(propylene imine) (PPI) dendrimers with generation 2 and 3 QPPI-NHAc (G2)/(G3) were prepared and characterized by FT-IR, NMR and MALDI-TOF spectral techniques and then demonstrated as potential and biocompatible drug carriers using piroxicam (PXM, an anti-inflammatory drug) as a model drug. The drug carrying potential of QPPI-NHAc (G2)/(G3) along with their commercial parent dendrimers PPI (G2)/(G3) and linear branched polyethyleneimine (PEI) were investigated in terms of aqueous solubility, *in vitro* release of PXM and cytotoxicity studies. Pharmacokinetic and biodistribution studies were also carried out in male albino Wistar rats and the efficiency of dendrimer formulation was tested by carrageenan induced paw edema model. The observed results showed that the aqueous solubility of PXM significantly increased in the presence of the higher generation modified dendrimer, namely QPPI-NHAc (G3), to the tune of 50 fold compared with its intrinsic solubility and it is noted that the complexation of PXM with QPPI-NHAc (G3) is responsible for the increased solubility. The degree of complexation was evidenced through UV-Vis and NMR (<sup>1</sup>H & 2D) spectral analyses. The *in vitro* release of PXM in the presence of modified dendrimers was remarkably slow as compared to PPI (G3)/(G2) as well as PEI and thus proves their ability for sustained/delayed release of PXM. The cytotoxicity study on Vero cell line was performed through MTT assay and the results reveal that the QPPI-NHAc (G2)/(G3) showed appreciable increase in cell viability and thus indicate the reduced cytotoxicity and improved biocompatibility. Pharmacodynamic results reveal that in the case of QPPI-NHAc (G3)-PXM, 75% of inhibition was observed at the 4<sup>th</sup> h and this was maintained above 50% until the 8<sup>th</sup> h compared to plain PXM and proves that when the drug is associated with the dendrimer carrier, it was retained for longer in the bio system. The 2<sup>nd</sup>, 4<sup>th</sup> and 8<sup>th</sup> h organ distribution data showed that a higher recovery of PXM was observed for rat paws with dendrimer-drug formulation QPPI-NHAc (G3)-PXM than plain PXM and it may be concluded that dendrimer-drug formulations not only enhanced the solubility but also controlled the delivery of bioactive material with localized action at the site of inflammation.

Received 6th October 2015  
Accepted 30th November 2015

DOI: 10.1039/c5ra20704e

www.rsc.org/advances

## 1. Introduction

Nonsteroidal anti-inflammatory drugs (NSAIDs) are the most widely prescribed medications in the world. As a therapeutic class, NSAIDs exhibit analgesic, anti-inflammatory, antipyretic and platelet inhibitory properties.<sup>1,2</sup> However, these drugs have serious side effects such as gastrointestinal (GI) toxicities,

gastric mucosal ulcerations and hemorrhage due to inhibition of prostaglandin production.<sup>3,4</sup> The mechanism of action of NSAIDs has been attributed to their ability to inhibit the cyclooxygenase enzyme (COX). Out of the 2 isoforms of COX, COX-1 is responsible for mediating the production of prostaglandin while COX-2 is primarily associated with inflammation, pain and fever.<sup>5,6</sup> The traditional NSAIDs are nonselective COX inhibitors. COX-2 selective NSAIDs are ideal anti-inflammatory drugs with minimum drug-related side effects, since they spare COX-1 activity. Piroxicam (PXM) is one such COX-2 selective acidic non-steroidal anti-inflammatory drug (NSAID) with analgesic and antipyretic properties. PXM is quite efficient in the short or long term treatment of rheumatoid arthritis, osteoarthritis and other painful inflammatory disorders.<sup>7-9</sup> However, there are number of aspects that limit its

<sup>a</sup>Department of Physical Chemistry, School of Chemical Sciences, University of Madras, Maraimalai Campus, Guindy, Chennai – 600 025, Tamilnadu, India. E-mail: dr.e.murugan@gmail.com; emurugan\_68@yahoo.co.in; Fax: +91 44 22352494; Tel: +91 44 22202818/22202819; +91 9444275889

<sup>b</sup>Department of Science and Technology, Nano Mission Division, Technology Bhavan, New Mehrauli Road, New Delhi – 110016, India

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5ra20704e

pharmaceutical application, the chief being its aqueous insolubility that immensely creates formulation tribulations.<sup>10</sup> Apart from this, oral administration shows low absorption and gastrointestinal tract associated adverse effects, which displays poor bioavailability.<sup>11</sup> This restricts their use in topical and parenteral applications. Further, poorly water-soluble drugs often require high doses in order to reach therapeutic plasma concentrations after oral administration and hence its solubility has to be increased by suitable means.

Dendrimers are new class of artificial macromolecules with tree-like topological structure, nano-scale size, excellent monodispersity and dense peripheral groups.<sup>12</sup> Dendrimers have generated much interest in diverse areas more specifically for drug delivery due to their unique structure and chemical versatility.<sup>13,14</sup> Surface functionalities of dendrimers can be functionalized with drugs, targeting moieties and biologically active components.<sup>15</sup> Drugs or genes can be either encapsulated within the dendrimers through non-covalent strategies such as hydrophobic, ionic and hydrogen bond interactions or conjugated to the peripheral groups of dendrimers *via* covalent methods.<sup>16</sup> Among the various dendrimers used for drug delivery, poly(amidoamine) (PAMAM) and poly(propylene imine) (PPI) dendrimers are the two commercially available and most extensively investigated for drug delivery applications.<sup>17</sup> A plethora of reports are available which evidence the drug carrying potential of dendrimers through properties like increased solubility of the hydrophobic drugs, sustained drug release behavior and increased efficiency of the drugs.<sup>18–20</sup> Dendrimer based drug-delivery systems provide an attractive platform for loading and release of conventional drug molecules which improve the pharmacodynamic and pharmacokinetic behaviors of several families of drugs. These reveal promising future of dendrimer based drug delivery systems. Dendrimers such as PAMAM and PPI possess cationic primary amine groups at the surface showed excellent drug delivery efficacy, but the cytotoxicity issue of these dendrimer is a serious setback which limits the clinical applications of most dendrimer based drug formulations.<sup>21</sup> Generally, the cytotoxicity of dendrimer is dependent on dendrimer generation, surface functionality/charge and concentration.<sup>22</sup> To overcome this, there are several studies undertaken by various researchers to neutralize the cationic surface group through acetylation,<sup>23</sup> PEGylation,<sup>24</sup> glycosylation<sup>25</sup> and hydroxylation reactions.<sup>26</sup> Nair *et al.* described the dendron based octaguanidine appended molecular transport using poly(propylene imine) for efficient delivery of doxorubicin in to cancer cells.<sup>27</sup> Saha *et al.* designed the peptide-docetaxel for targeted delivery with reduced toxicity and enhanced efficiency of docetaxel.<sup>28</sup> Surface acetylated dendrimers reduced the cytotoxicity significantly but at the same time such dendrimers failed to load the hydrophobic drugs due to the absence of cationic binding sites.<sup>29</sup> In a subsequent study, the neutral dendrimer namely, PAMAM-OH was employed to form polyplex with DNA, with the analogy that the surface neutral PAMAM-OH dendrimer may be advantageous in terms of cytotoxicity. However, PAMAM-OH dendrimer is unable to form DNA polyplex because of low  $pK_a$  value of its interior

tertiary amines.<sup>30</sup> With a view to alleviate this, internal quaternary ammonium salt was introduced to the tertiary amines of PAMAM-OH dendrimers by methylation and thereby provides cationic binding sites for negatively charged plasmid DNA.<sup>31</sup> In continuation of this study, synthesis of surface modified and internally quaternized PAMAM dendrimer was carried out and employed as efficient siRNA delivery systems.<sup>32</sup> However, to the best of our knowledge there are no studies reported so far for synthesis of new dendrimer drug carrier having increased solubility, biocompatibility with reduced cytotoxicity and improved pharmacokinetic properties which obtained from surface acetylation and quaternization of commercial PPI (G2) and PPI (G3).

## 2. Materials and methods

### 2.1 Materials

PPI dendrimer diamino cored (DAB) with terminal amino groups of generation 2 and generation 3 was purchased from Symo Chem (Netherlands) and branched poly ethylene imine (PEI) from Alfa Aesar were used as received. Acetyl chloride, triethylamine, chloroform, methyl iodide and dimethyl formamide were of analytical grade and were procured from SRL. MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide) was purchased from Sigma-Aldrich, India. Amberlite IRA-402 Cl anion resin was from Sd fine Chem. Dialysis membranes (Float-A-Lyzer) were obtained from Spectra Pore. Membrane filter of pore size 0.45  $\mu\text{m}$  was obtained from Himedia lab. Double distilled water was used for solubility and *in vitro* release studies. Piroxicam was obtained as gift sample from Mesha Pharma Ltd, Mumbai, India.

UV-Vis spectra were recorded on TECHCOMP 8500 instrument. Fourier transform infrared (FTIR) spectra were recorded on a Bruker Tensor-27 FTIR spectrometer.  $^1\text{H}$ ,  $^{13}\text{C}$  and 2D NMR spectra were recorded with Bruker 500 MHz UltraShield Plus instrument. MALDI-TOF spectra were taken in perceptive bio systems voyager-DE instrument in negative ion mode taking dithranol as the matrix. Cytotoxicity study was done in King Institute of Preventive Medicine, Chennai, India.

### 2.2 Synthesis and characterization of QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimer

Quaternised poly(propylene imine) dendrimers QPPI-NHAc (G2) and QPPI-NHAc (G3) was synthesized from generation 2 and generation 3 of poly(propylene imine) dendrimer PPI (G2) and PPI (G3) by performing acetylation followed by quaternization reaction by adopting the reported procedure,<sup>32</sup> with slightly modified and simplified. The typical synthetic route followed is shown in Fig. 1a. The commercial PPI (G2) (0.3 g, 0.388 mmol) and PPI (G3) dendrimer (0.3 g, 0.180 mmol) was taken in a 50 mL RB flask and dissolved in 10 mL chloroform. The temperature of this flask was reduced to 0 °C by placing it in freezing mixture. To this an excess amount of acetyl chloride (5 equivalents per surface primary amine) and triethylamine (1.2 equivalent of acetyl chloride) were added. The temperature of the container was brought back to room temperature. The

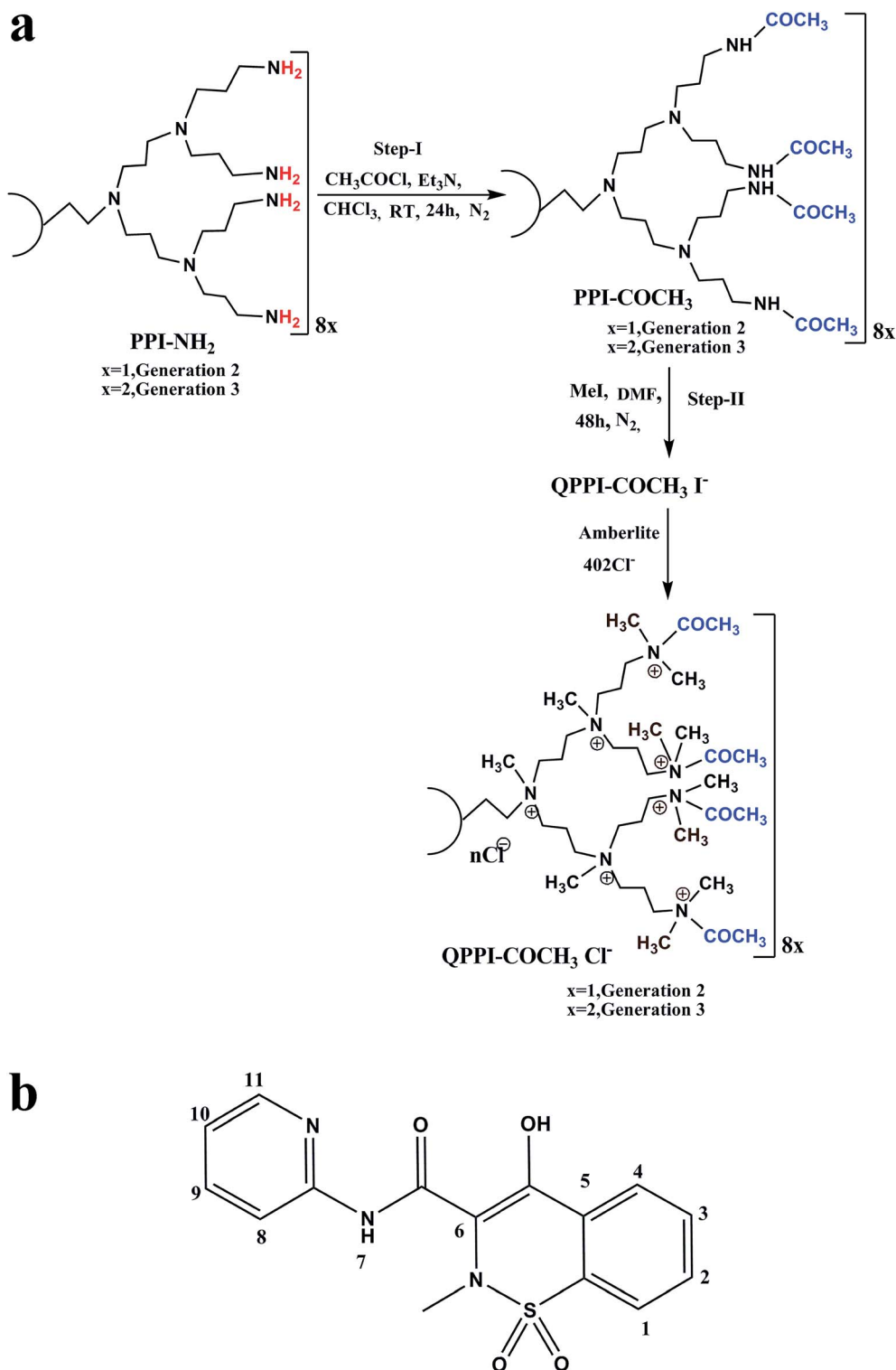


Fig. 1 (a) Synthetic route to get QPPI-NHAc (G2)/(G3) dendrimer from PPI (G2)/(G3) dendrimer. (b) Structure of piroxicam (PXM).

reaction mixture was stirred for 36 h under  $N_2$  atmosphere. The excess solvent from the container was evaporated under reduced pressure and the resulting residue was dissolved in 5 mL water. The mixture was dialyzed extensively (MW cutoff 1000 Da) against double distilled water for 24 h with frequent replacement (once in 2 h) of outer aqueous medium to remove

excess reactants and side products from the residual solution. The dialyzed solution was lyophilized to yield the corresponding reddish brown semi solid products *i.e.* PPI (G2)/(G3) dendrimers attached with  $-COCH_3$  as the surface group, *viz.*, PPI- $COCH_3$  (G2)/(G3). The acetylated product was further used for the quaternization reaction.

In this step, PPI-COCH<sub>3</sub> (G2) (0.1 g, 0.0917 mmol) and PPI-COCH<sub>3</sub> (G3) (0.1 g, 0.0428 mmol) was taken in a 50 mL RB flask and the same was dissolved in 2 mL DMF. To this, an excess amount of methyl iodide (1 mL) dissolved in DMF was added. The reaction mixture was stirred at 50 °C for 48 h. The resulting reddish brown clear/transparent reaction mixture was precipitated by the addition of diethyl ether and the residue was dried in vacuum. It was dissolved in double distilled water and I<sup>−</sup> exchanged for Cl<sup>−</sup> ions. The eluted solution was lyophilized and named as QPPI-NHAc (G2) and QPPI-NHAc (G3). This product was characterized by FTIR, <sup>1</sup>H & <sup>13</sup>C NMR and MALDI-TOF spectral techniques. The observed spectral results strongly confirmed its structure. The degree of functionalization of the product was determined by <sup>1</sup>H NMR and MALDI-TOF spectral techniques and the newly prepared quaternized dendrimers was assessed for its drug carrying potential using PXM as a model drug.

### 2.3 Solubility study of PXM using QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimer carrier

The newly prepared dendrimer based drug carrier *viz.*, QPPI-NHAc (G2) and QPPI-NHAc (G3) was tested for the solubility study of PXM under identical conditions. Additional control experiment was also carried out using poly ethylene imine (PEI) to compare for highlighting the significance of modified dendrimers under identical conditions. The degree of solubility of PXM was quantitatively estimated by adopting the previously described phase solubility analysis method.<sup>33</sup> The experimental procedure adopted for this solubility study involves 7 different concentrations of drug carrier in the range from 0.05 mM to 0.35 mM. Then from each concentration, 5 mL was taken in 20 mL vials and to which excess PXM (5 mg) was added. These vials were sonicated for 30 min and subsequently shaken mechanically for 24 h at 35 °C. In parallel, controlled experiment was also performed by taking 5 mg of drug in 5 mL of phosphate buffer solution (without dendrimer). The vials were allowed to stand for 24 h to attain equilibrium. Then these solutions were centrifuged at 5000 rpm for 20 min and filtered

using cellulose acetate membrane filter having pore size 0.45 μm. The respective filtrate was diluted appropriately with PBS and analyzed spectrophotometrically using UV-Vis spectrophotometer. The quantum of solubility of PXM was estimated by measuring the characteristic peak observed at 354 nm ( $\lambda_{\text{max}}$ ). Though the increase in absorption of PXM is the indication for the extent of drug solubilised, but to get the exact amount of the drug solubilised in the respective dendrimer carrier solution, a calibration curve was constructed within the drug concentration range of 3.5–35 μM which is described by the equation,  $y = (17\,545 \pm 224.34)x$ , ( $R^2 = 0.998$ ). The directional coefficient of the equation is equal to the absorption coefficient of PXM,  $\epsilon_{\text{max}} = 17\,545 \text{ M}^{-1} \text{ cm}^{-1}$ . Using this calibration curve, the amount of drug solubilised in each test solution was determined. The blanks were performed on the same concentrations of dendrimer delivery solutions so as to cancel any absorbance that may be exhibited by the dendrimer molecules. The results of phase solubility studies are shown in Table 1. The solubility experiments were repeated three times. The loading efficiency of drug in dendrimers was calculated by below mentioned formula.

$$\text{Loading efficiency(\%)} = \frac{\text{actual drug content(AC)}}{\text{theoretical drug content(TC)}}$$

AC – actual quantity of drug present in the carrier, TC – 100% theoretical quantity of drug present in the carrier.

### 2.4 UV-Vis spectroscopy studies

The association of PXM to QPPI-NHAc (G3) dendrimer carrier was confirmed by UV-Vis spectroscopy where PXM gives maximum absorbance at its characteristic wavelength (354 nm). The saturated solutions obtained from the solubility studies were diluted to a proper concentration. Since the dendrimers in the diluted solutions give extremely low absorbance between 250 and 700 nm, the absorbance obtained from PXM-QPPI-NHAc (G3) dendrimer solution would be solely from PXM. The absorbance of PXM at its characteristic wavelengths was

Table 1 Solubility of PXM in various concentrations of PEI and dendrimer carrier solutions<sup>a</sup>

S. No.	[Carrier] 10 <sup>−4</sup> M	[PXM] 10 <sup>−4</sup> M					Solubility increment $S_0/S_0^*$				
		PEI	Parent dendrimers		Modified dendrimers		PEI	Parent dendrimers		Modified dendrimers	
			PPI (G2)	PPI (G3)	QPPI-NHAc (G2)	QPPI-NHAc (G3)		PPI (G2)	PPI (G3)	QPPI-NHAc (G2)	QPPI-NHAc (G3)
1	0.50	8.51	5.38	9.51	3.82	7.783	12.25	7.75	13.69	5.50	11.20
2	1.00	9.31	8.75	16.97	5.02	12.83	14.31	12.60	24.30	7.23	18.46
3	1.50	10.12	12.2	23.54	7.30	19.27	14.57	17.57	33.87	10.51	27.74
4	2.00	13.29	15.4	29.46	9.04	22.58	19.14	22.18	42.43	13.02	32.49
5	2.50	17.18	18.8	36.33	10.98	27.94	24.75	27.08	52.30	15.81	40.21
6	3.00	18.46	20.8	40.15	12.92	31.79	26.59	29.96	57.83	18.61	45.76
7	3.50	19.18	22.8	45.70	15.48	34.72	27.63	32.84	65.83	22.30	49.97

<sup>a</sup>  $S_0^*$  is the intrinsic solubility of PXM (0.023 mg mL<sup>−1</sup>), absorbance measured at  $\lambda_{\text{max}}$  of PXM.

related with the amount/solubility of PXM in the dendrimer solutions or buffer solutions.

## 2.5 NMR spectral studies for interactions between PXM and QPPI-NHAc (G3) dendrimer carrier

**2.5.1  $^1\text{H}$  NMR studies.**  $^1\text{H}$  NMR experiments were conducted on a Bruker Advance 500.149 MHz NMR spectrometer at 298.2 K for QPPI (G3) dendrimer/PXM complex (2 mg of QPPI (G3) and 1 mg of PXM dissolved in 1 mL of  $\text{D}_2\text{O}$ ). The temperature was kept constant within 0.2 K by the use of a Bruker temperature control unit.

**2.5.2 2D NMR studies.** NOESY experiments were obtained for a QPPI-NHAc (G3) dendrimer/PXM solution (2 mg of QPPI-NHAc (G3) and 1 mg of PXM dissolved in 1 mL of  $\text{D}_2\text{O}$ ) acquired at 500.149 MHz, using a 300 ms mixing time and a  $8.2\ \mu\text{s}$   $^1\text{H}$   $90^\circ$  pulse width. The experiments were done with a 2 s relaxation delay and 205 ms acquisition time. Eight transients were averaged for each  $400 \times 1024$  complex  $t_1$  increment. The data were processed with Lorentz–Gauss window function and zero filling in both dimensions to display data on a  $2048 \times 2048$  2D matrix. All data were processed with NMR Pipe software on a Linux workstation.

## 2.6 *In vitro* drug release studies

The extent of drug released from the dendrimer carriers, QPPI-NHAc (G2), QPPI-NHAc (G3) and PEI were studied by an equilibrium dialysis method under *in vitro* conditions as described in the references. The drug–dendrimer formulation for *in vitro* study was as follows. That is, 2 mg of PXM was dissolved in 5 mL of dendrimer carrier (0.02 mM) to get the formulation having drug concentration 1.3 mM. Using this formulation, *in vitro* release study was performed by taking 5 mL of respective drug–dendrimer solution in dialysis bag (MW cutoff 1 kDa) and it was placed immediately in 150 mL beaker, containing 100 mL of phosphate buffer solution. The outer phase was stirred continuously to maintain sink conditions. At scheduled intervals, 3 mL of sample was withdrawn from outer phase and subsequently replenished with equal amount of fresh buffer solution. The withdrawn sample irrespective of type was analyzed spectrophotometrically with UV using the drug characteristic peak at 354 nm. Similarly, *in vitro* release of PXM from their parent dendrimers *viz.*, PPI-NH<sub>2</sub> (G2) & PPI-NH<sub>2</sub> (G3) and branched polyethyleneimine (PEI) was also studied by adopting the same experimental condition. The quantum of drug released by the parent dendrimer and PEI was determined as before. The control experiment was also carried out by taking 2 mg of PXM in 5 mL of phosphate buffer solution to get the same drug concentration as in the drug–dendrimer formulations. This was also studied for *in vitro* release by usual UV analysis. Three repeats were conducted for each sample. Irrespective of the dendrimer carriers, PEI and the control, plots were drawn for the % of drug released *vs.* time.

## 2.7 Cytotoxicity assay

The cytotoxicity of modified dendrimer drug complex QPPI-NHAc (G2)–PXM, QPPI-NHAc (G3)–PXM and its parent

dendrimers PPI-NH<sub>2</sub> (G2)–PXM, PPI-NH<sub>2</sub> (G3)–PXM complexes along with PXM (without dendrimers) were evaluated by the well-established 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT) colorimetric assay. Vero cell line (NCCS) was chosen for the MTT colorimetric assay. The Vero cell line used for this study is “normal cells” and it was isolated from kidney epithelial cells extracted from an African green monkey. One of the reasons for choosing the normal cell line is, while performing/discharging its assigned role; *i.e.* it should not disturb or cause toxicity to normal cells. Further, our objective in this study is to examine the compatibility of synthesized dendrimer carriers for the poorly soluble drugs (or) to prove that the therapeutic value for the formulation of QPPI-NHAc (G2)/ (G3)–piroxicam should be improved and thus have plan to recommend the same to human cell line. In this context, it is observed and known that the genetical similarities between human cell line and Vero cell line are very close and hence we decided and choose Vero cell line as a model for this study. The assay procedure involves the incubation of cell line at  $37^\circ\text{C}$  in 5%  $\text{CO}_2$  and minimum essential medium (MEM) with Earle's salts supplemented with 10% fetal calf serum. The cells were used when they reached 70–80% confluence. These cells were seeded in 96 well plates for 48 h before MTT assay to allow adherence of the cells to the plates. Then culture medium from the plates was replaced with fresh medium containing different concentrations of QPPI-NHAc (G3) and PPI-NH<sub>2</sub> (G3). The control study was performed by treating the cells only with the medium without the dendrimer samples. Following the standard protocol for MTT, the mediums were removed after 24 h from each of the well and the cells were incubated with MTT in PBS at a concentration of  $1\ \text{mg mL}^{-1}$  (100  $\mu\text{L}$ ) for 4 h. After this, in each well, the MTT containing medium was replaced with DMSO to dissolve the resulted purple formazan crystals which was generated from MTT by the living cells. Absorbance of the solutions in each well irrespective of carrier was recorded at 570 nm using a micro plate reader (Thermo multiscan EX). The viability of cells was calculated with regard to the untreated cell control, which was considered as 100% viability. The viability of cells treated with the different dendrimer carrier is expressed as a percentage of the untreated cell control. The obtained cell viability was plotted against the concentration of the carrier.

## 2.8 *In vivo* anti-inflammatory activity

The animal studies were performed in accordance with the protocol approved by the Ministry of Social Justices and Empowerment, Government of India and Institutional Animal Ethical Committee Guidelines (approval no. 02.01.2012) for the investigation of experimental pain in conscious animals. In this study, a pharmacodynamic study was performed using non-immunological carrageenan induced hind paw edema method.<sup>34,35</sup> In the anti-inflammatory activity 0.1 mL of 1% carrageenan was taken as phlogistic agent. Acute inflammatory activity was determined by measuring the change in volume of inflamed paw produced by injection of carrageenan using a plethysmometer. On the basis of solubilization efficacy as well as better *in vitro* performance of PPI dendrimers compared to

PEI (branched) it is selected for *in vivo* evaluation. Male albino Wistar rats weighing around 160–180 g were procured from Tamil Nadu veterinary and Animal Sciences University, Chennai and were housed under standard husbandry conditions and the rats were fasted for 12 h before dose administration. The animals were numbered and marked on the right hind paw on each animal. Each time the paw was dipped in the plethysmometer up to the fixed mark to ensure constant paw volume and the study was carried out at daytime to avoid any variation due to circadian rhythms. Animals were divided into five groups (including one control group), each group comprising six animals. The dendrimer–drug test formulations namely PPI (G2)–PXM, PPI (G3)–PXM, QPPI-NHAc (G2) and QPPI-NHAc (G3) which in turn were solubilized in PBS solution (pH 7.4) and used for the study. The plain drug PXM (control) was solubilized in PBS pH 7.4 using DMSO as cosolvent. The dose of  $2 \text{ mg kg}^{-1}$  (equivalent to PXM) body weight was administered through an intravenous route in albino rats of the respective group, precluding the control. A dose of 0.1 mL solution of carrageenan (1% w/v in normal saline) was injected in the right hind paw of the animals to be tested, 10 min post administration of the test formulation. The paw volume was measured every hour until the eighth hour, the last two observations were recorded at 12<sup>th</sup> h and 24<sup>th</sup> h, and from the result obtained a graph was plotted between percentage inhibitions of edema *vs.* time (h). Percentage inhibition of edema was calculated for respective group by following the equation<sup>36</sup> % inhibition of edema =  $(V_{\text{control}} - V_{\text{treated}}/V_{\text{control}}) \times 100$ , where  $V_{\text{control}}$  and  $V_{\text{treated}}$  are the mean edema volume of rats in control and test groups respectively.

## 2.9 Pharmacokinetics and biodistribution studies

Pharmacokinetic and organ distribution studies of dendrimer–drug complexes and free drug along with their parent dendrimers were carried out in male albino Wistar rats with carrageenan-induced inflammation in the right hind paw. Albino rats were divided in five groups each comprising six rats were numbered and marked. The test formulations of dendrimer–drug complexes QPPI-NHAc (G2)–PXM & QPPI-NHAc (G3)–PXM and plain drug PXM in the doses of  $2 \text{ mg kg}^{-1}$  body weight were administered through an intravenous route. The blood samples were collected from the retro-orbital plexus at regular intervals of time duration and the blood samples were allowed to stay for 15 min and then centrifuged at 3000 rpm for 15 min to separate RBCs and serum. The supernatant (serum) was collected with the help of micropipet and were analyzed for drug content using HPLC (Waters 2489 UV/Vis detector, USA) for the determination of various pharmacokinetic parameters.

For organ distribution studies, animals were divided into five groups, and each group was administered with the same intravenous dose ( $2 \text{ mg kg}^{-1}$ ) of dendrimer–drug formulations and plain PXM with their corresponding parent dendrimers. Five rats from each group were humanely sacrificed at 2, 4, and 8 h, and the organs, *viz.*, paw, kidney, liver and spleen were removed immediately and weighed then the organs were refrigerated at  $-20^\circ\text{C}$  by wrapping them in aluminum foil. The

organs were homogenized, centrifuged at 4000 rpm for 15 min followed by collection of supernatant and assayed for PXM by the HPLC method as reported earlier, with slight modifications<sup>37,38</sup> employing a C18 column, 5  $\mu\text{m}$  particle size (analytical column protected by a compatible guard column). The HPLC system (Waters 2489 UV/Vis detector, USA), and the elution was carried out by using water : methanol : acetic acid (44 : 50 : 6) as the mobile phase. The flow rate was kept at  $1.0 \text{ mL min}^{-1}$  throughout the process.

## 3. Results and discussion

### 3.1 Synthesis and characterization of QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimers

Earlier studies reveal that the commercial PAMAM and PPI dendrimer have been modified by surface functionalization using different moieties and explored the same for different applications and found convincing results.<sup>39,40</sup> In connection to our expertise in this field, in this study, we have modified the commercial PPI (G2) and PPI (G3) in to QPPI-NHAc (G2) & QPPI-NHAc (G3) *via* procedure followed in the earlier study. That is, in the first step acetyl chloride was surface functionalized on PPI (G2)/(G3) and the obtained surface acetylated product was then quaternized using methyl iodide and the product obtained in each reaction was characterized with FTIR,  $^1\text{H}$  and  $^{13}\text{C}$  NMR and MALDI-TOF techniques. The FTIR spectra of commercial PPI-NH<sub>2</sub> (G2) & PPI-NH<sub>2</sub> (G3) and hydroxylated PPI-NHAc (G2) & PPI-NHAc (G3) were shown in Fig. 2a, b, d and e respectively. The spectra for PPI-NH<sub>2</sub> (G2) (Fig. 2a) and for PPI-NH<sub>2</sub> (G3) (Fig. 2d) shows stretching vibration at  $3356 \text{ cm}^{-1}$  &  $3280 \text{ cm}^{-1}$  and  $3334 \text{ cm}^{-1}$  and  $3284 \text{ cm}^{-1}$  assigned for the surface amine groups. In the FTIR spectra of PPI-NHAc (G2) (Fig. 2b) & PPI-NHAc (G3) (Fig. 2e) these  $-\text{NH}_2$  peaks disappeared and also individual peaks at  $1717 \text{ cm}^{-1}$  and  $1715 \text{ cm}^{-1}$  appeared corresponding to the characteristic C=O (str) vibrations of carbonyl group in acetyl chloride. This observation confirms the alkylation of PPI (G2) and PPI (G3) with acetyl chloride. The FTIR spectrum of QPPI-NHAc (G2) (Fig. 2c) shows characteristic peaks at  $2958 \text{ cm}^{-1}$  and  $2740 \text{ cm}^{-1}$  for C–H (str),  $1434 \text{ cm}^{-1}$  for C–H (bend) and  $1160 \text{ cm}^{-1}$  for C–N (str) and the FTIR spectrum of QPPI-NHAc (G3) (Fig. 2f) shows characteristic peaks at  $2991 \text{ cm}^{-1}$  and  $2773 \text{ cm}^{-1}$  for C–H (str),  $1454 \text{ cm}^{-1}$  for C–H (bend) and  $1162 \text{ cm}^{-1}$  for C–N (str). The new intense peaks at  $1160 \text{ cm}^{-1}$  and  $1162 \text{ cm}^{-1}$  respectively which correspond to the  $\text{C–N}^+_{\text{str}}$  due to the quaternization of tertiary amines in PPI-NHAc (G2) and PPI-NHAc (G3). The chemical shift value from the  $^1\text{H}$  &  $^{13}\text{C}$  NMR spectra of QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimers further provide the convincing evidence for surface alkylation and internal quaternization. The  $^1\text{H}$  NMR for QPPI-NHAc (G2) (Fig. 3a) gives a sharp singlet at  $\delta$  2.734 due to  $-\text{COCH}_3$  (acetyl protons) and a sharp singlet at  $\delta$  3.353 which is due to  $(\text{N–CH}_3)$  methyl protons. Similarly,  $^1\text{H}$  NMR recorded for QPPI-NHAc (G3) (Fig. 3b) gives an intense singlet signal at  $\delta$  2.673 due to  $-\text{COCH}_3$  (acetyl protons) and a sharp singlet at  $\delta$  3.241 due to  $(\text{N–CH}_3)$  methyl protons. In  $^{13}\text{C}$  NMR for QPPI-NHAc (G2), methyl carbons  $(\text{N–CH}_3)$  show peak at  $\delta$  49.321, acetyl carbons

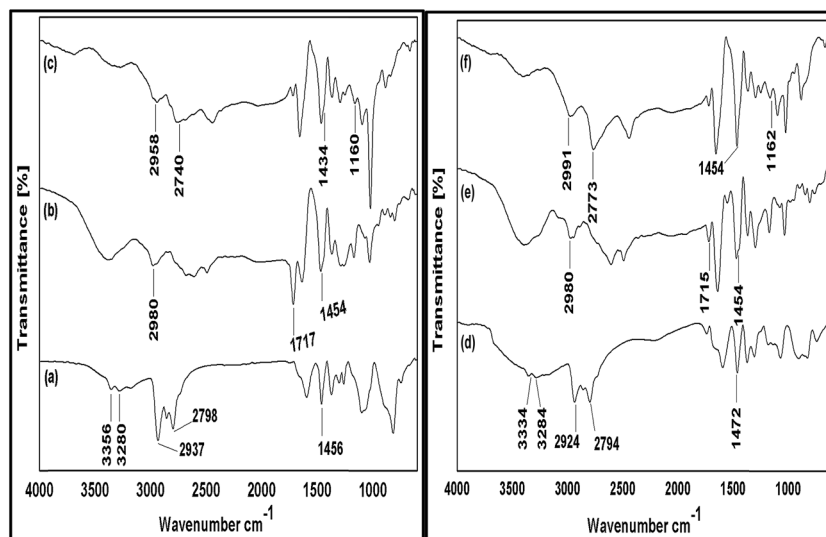


Fig. 2 FTIR Spectrum of (a) PPI-NH<sub>2</sub> (G2), (b) PPI-NHAc (G2) and (c) QPPI-NHAc (G2), (d) PPI-NH<sub>2</sub> (G3), (e) PPI-NHAc (G3) and (f) QPPI-NHAc (G3) respectively.

(-COCH<sub>3</sub>) show peak at  $\delta$  38.837 and carbonyl carbons (COCH<sub>3</sub>) show peak at  $\delta$  176.96 respectively. Likewise, <sup>13</sup>C NMR spectrum of QPPI-NHAc (G3), methyl carbons (N-CH<sub>3</sub>) show peak at  $\delta$  46.839, acetyl carbons (-COCH<sub>3</sub>) show peak at  $\delta$  36.096 and carbonyl carbons (COCH<sub>3</sub>) show peak at  $\delta$  174.34 respectively.

The structures of the QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimers were also evidenced through MALDI-TOF analysis and the observed spectra are shown in Fig. S1a and b.† The characteristic peaks corresponding to *m/z* values at 2154 and 4559 for QPPI-NHAc (G2) and QPPI-NHAc (G3). The obtained experimental *m/z* values agree well with the theoretical values for the modified dendrimers. The collective results observed from FT-IR, NMR (<sup>1</sup>H & <sup>13</sup>C) and MALDI-TOF spectra strongly confirm the respective structure of surface alkylated and internally quaternized PPI (G2) and PPI (G3) dendrimers. This modified dendrimer QPPI-NHAc (G2) and QPPI-NHAc (G3) was employed for drug delivery applications using PXM as a model drug.

### 3.2 Solubility studies of PXM using QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimer carriers

Drug solubility is one of the core factors, which influences the movement of a drug from the site of administration into the blood stream. It is widely acknowledged that insufficient drug solubility can lead to poor absorption.<sup>41</sup> The poor solubility of PXM can be addressed with newly developed QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimers by employing it as solubility enhancer. Additionally in order to project the significance of modified dendrimers in drug delivery the PPI (G2) & PPI (G3) and branched PEI also employed and studied. Solubility studies were performed by keeping the seven different concentrations in the range from 0.05 mM to 0.35 mM. The amount of solubilised drug was estimated in each concentration of dendrimer carriers from the UV absorption and the

calibration curve was drawn. The observed drug solubility with increasing trend is given in Table 1. To record precisely, at the maximum concentration (0.35 mM) the QPPI-NHAc (G2), QPPI-NHAc (G3) and PEI has promoted the solubility of PXM to the tune of 22, 50 and 27 folds respectively higher than with the intrinsic solubility of plain drug. In the case of commercial PPI (G2) & PPI (G3) under similar experimental conditions, the solubility of PXM has increased to 32 and 65 fold to its intrinsic solubility. However, the parent dendrimers are able to load the drug but at the same time its cytotoxicity is not tolerable. Therefore, the quaternized forms of modified dendrimers with reduced toxicity are reasonable in drug loading and releasing behaviour. This observation in turn ensures that the water solubility of PXM has been significantly increased proportionally to the concentration of dendrimer carriers. The comparative results reveal that the dendrimer derived from higher generation *viz.*, QPPI-NHAc (G3) has increased the solubility of PXM to the tune of 50 folds compare to its intrinsic solubility and 2 fold than with their corresponding lower generation dendrimer *viz.*, QPPI-NHAc (G2). Also it is clearly explained that the solubility of PXM in the presence of branched PEI is substantially lower than PPI (G3) and QPPI-NHAc (G3), due to poor ionic interaction and signifies that the higher generation of modified dendrimers is better way of delivering drug.

The linear equations for QPPI-NHAc (G2) and QPPI-NHAc (G3) carriers were obtained from their solubility plots (Fig. 4b) and it is calculated as  $y = (3.89 \pm 0.129)x + (1.452 \times 10^{-4} \pm 0.285 \times 10^{-4})$  ( $R^2 = 0.994$ ) and  $y = (9.10 \pm 0.415)x + (4.216 \times 10^{-4} \pm 0.928 \times 10^{-4})$  ( $R^2 = 0.988$ ) respectively. The directional coefficients of former one is  $n = 3.89 \pm 0.129$  and for later one is  $n = 9.10 \pm 0.415$  respectively and these values in turn can be interpreted as a number of drug molecules combined/bound by one molecule of dendrimeric unit. That is, the tentative number of drug molecules combined/bound by a single molecule of

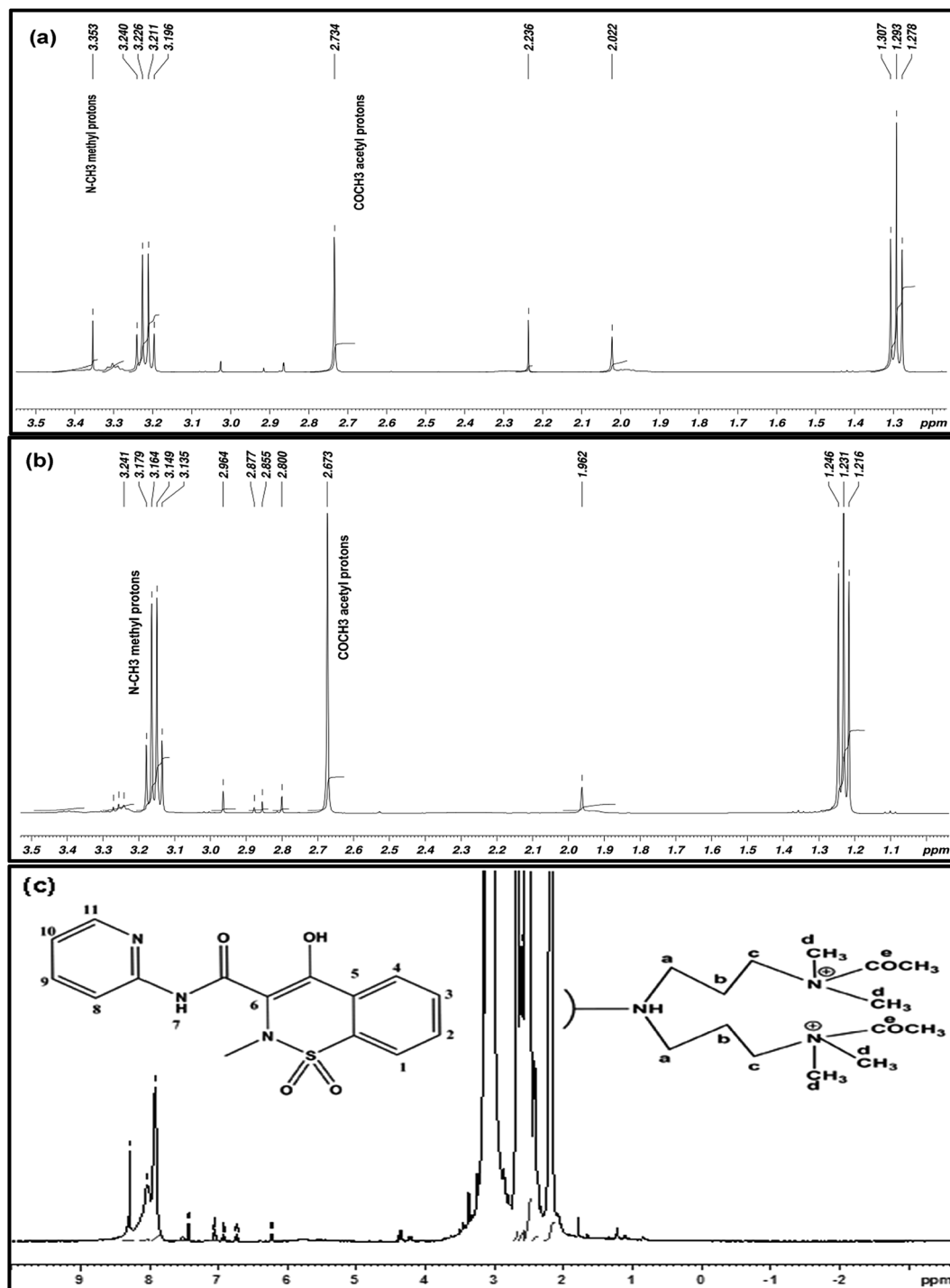


Fig. 3  $^1\text{H}$  NMR spectrum of (a) QPPI-NHAc (G2), (b) QPPI-NHAc (G3) and (c) QPPI-NHAc (G3)-PXM complex.

dendrimeric unit was determined as  $3.89 \pm 0.129$  for QPPI-NHAc (G2) and  $9.10 \pm 0.415$  for QPPI-NHAc (G3) carriers. Further, we also found that the drug loading efficiency of QPPI-OH (G2) is 64.3% and for QPPI-OH (G3) is 58.4% respectively. The increased solubility noticed in the presence of QPPI-NHAc (G2) and QPPI-NHAc (G3) is due to (i) the availability of internal positive binding sites with neutral outer surface. (ii) The multi charged cations *viz.*, QPPI-NHAc (G2)/(G3), attracted the anion

of PXM whose formation is facilitated by the generation of HCl, a stronger acid compared to PXM, (iii) thus PXM binds electrostatically with QPPI-NHAc (G2)/(G3) in aqueous medium and thus forming the complex *viz.*,  $(\text{PXM-O}^-)_n-(\text{QPPI-NHAc})^{n+}$ , (iv) thereby PXM gets internalized within the QPPI-NHAc (G2)/(G3) dendrimer branches. This interaction substantially promotes the solubility of PXM. The increased solubility has directly helps to increase its bioavailability.

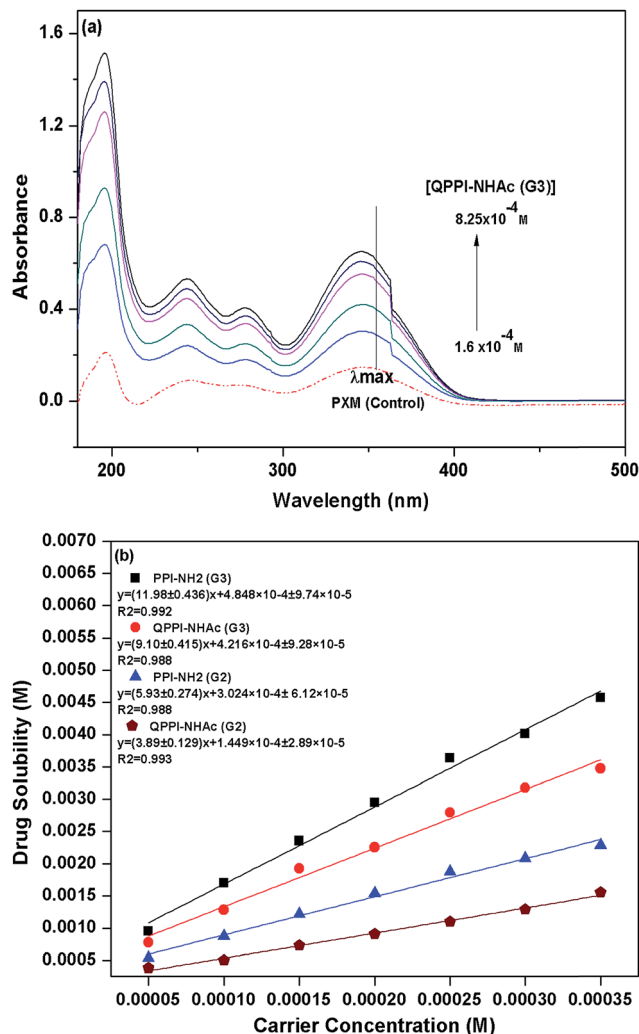


Fig. 4 (a) UV-Vis spectra of plain PXM (dotted line) and PXM with different concentrations of QPPI-NHAc (G3) (bold lines) and (b) solubility behaviour of PXM at various concentrations of PPI-NH<sub>2</sub> (G2), PPI-NH<sub>2</sub> (G3), QPPI-NHAc (G2) and QPPI-NHAc (G3) dendrimer solutions.

### 3.3 UV-Vis spectroscopy for the interactions between PXM and QPPI-NHAc (G3) dendrimer carrier

The use of dendrimer as a drug delivery carrier usually depends on its ability to form a complex with the drug. Interactions between PXM and dendrimer molecules were determined by UV-Vis spectroscopy. PXM gives maximum absorbance at its characteristic wavelength (354 nm). After the addition of increasing concentrations of QPPI-NHAc (G3) dendrimers, blue shifts were observed (Fig. 4a) (the  $\lambda_{\max}$  peak shifted about 8–10 nm from 354 nm). In general, any shift in  $\lambda_{\max}$  suggests the existence of interaction between the complex forming components.<sup>42</sup> In this study, electrostatic interaction between the anionic group of PXM and the cationic groups of QPPI-NHAc dendrimer can explain this blue shift. It is known that increase in polarity of the solvent generally shifts  $n-\pi$  bands to shorter wavelength. The hypsochromic groups diminish the color of chromogen *i.e.* they cause displacement to shorter wavelengths. In our study, acetylation of NH<sub>2</sub> groups *i.e.* COCH<sub>3</sub> and

NHCOCH<sub>3</sub> produces hypsochromic shift (blue shift). Further, saturated sulphoxide absorbs with high intensities. In our study also, SO<sub>2</sub> absorption is due to  $n-\pi$  transition present in the PXM drug. Therefore, the occurrence of shift noticed in ( $\lambda_{\max}$ ) supports the formation of complex through electrostatic interaction.

### 3.4 NMR spectral studies for interactions between PXM and QPPI-NHAc (G3) dendrimer carrier

**3.4.1 <sup>1</sup>H NMR studies.** It is well known that the <sup>1</sup>H NMR spectroscopy is a useful technique to investigate the intermolecular interactions in solutions because it gives information pertinent to the formation of aggregates, ion pairing, encapsulation and size variations.<sup>43</sup> The <sup>1</sup>H NMR technique was employed here to investigate the molecular interactions between QPPI-NHAc (G3) dendrimer and PXM drug molecules. The chemical shift assignment for each proton in the dendrimer, PXM and dendrimer-PXM complex is so critical because the shift of each signal is helpful to define the zone of interaction between the dendrimer and PXM.<sup>44</sup> The Fig. 3c shows the <sup>1</sup>H NMR spectra and chemical shift assignments for QPPI-NHAc (G3)/PXM complex. The protons of PXM and QPPI-NHAc (G3) dendrimer were labeled with numbers 1–11 and characters a–e, respectively as shown in Fig. 3c. It can be observed that PXM has scattered NMR peaks in the range of 2–8.5 ppm and that QPPI-NHAc (G3) dendrimer has seven NMR peaks in the range of 1.0–4.0 ppm. Having assigned the chemical shifts of the QPPI-NHAc (G3) dendrimer and PXM, we can analyze the complexation of the dendrimer and PXM by <sup>1</sup>H NMR. The QPPI-NHAc (G3)-PXM complex has shown two types of peaks which correspond to PXM molecule and dendrimer scaffold respectively. It is observed that significant changes were noticed in chemical shifts of methylene protons (c, d and e) in QPPI-NHAc (G3) dendrimer induced by the addition of PXM. The downfield chemical shift of these protons have localized at the outermost layer of the QPPI-NHAc (G3) dendrimer and thus proves the ionic interactions between quaternized ammonium groups of the dendrimer and deprotonated group of PXM.<sup>45</sup>

**3.4.2 2D-NOSEY analysis.** It is well known that 2D NOESY technique is proved to be an effective method for providing insights into the molecular features of host-guest interactions.<sup>46–48</sup> It is capable of revealing a spatial relationship between nuclei in a molecule or in a complex of molecules.<sup>46</sup> If the host and guest are bound, they should be in close proximity to each other and NOE cross-peaks should be seen in the corresponding spectral region. In other words, the absence of a NOE cross-peak in the region can be used to rule out the interaction between related nuclei.<sup>48</sup> In this study, <sup>1</sup>H-<sup>1</sup>H NOESY experiments were performed to determine the dipolar contacts for both intermolecular and intramolecular interactions of QPPI-NHAc (G3) dendrimer with PXM. The <sup>1</sup>H-<sup>1</sup>H NOESY spectrum of a QPPI-NHAc (G3)/PXM/DMSO solution at a mixing time of 300 ms is shown in Fig. 5a. Strong NOE cross-peaks are observed between the hydroxyl protons of PXM and

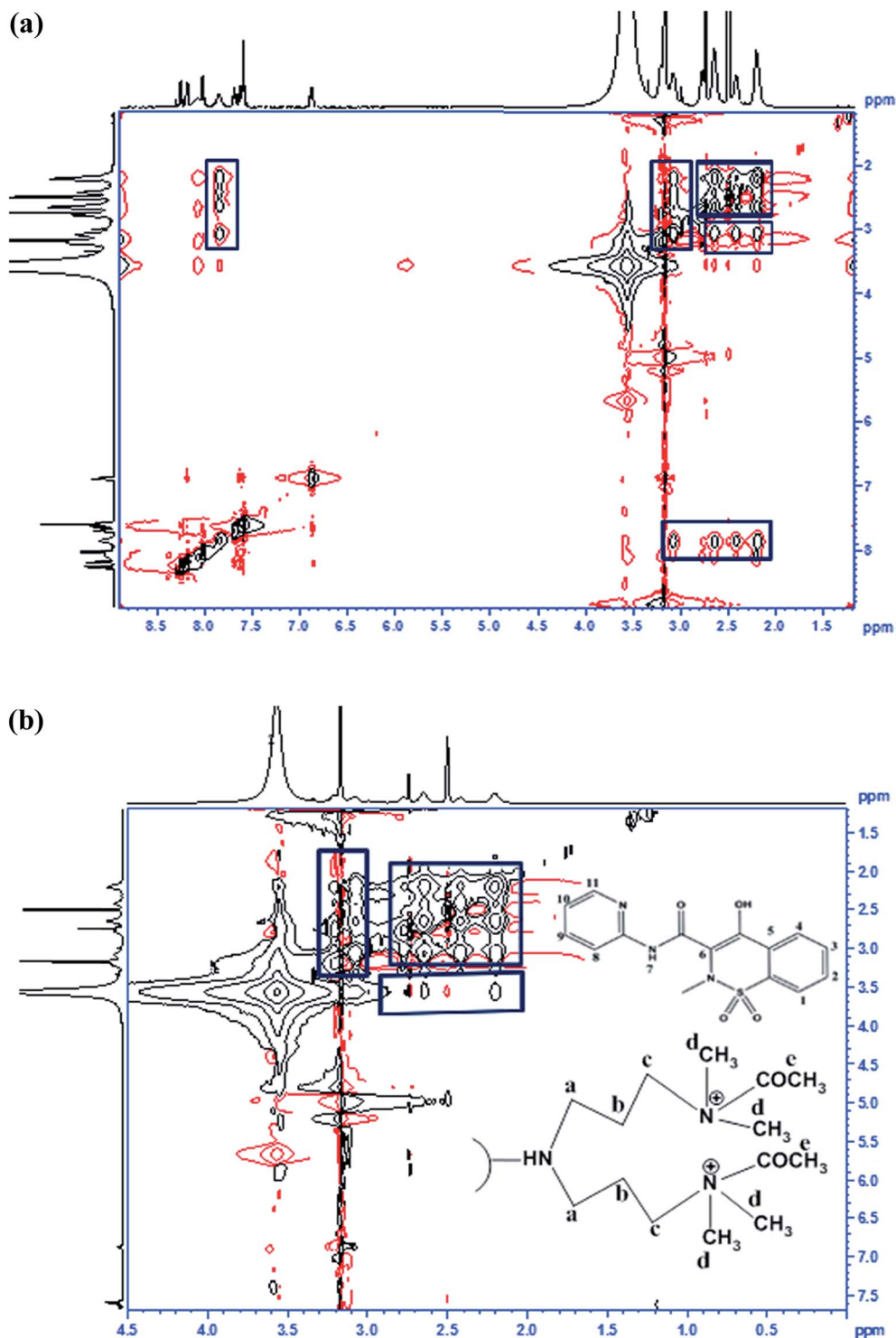


Fig. 5 (a)  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum of the QPPI-NHAc (G3)/PXM/DMSO solution at a mixing rate of 300 ms. (b) Expanded region of the  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum shown in (a), showing the cross-peaks between protons of drug and QPPI-NHAc (G3). The cross-peaks are indicated by rectangles.

scaffold protons of the QPPI-NHAc (G3) dendrimer, which indicates close proximity between these protons.

The partial  $^1\text{H}$ - $^1\text{H}$  NOESY spectrum (the region of the dendrimer proton signals with chemical shifts ranging from 1.0 to 4.5 ppm) for QPPI-NHAc (G3) dendrimer-PXM complex (Fig. 5b)

shows 5 peaks corresponding to three  $\text{CH}_2$  protons (a-c) in the interior of the dendrimer, one N- $\text{CH}_3$  proton (d) and one  $\text{COCH}_3$  proton (e) in the outermost layer of QPPI-NHAc (G3) dendrimer and these observations are in accordance with the  $^1\text{H}$  NMR spectrum of the QPPI-NHAc (G3)-PXM complex.

Several cross peaks were obtained between the protons of QPPI-NHAc (G3) and the protons of PXM. Intermolecular NOE peaks such as (c-8, 9), (d-8, 9) and (e-8, 9) as well as similar intramolecular NOE peaks (a-c), (b and c), (a-e) and (b-e) were also noticed. In addition to this, some intramolecular cross peaks (8, 9 and 10, 11) and (1, 2 and 8, 9) were also observed and this in turn indicates the interaction between the aromatic protons of PXM. Therefore, all these observations strongly confirm the interactions of PXM molecules with the scaffold protons of QPPI-NHAc (G3) dendrimer.

### 3.5 *In vitro* drug release studies

The very purpose of drug carrier is to carry the drugs and deliver them at the right site of action. One of the prerequisite of a good carrier is it should deliver the drug in a sustained/delayed manner. Thereby the side effect of drug can be minimized by preventing the fluctuation of the therapeutic concentration of the drug in the body. It also eliminates the chance of over or under dosing. Normally, one of the methods to achieve this sustained/delayed release is *via* diffusion through a matrix. According to the previous reports, dendrimers are able to pack hydrophobic drugs and these formulations are suitable for different administration routes. These polymeric delivery systems with sustained release behaviors can improve the biocompatibility of drugs loaded in the carriers, simplify the dosing schedule and improve patient compliance.<sup>40–44</sup> To study the extent of drug release from drug–dendrimer formulations, we investigated the release profiles of PXM from QPPI-NHAc (G2), QPPI-NHAc (G3) and its parent dendrimers PPI (G2) & PPI (G3) and as well as from PEI. The release pattern was plotted as percentage of drug released *vs.* time and the same is shown in Fig. 6. The obtained plot clearly shows that after 1 h, 90% of drug gets released from the control formulations (drug alone)

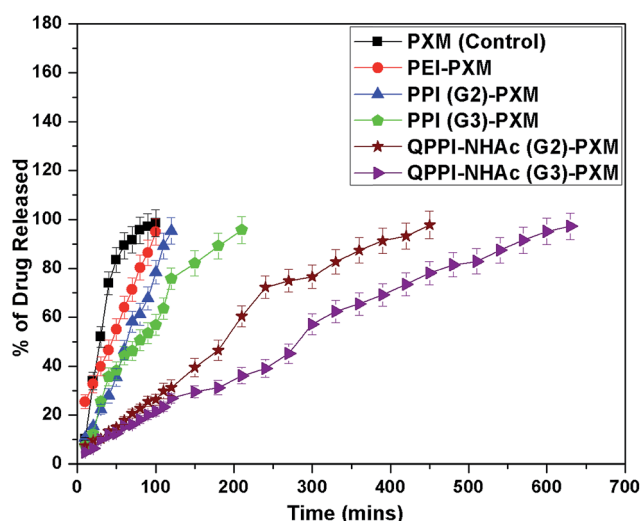


Fig. 6 Release behaviour of PXM alone (control) and also in presence of PEI, PPI (G2), PPI (G3) and modified dendrimers QPPI-NHAc (G2) & QPPI-NHAc (G3). PXM and dendrimer concentrations were 1.3 mM and 0.2 mM in dialysis bag respectively. Each data point represents mean  $\pm$  standard error (S.E) ( $n = 3$ ).

and 70% of drug gets released from the PEI–PXM formulations, whereas in commercial dendrimers 50% of drug was released from PPI-NH<sub>2</sub> (G2)–PXM formulations and it is 40% from PPI-NH<sub>2</sub> (G3)–PXM formulations. Further, in the case of QPPI-NHAc (G2)–PXM formulation the release gets slower to 20% and in the case of QPPI-NHAc (G3)–PXM formulation the release gets sustained to 13% in the same span of time (1 h). That is, as much as 90% drug gets released within 60 min from the control formulations, and it takes 90 min to get released from PEI–PXM. Whereas, it takes 110 min for 90% of drug release from PPI-NH<sub>2</sub> (G2)–PXM formulation and 180 min for the same amount of drug to get released from the PPI-NH<sub>2</sub> (G3)–PXM formulations.

However, important observation that has to be necessarily highlighted is that the same amount (90%) of drug released from quaternized form of dendrimers QPPI-NHAc (G2)–PXM is extended to 390 min and for higher generation QPPI-NHAc (G3)–PXM formulation it was extended to 570 min. These results suggested that the QPPI-NHAc (G3) dendrimer has significantly contributed for the sustained release of PXM than with their control, commercial and PEI formulations. In general, it is well established fact that the sustained/delayed release of a drug by any drug carrier is an indication of formation of stable complex due to the interaction between drug carrier and drug or it is an indication for effective internalization of drug with the drug carrier. In this study, we established that the complexes formed between the quaternary ammonium dendrimer drug carriers and the drugs should facilitate *via* electrostatic interaction, whose magnitude depends on the ionic strength and structure of QPPI-NHAc (G2)/(G3). The developed quaternary ammonium dendrimer drug carriers have sustained the drug release significantly, which is due to the formation of stable complexes in the form of HCl (PXM-O<sup>−</sup>)<sub>*n*</sub>–(QPPI-NHAc (G2)/(G3))<sup>*n*+</sup>. The formation of this stable complex and their internalization within the dendrimer branches greatly contributed towards the sustained/slower release of the drug. However, in this study it is observed that QPPI-NHAc (G2)/(G3) entraps the PXM more stably and thus reflected in the sustained release of PXM. The multi charged cations with hydrophilic character which made to interact well with the anionic drugs. The electrostatic interaction between them are intensively attracted the anionic PXM drug and which in turn favors more drug loading and thus forms stable complex which in turn reflected in the sustained release of respective drug molecule. The prolonged sustain release profile of PXM from the QPPI-NHAc (G2)/(G3) under both generation implied that quaternized poly(propylene imine) dendrimers are extremely important for effective encapsulation and retention of piroxicam.

### 3.6 Cytotoxicity assay

An ideal polymeric drug carrier should be biocompatible by performing its expected functions without the concomitants of toxic and immunologic effects to normal cells.<sup>49</sup> The effect of dendrimers and their complexes with PXM on the viability of Vero cell lines were measured by the MTT assay. The variation of percentage cell viability incubated in Vero cell line with

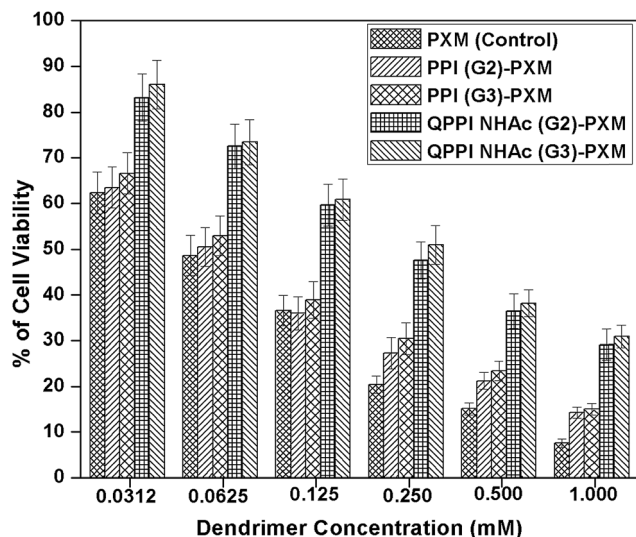


Fig. 7 Cytotoxicity of PXM (control), PPI (G2/G3) and QPPI-NHAc (G2/G3) dendrimers with PXM at different concentrations after 24 h of incubation as determined by MTT assay on Vero cell line. Each data point represents mean  $\pm$  standard error (S.E) ( $n = 3$ ).

different concentrations of QPPI-NHAc (G3) and PPI-NH<sub>2</sub> (G3) are shown in Fig. 7. From the plot, it is observed that, the modified dendrimer carrier *viz.*, QPPI-NHAc (G3) has proved to be significantly less cytotoxic than its parent dendrimer PPI (G3) and PXM. The presence of PXM with the dendrimer carriers did not reduce cell viability on incubation over the tested concentration. Photographs of Vero cells (100  $\mu$ m) shown in Fig. 8a–d, which reflect the cytotoxicity, caused by 0.125 mM PPI-NH<sub>2</sub> (G3) dendrimers on Vero cells after 24 h. The cells got distorted and a significant decrease in cell density can be found (Fig. 8c). However, cells treated with QPPI-NHAc (G3) dendrimers were observed by retaining the cell structure and also considerable cell density (Fig. 8d), suggesting that modified dendrimers were less toxic on Vero cells compare to PXM. Normally, under *in vitro* conditions, the degree of cytotoxicity usually been evaluated based on the IC<sub>50</sub> values (concentrations where 50% of cells are viable) and this value (IC<sub>50</sub>) increases significantly for QPPI-NHAc (G3) dendrimers. That is, IC<sub>50</sub> value for PXM (control) and PPI-NH<sub>2</sub> (G3) is less than 0.0625 mM, whereas for QPPI-NHAc (G3), it is 0.5 mM. These results show that the QPPI-NHAc (G3) dendrimer can effectively increase the tolerance concentration during its administration. Previously reported studies reveal that electrostatic interaction between positively charged polymers and negatively charged cell membranes is responsible for cytotoxicity.<sup>21,32,50,51</sup> Particularly, cationic dendrimers containing –NH<sub>2</sub> surface group, are prone to attract the negatively charged cell membrane that in turn damaged the cell lines and thus reflected more cytotoxic effect.

Therefore, it is necessary that in any delivery, the cationic surface group of the polymer/dendrimer must be neutralized so as to eliminate the electrostatic interaction of surface group with the negatively charged cell membranes and this in turn reduced the cytotoxicity. On par with this, QPPI-NHAc (G3) is a carrier containing neutral surface group *viz.* –COCH<sub>3</sub> and CH<sub>3</sub>,

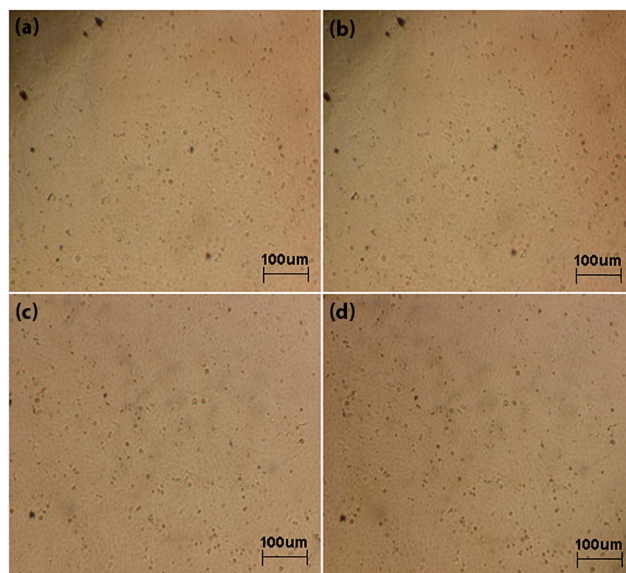


Fig. 8 Photographs of (a) Vero cell line and the same cell lines treated with (b) PXM, (c) PPI-NHAc (G3) and (d) QPPI-NHAc (G3).

and thereby their electrostatic interaction with Vero cell line was sizably reduced and hence lower cytotoxicity is noticed than its parent dendrimer and plain PXM. This proves that QPPI-NHAc (G3) dendrimer provides a potential platform to reduce the cytotoxicity and thus gives the pathway to increase its biocompatibility.

### 3.7 *In vivo* anti-inflammatory activity

The anti-inflammatory studies were performed using carrageenan-induced paw edema model. It is observed that the inhibition level of plain drug after 12 and 24 h was found to be 29.85% and 26.71% respectively (Fig. 9). But there was a marked observation in the case of dendrimer–drug formulations as in

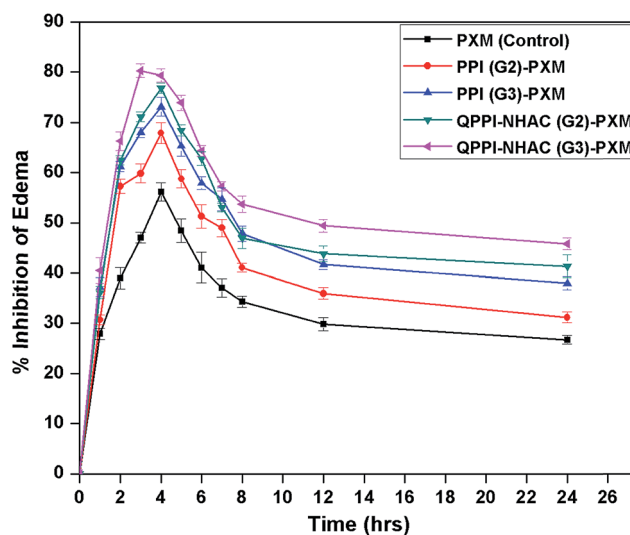


Fig. 9 Anti-inflammatory activity of PXM and various dendrimer–drug formulations ( $n = 6$ ).

PPI-NHAc (G2)-PXM 35.96% (12 h) & 31.18% (24 h), in PPI NHAc (G3)-PXM it was found to be 41.37% (12 h) & 37.96% (24 h) and in the modified dendrimer formulations namely QPPI-NHAc (G3)-PXM it was found as 43.82% (12 h) & 41.38% (24 h) and in the case of higher generation *viz.*, QPPI-NHAc (G3)-PXM displayed a maximum inhibition of 79.31% after 4 h and until the 8<sup>th</sup> h inhibition level was maintained above 50% and even after 12 and 24 h 49.42% and 45.83% inhibition was observed. However, in the case plain PXM drug the maximum inhibition was shown in 4<sup>th</sup> h with magnitude of 56% and just after 4<sup>th</sup> h it scored below 50%. The plausible reason for this may be the drug concentration in the body which was maintained for a longer duration in case of QPPI-NHAc (G3)-PXM when compared to plain PXM drug which signifies the localized action of QPPI-NHAc (G3)-PXM formulation in inflamed paw.

### 3.8 Pharmacokinetics and biodistribution studies

Pharmacokinetic results reveal that the blood levels of PXM in edema induced paw tissues were more with QPPI-NHAc (G3)-PXM formulation as compared with other formulations including the plain drug. The area under the plasma concentration time profile with QPPI-NHAc (G3)-PXM was  $298.54 \pm 2.07 \mu\text{g mL}^{-1} \text{ h}^{-1}$  which was higher compared to other formulations and plain drug which was  $281.28 \pm 1.98$  (QPPI-NHAc

(G2)-PXM),  $278.36 \pm 1.38$  (PPI-NHAc (G3)-PXM),  $275.24 \pm 1.33$  and  $268.11 \pm 1.29$  (plain PXM) respectively. The elimination rate constant for PXM observed in QPPI-NHAc (G3)-PXM formulation was 0.0149 as against QPPI-NHAc (G2)-PXM, PPI-NHAc (G3)-PXM, PPI-NHAc (G2)-PXM and plain drug as 0.0165, 0.0178, 0.0197 and 0.0186 respectively. Also the half-life of the drug PXM with QPPI-NHAc (G3)-PXM formulation was 38.1 h which is significantly higher compared to other formulation and the plain drug (QPPI-NHAc (G2)-PXM, PPI-NHAc (G3)-PXM, PPI-NHAc (G2)-PXM and plain drug as 33.4, 35.2, 32.6 and 31.7 respectively). All the results collectively suggested the extended plasma level of PXM.

From the organ distribution in the case of plain drug, percentage of drug recovered in paw increased in 4<sup>th</sup> h ( $2.32 \pm 0.38$ ) compared to 2<sup>nd</sup> h ( $1.49 \pm 0.24$ ) and diminished slightly after 8<sup>th</sup> h with  $0.30 \pm 0.64$  (Fig. 10). But in the case of QPPI-NHAc (G3)-PXM formulation it was observed that a continuous increase in percentage of drug recovered from paw with  $3.65 \pm 0.80$  (2 h),  $4.58 \pm 1.02$  (4 h) and  $4.92 \pm 1.24$  (8 h) respectively. The same trend of drug recovery in paw was also observed in other formulations *i.e.* in QPPI-NHAc (G2)-PXM the trend was  $3.18 \pm 0.62$  (2 h),  $3.97 \pm 0.81$  (4 h) &  $4.31 \pm 0.94$  (8 h), in the case of parent formulations PPI-NHAc (G3)-PXM the trend was  $2.87 \pm 0.24$ ,  $3.36 \pm 0.75$  and  $4.39 \pm 1.09$  and similarly for PPI-NHAc (G2)-PXM it was  $2.21 \pm 0.27$ ,  $2.92 \pm 0.34$  and 3.45

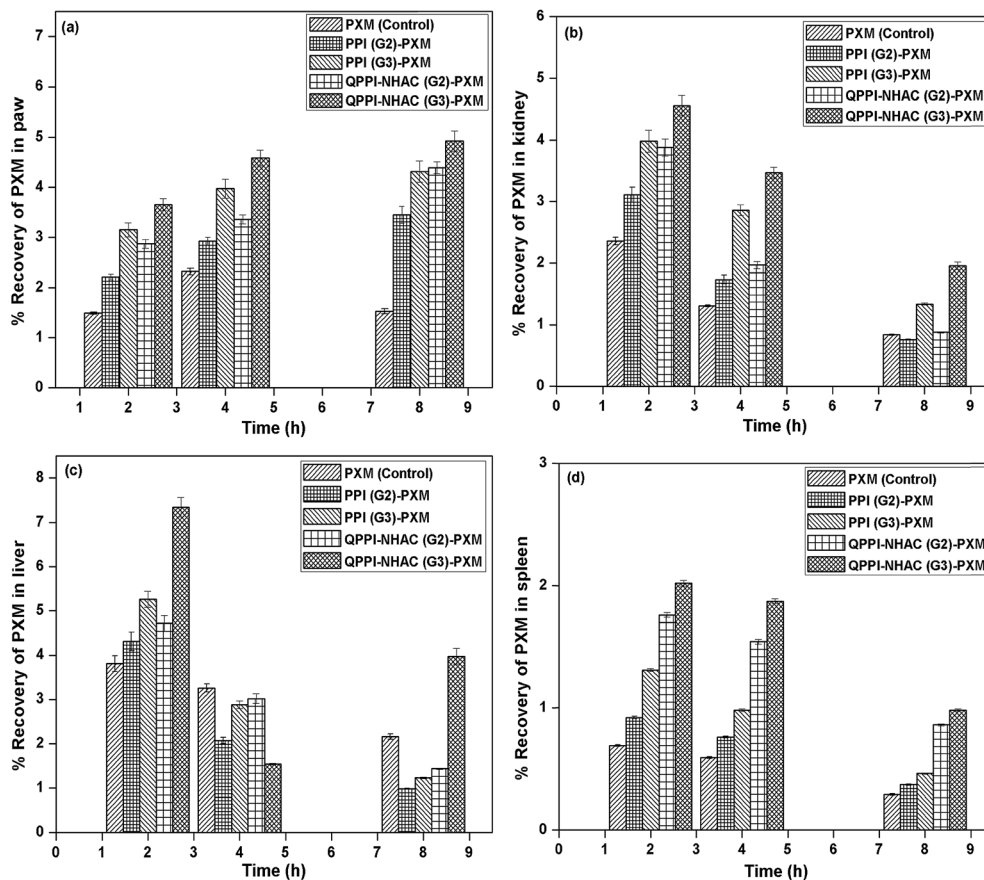


Fig. 10 Percentage of drug recovery in various organs after intravenous administration of the PXM with different formulations (a) paw, (b) kidney, (c) liver and (d) spleen ( $n = 6$ ).

$\pm 0.78$  respectively. Further it was also observed that the percentage of drug recovered in paw with QPPI-NHAc (G3)-PXM formulation was higher compared to others. It is noted that after 2 h of administration of plain drug, higher concentration of drug was observed in kidney and liver than in paw and also similar observations were made with other formulations. After 4<sup>th</sup> h, the plain PXM, QPPI-NHAc (G2)-PXM and their parent dendrimers continued to show higher recovery in liver ( $3.26 \pm 0.92$ ,  $2.07 \pm 0.28$ ,  $2.88 \pm 0.49$  &  $3.02 \pm 0.51$  respectively) than in paw ( $2.32 \pm 0.38$ ,  $2.92 \pm 0.34$ ,  $3.97 \pm 0.81$  &  $3.36 \pm 0.75$  respectively). However, it is worth to mention here that, from the formulation QPPI-NHAc (G3)-PXM, after 4<sup>th</sup> h higher concentration of PXM was recovered from paw  $4.58 \pm 1.02$  against  $3.47 \pm 0.84$  and  $3.98 \pm 0.90$  in kidney and liver respectively. From the collective results, it is evident that the preferred sites of drug associated with dendrimer formulation are the liver and inflamed paw.

In general, hydrophobic particles are immediately recognized as “foreign” and are generally covered by plasma proteins known to function as opsonins, which facilitate phagocytosis, thus in the present strategy, surface characteristics increase hydrophilicity and hence decrease macrophage phagocyte systems (MPS) clearance.<sup>52</sup> Because of this the drug levels observed in edema-induced paw tissues with dendritic formulation were much higher than the free drug at all-time intervals of study. As the dendritic formulation exhibits prolonged retention in blood, it may be proposed that the greater perfusion of blood at the inflamed site could bring a larger amount of PXM from QPPI-NHAc (G3)-PXM formulation in paw tissues. Also, as there is considerable change in vasculature at the site of inflammation, the chances of QPPI-NHAc (G3)-PXM localization increase further owing to the enhanced permeability and retention (EPR) effect at such sites as reported for macromolecular polymeric structures like PAMAM.<sup>53,54</sup> The other more plausible reason could be the interaction of PAMAM dendrimer with albumin protein as reported with bovine serum albumin.<sup>55</sup>

## 4. Conclusions

Dendrimer based drug delivery system provides an attractive platform to load and release of poorly soluble drug molecules and thus improves the pharmacodynamic and pharmacokinetic behavior of those drugs. In the present study, for the first time an efficient biocompatible dendrimer based drug carrier *viz.*, QPPI-NHAc (G2) and QPPI-NHAc (G3) were synthesized through simple reactions like surface acetylation and internal quaternization of their commercial PPI (G2) & PPI (G3) dendrimers. The presence of acetyl group on the surface and quaternary onium in the internal cavities of QPPI-NHAc (G2)/(G3) were established through spectral techniques. The drug carrying potential of these surface acetylated and interior quaternized dendrimers *viz.*, QPPI-NHAc (G2) & QPPI-NHAc (G3) with respect to piroxicam drug was investigated through aqueous solubility, *in vitro* release and cytotoxicity studies and proved these dendrimers are more active as compared with their PPI (G2), PPI (G3) and PEI. The observed result reveals that the higher generation dendrimer *i.e.*, QPPI-NHAc (G3) has

significantly solubilizing the PXM to the tune of 50 folds compared to its intrinsic solubility and ensures the promising loading of PXM. It is proved that the complexation of PXM with QPPI-NHAc (G3) is responsible for drug loading and in turn for increased solubility. The drug release noticed under *in vitro* conditions have suggested that QPPI-NHAc (G3) is a potential one to release the PXM for long time up to 570 min and hence it is more suitable for sustained/delayed release. The extent of sustained release observed in this study has reflected the magnitude of interaction between the dendrimer carrier and the drug molecule. Further, QPPI-NHAc (G3) dendrimers have showed increased cell viabilities on Vero cell lines which is correlated with its reduced cytotoxicity over the parent dendrimer as well as plain PXM and hence they are more biocompatible. The increased IC<sub>50</sub> value observed for QPPI-NHAc (G3) implies that it can effectively increase their tolerance concentration during the preparation of QPPI-NHAc-PXM (G3) dendrimer formulations. To suggest real time applications, the newly synthesized dendrimer based drug carriers were employed for *in vivo* studies namely (i) *in vivo* anti-inflammatory activity and (ii) biodistribution. The pharmacodynamic results in the case of dendrimer-drug formulation QPPI-NHAc-PXM (G3) which displays a maximum inhibition after 4<sup>th</sup> h (79.31%) and significant level was observed even after 12<sup>th</sup> and 24<sup>th</sup> h (49.42% and 45.83%). The 2<sup>nd</sup>, 4<sup>th</sup> and 8<sup>th</sup> h organ distribution results show higher recovery of PXM in rat paw with QPPI-NHAc-PXM (G3) formulation. In conclusion, dendrimer-based formulation may not only improve the solubility of PXM but also can certainly helpful to localization of drug at the site of inflammation and hence provide better therapeutic efficacy at a lower dose. Therefore, considering all the observed results, it is reported that the surface acetylated and internally quaternized dendrimer, QPPI-NHAc (G3) is undisputedly a promising candidate in the design of polymeric drug delivery system by selectively inhibiting cyclooxygenase (COX) located at the inflammation site.

## Conflict of interest

The opinions expressed are solely those of the author, and the authors state no conflict of interest and have received no payment in preparation of this manuscript.

## Acknowledgements

One of the authors Mr V. Yogaraj wishes to thank the UGC for providing financial assistance under UGC-UPPE Phase-II Programme and Mrs D. P. Geetha Rani would like to record her thanks to University of Madras and Govt. of Tamilnadu for providing financial support in the form of TNJRF.

## References

- 1 N. E. Lane and J. M. Thompson, *Am. J. Med.*, 1997, **103**(6A), 25–35.
- 2 P. Needleman and P. C. Isakson, *J. Rheumatol.*, 1997, **24**(49), 6–8.

- 3 W. W. Bolten, *J. Rheumatol.*, 1998, **25**(51), 2–7.
- 4 M. Depre, E. Ehrich, A. van Hecken, I. de Lepeleire, A. Dallob, P. Wong, A. Porras, B. J. Gertz and P. J. de Schepper, *Eur. J. Clin. Pharmacol.*, 2000, **56**, 167–174.
- 5 T. Hla and K. Neilson, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 7384–7388.
- 6 S. Silas and D. O. Clegg, *Bull. Rheum. Dis.*, 1999, **40**(2), 1–4.
- 7 T. W. Twomey, S. R. Banolucci and D. C. Hobbs, *J. Chromatogr.*, 1980, **183**, 104–108.
- 8 S. C. Chi and H. W. Jun, *J. Pharm. Sci.*, 1990, **79**, 974–977.
- 9 S. Z. Lin, D. Wouessidjewe, M. C. Poelman and D. Duchone, *Int. J. Pharm.*, 1994, **106**, 63–67.
- 10 S. C. Sweetman, *Martindale: The Complete Drug Reference*, Pharmaceutical Press, London, 34th edn, 2005.
- 11 S. Charumanee, S. Okonki and J. Sirithunyalug, *Chiang Mai J. Sci.*, 2004, **3**, 77–84.
- 12 E. R. Gillies and J. M. J. Fréchet, *Drug Discovery Today*, 2005, **10**, 35–42.
- 13 S. Svenson and D. A. Tomalia, *Adv. Drug Delivery Rev.*, 2005, **57**, 2106–2129.
- 14 E. Roseita and D. A. Tomalia, *Drug Discovery Today*, 2001, **6**, 427–436.
- 15 R. Duncan and L. Izzo, *Adv. Drug Delivery Rev.*, 2005, **57**, 2215–2237.
- 16 A. D. Emanuele and D. Attwood, *Adv. Drug Delivery Rev.*, 2005, **57**, 2147–2162.
- 17 A. R. Menjoge, R. M. Kannan and D. A. Tomalia, *Drug Discovery Today*, 2010, **15**, 171–185.
- 18 C. C. Lee, J. A. Mackay, J. M. J. Fréchet and F. C. Szoka, *Nat. Biotechnol.*, 2005, **23**, 1517–1526.
- 19 B. Boisselier, L. Liang, M. Dalko-Csiba, J. Ruiz and D. Astruc, *Chem.–Eur. J.*, 2010, **16**, 6056–6068.
- 20 U. Gupta, H. B. Agashe, A. Asthana and N. K. Jain, *Biomacromolecules*, 2006, **7**, 649–658.
- 21 R. Duncan and L. Izzo, *Adv. Drug Delivery Rev.*, 2005, **57**, 2215–2237.
- 22 R. Jevprasesphant, J. Penny, R. Jalal, D. Attwood, N. McKeown and D. D'Emanuele, *Int. J. Pharm.*, 2003, **252**, 263–266.
- 23 R. B. Kolhatkar, K. M. Kitchens, P. W. Swaan and H. Ghandehari, *Bioconjugate Chem.*, 2007, **18**, 2054–2060.
- 24 Y. Zhang, Y. Sun, X. Xu, H. Zhu, L. Huang, X. Zhang, Y. Qi and Y. Shen, *Bioorg. Med. Chem. Lett.*, 2010, **20**, 927–931.
- 25 I. J. Majoros, T. P. Thomas, C. B. Mehta and J. R. Baker, *J. Med. Chem.*, 2005, **48**(19), 5892–5899.
- 26 R. Roy and M. G. Baek, *Rev. Mol. Biotechnol.*, 2002, **90**, 291–309.
- 27 J. B. Nair, S. Mohapatra, S. Ghosh and K. K. Maiti, *Chem. Commun.*, 2015, **51**, 2403–2406.
- 28 A. Saha, S. Mohapatra, P. Kurkute, B. Jana, J. Sarkar, P. Mondal and S. Ghosh, *RSC Adv.*, 2015, **5**, 92596–92601.
- 29 K. Yang, L. Weng, Y. Y. Cheng, H. H. Zhang and J. H. Zhang, *J. Phys. Chem. B*, 2011, **115**, 2185–2195.
- 30 D. A. Tomalia, H. Baker, J. Dewald, M. Hall and G. Kallos, *Polym. J.*, 1985, **17**, 117–132.
- 31 Y. Zhang, Y. Sun, X. Xu, H. Zhu, L. Huang, Y. Qi and Y. Shen, *J. Med. Chem.*, 2010, **53**, 3262–3272.
- 32 M. L. Patil, M. Zhang, S. Betigeri, O. Taratula, H. He and T. Minko, *Bioconjugate Chem.*, 2008, **19**(7), 1396–1403.
- 33 T. Higuchi and K. A. Connors, Phase-Solubility Techniques, in *Advances in Analytical Chemistry and Instrumentation*, ed. C. N. Reilly, Wiley-Interscience, New York, 1965, vol. 4, pp. 117–212.
- 34 A. Asthana, A. S. Chauhan, P. Diwan and N. K. Jain, *AAPS PharmSciTech*, 2005, **6**, E536–E542.
- 35 A. S. Chauhan, N. K. Jain, P. V. Diwan and A. J. Khopade, *J. Drug Targeting*, 2004, **12**, 575–583.
- 36 N. Prajapati, K. R. Tekade, U. Gupta, V. Gajbhiye and K. Narendra, *Mol. Pharmaceutics*, 2009, **6**(3), 940–950.
- 37 M. Amanlou and A. R. Dehpour, *J. Chromatogr. B: Biomed. Sci. Appl.*, 1997, **696**, 317–319.
- 38 T. M. Manuela, P. P. Joao and A. M. Jose, *J. Pharm. Biomed. Anal.*, 1995, **13**, 319–322.
- 39 E. Murugan, D. P. Geetha Rani, K. Srinivasan and J. Muthumary, *Expert Opin. Drug Delivery*, 2013, **10**(10), 1319–1334.
- 40 E. Murugan, D. P. Geetha Rani and V. Yogaraj, *Colloids Surf., B*, 2014, **114**, 121–129.
- 41 Y. H. Zhao, M. H. Abraham, J. Lee, A. Hersey, C. N. Luscombe, G. Beck, B. Sherborne and I. Cooper, *Pharm. Res.*, 2002, **19**, 1446–1457.
- 42 Y. Cheng and T. Xu, *Eur. J. Med. Chem.*, 2008, **43**, 2291–2297.
- 43 E. Biosselier, C. Ornelas, I. Pianet, J. R. Aranzaes and D. Astruc, *Chem.–Eur. J.*, 2008, **14**(18), 5577–5587.
- 44 M. Santo and M. A. Fox, *J. Phys. Org. Chem.*, 1999, **12**, 293–307.
- 45 M. A. C. Broeren, B. F. M. de Waal, M. H. P. van Genderen, H. M. H. F. Sanders, G. Fytas and E. W. Meijer, *J. Am. Chem. Soc.*, 2005, **127**, 10334–10343.
- 46 M. Chai, Y. Niu, W. J. Youngs and P. L. Rinaldi, *J. Am. Chem. Soc.*, 2001, **123**, 4670–4678.
- 47 M. H. Chai, A. K. Holley and M. Kruskamp, *Chem. Commun.*, 2007, **2**, 168–170.
- 48 Y. Cheng, Y. Li, Q. Wu and T. Xu, *J. Phys. Chem. B*, 2008, **112**, 12674–12680.
- 49 Q. Wu, Y. Cheng, J. Hu, L. Zhao and T. Xu, *J. Phys. Chem. B*, 2009, **113**, 12934–12943.
- 50 J. Hu, M. Fang, Y. Cheng, J. Zhang, Q. Wu and T. Xu, *J. Phys. Chem. B*, 2010, **114**, 7148–7157.
- 51 J. H. Lee, Y. B. Lim, J. S. Choi, Y. Lee and T. Kim, *Bioconjugate Chem.*, 2003, **14**, 1214–1221.
- 52 A. Asthana, A. S. Chauhan, P. V. Diwan and N. K. Jain, *AAPS PharmSciTech*, 2005, **6**(3), E536–E542.
- 53 M. Kissel, P. Peschke and V. Subr, *J. Pharm. Sci. Technol.*, 2001, **55**, 191Y201.
- 54 D. J. A. Crommelin, W. E. Hennik and G. Strom, Drug targeting systems: fundamentals and applications to parenteral drug delivery, in *Drug Delivery and Targeting*, ed. A. M. Hillery, A. W. Lloyd and J. Swarbrick, Taylor and Francis Inc., New York, 2001, p. 119Y125.
- 55 B. Klajnert, L. Stanislawski, M. Bryszewska and B. Palecz, *Biochim. Biophys. Acta*, 2003, **1648**, 115Y126.



**Department of  
Master of Business  
Administration**



Vol 5 Issue 5 Feb 2016

ISSN No : 2249-894X

---

*Monthly Multidisciplinary  
Research Journal*

*Review Of  
Research Journal*

Chief Editors

---

**Ashok Yakkaldevi**  
A R Burla College, India

**Flávio de São Pedro Filho**  
Federal University of Rondonia, Brazil

**Ecaterina Patrascu**  
Spiru Haret University, Bucharest

**Kamani Perera**  
Regional Centre For Strategic Studies,  
Sri Lanka

## Welcome to Review Of Research

**RNI MAHMUL/2011/38595**

**ISSN No.2249-894X**

Review Of Research Journal is a multidisciplinary research journal, published monthly in English, Hindi & Marathi Language. All research papers submitted to the journal will be double-blind peer reviewed referred by members of the editorial Board readers will include investigator in universities, research institutes government and industry with research interest in the general subjects.

## Regional Editor

Manichander Thammishetty

Ph.d Research Scholar, Faculty of Education IASE, Osmania University, Hyderabad.

## Advisory Board

Kamani Perera Regional Centre For Strategic Studies, Sri Lanka	Delia Serbescu Spiru Haret University, Bucharest, Romania	Mabel Miao Center for China and Globalization, China
Ecaterina Patrascu Spiru Haret University, Bucharest	Xiaohua Yang University of San Francisco, San Francisco	Ruth Wolf University Walla, Israel
Fabricio Moraes de Almeida University of Rondonia, Brazil	Karina Xavier Massachusetts Institute of Technology (MIT), USA	Jie Hao University of Sydney, Australia
Anna Maria Constantinovici AL. I. Cuza University, Romania	May Hongmei Gao Kennesaw State University, USA	Pei-Shan Kao Andrea University of Essex, United Kingdom
Romona Mihaila Spiru Haret University, Romania	Marc Fetscherin Rollins College, USA	Loredana Bosca Spiru Haret University, Romania
	Liu Chen Beijing Foreign Studies University, China	Ilie Pinte Spiru Haret University, Romania
Mahdi Moharrampour Islamic Azad University buinzahra Branch, Qazvin, Iran	Nimita Khanna Director, Isara Institute of Management, New Delhi	Govind P. Shinde Bharati Vidyapeeth School of Distance Education Center, Navi Mumbai
Titus Pop PhD, Partium Christian University, Oradea, Romania	Salve R. N. Department of Sociology, Shivaji University, Kolhapur	Sonal Singh Vikram University, Ujjain
J. K. VIJAYAKUMAR King Abdullah University of Science & Technology, Saudi Arabia.	P. Malyadri Government Degree College, Tandur, A.P.	Jayashree Patil-Dake MBA Department of Badruka College Commerce and Arts Post Graduate Centre (BCCAPGC), Kachiguda, Hyderabad
George - Calin SERITAN Postdoctoral Researcher Faculty of Philosophy and Socio-Political Sciences AL. I. Cuza University, Iasi	S. D. Sindkhedkar PSGVP Mandal's Arts, Science and Commerce College, Shahada [ M.S. ]	Maj. Dr. S. Bakhtiar Choudhary Director, Hyderabad AP India.
REZA KAFIPOUR Shiraz University of Medical Sciences Shiraz, Iran	Anurag Misra DBS College, Kanpur	AR. SARAVANAKUMARALAGAPPA UNIVERSITY, KARAIKUDI, TN
Rajendra Shendge Director, B.C.U.D. Solapur University, Solapur	C. D. Balaji Panimalar Engineering College, Chennai	V.MAHALAKSHMI Dean, Panimalar Engineering College
	Bhavana vivek patole PhD, Elphinstone college mumbai-32	S.KANNAN Ph.D , Annamalai University
	Awadhesh Kumar Shirotriya Secretary, Play India Play (Trust), Meerut (U.P.)	Kanwar Dinesh Singh Dept.English, Government Postgraduate College , solan

More.....

**Address:-Ashok Yakkaldevi 258/34, Raviwar Peth, Solapur - 413 005 Maharashtra, India**  
**Cell : 9595 359 435, Ph No: 02172372010 Email: ayisrj@yahoo.in Website: www.ror.isrj.org**

# Review of Research

International Online Multidisciplinary Journal

ISSN: 2249-894X

Impact Factor : 3.1402(UIF)

Volume - 5 | Issue - 5 | Feb - 2016



## AN EMPIRICAL STUDY ON IMPACT OF ORGANIZATIONAL CITIZENSHIP BEHAVIOR AMONG EMPLOYEES OF PRIVATE SECTOR BANKS IN TELANGANA STATE



A. Satish Kumar<sup>1</sup> And K. Veeraiah<sup>2</sup>

<sup>1</sup>Assoc. Professor-cum-HOD, Dvr College Of Engineering And Technology, Kashipur Village, Kandi, Sangareddy Mandal, Medak.

<sup>2</sup>professor-cum-HOD, Malla Reddy Engineering College (autonomous), Maissmaguda, Dhulapally (v), Secunderabad.

### ABSTRACT :

Organizational citizenship behavior (OCB) is referred to as a set of discretionary workplace behaviors that exceed one's basic job requirements. They are often described as behaviors that go beyond the call of duty. In this research the various factors that influences the OCB among middle level employee's of Private sector banks were identified. Primary data collected with the help of structured questionnaire administrated to middle level employees of top 2 private sector banks in Telangana state.

Data was collected from respective banks main branch. Totally there were 215 middle level employees in all the two banks. In that 183 respondents were selected by adopting simple random sampling using lottery method. To find out the suitability of the questionnaire a pilot study was carried out among 15 respondents and necessary additions and deletions were made in the questionnaire. To test the reliability and validity of the data collected cronbach's alpha test was used and values of Coefficient alpha (Cronbach's Alpha) have been obtained, the minimum value of Coefficient alpha obtained was 0.845. This shows data has satisfactory internal consistency reliability. Using Statistical Package for Social Science (SPSS) the following relevant tools were administered 1) Factor Analysis 2) Multiple Regression and 3) Reliability. Some of the relevant findings were derived that will be significant to present Indian Scenario.



**KEY WORDS :** Organization, Citizenship, Banks, Behaviour etc

## INTRODUCTION

Organizational Citizenship Behavior (OCB) is presumed as one of the emerging management concepts that are being emphasized for the organizational effectiveness. Organizational citizenship behavior (OCB) is a term that encompasses anything positive and constructive that employees do, of their own volition, which supports co-workers and benefits the company. Typically, employees who frequently engage in OCB may not always be the top performers (though they could be, as task performance is related to OCB), but they are the ones who are known to 'go the extra mile' or 'go above and beyond' the minimum efforts required to do a merely satisfactory job. Organizational citizenship behavior (OCB) has garnered much academic attention since its conception.

There are many factors that can contribute to the determination of Organizational Citizenship Behavior (OCB) which include Altruism, Conscientiousness, Civic Virtue, Sportsmanship, Courtesy, etc. OCBs may enhance coworker productivity. Employees who help another coworker "learn the ropes" may help them to become more productive employees faster. Over time, helping behavior can help to spread "best practices" throughout the work unit or group.

Employees who attend and actively participate in meetings may aid the dissemination of information in an organization, thus enhancing its responsiveness. Employees, who exhibit sportsmanship, by demonstrating a willingness to take on new responsibilities or learn new skills, enhance the organization's ability to adapt to changes in its environment.

## Review of Literature

**Abubakr suliman and Hanan al obaidly, (2013)** This research aims at investigating, for the first time in the Arab World, the influence of leadership behaviors on organizational citizenship behaviors (OCB) in the Islamic banking sector. Also, it explores the role of OCB in affecting work outcomes. The study consists of two core concepts: leadership behaviors and OCB. The sample population for the study was drawn from 150 employees working for several Islamic banks in the United Arab Emirates (UAE). A self-administered questionnaire was developed by combining two instruments. The first one was Bass and Avolio (1995) MLQ x5 scales, and the second one was Konovsky and Organ's (1996) OCB questionnaire. Participants, who were Islamic banks employees, were invited by e-mail to complete a web-based survey. The data was collected through a self-response questionnaire.

**Zahra Jafari Karfestani et al., (2013)** In the current challenging world, the organizations to compete in the world level, meet the needs and expectations of customers and compatibility with the transforming nature of the job attempt to engage the employees that act beyond the task and role determined in their job description. Against the past that the employees were expected to act to the extent of formal roles, in the new psychological contracts, the behaviors beyond the role is expected. Today, these efforts beyond the expectation are referred to as behaviors in addition to the role or organizational citizenship behaviors in the literature of organization and management science. Organizational citizenship behavior is a completely voluntary behavior that the rewarding system is not able to identify it directly but altogether it increases the effective performance in the organization. In this paper, in addition to presenting the definitions of organizational citizenship behavior, the classifications and formation of this behavior approaches are studies, as well as its nature in the various cultures are discussed, ultimately the factors affecting and forming the citizenship behavior together with its consequences are explained.

**Seyed Ali Vaziri et al., (2013)** Nowadays, each organization is trying to increase the productivity and

effectiveness of its staff by paying attention to issue of management personality. According to the role characteristics of organizational citizenship in management personality, this research aims to study the relationship between the two items. This study is applicable and the nature of study is descriptive. All managers and population are 315 persons in Khorasan Agriculture Organization (KAO). The sample selected randomly and there and included 174 persons. Data collection –tool was questionnaire and used descriptive statistics and inferential statistics for data analysis. The results showed that the positive and significant relation exist between characteristic of managers for entrepreneurial and organizational citizenship behavior as much as 95%.

**Lee Kim Lian and Low Guan Tui (2012)** The objective of this study is to test a theory-based model predicting the relationships between leadership styles, subordinates' competence, downward influence tactics and outcome of organizational citizenship behavior in Malaysian-based organizations. Data was collected from 347 respondents that represent major industries like services, manufacturing, mining and construction companies. Path analysis technique was used to test the model developed. The results show that the transformational leadership style has significant positive relationship with subordinates' organizational citizenship behavior, whereas the transactional leader style is negatively related to organizational citizenship behavior. This result illustrates the direct effects of leadership styles on the subordinates' outcome. In addition, inspirational appeals and consultation tactics, as downward influence tactics, were found to mediate the relationship between transformational leadership and organizational citizenship behavior. Likewise, subordinates' competence mediates the relationship between transformational leadership and consultation tactics. These results only partially support the efficacy of the influence theory, and therefore lend support to contingency theories of leadership. Implications for research and direction for future research are also discussed. The findings revealed that transformational and transactional leadership styles tend to play a significant role in employees' OCB. Nonetheless, passive/avoidant leadership style plays no role of statical evidence in the relationship. The findings of this study also revealed that transactional leadership behaviors have a stronger positive influence on OCB than transformational leadership behaviors. Though, the difference of the relationship between the two styles is very small. The findings are discussed in details in the research along with some recommendations for supervisors and researchers.

**Zahra Jafari Karfestani et al., (2012)** Organizations want and need employees who will do those things that aren't in any job description. And the evidence indicates that those organizations that have such employees outperform those that don't. As a result, some human subject studies are concerned with organizational citizenship behavior as a dependent variable. Organizational citizenship behavior (OCB) means individual behaviors that are beneficial to the organization but not directly recognized by the formal reward system. Though there are research studies that provide the domain of OCB and its effects on the organizational performance but they have varied viewpoints and are inadequate. The present paper is an effort towards this direction. The main objective of the paper is to build grounding for analyzing the impact of OCB on various organizational performance measures through various propositions based on antecedents and consequences of organizational citizenship behavior. The authors have described the antecedents of OCB from comprehensive perspective, which include role clarity, leadership, organizational commitment, organizational justice and individual traits. The impact of these antecedents is correlated with five organizational performance parameters namely reduced turnover, reduced absenteeism, employee satisfaction and loyalty, consumer satisfaction and consumer loyalty. The paper also stressed on the need to asses the influence of age, gender and

experience on OCB, which have been the main lacunae in the existing literature on OCB. Moreover, the present paper proposed an endeavor to incorporate three dimensions viz: age, gender and experience to assess their influence on OCB.

**Madiha Rehman Farooqui., (2012)** The present study at hand aims to explore different dimensions of Organizational Climate (OC) and establishes its relationship with OCB. The effect of gender is also taken into account. This is an explanatory study based on 114 faculty members (lecturers) randomly selected from public sector universities of Lahore. Data is collected through an online questionnaire designed on five point likert scale. Data has been analyzed through both descriptive and inferential statistics. All the dimensions of the OC are found to be significantly related to OCB and gender has also an explanatory power towards OCB. The practical implications along with the recommendations for the future researchers are provide at the end of the study.

**Dr.Nadeem Ahmed and Anwar Rasheed., (2012)** point out Organizational Citizenship Behavior(OCB) is presumed as one of the emerging management concepts that are being emphasized for the organizational effectiveness. This study aims to explore OCB and its significance for the organizations in present scenario particularly banking sector. Key predictors of OCB are identified through comprehensive literature review whereas qualitative research method is employed to explore the association. A model has presented by the researchers organizational citizenship behavior and its significant relation with Job satisfaction and commitment, employee engagement and human resource development climate (HRDC). It is proposed that well established predictors of OCB may lead to promote required behaviors among employees for improved performance and negative voluntary intentions

#### Research Questions

1. To study the factors that influences the OCB among the middle level employee's in Private sector banks in Telangana State.
2. To know the influence of Sportsmanship Behavior on OCB in Private Banks.

#### Methodology

The study is a descriptive one. Primary data collected with the help of structured questionnaire administrated to middle level employees of top 2 private sector banks in Telangana state. Data was collected from respective banks main branch. Totally there were 215 middle level employees in all the two banks. In that 183 respondents were selected by adopting simple random sampling using lottery method. To find out the suitability of the questionnaire a pilot study was carried out among 15 respondents and necessary additions and deletions were made in the questionnaire. To test the reliability and validity of the data collected cronbach's alpha test was used and values of Coefficient alpha (Cronbach's Alpha) have been obtained, the minimum value of Coefficient alpha obtained was 0.845. This shows data has satisfactory internal consistency reliability. Using Statistical Package for Social Science(SPSS) the following relevant tools were administered 1) Factor Analysis 2) Multiple Regression and 3) Reliability.

#### KMO and Bartlett's Test

The individual statements on OCB was examined using factor analysis based on 50 individual statements and the reliability of the samples collected was tested for internal consistency of the

grouping of the items

**Table 1**

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.667
Bartlett's Test of Sphericity	Approx. Chi-Square	10145.001
	Df	1225
	Sig.	.000

KMO measure of sampling adequacy is an index to examine the appropriateness of factor analysis. High values between 0.5 and 1.0 indicate factor analysis is appropriate. Values below 0.5 imply that factor analysis may not be appropriate. From the above table it is seen that Kaiser – Meyer – Olkin measure of sampling adequacy index is 0.667 and hence the factor analysis is appropriate for the given data set. Bartlett's Test of Sphericity is used to examine the hypothesis that the variables are uncorrelated. It is based on chi- Square transformation of the determinant of correlation matrix. A large value of the test statistic will favor the rejection of the null hypothesis. In turn this would indicate that factor analysis is appropriate. Bartlett's test of Sphericity Chi-square statistics is 10145.001, that shows the 50 statements are correlated and hence as inferred in KMO, factor analysis is appropriate for the given data set.

**Table 2**

Component	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.658	17.317	17.317	8.658	17.317	17.317	3.944	7.888	7.888
2	5.280	10.561	27.877	5.280	10.561	27.877	3.854	7.708	15.596
3	3.847	7.694	35.572	3.847	7.694	35.572	3.562	7.125	22.721
4	2.471	4.943	40.514	2.471	4.943	40.514	3.388	6.776	29.497
5	2.121	4.242	44.756	2.121	4.242	44.756	3.291	6.582	36.079
6	1.904	3.807	48.563	1.904	3.807	48.563	2.579	5.157	41.236
7	1.851	3.702	52.264	1.851	3.702	52.264	2.236	4.472	45.708
8	1.608	3.217	55.481	1.608	3.217	55.481	2.187	4.374	50.082
9	1.386	2.773	58.254	1.386	2.773	58.254	2.051	4.101	54.183
10	1.371	2.743	60.996	1.371	2.743	60.996	1.943	3.886	58.069
11	1.258	2.516	63.512	1.258	2.516	63.512	1.786	3.573	61.642
12	1.127	2.254	65.767	1.127	2.254	65.767	1.557	3.114	64.755
13	1.025	2.051	67.817	1.025	2.051	67.817	1.531	3.062	67.817
14	0.958	1.917	69.734						
15	0.906	1.812	71.546						
16	0.841	1.682	73.228						
17	0.808	1.615	74.844						
18	0.773	1.546	76.390						
19	0.732	1.463	77.854						
20	0.698	1.395	79.249						
21	0.638	1.276	80.525						
22	0.601	1.203	81.728						
23	0.578	1.156	82.884						
24	0.561	1.122	84.006						
25	0.556	1.112	85.118						
26	0.536	1.071	86.190						
27	0.489	0.978	87.168						
28	0.459	0.919	88.086						
29	0.452	0.904	88.990						
30	0.433	0.865	89.855						
31	0.415	0.831	90.686						
32	0.390	0.781	91.466						
33	0.372	0.745	92.211						
34	0.368	0.736	92.947						

AN EMPIRICAL STUDY ON IMPACT OF ORGANIZATIONAL CITIZENSHIP BEHAVIOR AMONG EMPLOYEES OF.....

35	0.345	0.690	93.637						
36	0.320	0.639	94.277						
37	0.316	0.632	94.908						
38	0.294	0.587	95.495						
39	0.281	0.563	96.058						
40	0.263	0.526	96.584						
41	0.249	0.498	97.082						
42	0.239	0.478	97.560						
43	0.207	0.415	97.975						
44	0.199	0.397	98.372						
45	0.187	0.373	98.745						
46	0.154	0.307	99.052						
47	0.144	0.289	99.341						
48	0.131	0.261	99.602						
49	0.116	0.232	99.835						
50	0.083	0.165	100.000						
Extraction Method: Principal Component Analysis.									

Eigen Value represents the total variance explained by each factor. Percentage of the total variance attributed to each factor. One of the popular methods used in Exploratory Factor Analysis is Principal Component Analysis, Where the total variance in the data is considered to determine the minimum number of factors that will account for maximum variance of data.

Table 3

Rotated Component Matrix													
ITEM	Component												
	1	2	3	4	5	6	7	8	9	10	11	12	13
Brought work home to prepare for next day	0.293	0.761	0.025	-0.154	-0.089	0.098	-0.191	0.122	-0.107	-0.064	-0.043	0.007	-0.074
Your supervisor deal with you in a truthful manner	0.285	<b>0.767</b>	0.002	-0.152	-0.066	0.129	-0.205	0.155	-0.083	-0.051	0.019	-0.005	-0.095
Obeay bank rules and regulation evenno one watching	0.424	0.623	0.000	0.011	0.022	0.019	0.201	0.154	-0.010	-0.110	0.012	0.115	-0.151
Am one of the most conscientious employees in this bank	0.747	0.124	-0.157	0.115	-0.097	-0.028	-0.055	0.190	0.133	-0.056	-0.005	-0.023	0.121
Giving an honest day's work for an honest day's pay	0.739	0.140	0.046	0.020	0.011	0.088	-0.199	0.193	-0.047	0.136	-0.021	0.186	0.130
Would not be as attached to its as i am currently	0.761	0.175	0.227	-0.166	0.009	0.104	0.142	-0.006	-0.181	0.018	-0.072	0.022	-0.037
Personal values and those of the bank's have become more similar	<b>0.792</b>	0.201	-0.036	-0.100	0.024	0.179	-0.032	-0.067	0.028	0.074	0.068	-0.044	-0.071
Prefer this bankto others is because of what it stands for its values	0.346	0.190	0.232	-0.196	0.477	0.259	0.156	-0.016	0.060	-0.076	0.038	-0.005	0.113
Primarily based on the similarity of my values and those respected by the bank	0.095	0.029	0.015	-0.144	0.037	0.783	-0.112	0.081	-0.009	0.005	0.007	0.073	0.134
Proud to tell others that am part of this bank	0.086	0.147	0.080	-0.054	0.056	0.747	0.241	0.043	-0.067	-0.160	-0.024	-0.021	-0.160
Volunteered for extra work assignments	0.107	0.067	-0.060	-0.095	-0.102	<b>0.846</b>	-0.068	0.060	-0.029	-0.023	0.024	-0.022	-0.033
volunteered to attend meetings or work on committees on own time	0.008	0.117	0.137	-0.022	0.094	0.362	<b>0.663</b>	0.084	0.224	-0.073	-0.065	0.032	-0.058
Decorated, straightened up, or otherwise beautified common work space	-0.081	-0.129	0.156	0.063	-0.023	-0.083	0.007	0.008	<b>0.750</b>	0.000	-0.135	-0.121	-0.086
Functions that are not required, but help the bank's image	0.072	-0.105	-0.284	0.067	-0.417	0.032	0.009	-0.023	0.524	0.169	-0.263	0.097	0.027
Read and keep up with banking announcements, memos and so on	0.088	0.000	0.138	0.025	-0.012	0.040	0.005	0.029	0.484	0.326	-0.452	0.093	-0.059
No reason to expend extra effort on behalf of this bank	-0.034	0.132	0.595	0.063	0.221	0.032	0.083	-0.007	0.258	0.174	-0.170	-0.178	0.149
Work for the bank is directly linked to how much i am rewarded	-0.005	-0.080	-0.385	-0.011	-0.436	0.030	-0.322	0.007	0.076	0.025	0.483	-0.271	-0.075

# AN EMPIRICAL STUDY ON IMPACT OF ORGANIZATIONAL CITIZENSHIP BEHAVIOR AMONG EMPLOYEES OF.....

It is necessary to express the right attitude	0.047	0.004	-0.389	-0.007	-0.083	0.107	-0.204	0.113	-0.182	0.111	0.541	0.433	-0.049
Abreast of changes in the bank	0.023	0.120	0.190	-0.041	0.107	-0.017	0.083	0.073	-0.154	0.066	<b>0.822</b>	0.016	0.007
Attend meetings are not mandatory but are considered important	0.035	0.038	-0.121	-0.006	0.090	0.030	0.068	0.026	-0.049	0.014	-0.018	<b>0.850</b>	0.082
Lent a compassionate ear when someone had a work problem	0.113	0.159	<b>0.676</b>	-0.014	0.130	0.017	-0.054	-0.101	0.110	0.004	0.147	-0.154	0.076
Lent a compassionate ear when someone had a personal problem	0.055	0.009	-0.632	0.055	-0.422	-0.037	0.095	-0.024	0.186	0.209	0.086	0.085	0.109
contributed and/or sent cards for co-workers birthday and special occasions	0.008	-0.122	-0.012	0.177	-0.059	-0.129	0.177	0.024	0.061	<b>0.805</b>	0.017	0.035	0.045
Changed vacation schedule, work days, or shifts to accommodate co-worker's needs	0.135	-0.071	-0.066	0.176	-0.074	0.014	-0.147	0.135	0.026	0.796	0.074	-0.052	-0.145
Worked weekends or other days off to complete a project or task	-0.064	0.076	0.586	0.068	0.515	-0.008	0.364	-0.021	0.045	0.070	-0.052	0.175	0.041
Developed extracurricular activities for co-workers	0.089	-0.082	-0.056	-0.116	-0.172	0.354	-0.697	0.032	0.107	-0.057	-0.025	-0.073	-0.177
Gave up meal and other breaks to complete work	0.177	0.201	-0.118	-0.198	0.035	0.282	0.197	0.248	-0.001	-0.442	0.129	-0.049	-0.148
Try to avoid creating problems for co-workers	0.231	0.010	-0.043	-0.085	0.175	0.153	0.143	0.582	0.120	-0.098	0.211	0.026	0.042
Considered impact of my actions on co-workers	0.042	0.035	-0.096	-0.087	-0.049	0.056	-0.027	<b>0.813</b>	0.014	0.005	0.050	-0.050	0.062
Always ready to lend a helping hand to those around me	0.041	0.179	-0.001	-0.151	-0.156	0.038	-0.086	0.752	-0.157	0.169	-0.097	0.053	-0.059
Happy to share meals with my co-workers	-0.033	0.288	0.476	-0.176	0.101	-0.005	0.299	0.481	-0.090	0.063	-0.036	0.115	0.065
Engage bank guests as part of my duty	0.104	-0.017	-0.143	0.102	-0.683	0.098	0.016	-0.013	0.008	0.134	-0.050	-0.049	0.105
Time to advise, coach or mentor of co-worker	0.186	-0.071	0.173	0.219	0.120	0.137	-0.091	-0.048	0.299	0.011	0.105	-0.107	-0.471
Helped co-worker learn new skills or shared job knowledge	0.307	0.017	-0.185	0.232	-0.454	0.103	-0.035	-0.041	0.163	-0.078	0.076	0.133	-0.224
Offered suggestions to improve how work is done	0.232	0.067	0.517	0.135	-0.052	-0.073	0.101	-0.081	0.264	-0.093	0.079	0.462	-0.133
Offered suggestions to improving work environment	0.028	0.071	0.411	0.186	0.155	-0.051	0.117	-0.144	0.466	-0.092	0.128	0.112	0.213
Finished something for co-worker who had to leave early	-0.068	-0.100	-0.239	0.576	-0.069	-0.058	-0.045	-0.094	0.337	0.023	-0.097	-0.030	0.096
Helped less capable co-worker	-0.026	-0.076	-0.111	0.717	0.036	-0.024	-0.023	-0.035	0.156	0.120	-0.135	0.025	-0.117
Help others who have been absent for long time	-0.015	0.019	-0.015	<b>0.786</b>	-0.211	-0.022	-0.020	-0.120	-0.063	0.145	-0.020	-0.034	0.011
Help orient new people even though it is not required	0.016	-0.114	0.152	0.785	-0.046	-0.079	0.075	-0.065	-0.115	0.062	0.017	0.083	-0.004
Consume lot of complaining about trivial matters	-0.029	-0.123	0.116	0.738	-0.014	-0.198	0.066	-0.062	0.081	0.079	0.152	-0.063	0.035
Tend to make problems bigger than they are	-0.047	-0.116	0.132	0.421	0.406	-0.092	-0.348	-0.074	0.078	0.030	-0.076	0.236	0.065
Always focus on what is wrong with my situation, rather than positive side of it	0.033	0.733	0.177	-0.088	0.116	0.050	0.175	-0.025	-0.007	-0.060	0.064	-0.072	0.298
Talk up this bank to my friends as a great place to work	0.034	0.764	0.175	-0.093	0.237	0.032	0.184	0.013	0.015	-0.056	0.053	0.001	0.226
Take steps to try to prevent problems with other workers	0.115	0.644	0.146	-0.052	0.389	0.048	0.284	0.031	-0.057	-0.068	0.075	0.117	-0.041
Volunteered to work at after-hours or out-of-town events	0.438	0.232	0.084	0.078	0.387	0.020	0.023	0.089	0.144	-0.043	-0.012	-0.068	0.466
Volunteered to help co-worker deal with difficult customer, vendor, or co-worker	0.391	0.029	0.299	0.112	0.073	0.007	0.021	0.048	0.043	-0.051	0.059	0.066	<b>0.628</b>
Mindful of how my behavior affect other people's job	0.418	0.180	0.420	-0.098	0.364	0.054	0.308	0.005	-0.084	0.019	0.047	0.008	0.139
Help others who have heavy work loads	0.401	0.201	0.315	-0.023	0.428	0.129	0.241	-0.155	-0.002	0.002	0.116	-0.037	0.292
Willing to help others who have work related problems	0.284	0.310	0.118	-0.031	<b>0.622</b>	0.097	0.210	-0.109	0.024	0.028	0.027	0.044	0.081
Extraction Method: Principal Component Analysis.													
Rotation Method: Varimax with Kaiser Normalization.													
a. Rotation converged in 20 iterations.													

Interpretation of factors is facilitated by identifying the statements that have large loadings in the same factor. The factor can be interpreted in terms of the statement that loads high on it.

The factors of a study on factors influencing the OCB comprises of 50 individual statements. Out of that, 13 individual statements contribute more towards the study (67.817%).

The statements are:

- 1. Your supervisor deal with you in a truthful manner
- 2. Personal values and those of the bank's have become more similar
- 3. Volunteered for extra work assignments
- 4. Volunteered to attend meetings or work on committees on own time
- 5. Decorated, straightened up, or otherwise beautified common work space
- 6. Abreast of changes in the bank
- 7. Attend meetings are not mandatory but are considered important
- 8. Lent a compassionate ear when someone had a work problem
- 9. Contributed and/or sent cards for co-workers birthday and special occasions
- 10. Considered impact of my actions on co-workers
- 11. Help others who have been absent for long time
- 12. Volunteered to help co-worker deal with difficult customer, vendor, or co-worker
- 13. Willing to help others who have work related problems

Influence of OCB on Sportsmanship  
Table 4 Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.537	0.288	0.270	0.517
a. Predictors: (Constant), Willing to help others who have work related problems, Mindful of how my behavior affect other people's job, Consume lot of complaining about trivial matters, Volunteered to help co-worker deal with difficult customer, vendor, or co-worker, Volunteered to work at after-hours or out-of-town events, Talk up this bank to my friends as a great place to work, Trend to make problems bigger than they are, Help others who have heavy work loads, Take steps to try to prevent problems with other workers, Always focus on what is wrong with my situation, rather than positive side of it				

The multiple regressions are shown in the above table. The model summary table shows R-Square for this model is .288. This means that 28.8 percent of the variation in overall Organizational effectiveness (dependent variable) can be explained from the 10 independent variables. The table also shows the adjusted R-square for the model as .270.

Any time another independent variable is added to a multiple regression model, the R-square will increase (even if only slightly). Consequently, it becomes difficult to determine which models do the best job of explaining variation in the same dependent variable. The adjusted R-Square does just what its name implies. It adjusts the R-square by the number of predictor variables in the model. This adjustment allows the easy comparison of the explanatory power of models with different numbers of predictor's variable. It also helps us decide how many variables to include in our regression model.

TABLE 5 ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	44.274	10	4.427	16.533	.000
	Residual	109.526	409	0.268		
	Total	153.800	419			
Predictors: (Constant), Willing to help others who have work related problems, Mindful of how my behavior affect other people's job, Consume lot of complaining about trivial matters, Volunteered to help co-worker deal with difficult customer, vendor, or co-worker, Volunteered to work at after-hours or out of-town events, Talk up this bank to my friends as a great place to work, Trend to make problems bigger than they are, Help others who have heavy work loads, Take steps to try to prevent problems with other workers, Always focus on what is wrong with my situation, rather than positive side of it Dependent Variable: Organizational citizenship behavior						

The ANOVA table, as displayed in the above table shows the F ratio for the regression model that indicates the statistical significance of the overall regression model. The F ratio is calculated the same way for regression analysis as it was for the ANOVA technique. The variance Independent variable that is associated with dependent variable (Overall Organizational citizenship behavior) is referred to as explained variance. The remainder of the total variance in Independent variable that is not associated with dependent variable is referred as unexplained variance.

The larger the F ratio the more will be the variance in the dependent variable that is associated with the independent variable. The F ratio = 16.533. The statistical significance is .000 - the "Sig". So we can reject the null hypothesis that no relationship exists between the two variables. There is relationship between independent and dependent variables.

TABLE 6 Coefficients

Model	Unstandardised Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	2.224	0.223		9.982	0.000
Consume lot of complaining about trivial matters	0.159	0.047	0.198	3.413	<b>0.001</b>
Trend to make problems bigger than they are	0.056	0.039	0.088	1.434	0.152
Always focus on what is wrong with my situation, rather than positive side of it	-0.131	0.041	-0.210	-3.177	<b>0.002</b>
Talk up this bank to my friends as a great place to work	0.060	0.034	0.097	1.766	0.078
Take steps to try to prevent problems with other workers	-0.095	0.042	-0.149	-2.279	<b>0.023</b>
Volunteered to work at after-hours or out-of-town events	0.195	0.038	0.312	5.195	<b>0.000</b>
Volunteered to help co-worker deal with difficult customer, vendor, or co-worker	-0.003	0.028	-0.006	-0.109	0.913
Mindful of how my behavior affect other people's job	-0.001	0.035	-0.002	-0.041	0.968
Help others who have heavy work loads	0.053	0.035	0.095	1.545	0.123
Willing to help others who have work related problems	0.131	0.029	0.244	4.563	<b>0.000</b>

Dependent Variable: Overall Organizational citizenship behavior

To determine if one or more of the independent variables are significant predictors of overall Organizational citizenship behavior, we examine the information provided in the coefficient table. Out of ten independent statements five statements are statistically significant.

Consume lot of complaining about trivial matters has a beta coefficient 0.198, which is significant (0.001). Trend to make problems bigger than they are has a beta coefficient 0.088, which is not significant (0.152). Always focus on what is wrong with my situation, rather than positive side of it has a beta coefficient -0.210, which is significant (0.002). Talk up this bank to my friends as a great place to work has a beta coefficient 0.097, which is not significant (0.078). Take steps to try to prevent problems with other workers has a beta coefficient -0.149, which is significant (0.023). Volunteered to work at after-hours or out-of-town events has a beta coefficient 0.312, which is significant (0.000). Volunteered to help co-worker deal with difficult customer, vendor, or co-worker has a beta coefficient -0.006, which is not significant (0.913). Mindful of how my behavior affects other people's job has a beta coefficient -0.002, which is not significant (0.968). Help others who have heavy work loads has a beta coefficient 0.095, which is not significant (0.123). Willing to help others who have work related problems has a beta coefficient 0.244, which is significant (0.000).

## CONCLUSION

OCB plays a significant role in the organization's overall success; it has such a significant impact on the productivity and efficiency of the organization. OCB should be considered an efficient way of improving organizational profitability and reducing costs through, for example, lowering rates of absenteeism and turnover. At the same time it increases employee performance and wellbeing, as cooperative workers are more productive, and OCB enhances the social environment in the workplace.

## REFERENCES

1. Abubakr Suliman and Hanan AL Obaidly (2013). Leadership and Organizational Citizenship Behavior (OCB) in the Financial Service Sector: The Case of the UAE, Asia-Pacific Journal of Business Administration, Vol. 5 Iss: 2.
2. Zahra Jafari Karfestani et al., (2013). Organizational Citizenship Behavior as an Unavoidable Necessity for increasing the Effectiveness of Organizations, Interdisciplinary Journal of Contemporary Research in Business, January 2013, VOL 4, NO 9, Pp.827-850.
3. Seyed Ali Vaziri et al., (2013). Investigation Relationship between Organizational Citizenship Behavior and personality of Organizational Entrepreneurial Managers, Life Sci J 2013;10(1s):270-277.
4. Dr.Nadeem Ahmed and Anwar Rasheed, (2012). An Exploration of Predictors of Organizational Citizenship Behaviour and its Significant Link to Employee Engagement, International Journal of Business, Humanities and Technology, Vol. 2 No. 4; June 2012. pp. 99-100.
5. Zeinab Sheikhi et al., (2012). Consideration on Effects of Organizational Citizenship Behavior on Improvement of Electronic Customer Relationship Managements Through Website of Bank, Journal of Basic and Applied Scientific Research, J. Basic. Appl. Sci. Res., 2(6) 5806-5813, 2012 TextRoad Publication.
6. Muhammad Shakeel Aslam et al., (2012). Job Burnout and Organizational Citizenship Behaviors: Mediating Role of Affective Commitment, Journal of Basic and Applied Scientific Research J. Basic.

Appl. Sci. Res., 2(8)8120-8129, 2012 TextRoad Publication.



**A. Satish Kumar**

Assoc. Professor-cum-HOD, Dvr College Of Engineering And Technology,  
Kashipur Village, Kandi, Sangareddy Mandal, Medak.

# **Publish Research Article International Level Multidisciplinary Research Journal For All Subjects**

Dear Sir/Mam,

We invite unpublished Research Paper, Summary of Research Project, Theses, Books and Books Review for publication, you will be pleased to know that our journals are

## **Associated and Indexed, India**

- ★ Directory Of Research Journal Indexing
- ★ International Scientific Journal Consortium Scientific
- ★ OPEN J-GATE

## **Associated and Indexed, USA**

- DOAJ
- EBSCO
- Crossref DOI
- Index Copernicus
- Publication Index
- Academic Journal Database
- Contemporary Research Index
- Academic Paper Database
- Digital Journals Database
- Current Index to Scholarly Journals
- Elite Scientific Journal Archive
- Directory Of Academic Resources
- Scholar Journal Index
- Recent Science Index
- Scientific Resources Database

Review Of Research Journal  
258/34 Raviwar Peth Solapur-413005, Maharashtra  
Contact-9595359435  
E-Mail-ayisrj@yahoo.in/ayisrj2011@gmail.com  
Website : [www.ror.isrj.org](http://www.ror.isrj.org)

**CONFERENCES**



## MALLA REDDY ENGINEERING COLLEGE(AUTONOMOUS)

### Faculty Published List of Conference Publications during the Academic year 2015-2016

S.No.	Names of the Authors	Department	Title	Name of Conference	Organizer	Venue	Page No's	Year
1	Vamsi Krishna Basava Dr. M Kameswara Rao	Civil	A Comparative And Experimental Study On The Mechanical Properties Of Various Steel And Glass Fiber Reinforced High Strength Concrete	Three-Day International Conference On Innovations In Structural Engineering (IC-ISE 2015)	Osmania University, Hyderabad	Osmania University	1370-1378	2015
2	Dr. M Kameswara Rao Vamsi Krishna Basava	Civil	An Experimental Investigation On Strength Properties Of Plain Concrete Using Waste Foundry Sand (PAPER ID : ICES-EN007)	2 <sup>nd</sup> International Conference On Sustainable Energy And Built Environment	VIT Univeristy, Vellore	VIT Univeristy	Accepted	
3	R.N.Sujatha	Civil	Assessment of Musi River Water and Near by Ground Water, Impacts on Health of Down Stream Villages of Hyderabad (PAPER ID : ICES-EN006)	2 <sup>nd</sup> International Conference On Sustainable Energy And Built Environment	VIT Univeristy, Vellore	VIT Univeristy	Accepted	
4	Mr.P.Sankarbabu Mrs.V.Sumadeepthi	EEE	Comparison Of Pi And Rst Controllers In Double -Fed Induction Generator Wind Energy Conversion Systems	ICETSTEM	MREC	Hyderabad	97-104	2015
5	Mrs.Y.Sudha, Mrs.S.Bharathi	EEE	Synchronous Static Series Compensator(Sssc) On Enhancement Of Voltage Stability And Power Oscillation Damping For Unbalanced System	ICETSTEM	MREC	Hyderabad	111-116	2015
6	Mr.T.Sanjeeva Rao	EEE	Fuzzy Logic Controller Based Blde Motor Drive With Source Side Power Factor Correction	ICETSTEM	MREC	Hyderabad	136-141	2015
7	Mr.T.Sanjeeva Rao	EEE	Simulation Of Standalone Inverter For Distribution Generation System With Fuzzy Control	ICETSTEM	MREC	Hyderabad	142-148	2015

S.No.	Names of the Authors	Department	Title	Name of Conference	Organizer	Venue	Page No's	Year
8	Mrs.M.Laxmiswarupa Mr.P.Ganesh, Ms.K.Chetaswi	EEE	Optimal Location Of Facts Devices Considering Cost Of Installation And System Loadability By Using Partical Swarm Optimization Technique	ICETSTEM	MREC	Hyderabad	124-127	2015
9	Mrs.M.Laxmiswarupa	EEE	Fuzzy Logic Based Maximum Torque Control For 1-Phase Induction Motor	ICETSTEM	MREC	Hyderabad	128-132	2015
10	Mr.P.Ganesh	EEE	Dual - Mode Mechanism With Input Voltage Feed-Forward For The Two-Switch Buck-Boost Dc-Dc Converter	ICETSTEM	MREC	Hyderabad	91-96	2015
11	Ms.K.Chetaswi	EEE	Lyapunov Design Based Mit Rule For Model Refernce Adaptive Control Systems	ICETSTEM	MREC	Hyderabad	117-120	2015
12	Ms.K.Chetaswi	EEE	Optimal Design Technique Based On Linear Quadratic Gaussian Theory	ICETSTEM	MREC	Hyderabad	121-123	2015
13	Mr.K.Ramesh	EEE	Analysis,Design And Implementation Of A Bidirectional Double - Boost Dc-Dc Converter With Coupled Inductor Technique	ICETSTEM	MREC	Hyderabad	88-90	2015
14	Mrs.K.Anitha Reddy	EEE	Adaptive Fuzzy Logic Controller For Permanent Magnet Synchronous Motor With Non-Linear Friction	ICETSTEM	MREC	Hyderabad	105-110	2015
15	D Praveen kumar	ECE	Suboptimal Comparison of AF and DF Relaying For Fixed Target Error Probability	IEEE International Conference on Computer Communication and Informatics	IEEE	Coimbatore	466-469	2015
16	J. Sireesha	CSE	Document Clustering with Concept based Vector Suffix Tree Document Model	NCRTCST-2015	CMRCET	Kompally	35-38	2015
17	Dr. Md. Mastan	CSE	An approach for secured Authorized deduplication on Hybrid Cloud	ICETSTEM 2015	MREC	Dhulapally	358-363	2015

S.No.	Names of the Authors	Department	Title	Name of Conference	Organizer	Venue	Page No's	Year
18	Dr. Md. Mastan	CSE	An Efficient FDM Based Association Rule Mining in Horizontally Distributed Databases	ICETSTEM 2015	MREC	Dhulapally	364-367	2015
19	S .Ajay Kumar	CSE	Enabling Public Verification And Privacy Preserving Audit For Secure Cloud Storage	ICETSTEM 2015	MREC	Dhulapally	373-379	2015
20	V.Sathish Kumar	CSE	Enabling Privacy Data Storage By Using Fragmentation And Encryption For the Cloud	ICETSTEM 2015	MREC	Dhulapally	374-378	2015
21	Dr.U.Ravi BaBu	CSE	Implementation of Histogram Shifting for Reversible watermarking Technique	ICETSTEM 2015	MREC	Dhulapally	383-386	2015
22	R.Chandra Shekar	CSE	Effective Enhanced Adaptive Acknowledge	ICETSTEM 2015	MREC	Dhulapally	390-394	2015
23	M.Ahalya Rani	CSE	A Survey On Effect Of Node Mobility And Time Integration In Underwater Wireless Sensor Networks	ICETSTEM 2015	MREC	Dhulapally	398-402	2015
24	R.Chandra Shekar	CSE	Enabling Indirect Mutual Trust And Dynamic Data For Storage Services In Cloud Computing	ICETSTEM 2015	MREC	Dhulapally	415-420	2015
25	Dr.U.Ravi Babu	CSE	Textured Based Combined Approach For Plant Leaf Classification	ICETSTEM 2015	MREC	Dhulapally	387-389	2015
26	Mr.K.Vijaykrupa Vatsal	CSE	Geo-Social Application Oriented Privacy Preserving Location	ICETSTEM 2015	MREC	Dhulapally	459-464	2015
27	G.Sandhya Rani	CSE	Mining Order-Preserving Submatrices From Data With Multi-Repeated Measurements	ICETSTEM 2015	MREC	Dhulapally	465-468	2015

S.No.	Names of the Authors	Department	Title	Name of Conference	Organizer	Venue	Page No's	Year
28	R.S Muralinath	CSE	Shared Data In The Cloud With Enabled Privacy-Preserving Public Auditing	ICETSTEM 2015	MREC	Dhulapally	469-472	2015
29	P.V.Ramanamurthy	CSE	Decentralized Authentication of data storage using key policy Attribute Based Encryption	ICETSTEM 2015	MREC	Dhulapally	478-484	2015
30	K.V. Raghavender	CSE	Review on Existing Load Balancing Techniques of Cloud Computing	Big Data Analytics and Cloud Computing	Osmania University	Hyderabad	225-229	2015
31	M Swami Das	CSE	QoS web service Security Access Control case study using HTTP Secured Socket Layer Approach	ICEMIS2015	IARES	İstanbul Kemerburgaz University in Istanbul	252-259	2015
32	S .Ajay Kumar	CSE	Implementation Of Attribute Based Encryption For Secure Data Sharing In Cloud	ICETSTEM 2015	MREC	Dhulapally	473-477	2015
33	K. Rajeshwar Rao	CSE	Big data technology Predictions and trends	NCRTCST-2015	CMRTC	Kompally	1 - 4	2015
34	K. Rajeshwar Rao	CSE	Implementenation of a Robust Authenticate for Cloud Services	ICRAESIT - 2015	BVRIT	Narsapur	337-338	2015
35	Mr. Ch. Ramesh Babu	CSE	A Refuge Business Model for Cloud Computing Using Novel AES Approach	ICATACSE-2015	MRECW	Dhulapally	1-8	2015
36	Dr.U.Ravi Babu	CSE	Textured Based Combined Approach for Plant Leaf Classification	ICETESM-2015	MREC	Dhulapally	387-389	2015
37	M Sabhapathy	CSE	Prediction of the post-operative life expectancy using Data Mining techniques	ICCEEDSP	SUNFLOWER	Vijaywada	1-6	2015

S.No.	Names of the Authors	Department	Title	Name of Conference	Organizer	Venue	Page No's	Year
38	Dr. K Sharath Babu	H&S	Numerical Integration (Quadrature) method for steady - state convection diffusion problems	International Conference on Applied Sciences	Delhi University	Delhi University	726-733	2015
39	Dr.K.Veeraiah	MBA	Corporate restructuring dynamics	National confrence on Marketing	Rk publisher	vijayawada	77 -90	2015
40	Dr.K.Veeraiah	MBA	A Study on Technological Impact on Rural Market in India	ICETSTEM	MREC	Hyderabad	747 -783	2015
41	V.V.L.Sameera	MBA	A Study on comparative analysis of select ETF's and Index based mutual funds	GE-International Journal	AARF	Haryana	60-66	2016
42	P.Rajitha ,A.Latha, G.V.Venela	MBA	A Study on Women Entrepreneurs in Ranga Reddy District, A.P., India.	ICETSTEM	MREC	Hyderabad	714 -718	2015
43	A.Latha	MBA	Women Empowerment through Microfinance: A Study on SHGs In RangaReddy District, Telangana.	ICESBA	MRECW	Hyderabad	525-536	2015
44	A.Latha , P.Buela Prasanna Kumari , V.Venu Kumar Reddy	MBA	A Study on Risk-Return Analysis of Banking sector with respect to public and private sector	ICETSTEM	MREC	Hyderabad	729-734	2015
45	S.Eswara Reddy	MBA	Customer response Credit card usage	ICETSTEM	MREC	Hyderabad	732-737	2015
46	V.V.L.Sameera	MBA	A Study on Impact of Select Exchange Rate Fluctuations On Indian Stock Market	ICETSTEM	MREC	Hyderabad	709-713	2015



**Department of  
Civil Engineering**



*Innovations in*  
**STRUCTURAL  
ENGINEERING**

*Editors*

**Dr V Bhikshma**

**Dr D Rupesh Kumar**



Department of Civil Engineering  
UNIVERSITY COLLEGE OF ENGINEERING (A)  
Osmania University  
Hyderabad 500 007, T.S., India.



**Universities Press**

UNIVERSITIES PRESS (INDIA) PRIVATE LIMITED

*Registered Office*

3-6-747/1/A & 3-6-754/1, Himayatnagar, Hyderabad 500 029, Telangana, India  
info@universitiespress.com; www.universitiespress.com

*Distributed by*

Orient Blackswan Private Limited

*Registered Office*

3-6-752 Himayatnagar, Hyderabad 500 029, Telangana, India

*Other Offices*

Bengaluru / Bhopal / Chennai / Ernakulam / Guwahati / Hyderabad / Jaipur  
Kolkata / Lucknow / Mumbai / New Delhi / Noida / Patna / Vijayawada

© IC-ISE 2015, Department of Civil Engineering, Osmania University, Hyderabad

First published by Universities Press (India) Limited 2016

ISBN 978 81 7371 986 8

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage or retrieval systems, without permission in writing from the publisher

*Cover and book design*

© Universities Press (India) Private Limited 2016

*Typeset in Times New Roman 10/12 by*

Techastra Solutions Pvt Ltd, Hyderabad

*Printed at*

Graphica Printers and Binders  
Hyderabad 500 013

*Published by*

Universities Press (India) Private Limited  
3-6-747/1/A & 3-6-754/1, Himayatnagar,  
Hyderabad 500 029, Telangana, India.

*Disclaimer*

The authors are solely responsible for the contents of the papers presented in this volume. The publisher or the editors do not take any responsibility for the same in any manner. Errors, if any, are purely unintentional and readers are requested to address communication in this regard to the editors.

<b>115. Evaluation of Durability Characteristics of Self Compacting Concrete with Controlled Permeable Formwork Liner</b> <i>S. Kandasamy, S. Kothandaraman and K. Sivaraman</i>	1238
<b>116. Study of Use of Discontinuous Glass Fibres in PSC Spun Pole</b> <i>M.B. Varma and S.T. Patil</i>	1246
<b>117. Fresh and Hardened Properties of Plain Rubberized Concrete Using Fine Rubber</b> <i>S. Padhi and K.C. Panda</i>	1256
<b>118. Effect of Basalt Fibre on Mechanical Properties of Silica Fume Concrete</b> <i>S.A. Bhalchandra, A.V. Pathak and S.N. Pathak</i>	1268
<b>119. Comparison of 28 Days Concrete Compressive Strengths by Accelerated Curing and Normal Moist Curing</b> <i>Abhijit, J. Pawar, Jagruti, S. Nikam, Pravinchandra and D. Dhake</i>	1276
<b>120. Impact Resistance of Steel Fibre Reinforced Concrete</b> <i>P.B. Murnal and R.N. Chatorikar</i>	1283
<b>121. Strength and Durability Properties of Concrete Containing Manufactured Sand as Fine Aggregate</b> <i>B. Vijaya and S. Senthil Selvan</i>	1293
<b>122. Strength and Elastic Properties of Fibrous Self Compacting Concrete</b> <i>S. Vijaya Kumar and B.L.P. Swami</i>	1304
<b>123. Studies on Fibrous Recycled Aggregate Triple Blended High Strength Concrete</b> <i>M.V.S.S. Sastri, K. Jagannadha Rao and V. Bhikshma</i>	1314
<b>124. Influence of Ceramic Waste Aggregate Properties on Strength of Ceramic Waste Aggregate Concrete</b> <i>V. Giridhar, H. Sudarsana Rao and P. Suresh Praveen Kumar</i>	1325
<b>125. Optimization of Fly Ash Content in Different Grades of Magnetic Water Concrete</b> <i>A. Venkat Reddy and Divya Kolasani</i>	1337
<b>126. Studies of Sulphate Attack on Sugar Cane Bagasse Ash Concrete</b> <i>R. Ramya Swetha, V. Srinivasa Reddy and P. Sai Sujan</i>	1346
<b>127. Study on Elastic Property of Mixed Fibre Reinforced Concrete</b> <i>Urooj Masood, B.L.P. Swami and A.K. Asthana</i>	1355
<b>128. A Comparative and Experimental Study on the Mechanical Properties of Various Steel and Glass Fibre Reinforced High Strength Concrete</b> <i>B. Vamsi Krishna and M. Kameswara Rao</i>	1370
<b>129. Experimental Study on the Torsional Behaviour of Recycled Aggregate Concrete Beams</b> <i>D.V. Lokesh, I. Jaagruthi and Ch. Naga Satish Kumar</i>	1380

# A COMPARATIVE AND EXPERIMENTAL STUDY ON THE MECHANICAL PROPERTIES OF VARIOUS STEEL AND GLASS FIBRE REINFORCED HIGH STRENGTH CONCRETE

**B. Vamsi Krishna<sup>1</sup> and M. Kameswara Rao<sup>2</sup>**

<sup>1</sup>Assistant Professor, <sup>2</sup>Professor & HOD, Department of Civil Engineering  
Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

## Abstract

Cement concrete is the most extensively used construction material in the world. It has been found that different type of fibers added in specific percentage to concrete improves the mechanical properties, durability and serviceability of the structure. It is now established that one of the important properties of hooked steel, crimped steel & glass Fiber Reinforced Concrete is its superior resistance to cracking and crack propagation. Concrete is most widely used construction material in the world. Fiber reinforced concrete (FRC) is a concrete in which small and discontinuous fibers are dispersed uniformly. The fibers used in FRC may be of different materials like steel, G.I., carbon, glass, aramid, asbestos, polypropylene, jute etc. The addition of these fibers into concrete mass can dramatically increase the compressive strength, tensile strength, flexural strength and impact strength of concrete. FRC has found many applications in civil engineering field. Based on the laboratory experiment on fiber reinforced concrete (FRC), cube and cylinders specimens have been designed with steel fiber reinforced concrete (SFRC) and Glass fiber reinforced concrete (GFRC) containing fibers of 0% and 0.5% volume fraction of hook end Steel fibers of 53.85, 50 aspect ratio and alkali resistant glass fibers containing 0% and 0.25% by weight of cement of 12 mm cut length were used without admixture

In this paper effect of fibers on the different mechanical properties of grade M 80 have been studied. It optimizes 1.5% for steel Fiber content and 1% for glass fiber content by the volume of cement is used in concrete. The percentage increase in compressive strength at 28 days for hooked end steel fiber when compared to conventional concrete is 7.3%, crimped steel fiber with 6.08%, glass fiber with 4.3. The percentage increase in split tensile strength at 28 days hooked end steel fiber when compared to conventional concrete is 4.54%, crimped steel fiber with 3.40%, glass fiber with 2.27% and also The percentage increase of flexural strength at 28 days for hooked end steel fiber when compared to conventional concrete is 3.57%, crimped steel fiber with 2.380%, glass fiber with 2.140%.

**Keywords:** Steel fiber reinforced concrete (SFRC) and Glass fiber reinforced concrete (GFRC), High strength concrete, M-80 Grade, IS:1386, IS:383

## INTRODUCTION

Concrete is a composite material containing hydraulic cement, water, coarse aggregate and fine aggregate. The resulting material is a stone like structure which is formed by the chemical reaction of the cement and water. This stone like material is a brittle material which is strong in compression but very weak in tension. This weakness in the concrete makes it to crack under small loads, at the tensile end. These cracks gradually propagate to the compression end of the member and finally, the member breaks. The formation of cracks in the concrete may also occur

# **A COMPARITIVE AND EXPERIMENTAL STUDY ON THE MECHANICAL PROPERTIES OF VARIOUS STEEL AND GLASS FIBER REINFORCED HIGH STRENGTH CONCRETE**

**B. VAMSI KRISHNA**

M.S, Structural Engineering, Assistant Professor, Department of Civil Engineering  
Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

**Dr. M. KAMESWARA RAO**

Structural Engineering, Professor & HOD, Department of Civil Engineering,  
Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

## **ABSTRACT**

Cement concrete is the most extensively used construction material in the world. It has been found that different type of fibers added in specific percentage to concrete improves the mechanical properties, durability and serviceability of the structure. It is now established that one of the important properties of hooked steel, crimped steel & glass Fiber Reinforced Concrete is its superior resistance to cracking and crack propagation. Concrete is most widely used construction material in the world. Fiber reinforced concrete (FRC) is a concrete in which small and discontinuous fibers are dispersed uniformly. The fibers used in FRC may be of different materials like steel, G.I., carbon, glass, aramid, asbestos, polypropylene, jute etc. The addition of these fibers into concrete mass can dramatically increase the compressive strength, tensile strength, flexural strength and impact strength of concrete. FRC has found many applications in civil engineering field. Based on the laboratory experiment on fiber reinforced concrete (FRC), cube and cylinders specimens have been designed with steel fiber reinforced concrete (SFRC) and Glass fiber reinforced concrete (GFRC) containing fibers of 0% and 0.5% volume fraction of hook end Steel fibers of 53.85, 50 aspect ratio and alkali resistant glass fibers containing 0% and 0.25% by weight of cement of 12mm cut length were used without admixture

In this paper effect of fibers on the different mechanical properties of grade M 80 have been studied. It optimizes 1.5% for steel Fiber content and 1% for glass fiber content by the volume of cement is used in concrete. The percentage increase in compressive strength at 28 days for hooked end steel fiber when compared to conventional concrete is 7.3% , crimped steel fiber with 6.08%, glass fiber with 4.3. The percentage increase in split tensile strength at 28 days hooked end steel fiber when compared to conventional concrete is 4.54%, crimped steel fiber with 3.40%, glass fiber with 2.27% and also The percentage increase of flexural strength at 28 days for hooked end steel fiber when compared to conventional concrete is 3.57%, crimped steel fiber with 2.380%, glass fiber with 2.140%.

**Keywords**— Steel fiber reinforced concrete (SFRC) and Glass fiber reinforced concrete (GFRC), High strength concrete, M-80 Grade, IS:1386, IS:383

## 1. INTRODUCTION

Concrete is a composite material containing hydraulic cement, water, coarse aggregate and fine aggregate. The resulting material is a stone like structure which is formed by the chemical reaction of the cement and water. This stone like material is a brittle material which is strong in compression but very weak in tension. This weakness in the concrete makes it to crack under small loads, at the tensile end. These cracks gradually propagate to the compression end of the member and finally, the member breaks. The formation of cracks in the concrete may also occur due to the drying shrinkage. These cracks are basically micro cracks. These cracks increase in size and magnitude as the time elapses and the finally makes the concrete to fail.

The formation of cracks is the main reason for the failure of the concrete. To increase the tensile strength of concrete many attempts have been made. One of the successful and most commonly used methods is providing steel reinforcement. Steel bars, however, reinforce concrete against local tension only. Cracks in reinforced concrete members extend freely until encounter ring are bar. Thus need for multidirectional and closely spaced steel reinforcement arises. That cannot be practically possible. Fibre reinforcement gives the solution for this problem. So to increase the tensile

strength of concrete a technique of introduction of fibres in concrete is being used. These fibres act as crack arrestors and prevent the propagation of the cracks. These fibres are uniformly distributed and randomly arranged. This concrete is named as fibre reinforced concrete. The main reasons for adding fibres to concrete matrix is to improve the post cracking response of the concrete, i.e., to improve its energy absorption Fibre Reinforced Concrete can be defined as a composite material consisting of mixtures of cement, mortar or concrete and discontinuous, discrete, uniformly dispersed suitable fibres. Continuous meshes, woven fabrics and long wires or rods are not considered to be discrete fibres FRC increases the tensile strength of the concrete, it reduce the air voids and water voids the inherent porosity of gel. It increases the durability of the concrete. Fibres such as graphite and glass have excellent resistance to creep. The addition of small closely spaced and uniformly dispersed fibres to concrete would act as crack arrester and would substantially improve its static and dynamic properties. Fibre reinforced concrete is in use since many years in India, but the structural applications are very much limited. However, its application is picking up in the recent days.

## 2. FIBRE REINFORCED CONCRETE

Fiber reinforced concrete (FRC) is concrete containing fibrous material which increases its structural integrity. So we can define fibre

reinforced concrete as a composite material of cement concrete or mortar and discontinuous discrete and uniformly dispersed fibre. Fibre is discrete material having some characteristic properties. The fibre material can be anything. But not all will be effective and economical. Some fibres that are most commonly used are:

- Steel              • Glass              • Carbon
- Natural           • NBD

Steel fibre is one of the most commonly used fibre. Generally round fibres are used. The diameter may vary from 0.25 to 0.75mm. The steel fibre sometimes gets rusted and lose its strength. But investigations have proved that fibres get rusted only at surfaces. It has high modulus of elasticity. Use of steel fibres makes significant improvements in flexure, impact and fatigue strength of concrete. It has been used in various types of structures. Glass fibre is a recently introduced fibre in making fibre concrete. It has very high tensile strength of 1020 to 4080Mpa. Glass fibre concretes are mainly used

### **3. PROPERTIES OF FIBRE REINFORCED CONCRETE**

Properties of concrete are affected by many factors like properties of cement, fine aggregate, coarse aggregate. Other than this, the fibre reinforced concrete is affected by following factors:

- Type of fibre
- Aspect ratio
- Quantity of fibre
- Orientation of fibre

#### **3.1 Type of Fibre**

A good fibre is the one which possess the following qualities:

- Good adhesion within the matrix.
- Adaptable elasticity modulus (sometimes higher than that of the matrix)
- Compatibility with the binder, which should not be attacked or destroyed in the long term
- An accessible price, taking into account the proportion within the mix
- being sufficiently short, fine and flexible to permit mixing, transporting and placing
- Being sufficiently strong, yet adequately robust to withstand the mixing process.

#### **3.2 Aspect ratio**

Aspect ratio is defined as the ratio of length to width of the fibre. The value of aspect ratio varies from 30 to 150. Generally the increase in aspect ratio increases the strength and toughness till the aspect ratio of 100. Above that the strength of concrete decreases, in view of decreased workability and reduced compaction. From investigations it can be found out that good results are obtained at an aspect ratio around 80 for steel fibres. Keeping that in view we have considered steel hooked end fibres with aspect ratio of 80 (Length 60 mm and Diameter 0.75 mm).

#### **3.3 Fibre quantity**

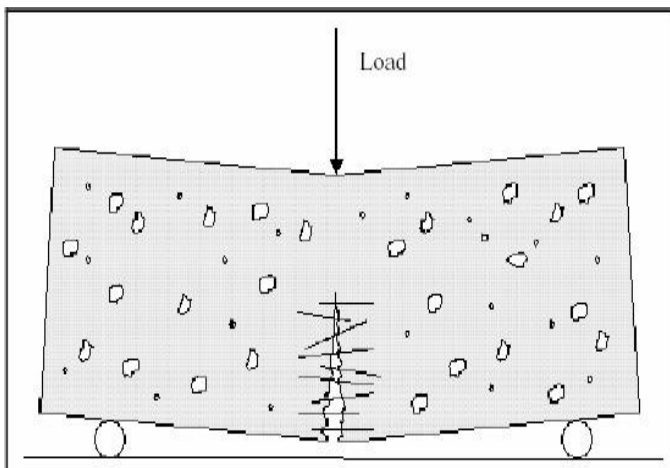
Generally quantity of fibres is measured as percentage of cement content. As the volume of fibres increase, there should be increase in strength and toughness of concrete. Regarding our fibre, we hope that there will be an increase in strength, with increase in fibre content. We are going to test for percentages of 1.0, 2.0 and 3.0.

### 3.4 Orientation of fibre

The orientations of fibres play a key role in determining the capacity of concrete. In RCC the reinforcements are placed in desired direction. But in FRC, the fibres will be oriented in random direction. The FRC will have maximum resistance when fibres are oriented parallel to the load applied.

### 3.5 Bridging Action

Pullout resistance of fibres (dowel action) is important for efficiency. Pullout strength of fibres significantly improves the post-cracking tensile strength of concrete. As an FRC beam or other structural element is loaded, fibres bridge the cracks. Such bridging action provides the FRC specimen with greater ultimate tensile strength and, more importantly, larger toughness and better energy absorption. An important benefit of this fibre behavior is material damage tolerance. Bayasi and Kaiser (2001) performed a study where damage tolerance factor is defined as the ratio of flexural resistance at 2-mm maximum crack width to ultimate flexural capacity. At 2% steel fibre volume, damage tolerance factor according to Bayasi and Kaiser was determined as 93%.



***Fig:1 Pullout Mechanism***

### 3.6 Workability

A shortcoming of using fibres in concrete is reduction in workability. Workability of FRC is affected by fibre aspect ratio and volume fraction as well the workability of plain concrete. As fibre content increases, workability decreases. Most researchers' limit volume of fibres to 4.0% and aspect ratio to 100 to avoid unworkable mixes. In addition, some researchers have limited the fibre reinforcement index [volume of fibres as %  $\times$  aspect ratio] to 1.5 for the same reason. To overcome the workability problems associated with FRC, modification of concrete mix design is recommended. Such modifications can include the use of additives.

## 4. EXPERIMENTAL INVESTIGATION

The materials used in the experimental investigation are locally available cement, sand, coarse aggregate, mineral and chemical admixtures. The chemicals used in the present investigation are of commercial grade.

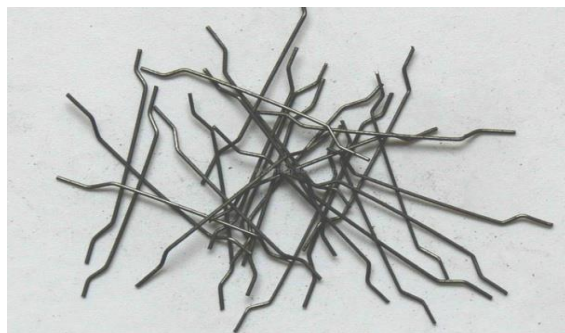
### 4.1 STEEL FIBRES

Steel fibre-reinforced concrete (SFRC) is concrete (spray concrete) with steel fibres added. It has higher tensile strength than unreinforced concrete and is quicker to apply than weld mesh reinforcement. It has often been used for tunnels

- Addition of steel fibers into the concrete improves the crack resistance (or ductility) capacity of the concrete. Traditional rebars are generally used to improve the tensile strength of the concrete in a particular direction, whereas steel fibers are useful for multidirectional reinforcement. This is one

of the reasons why steel fiber reinforced (shotcrete form) concrete successfully replaced weldmesh in lining tunnels.

- Less labour is required.
- Less construction time is required



**Fig:2 crimped and hooked steel fibres**

## 4.2 GLASS FIBRES

Glass fiber reinforced concrete, also known as GFRC or GRC, is a type of fiber reinforced concrete. Glass fiber concretes are mainly used in exterior building façade panels and as architectural precast concrete. Somewhat similar materials are fiber cement siding and cement boards. The photograph of glass fibres is shown in Fig. 3



**Fig:3 Glass fibre**

## 4.3 INITIAL TESTS

Material test	Result
Specific gravity of cement	3.12
Specific gravity of fly ash	2.24
Specific gravity of silica fume	2.21
Specific gravity of coarse aggregate	2.74
Specific gravity of fine aggregate	2.7
Slump cone test	2 inches
Dry rod unit weight of fine aggregate	107.7 lb/ft <sup>3</sup>
Dry rod unit weight of coarse aggregate	101 lb/ft <sup>3</sup>
Initial and final setting time	96 min & 207 min

**Table: 1 Initial tests**

## 4.4 MIX PROPORTION

Cement	Fly Ash	Silica Fume	Fine Aggregate	Coarse Aggregate	Water	Super Plasticizer
1	0.28	0.28	1.38	2.38	0.23	0.01

**Table: 2 Mix Proportions**

## 5. TESTING OF SPECIMENS

Different tests were conducted on the specimens to determine and compare the mechanical properties between crimped steel fibres, hooked steel fibres and glass fibres.

## 5.1 COMPRESSIVE STRENGTH



**Fig: 4 Cube specimens under test**

Sl no	cubes casted day	Conventi onal concrete (N/mm <sup>2</sup> )	Hooked end steel fiber (N/mm <sup>2</sup> )	Crimped steel fiber(N/ mm <sup>2</sup> )	Glass fiber(N/ mm <sup>2</sup> )
1	3 <sup>rd</sup> day	27	31.75	29.3	29
2	7 <sup>th</sup> day	44	48	46.4	49
3	28 <sup>th</sup> day	69	74	73.2	72

**Table: 3 Compressive strength results**

## 5.2 FLEXURAL TEST



**Fig: 5 Flexural beam specimens under test**

Sl no	Beams casted day	Conventi onal concrete	Hooked end steel fiber	Crimped steel fiber	Glass fiber
1	3 <sup>rd</sup> day	3.7	4.6	4.2	4.1
2	7 <sup>th</sup> day	6.7	7.2	6.8	7
3	28 <sup>th</sup> day	8.4	8.7	8.6	8.5

**Table: 4 Flexure test results**

## 5.3 SPLIT TENSILE TEST



**Fig: 6 Cylinder specimens under test**

Sl no	Cylinder s casted day	Conventi onal concrete (N/mm <sup>2</sup> )	Hooked end steel fiber (N/mm <sup>2</sup> )	Crimped steel fiber (N/mm <sup>2</sup> )	Glass fiber (N/mm <sup>2</sup> )
1	3 <sup>rd</sup> day	1.9	2.6	2.4	2.3
2	7 <sup>th</sup> day	5.2	5.7	5.3	5.5
3	28 <sup>th</sup> day	8.8	9.2	9.1	9

**Table: 5 Split Tensile Strength results**

## CONCLUSIONS

The present study is about using different fibres i.e., crimped steel fibres hooked steel fibres and glass fibres after optimising them; comparison is made between the three fibres for different mechanical properties

- The increasing percentage of compressive strength of hooked end steel fiber reinforced concrete cubes when compared to the conventional concrete cubes at 28 days is 7.3%. And the increasing percentage of compressive strength of crimped steel fiber reinforced concrete cubes when compared to the conventional concrete cubes at 28 days is 6.08%. And The increasing percentage of compressive strength of hooked end steel fiber reinforced concrete cubes when compared to the conventional concrete cubes at 28 days is 4.34%.
- The increasing percentage of split tensile strength of hooked end steel fiber reinforced concrete cylinders when compared to the conventional concrete cylinders at 28 days is 4.54%. And the increasing percentage of split tensile strength of crimped steel fiber reinforced concrete cylinders when compared to the conventional concrete cylinders at 28 days is 3.40%. And The increasing percentage of compressive strength of hooked end steel fiber reinforced concrete cylinders when compared to the conventional concrete cylinders at 28 days is 2.27%.
- The increasing percentage of flexural strength of hooked end steel fiber reinforced concrete beams when compared to the conventional concrete beams at 28 days is 3.57%. And the increasing

percentage of flexural strength of crimped steel fiber reinforced concrete beams when compared to the conventional beams at 28 days is 2.380%. And The increasing percentage of flexural strength of glass fiber reinforced concrete beams when compared to the conventional concrete beams at 28 days is 2.140%.

## REFERENCES

1. ACI 234R-06 "Guide for the use of Silica Fume in Concrete". American Concrete Institute.
2. ACI 544.1R-96, (Reapproved 2009) "Report on fiber reinforced concrete". American Concrete Institute.
3. ACI 544.4R-88, "Design Considerations for Steel Fiber Reinforced Concrete". American Concrete Institute.
4. ASTM A 820M-06, "Specification for Steel Fibers for Fiber Reinforced Concrete". ASTM International.
5. Balendran.R.V, Rana T.M., Maqsood T, Tang W.C., "Strength and durability performance of HPC incorporating pozzolans at elevated temperatures", Structural Survey, Vol. 20 ,2002 pp.123 – 128.
6. Brooks, J.J. et al "Effect of admixtures on the setting times of high-strength concrete" Cement Concrete Compos, vol 22, 2000, pp293-301.
7. Caldarone M.A and Gruber K.A , "High Reactivity Metakaolin (HRM) for High Performance Concrete", special publications, vol.153, june1995, pp:815-828.
8. Caldarone, M.A. et al "High reactivity metakaolin: a new generation mineral

- admixture". Concrete Int, vol.34, November 1994, pp: 37-40.
9. Curcio, F. et al "Metakaolin as a pozzolanic microfiller for high-performance mortars".
10. Ghosh.S, . Bhattacharjya S, Chakraborty S "Compressive behaviour of Short Fibre Reinforced Concrete", Magazine of Concrete Research, vol.59(8), 2007, pp 567 –574.
11. Ghosh.S, . Bhattacharjya S, Chakraborty S "Mechanics of Steel Fibre Reinforced Composite in Flexural Shear", Int. Conf. CENeM –2007, Bengal Engg. and science University, Shibpur, India, Jan 11-14, 2007.
12. I.S. 383-1970, "Specification for coarse and fine aggregate from natural sources for concrete". BIS
13. I.S. 456-2000, "Code of practice of plain and reinforced concrete". BIS.
14. I.S. 516-1959, "Method of test for strength of concrete", BIS.
15. I.S. 1344-1968 "India standard specification for pozzolanas" bureau of Indian Standards.
16. I.S. 2386 (Part 1) 1963 "Methods of test for Aggregates for Concrete, Part 1 Particle Size and Shape", BIS.
17. I.S. 6461 (Part 7) 1973 "Mixing, laying, compaction, curing and other construction aspects", BIS.
18. I.S. 7246 1974 "Recommendations for use of table vibrators for consolidating concrete", BIS.
19. I.S. 9103-1999, "Specification for admixtures for concrete". BIS
20. I.S. 10262-1982, "Recommend guidelines for concrete mix design". BIS.
21. I.S. 10262-2009, "Recommended guidelines for concrete mix design". BIS.
22. I.S. 12269-1987, "Specification for 53 grade ordinary Portland cement". BIS.
23. I.S. 7869(part 2)-1981: "Indian standard specification for admixtures for concrete", BIS.
24. Justice J.M, Kennison L.H, Mohr B.J, Beckwith B.L, McCormick L.E, Wiggins B, Zhang Z.Z, and Kurtis K.E, "Comparison of Two Metakaolins and a Silica Fume Used as Supplementary Cementitious Materials", Proc. Seventh International Symposium on Utilization of High-Strength/HighPerformance Concrete, held in Washington D.C., June 20-24, 2005
25. Khatib J.M, Wild S : "Size distribution of metakaolin paste" Cement Concrete Res, vpl.26(10), 1996, pp: 1545-53.
26. Khatib J M and Wild S, "Sulphate Resistance of Metakaolin Mortar", Cement and Concrete Research Journal, Vol. 28, 1998, pp 83-92.
27. Kostuch, J.A. et al "High performance concrete incorporating metakaolin - a review", Concrete 2000. University of Dundee, September 1993, pp: 1799-1811.
28. Krishnaraju.N, "design of concrete mix"—CBS publisher—1985.
29. Malvin Sandvik and Odd Gjorv E, "Effect of Condensed Silica Fume on the Strength Development of Concrete", special publications, vol.93, feb1986.

# AN EXPERIMENTAL INVESTIGATION ON STRENGTH PROPERTIES OF PLAIN CONCRETE USING WASTE FOUNDRY SAND

M. Kameswara Rao<sup>1</sup> and B. Vamsi Krishna<sup>2</sup>

<sup>1</sup>Professor, HOD, Civil Engineering, Malla Reddy Engineering College (Autonomous), Secunderabad  
<sup>2</sup>Assistant Professor, Civil Engineering, Malla Reddy Engineering College (Autonomous), Secunderabad

## ABSTRACT

**Objectives:** To evaluate the strength properties of concrete mixtures, in which river sand was partially replaced with Waste Foundry Sand by weight.

**Methods/Analysis:** Compression test was carried out at the age of 28 days of curing. Split tensile test was performed at the age of 28 days. Flexural strength was tested at 28 days of curing. Test results indicate an increase in compressive strength of plain concrete by inclusion of WFS as a partial replacement of fine aggregate

**Findings:** The maximum strength was achieved at 40% replacement, after which there was loss in compressive strength, split tensile strength and Flexural strength decreased. The results indicate in concrete. However, the partial replacement should not exceed 40% in plain concrete.

**Novelty/Improvement:** Foundry sand will replace the river & Robo sand in future because of huge scarcity. The mechanical properties will also improve for the better serviceability of the structures.

**Keywords:** Plain Cement Concrete, Foundry Sand, bentonite, Compression strength, Split Tensile strength, Flexural strength.

## 1. Introduction

An acute shortage of river sand which is generally used as a fine aggregate in concrete has been affecting the construction sector. The scarcity has led to the skyrocketing price of sand, escalating construction costs. The situation has dashed the dreams of many in the lower- and middle-income groups to own a house. There were studies about the depletion of river sand and the need for

scientific management and exploitation of the available resource. Following the shortage of river sand, some research institutions are searching alternatives that can be used for construction. Ferrous and non ferrous metal casting industries produce several million tons of byproduct in the world. In India, approximately 2 million tons of waste foundry sand is produced yearly. WFS is a major byproduct of metal casting industry and successfully used as a land filling material for many

years. In an effort to use the WFS in large volume, research is being carried out for its possible large scale utilization in making concrete as partial replacement of fine aggregate. Foundry sand consists primarily of silica sand, coated with a thin film of burnt carbon, residual binder (bentonite, sea coal, resins) and dust from [Reference \[1\]](#). Foundry sand can be used in concrete to improve its strength and other durability factors. Foundry Sand can be used as a partial replacement of fine aggregates or total replacement of fine aggregate and as supplementary addition to achieve different properties of concrete.

Solid waste management has become one of the global environmental issues, as there is continuous increase in industrial by-products and waste materials. Due to lack of land filling space and its ever increasing cost, utilization of waste material and by-products has become an attractive alternative to disposal. Waste foundry sand (WFS) is one of such industrial by-product.

Ferrous and non ferrous metal casting industries produce several million tons of by-product in the world. In India, approximately 2 million tons of waste foundry sand is produced yearly. WFS is major by-product of metal casting industry and successfully used as a land filling material for many years. But use of waste foundry sand for land filling is becoming a problem due to rapid increase in disposal cost. Metal foundries use large amounts of sand as part of the metal casting process. Foundries successfully recycle and reuse the sand many times in a foundry. When the sand can no longer be

reused in the foundry, it is removed from the foundry and is termed “Waste Foundry Sand”.

Foundry industry produces a large amount of by-product material during casting process. The ferrous metal casts in foundry are cast iron and steel, non ferrous metal are aluminum, copper, brass and bronze. Over 70% of the total by-product material consists of sand because moulds consist usually of molding sand, which is easily available, inexpensive, resistance to heat damage and easily bonded with binder and other organic material in mould. Foundry industry use high quality specific size silica sand for their molding and casting process. These WFS is black in color and contain large amount of fines. The typical physical and chemical property of WFS is dependent upon the type of metal being poured, casting process, technology employed, type of furnaces (induction, electric arc and cupola) and type of finishing process (grinding, blast cleaning and coating)

[Reference \[1-2\]](#).

## **2. Chemical Composition**

Chemical Composition of the foundry sand relates directly to the metal molded at the foundry. This determines the binder that was used, as well as the combustible additives. Typically, there is some variation in the foundry sand chemical composition from foundry to foundry from [Reference \[3\]](#). Sands produced by a single foundry, however, will not likely show significant variation over time. Moreover, blended sands produced by consortia of foundries often produce consistent sands. The chemical composition of the foundry sand can impact its performance. Waste foundry sand

consists primarily of silica sand, coated with a thin film of burnt carbon, residual binder (bentonite, sea coal, resins) and dust from [Reference \[3\]](#).

### 3. Objectives of Present Investigation

- To investigate the effect of waste foundry sand as a partial replacement of fine aggregate on strength properties of  $M_{20}$ ,  $M_{40}$  and  $M_{60}$  grades of concrete
- To reduce the problem of disposal of industrial waste.

### 4. Scope of the Present Work

The present experiment is carried out to investigate on strength properties of concrete mixes of grade  $M_{20}$ ,  $M_{40}$  and  $M_{60}$  in which fine aggregate (river sand) is to be partially replaced with Waste Foundry Sand. Fine aggregate will be replaced with six percentages (0%, 20%, 40%, 60%, 80% and 100%) of WFS by weight. Some of the strength properties such as Compressive strength, Split tensile strength and Flexural strength of Plain Concrete [Reference \[5-10\]](#).

### 5. Preliminary Investigation on Foundry Sand

Chemical Composition of Waste Foundry Sand, Sieve Analysis Chart for Waste Foundry Sand, Physical Properties of Waste Foundry Sand was investigated.

### 6. Concrete Mix Design

In the present investigation using the properties of cement, aggregate concrete mix of  $M_{20}$ ,  $M_{40}$  and  $M_{60}$  grade was designed as per IS 10262-1982 the

mix design procedure and calculations are presented in Appendix A the following proportions by weight were obtained.

## 7. Results and Discussions

### 7.1 Compressive Strength of Concrete

Cube specimens were tested for compression and the ultimate compressive strength was determined from failure load, measured using compression testing machine. The average values of compressive strength of 3 specimens for each category at the age of 28 days are tabulated in the Table 4.2. The relative compressive strength of various concrete mixes (0%, 20%, 40%, 60%, 80% and 100%) for different grades ( $M_{20}$ ,  $M_{40}$  and  $M_{60}$ ) of concrete.

### 7.2 Split Tensile Strength of Concrete

Cylinder specimens were tested for split tensile strength and strength was determined from failure load, measured using compression testing machine. The average values of split tensile strength of 3 specimens for each category at the age of 28 days are tabulated in the Table 4.3 and Figure 4.3 show the graphical representation of variation of split tensile strength of plain concrete of various concrete mixes (0%, 20%, 40%, 60%, 80% and 100%) for different grades ( $M_{20}$ ,  $M_{40}$  and  $M_{60}$ ) of concrete.

### 7.3 Flexural Strength of Concrete

Beam specimens were tested for flexural strength using universal testing machine. The tests were carried out confirming to IS 516-1959; the specimens were tested under two point loading. The average value of 3 specimens for each category at the age of 28 days is tabulated in the Table 4.4.

Figure 4.4 shows the graphical representation of variation of flexural strength of plain concrete of various concrete mixes (0%, 20%, 40%, 60%, 80% and 100%) for different grades (M20, M40 and M60) of concrete.

## Conclusions

- Increase in compressive strength of the concrete with increases in waste foundry sand up to 40% and the maximum compressive strength is achieved at 40% replacement of natural fine aggregate with waste foundry sand which comes to be 33.86 MPa for (M20 grade), 54.65 MPa for ( M40 grade) and 73.83 MPa for (M60 grade)respectively and then there was a considerable decrease in the strength.
- Replacement of fine aggregate with waste foundry sand showed increase in the split tensile strength of plain concrete of grade M20, M40 and M60 up to 40% and then there was a considerable decrease in the strength. Maximum strength was achieved at 40% i.e. 4.34 MPa , 6.37 MPa and 9.06 MPa respectively.
- Replacement of fine aggregate with waste foundry sand showed increase in the Flexural strength of plain concrete of grade M20, M40 and M60 up to 40% and then there was a considerable decrease in the strength. Maximum strength was achieved at 40% i.e. of 6.02 MPa, 9.05 MPa and 11.64 MPa respectively.
- When percentage of waste foundry sand was increased beyond 40% the mix started losing its workability.

- Use of foundry sand in concrete can save the ferrous and non-ferrous metal industries disposal, cost and produce a ‘greener’ concrete for construction.
- Environmental effects from wastes and disposal problems of waste can be reduced through this research.
- A better measure by an innovative Construction Material is formed through this research.

The used foundry sand can be innovative Construction Material but judicious decisions are to be taken by engineers

## Future scope of work

- Further research can be carried out to study the durability properties of concrete incorporating waste foundry sand as a partial replacement of fine aggregate.
- The investigation of concrete incorporating waste foundry sand can be carried out with addition of different types of fibers like steel fibers, recron fibers, synthetic fibers, dura fibers, natural fibers and glass fibers and with different aspect ratio.
- Further research can be carried out to study the properties of concrete with partial replacement of fine aggregate with waste foundry sand and partial replacement of cement with different mineral admixtures like GGBS, flyash, metakaolin, micro silica, rice husk ash etc, with addition of different percentages of fibers.

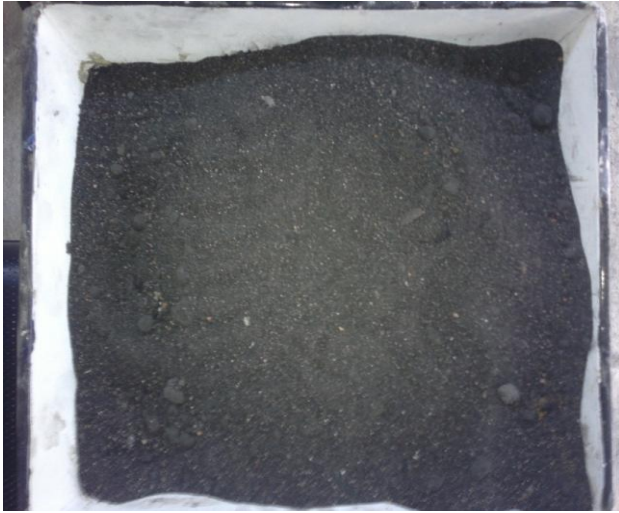
## REFERENCES

- 1) Rafat Siddique, Geert de Schutter and Albert Noumowe, (2008), “Effect of used-foundry sand on the mechanical properties of concrete”,

- Construction and Building Materials, vol. 23, pp 976–980.
- 2) D.Lawrence and M.Mavroulidou,(2009), “Properties of concrete containing waste foundry sand”, 11th International conference on environmental science and technology, Greece 3-5 September.
  - 3) J.M. Khatb, S. Baig, A Bougara, and C Booth, (2010), “Foundry sand utilization in concrete production”, Sustainable construction materials and technologies June 28-June 30.
  - 4) Yogesh Aggarwal, Paratibha Aggarwal, Rafat Siddique, El-Hadj Kadri and Rachid Bennacer, (2010), “Strength, durability, and micro-structural properties of concrete made with used-foundry sand (UFS)”, Construction and Building Materials, vol. 25, pp 1916–1925.
  - 5) Gurpreet Singh and Rafat Siddique, (2011), “Effect of waste foundry sand (WFS) as partial replacement of sand on the strength, ultrasonic pulse velocity and permeability of concrete”, Construction and Building Materials, vol. 26, pp 416–422.
  - 6) Khatib JM and Ellis DJ., (2001), “Mechanical properties of concrete containing foundry sand”. ACI special publication (SP-200), American Concrete Institute; pp. 733–748.
  - 7) Siddique R., Gupta R and Kaur I., (2007), “Effect of spent foundry sand as partial replacement of fine aggregate on the properties of concrete”. In: 22<sup>nd</sup> International conference on solid waste technology and management, Widener University, Philadelphia, USA.
  - 8) Gurpreet Singh and Rafat Siddique, (2011), “Abrasion resistance and strength properties of concrete containing waste foundry sand (WFS)”, Construction and Building Materials, vol. 28, pp 421–426.
  - 9) Rafat Siddique and El-Hadj Kadri, (2011), “Effect of metakaolin and foundry sand on the near surface characteristics of concrete”, Construction and Building Materials, vol. 25, pp 3257–3266.
  - 10) Neelam Pathak and Rafat Siddique, (2012), “Effects of elevated temperatures on properties of self-compacting-concrete containing fly ash and spent foundry sand”, Construction and Building Materials, vol. 34, pp 512–521.
  - 11) Naik, T.R., Patel V.M., Parikh D.M. and Tharaniyil M.P., (1994), “Utilization of used foundry sand in concrete”. Journal of Materials in Civil Engineering, Vol. 6, No. 2, pp. 254-263.

## ( Introduction)

**Fig: 1 Waste Foundry Sand**



**Fig: 2 Deleterious Materials in WFS**



## Preliminary Investigation

**Table No: 1 Chemical Composition of Waste Foundry Sand**

S.NO.	Constituent	Percentage
1	SiO <sub>2</sub>	83.8
2	Al <sub>2</sub> O <sub>3</sub>	0.81
3	TiO <sub>2</sub>	0.22
4	CaO	1.42
5	MgO	0.87
6	Fe <sub>2</sub> O <sub>3</sub>	5.39
7	Na <sub>2</sub> O	0.87
8	K <sub>2</sub> O	1.14

9	SO <sub>3</sub>	0.21
10	Mn <sub>3</sub> O <sub>4</sub>	0.047

**Table No:2 Sieve Analysis Chart for Waste Foundry Sand**

IS sieve size	Weight retained (gms)	Cumulative weight retained (gms)	Cumulative percentage weight retained	Cumulative percentage passing
4.75mm	7	7	0.70	99.30
2.36mm	10	17	1.71	98.29
1.18mm	10	27	2.72	97.28
600μ	80	107	10.77	89.23
300μ	493	600	60.42	39.58
150μ	293	893	89.92	10.08
Pan	100	993	-----	-----
		ΣF=	166.24	

**Table No: 3 Physical Properties of Waste Foundry Sand**

S.No.	Property	Test Method	Test Results
1	Fineness modulus	Sieve analysis (IS 2386-1963 Part 2)	1.66
2	Specific gravity	Pycnometer (IS 2386-1963 Part 3)	2.35
3	Bulk density (kg/m <sup>3</sup> )	(IS 2386-1963 Part 3)	1350

**Table No: 4&5 Mix Proportion for different Grades**

GRADE OF CONCRETE	CEMENT	F.A	C.A	W/C Ratio
M <sub>20</sub>	1	1.85	3.42	0.5
M <sub>40</sub>	1	1.33	2.55	0.38
M <sub>60</sub>	1	1.01	2.06	0.32

GRADE OF CONCRETE	CEMENT	F.A	C.A	WATER
M <sub>20</sub>	350	650	1200	175
M <sub>40</sub>	430	575	1100	164
M <sub>60</sub>	520	530	1075	166

### Results & Discussion

**Table No: 6 Compressive Strength of Various Concrete Mixes with Replacement of Fine Aggregate over Waste Foundry Sand for Different Grades of Concrete**

Sl. No.	Mix ID	Compressive Strength (MPa)		
		M 20 Grade	M 40 Grade	M 60 Grade
1	WFS0	26.89	48.69	69.76
2	WFS20	30.37	52.76	72.23
3	WFS40	33.86	54.65	73.83
4	WFS60	26.31	45.63	61.62
5	WFS80	21.95	41.57	57.41
6	WFS100	19.77	34.59	43.45

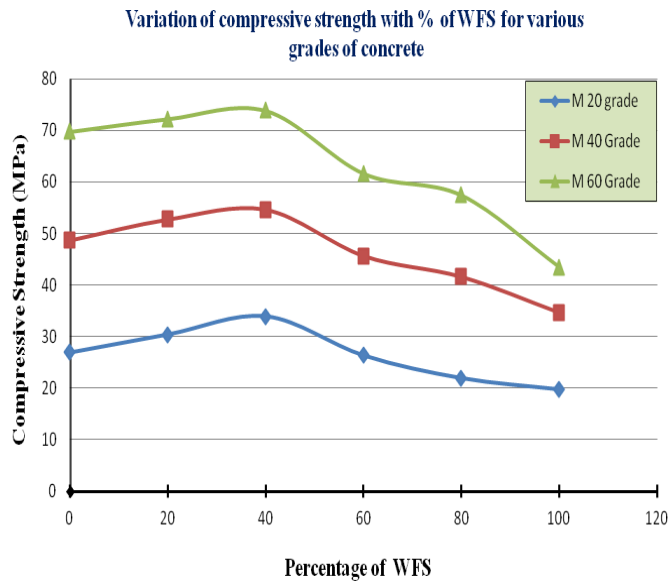
**Table No: 7 Split Tensile Strength of Various Concrete Mixes with Replacement of Fine Aggregate over Waste Foundry Sand for Different Grades of Concrete**

Sl. No.	Mix ID	Split Tensile Strength (MPa)		
		M 20 Grade	M 40 Grade	M 60 Grade
1	WFS0	3.26	5.05	7.12
2	WFS20	3.63	5.61	7.88
3	WFS40	4.34	6.37	9.06
4	WFS60	3.16	4.91	6.89
5	WFS80	2.93	4.25	6.28
6	WFS100	2.22	3.82	4.86

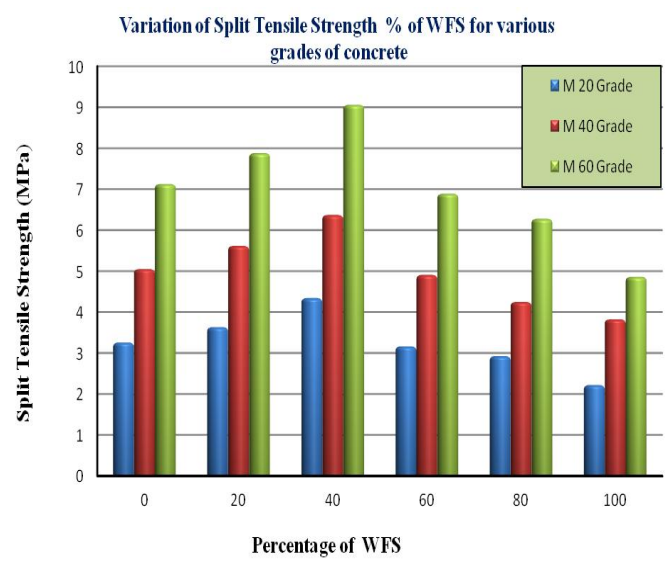
**Table No: 8 Flexural Strength of Various Concrete Mixes with Replacement of Fine Aggregate over Waste Foundry Sand for different Grades of Concrete**

Sl. No.	Mix ID	Flexural Strength (MPa)		
		M 20 Grade	M 40 Grade	M 60 Grade
1	WFS0	4.73	7.30	9.57
2	WFS20	5.34	8.08	10.49
3	WFS40	6.02	9.05	11.64
4	WFS60	4.50	6.80	9.00
5	WFS80	4.08	5.47	7.04
6	WFS100	3.45	4.71	6.15

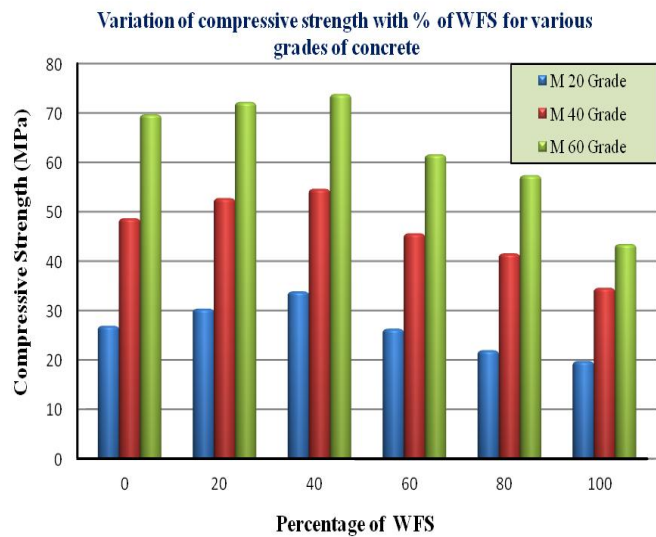
**Graph No: 1**



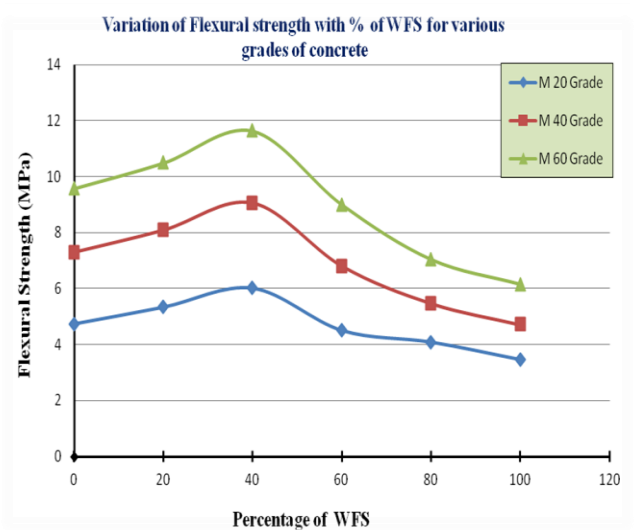
**Graph No: 4**



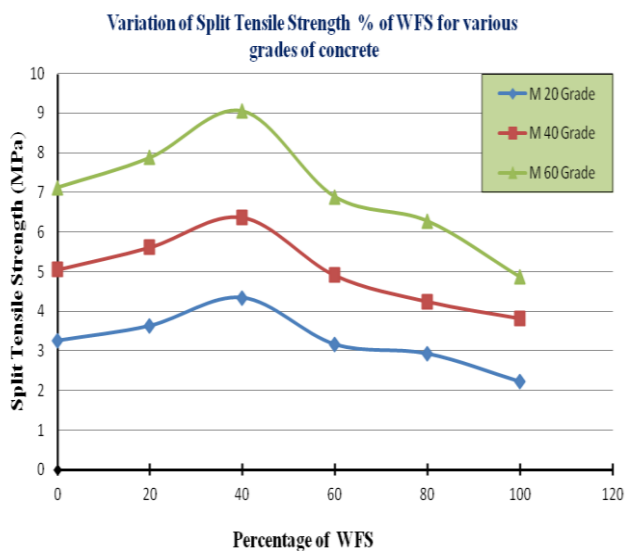
**Graph No: 2**



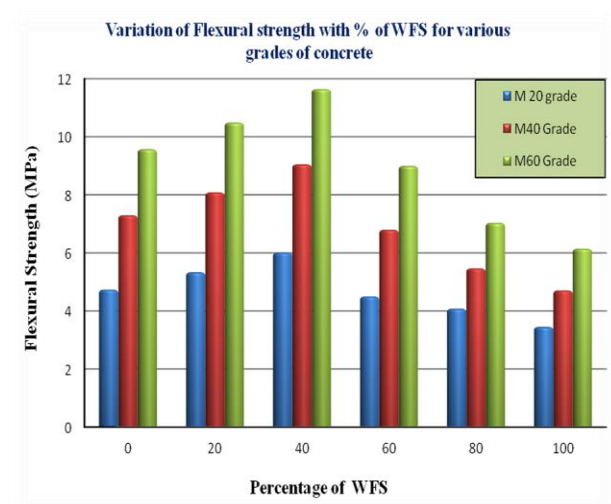
**Graph No: 5**



**Graph No: 3**



**Graph No: 6**



Dear K.N.Sujatha,

Thank You for showing interest in the 2nd International Conference on Sustainable Energy and Built Environment. We are pleased to tell you that your research paper **Assessment of Musi River Water and Near by Ground Water, Impacts on Health of Down Stream Villages of Hyderabad (PAPER ID : ICES-EN006)** has been selected for Oral Presentation.

You are kindly requested to register for the Conference by completing the payment process by **17th February, 2016**. The payment can be done through both online/offline (demand draft) mode. For further information click on the given link

<http://www.asceis-ices.com/conference.php>

Regards,  
Team ICES '16.

# Assessment of Musi River Water and Nearby Ground Water, Impacts on Health of Down Stream Villages of Hyderabad

K.N.Sujatha,

Malla Reddy Engineering College (A), Civil Engineering department, Telangana ,  
India.

suji.nissi07@gmail.com

---

## Abstract

**Objective:** The present study focused on the MUSI River contamination and its impacts on health of downstream villages of MUSI in Hyderabad city. These pollutants are responsible for the degradation of aquatic ecosystems and groundwater resources.

**Methods and Analysis:** The areas from PEERIZADIGUDA to PRATAPASINGARAM selected for analysis which are located at down streams of MUSI River. Ground water samples are collected from each location and Physical, chemical and biological parameters are analyzed by using suitable methodologies of  $P^H$ , Total dissolved salts, Turbidity, Total Hardness, chlorides sulphates and Biological oxygen demand water quality parameters are assessed. A study has made on these selected areas of population for health hazards due water born diseases by sampling method.

**Finding:** This paper discuss about Environmental sanitation, water pollution constitutes a serious problem particularly in the mega cities. City like Hyderabad has seen rapid economic development and urbanization which lead to the burden of illness and severity of skin and communicable diseases in and around MUSI River due to its worst contamination in recent days due to poor hygienic and to explore the importance of treated water supply.

**Novelty /Improvement:** Poor hygiene and polluted ground waters interventions are causing health hazards in these of areas. So, safe treatment of wastewater before disposal may sustain its natural environment. It protects the life in our rivers and ensures that all water sources are clean and may be easily used for the public supply.

---

**Key words:** water pollution, hygiene, and sanitary, health problems, waste water treatment

## 1. Introduction

Water is the most essential commodity for all living creatures. Organisms cannot survive without water. Water is one of the most essential constituents of the human environments. The expansion of agriculture and industrial development has not only increased water consumption considerably but has also affected water quality. Reference <sup>[3]</sup>. Water is easily polluted because of its great ability to dissolve substances. Even before raindrops touch the earth, they stand picking up pollutants. Once on the ground, water picks up things rapidly, and becomes contaminated. Reference <sup>[4]</sup>

The Hyderabad city discharges about 600 million liters per day untreated sewerage water into Musi River. The drinking water in entire area is brought from distant places; they were spending lots of money. There are at least 30 villages with a population of 1.00.000 that are directly affected in this region. Reference <sup>[2, 4]</sup> The quality of water in Musi is beyond description. Health problems can ensue for agricultural workers due to pathogenic bacteria, viruses and parasites present in the wastewater as well as for consumers of wastewater-irrigated produce particularly if the produce is not cooked before it is consumed. Hookworm infections are more common in agricultural workers who go barefoot in wastewater- irrigated fields Reference <sup>[8]</sup>.

This review was done to explore the impact of poor water treatment, hygiene, and sanitary interventions on health conditions, to reduce new and emerging infectious diseases have brought important new challenges to public health. In this research a systematic study has carried out on surface and ground waters pollution in the Musi River Basin, to evaluate its contamination,

and causes. Samples of water were obtained from individual water sources (wells) from villages and from surface water sources. Reference <sup>[5]</sup>. Contamination from different sources reaches groundwater through leaching of organic and inorganic fertilizers, animal waste, domestic effluents and industry. The untreated municipal sewage and solid waste into water bodies increases nitrogen concentrations and sometimes can reach more than 60mg/l. Reference <sup>[6]</sup>. The present study is an attempt to report a comparative account of health effects due to untreated municipal sewages, untreated effluent and water quality of the River Musi surrounding water ground Reference <sup>[7]</sup>.

Groundwater quality comprises the physical, chemical and biological characteristics of groundwater. The suitability of groundwater for various uses majorly depends on quality of groundwater. Hence protecting the quality of groundwater is a major concern Reference <sup>[1]</sup>. Groundwater pollution has been reported in many aquifers because of high concentration of organic and inorganic compounds in groundwater. The differences in health risks that these variations represent lead to different priorities for the treatment and provision of drinking water. Reference <sup>[11]</sup> Microbial contamination of drinking water remains a significant threat and constant vigilance is essential, even in the most developed countries, Reference <sup>[6]</sup> More recent research has suggested a possible association between disinfection by-products. But potential risks are largely outweighed by the benefits of drinking water with a low microbial load. Reference <sup>[12]</sup>

## 2. Study area

The study area is the upper part of the Musi basin a tributary of river Krishna of Telangana in India and downstream of Hyderabad. The area

falls in the Survey of India Topographic map number 56K/10 to a scale of 1:50,000. The area lies between to 17 ° 22 ' 58.8 "N latitude and to 78 ° 39 ' 39.276 "E longitude. The area is located between PEERIZADIGUDA to PRATAPASINGARAM of RANGA REDDY District, TELANGANA. The groundwater and river water quality analysis has been carried out for the water sample

Collected from the five locations located PEERIZADIGUDA bore well, RTC colony bore well, PARVATHAPURAM bore well, MUTHYLAGUDA bore well, and PRATAPASINGARAM bore well. In the present work attempts have been made to detect the quality by using conventional hydro geochemical methods.

### 3. Methodology

Ground water samples are collected from each location and Physical Properties and chemical properties and biological parameters are analyzed by using suitable methodologies for, PH, TDS, Turbidity, Total Hardness, chlorides sulphates and BOD water quality parameters are assessed.

A study has made on these selected areas of population for health hazards due water born diseases by sampling method

### 4. Results and Interpretation

**P<sup>H</sup>** of the water of the study area between 8.51 to 7.81 the PARVATHAPURAM and PRATAPASINGARAM ground water samples are not within normal range. Hence they are not suitable to be consumed as drinking water, why because due to the industrial effluents, the pH is also altering with the different time intervals at the same stations. From Table [1]

**Total Dissolved Solids**, the presence of dissolved salts in the water sample and From Table [1] the study area shows very high amount of TDS which indicates the water is not suitable for direct consumption and different components are responsible for the different types of the health hazards. Like diarrhea, joint pains, skin allergies, gastrointestinal disturbances, and vomiting. These symptoms also persist even by the consumption of agriculture productivity of those areas.

**Turbidity and Colour** is slightly varying in the study area, From Table [1]. It is due to leaching of Musi River which contain high amount of organic and inorganic wastes.

**Total hardness** is less than total alkalinity then the difference indicates the presence of sodium bicarbonate, it also influences the soil fertility. Excess Alkalinity in water is not good to consume and relatively not better and can cause Gastrointestinal problems.

**Chlorides:** Chloride levels in the study are slightly beyond the limits From Table [1]. It indicates that excess presence of Chlorine in water leads to gastrointestinal, diarrheic, and skin allergies

**Sulphates** induces the formation of sulphuric acid, Hydrogen sulphate, its may be a cause for the gastrointestinal and skin allergies. And study shows high amounts of sulphates as per From Table [1]

**BOD** is within the limits for the ground water samples of the study area From Table [1].

From Table [2] It has been found that the pollution has been given rise four major problems namely, pollution of drinking large incidence of diseases like arthritis, diarrheic, skin allergies, stomach pain, malaria, food poison, eye diseases, pediatric problems and jaundices diseases suffered by the people, are rich in these ground water and due to this has force people to buy water from outside, resulting extra burden on their family budget. But people of these areas are with less economic standards. So, they have satisfy with available ground waters and impact of pollution of Musi River surrounding ground water on these selected areas as health effects

## 5. Conclusions

The city of Hyderabad disposes very large quantities of untreated domestic sewage into the dry bed of the Musi River. Sewage disposal had a mixed impact on downstream users. Poor water quality had a negative impact on farmer health and undesirable. Ground water is the primary source of water used in these areas, thus knowledge about its availability and sustainability are essential for the successful

Management and future development of this limited resource. Ground-water availability and sustainability are influenced by many factors, one of which is water quality. Water quality generally has been over looked in these because the primary focus has been on obtaining a sufficient water supply.

As a result, people in poor areas are falling victim to water-borne diseases. Non-implementation of the environmental laws and hazard planning and growth of Hyderabad city have reduced the Musi river to a sewer drain carrying the domestic and industrial waste generated in Hyderabad city adversely impacting on the river ecology.

Drinking water treatment as applied to public water supplies consists of a series of barriers in a treatment train that will vary

according to the requirements of the supply and the nature and vulnerability of the source. The Guidelines are now based on Water Safety Plans that encompass a much more proactive approach to safety from source-to-tap.

The contamination of drinking water by pathogens causing diarrheal disease is the most important aspect of drinking water quality. The problem arises as a consequence of contamination of water by faecal matter, particularly human faecal matter, containing pathogenic organisms. The areas from PEERIZADIGUDA to PRATAPSINGARAM selected for analysis are located at down streams of Musi and the Musi conditions in these regions made worst due to heavy dumping of the organic waste, sewage from various sources from industrial and municipal waste. And the safe water supply for these areas is very less and people of these areas depend on ground water and poor hygiene, sanitary and polluted ground waters, interventions are causing health hazards in these of PRATAPSINGARAM. So, safe treatment of our wastewater and its return to the natural environment (rivers or the sea) is a key part of the water cycle. It protects the life in our rivers and ensures that all water sources are clean and may be easily used for the public. Thus, anthropogenic activity only can be a source of contaminants to the water table.

Sustainable developments like waste water treatment systems and natural techniques of filtration methods can sustain or help the environment surrounding of Musi, and from various skin and health hazards through water pollution.

## 6. References

1. S. Packialakshmi<sup>1</sup>, Meheli Deb<sup>2</sup> and Hrituparna Chakraborty Assessment of Groundwater Quality Index in and Around Sholinganallur Area, Tamil Nadu Indian Journal of Science and Technology, Vol 8(36), DOI: 10.17485/ijst/2015/v8i36/87645, December 2015
2. Pullaiahcheepi "Impact of pollution of Musi River Water in Down Stream Villages-A Study". IOSR Journal of Environmental Science, Toxicology and Food Technology ISSN: 2319-2402, ISBN: 2319-2399. Volume 1, Issue 4 (Sep-Oct. 2012), PP 40-51
3. Salve, V. B. and Hiware C. J. Study of water quality of Wanparakalpa reservoir Nagpur, Near ParliVaijnath, District Beed. Marathwada region, J. Aqua. Biol.(2008), 21(2): 113-117
4. Hujare, M. S. Seasonal variation of physico-chemical parameters in the perennial tank of Talsande, Maharashtra. Ecotoxicol. Environ. Monitor.(2008) 18(3): 233-242
5. Jeroen H. J. Ensink & Christopher A. Scott Simon Brooker & Sandy Cairncross. Sewage disposal in the Musi-River, India: water quality,(2009) Irrig Drainage Sys Springer DOI 10.1007/s10795-009-9088-4
6. Kumar RM "Predatory industrialisation and environmental degradation: a case study of Musi-river. Telangana: Dimensions of underdevelopment, Center for Telangana studies Hyderabad,(2003), pp 203–212.
7. Van der Hoek, Mehmood et al, Urban Wasterwater: A Valuable Resource for Agriculture- A Case Study from Haroonabad, Pakistan. (2002). IWMI –Research Report 63
8. APHA-WWE-WEF (1998) Standard methods for the examination of water and wastewater. American Public Health Association, Washington D.C Ayres RM, Mara DD (1996)
9. Swaranlatha, S. and A. Narsingrao Ecological studies of Banjara lake with reference to water pollution (1998). J. Envi. Biol. 19(2): 179-186. 13.
10. Rama Rao, Physico-chemical and biological characteristics of Husain sagar, an industrially polluted lake, Hyderabad (1993), Environ. Biol
11. Krishnamurthy R, Hydro-biological studies of Wohar reservoir Aurangabad (Maharashtra State) India.(1990). J. Environ. Biol. 11(3), 335-343
12. Karanth, K.R Groundwater Assessment Development and Management (1987): Tata McGraw Hill publishing company Ltd., New Delhi, 725p
13. Venkateswarlu V, Sponge city: water balance of mega-city water use and wastewater use in Hyderabad. (1969), India Irrig and Drain 54:S81–S9, An ecological study of algae of the river Moosi, Hyderabad (India) with special reference to water pollution I. Physico-chemical complexes. Hydrobiologica 33:117–143.

**Table no: 1**

PARAMETER S	PEERIZADIGUDA	RTC COLONY	PARVATHAPURAM	MUTHYLAGUDA	PRATAPSINGARAM	INDIAN STANDARD OF DRINKING WATER
p <sup>H</sup>	7.97	7.83	8.23	7.89	8.51	6-8
TDS mg/l	1250	1820	2410	2175	2250	500
Turbidity NTU	8 NTU	7	7.3	7.56	8.2	10
Total hardness mg/l	340	355	328	329	342	100
Chlorides mg/l	140	138	135	209	202	200
Sulphates mg/l	345	231	246	302	359	200
BOD mg/l	2.2	1.2	2.4 mg/l	2.3	2.6	3

**Ground water quality parameters of the study area**

**Table no: 2**

Study Areas	Sample distribution	Skin diseases %	Diarrheic %	Arthis %	Gastro intestinal problems%	Skin allergies %	Malaria %
<b>PEERIZADIGUDA</b>	57	55	35	45	45	65	35
<b>RTC COLONY</b>	63	68	46	52	50	75	45
<b>PARVATHAPURAM</b>	43	46	32	56	44	68	48
<b>MUTHYLAGUDA</b>	65	54	37	55	34	85	55
<b>PRATAPSINGARAM</b>	85	86	52	65	48	83	58

**Field survey conducted randomly on Jan15 to April-15 in the study area for health hazards**



**Department of  
Electrical and Electronics  
Engineering**





# COMPARISON OF PI AND RST CONTROLLERS IN DOUBLE-FED INDUCTION GENERATOR FOR WIND ENERGY CONVERSION SYSTEMS

Mrs.V.Sumadeepthi<sup>1</sup>, Mr. P.Sankar babu<sup>2</sup>

<sup>1</sup>Assistant Professor & MREC(A), <sup>2</sup>Associate Professor & MREC(A)

<sup>1</sup>sumadeepthi.veeraganti@gmail.com

[pscrbabu@gmail.com](mailto:pscrbabu@gmail.com)

**Abstract-** This paper deals with a variable speed device to produce electrical energy on a power network based on a doubly-fed induction machine used in generating mode (DFIG). This device is intended to equip nacelles of wind turbines. First, a mathematical model of the machine written in an appropriate d-q reference frame is established to investigate simulations. In order to control the power flowing between the stator of the DFIG and the power network, a control law is synthesized using two types of controllers: PI and RST. Their respective performances are compared in terms of fault. After modeling the DFIG and choosing the appropriate d-q reference frame, the transfer function of DFIG is calculated and with that transfer function PI and RST are synthesized mathematically. Using MATLAB SIMULINK software LG and LLLG faults are created without and with controllers like PI and RST controllers. Response of the system is studied. It is found that the RST controller offers better performance.

**Keywords:** DFIG, Parks transformation, RST controller, PI controller, wind turbine, d-q reference frame.

## NOMENCLATURE:

V<sub>ds</sub>, V<sub>qs</sub>: Two-phase statoric Voltages

V<sub>dr</sub>, V<sub>qr</sub>: Two-phase rotoric Voltages

Ψ<sub>ds</sub>, Ψ<sub>qs</sub>: Two-phase statoric fluxes.

Ψ<sub>dr</sub>, Ψ<sub>qr</sub>: Two-phase rotoric fluxes.

I<sub>ds</sub>, I<sub>qs</sub> : Two-phase statoric currents.

I<sub>dr</sub>, I<sub>qr</sub>: Two-phase rotoric currents

θ<sub>s</sub>, θ<sub>r</sub>: Statoric flux and rotoric mechanical position.

Ω: Mechanical speed

Γ<sub>m</sub>, Γ<sub>e</sub>: Prime mover and electromagnetic torque.

P : Number of pole pairs.

R<sub>s</sub>, R<sub>r</sub> : Per phase statoric and rotoric resistances.

M: Magnetizing inductance.

L<sub>s</sub> : Total cyclic statoric inductance.

L<sub>r</sub>: Total cyclic rotoric inductance.

g : generator slip.

J, F: Inertia rotoric inductances viscous friction.

p : Laplace operator

## I. INTRODUCTION

In order to meet power needs, taking into account economical and environmental factors, wind energy conversion is gradually gaining interest as a suitable source of renewable energy. The electromagnetic conversion is usually achieved by induction machines or synchronous and permanent magnet generators. Squirrel cage induction generators are widely used because of their lower cost, reliability, Instruction and simplicity of maintenance [1]. But when it is directly connected to a power network, which imposes the frequency, the speed must be set to a constant value by a mechanical device on the wind turbine. Then, for a high value of wind speed, the totality of the theoretical power cannot be extracted. To overcome this problem, a converter, which must be dimensioned for the totality of the power exchanged, can be placed between the stator and the network. In order to enable variable speed operations with a lower rated power converter, doubly-fed induction generator (DFIG) can be used. The stator is directly connected to the grid and the rotor is fed to magnetize the machine.

In this paper, the control of electrical Power exchanged between the stator of the DFIG and the power network by controlling independently the torque (consequently the active power) and the reactive power is presented. Several

investigations have been developed in this direction using cyclo-converters as converters and classical proportional-integral regulators. In our case, after modeling the DFIG and choosing the appropriate d-q reference frame, active and reactive powers are controlled using respectively Integral-Proportional (PI) and an RST controller based on pole placement theory. Their performances are compared in terms of fault.

## II. DOUBLE-FED INDUCTION GENERATOR

DFIG is an abbreviation for Double Fed Induction Generator, a generating principle widely used in wind turbines. It is based on an induction generator with a multiphase wound rotor and a multiphase slip ring assembly with brushes for access to the rotor windings. It is possible to avoid the multiphase slip ring assembly. The principle of the DFIG is that rotor windings are connected to the grid via slip rings and back-to-back voltage source converter that controls both the rotor and the grid currents. By controlling the rotor currents by the converter it is possible to adjust the active and reactive power fed to the grid from the stator independently of the generators turning speed

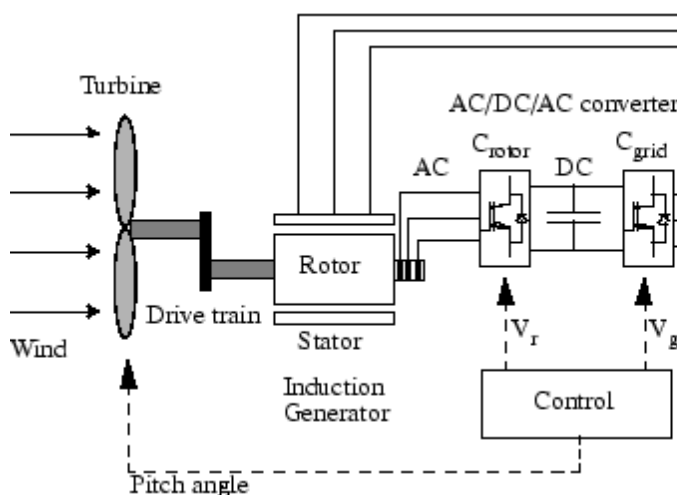


Fig.1 :Double-fed Induction Generator

## A. MATHEMATICAL MODEL OF DFIG

For a doubly-fed induction machine, the Concordia and Park transformation's application to the traditional a,b,c model allows you to write a dynamic model in d-q reference frame.

$$V_{ds} = R_s I_{ds} + \frac{d}{dt} \Psi_{ds} - \theta_s \Psi_{qs} \quad \text{---Equation1}$$

$$V_{qs} = R_s I_{qs} + \frac{d}{dt} \Psi_{qs} + \theta_s \Psi_{ds} \quad \text{---Equation2}$$

$$V_{dr} = R_r I_{dr} + \frac{d}{dt} \Psi_{dr} - \theta_r \Psi_{qr} \quad \text{---Equation3}$$

$$V_{qr} = R_r I_{qr} + \frac{d}{dt} \Psi_{qr} + \theta_r \Psi_{dr} \quad \text{---Equation4}$$

$$\Psi_{ds} = L_s I_{ds} + M I_{dr} \quad \text{---Equation5}$$

$$\Psi_{qs} = L_s I_{qs} + M I_{qr} \quad \text{---Equation6}$$

$$\Psi_{dr} = L_r I_{dr} + M I_{ds} \quad \text{---Equation7}$$

$$\Psi_{qr} = L_r I_{qr} + M I_{qs} \quad \text{---Equation8}$$

$$\Gamma m = \Gamma_e + J \frac{d\Omega}{dt} + f\Omega \quad \text{---Equation9}$$

$$\Gamma_e = -P M L_s (\Psi_{qs} I_{dr} - \Psi_{ds} I_{qr}) \quad \text{---Equation10}$$

By choosing this reference frame, statoric voltages and fluxes can be rewritten as follows:

$$V_{ds} = 0 \quad \text{---Equation11}$$

$$V_{qs} = V_s \sin \theta_s \quad \text{---Equation12}$$

$$\Psi_{ds} = 0 = L_s I_{qs} + M I_{qr} \quad \text{---Equation13}$$

$$\Psi_{dr} = L_r I_{dr} + M I_{ds} \quad \text{---Equation14}$$

$$\Psi_{qs} = \Psi_{ds} = L_s I_{ds} + M I_{dr} \quad \text{---Equation15}$$

$$\Psi_{qr} = L_r I_{qr} + M I_{qs} \quad \text{---Equation16}$$

The statoric active and reactive power, the rotoric fluxes and voltages can be written versus rotoric currents as:

$$P = -V_s (M L_s) I_{qr} \quad \text{---Equation17}$$

$$Q = V_s \sin \theta_s / L_s - (V_s M / L_s) I_{dr} \quad \text{---Equation18}$$

$$\Psi_{dr} = (L_r - (M^2 / L_s)) I_{dr} + M V_s / \sin \theta_s L_s \quad \text{---Equation19}$$

$$\Psi_{qr} = (L_r - (M^2 / L_s)) I_{qr} \quad \text{---Equation20}$$

$$V_{dr} = R_r I_{dr} + (L_r - (M^2 / L_s)) \frac{dI_{dr}}{dt} - g \omega_s (L_r - (M^2 / L_s)) I_{qr} \quad \text{Eq(21)}$$

$$V_{qr} = R_r I_{qr} + (L_r - (M^2 / L_s)) \frac{dI_{qr}}{dt} + g \omega_s (L_r - (M^2 / L_s)) I_{dr} \quad \text{Eq(22)}$$

In steady state, the second derivative terms of the two equations are nil. The third terms, which constitute cross-coupling terms, can be neglected because of their small influence. Knowing relations, it is possible to synthesize the regulators and establish the global block-diagram of the controlled system. To realize

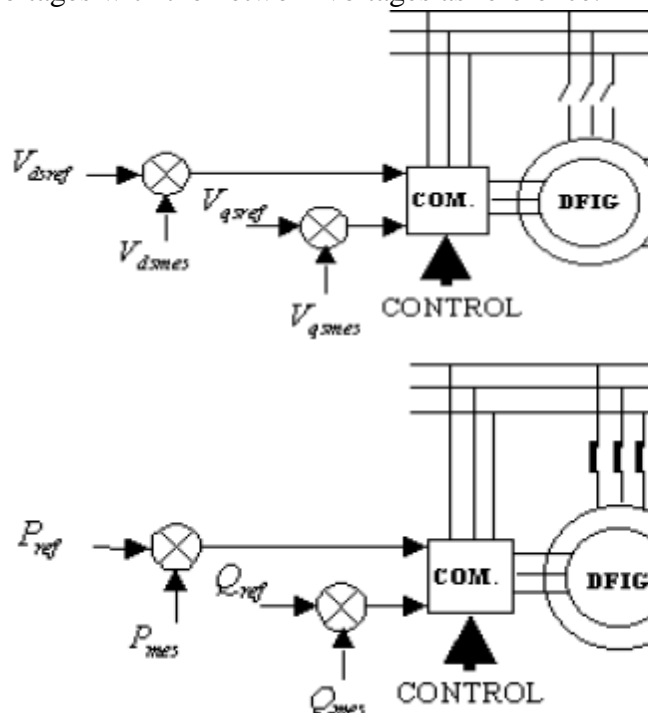
these objectives, two types of regulators are studied: Proportional Integral and RST controller based on pole placement theory [8]. Assuming that the cross-coupling terms are neglected, rotor currents are then directly related to active and reactive power by constant terms. Then, internal current control loops are not necessary. The synthesis of proportional-integral controller is achieved by the classical method of pole compensation and RST controller is achieved by robust pole placement method.

### III. DFIG CONTROL

#### A. Aim Of The Control

When the DFIG is connected to an existing network, this connection must be done in three steps:

1) The first step is the regulation of the statoric voltages with the network voltages as reference.



**Fig. 2 : First & Third step of the DFIG connection**

2) The second step is the stator connection to this network. As the voltages of the two devices are synchronized, this connection can be done without problem.

3) Once this connection is achieved, the third step is the power regulation between the stator and the network.

#### B. PI Controller Synthesis

In most systems of practical interest the problem is to reduce the steady-state error, while maintaining satisfactory transient response. a PI Controller (proportional-integral controller) is a feedback controller which drives the plant to be controlled with a weighted sum of the error (difference between the output and desired set point) and the integral of that value. This pi controller controls DFIG by classical method of pole compensation. Pole placement is the most straightforward means of controller design.

1. The design starts with assumption of what form the controller must take in order to control the given plant.
2. From that assumption a symbolic characteristic equation is formed.
3. At this point the desired closed loop poles must be determined.
4. Typically specifications designate overshoot, rise time, etc. this leads to the formation of second order equation.
5. Once the closed loop poles are decided a desired characteristic equation is formed.
6. The coefficients of each power of  $s$  are equated from the symbolic characteristic equation to the desired.
7. Algebra is used to determine the controller coefficients necessary to achieve the desired closed loop poles with the assumed controller form.

As Dfig has first order equation pi is also taken in first order equation. The block diagram containing Dfig and pi controller with unity feedback is shown below here  $G_c(s)$  is the pi controller and  $G(s)$  is the Dfig.

$R(s)$  is the step input and  $Y(s)$  is the output.

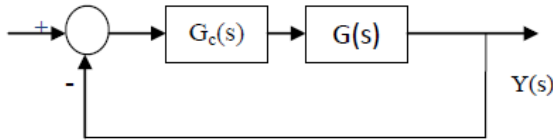


Fig.3: Block diagram of pi controller  
Transfer function of Dfig is  $120.000262s+0.019$   
---Equation23

The controller output is given by  
 $K_p \Delta + K_i \Delta dt$  -----Equation24  
 $\Delta = SP - PV$ . where  $\Delta$  is the error or deviation of actual measured value ( $PV$ ) from the set-point ( $SP$ ).

A PI controller can be modelled easily in software such as Simulink using a "flow chart" box involving Laplace operators:

$C = G(1+s\tau)s\tau$  -----Equation25 Where  $G = KP$  = proportional gain,  $G / \tau = KI$  = integral gain

The transfer function of the above block diagram is

$K(s) = G_c s G(s) / (1 + G_c s G(s))$  -----Equation26

Here we have to find  $G_c(s)$

$K(s) = K_p s + K_i = sK_p + K_i$  -----Equation27

By multiplying  $K(s)$  with  $B/A$  and comparing with second order equation comparing the setting peak overshoot by 90% and time setting as 0.2sec the values of  $k_p$  and  $k_i$  are 0.000768 and  $5.6e-006$  respectively. the total transfer function of the block diagram containing Dfig and pi controller with unity feedback is

$0.009216s+6.805e-50.000262s^2+0.02816s+6.805e-5$  ----Equation28

The plot obtained for this transfer function by giving a step input is shown below:

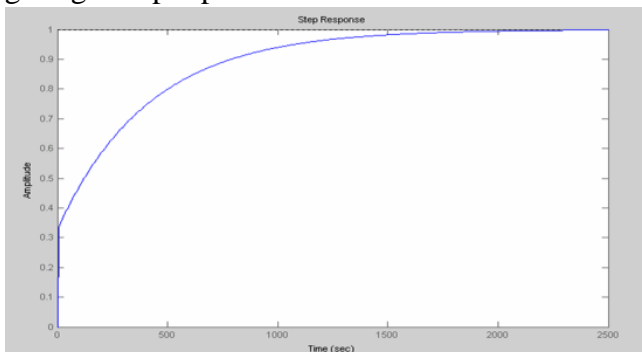


Fig.4: plot for step input

### C. Mathematical Synthesis Of Rst Controller To Dfig

The block-diagram of a system with its RST controller is presented as follows: The RST controller is also called two-degree of freedom why because in the feedback path there is one transfer function and it acts as a controller and in the feed-forward path there is another transfer function which also acts as an another controller. So it is called as two-degree of freedom. The system with the transfer-function  $B/A$  has  $Y_{ref}$  as reference and is disturbed by the variable  $\gamma$ .  $R$ ,  $S$  and  $T$  are polynomials which constitutes the controller. In our case, we have:

$A = Ls R_r + pLs(Lr - (M^2 / Ls))$  and  $B = MVs$  ----Equation29

Where  $p$  is Laplace operator. In this part, simulations are investigated with a 13 kW generator connected to a 220V/50Hz grid. The machine's parameters are presented below :

$R_s = 0.05\Omega$ ;  $R_r = 0.38\Omega$ ;  $M = 47.3mH$  ;  $L_s = 50mH$  ;  $L_r = 50mH$  ;  $J = 0.5kg.m^2$  ;  $f = 0.0035N.m.s-1$

By substitution,  $A = 0.09 + p0.000262$  ;  $B = 12$

By applying Besout equation, we put:  $D = AS + BR = CF$

Where  $C$  is the command polynomial and  $F$  is the filtering polynomial.

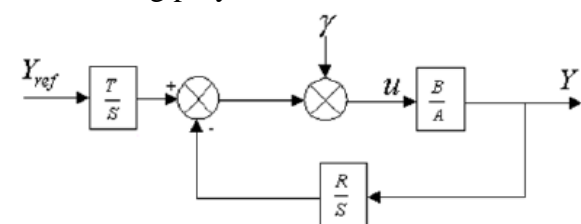


Fig.5: Block diagram of RST controller

Command polynomial is for reduce Robustness and the filtering polynomial is to maintain regulation. In order to have good adjustment accuracy, we choose a strictly proper regulator. So if  $A$  is a polynomial of  $n$  degree.

$(\deg(A)=n)$  we must have :  $\deg(D)=2n+1$ ,

$\deg(S)=\deg(A)+1$

$\deg(R)=\deg(A)$ . In our case

$A = a_1p + a_0$  -----Equation30

$$R=r_1p+r_0 \text{ -----Equation31}$$

$$B=b_0 \text{ -----Equation32}$$

$$S=s_2p^2+s_1p+s_0 \text{ -----Equation33}$$

$$D=d_3p^3+d_2p^2+d_1p+d_0 \text{ ---Equation34}$$

By robust pole placement method,

$$P_c = -5L_s R_r L_s (L_r - (M^2 / L_s)) = -361.615$$

$$P_f = -P_c = -120.538$$

$$D = (p - 120.538)^2 (p - 361.615) = p^3 + 86573.99 p^2 + 14529.409 p - 5254052.23.$$

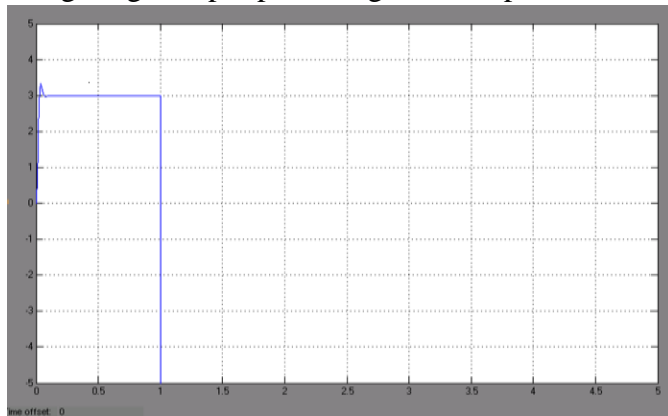
Here we are two poles of command polynomial and one pole of filtering polynomial, we need to reduce more Robustness than the Regulation. The Bezout equation leads to four equations with four unknown terms where the coefficients of D are related to the coefficients of polynomials R and S.

$$d_3 = a_1 s_2 \quad d_2 = a_1 s_1 \quad d_1 = a_0 s_1 + b_0 r_1 \quad d_0 = b_0 r_0$$

By calculating coefficients of R and S, now

$$RS = 2792.51s + 504908.843806s^2 + 2294132.94s \quad TS = 504908.843806s^2 + 2294132.94s$$

By placing the values in the RST block diagram and giving a step input, we get the scope as

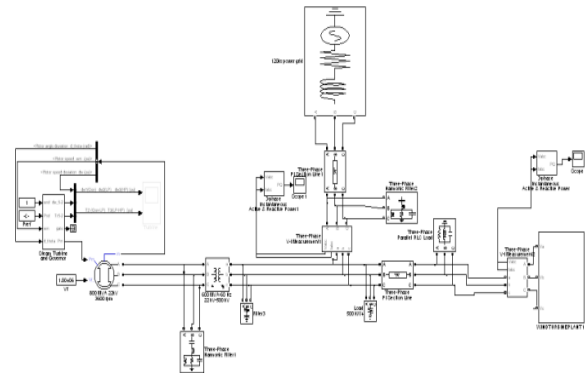


**Fig.6: Response to step input**

#### IV. SIMULATION

##### A. Construction Of DFIG To Grid

The double fed induction generator is synchronized to the grid as shown the diagram:



**Fig.7: Simulation diagram representing fault at wind side**

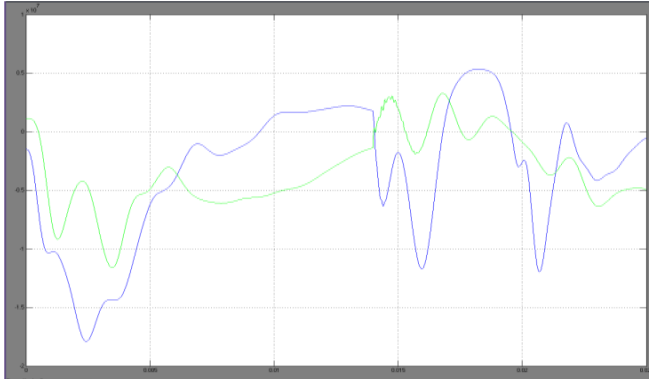
The grid parameters (voltage, frequency) are: 120kv, 60Hz. The total system consists of a thermal plant consisting a steam governor and an induction generator and wind plant consisting of double fed induction generator. The active and reactive power measurements are taken at the wind side and grid side with the help of three phase instantaneous active and reactive power measuring device. The system is connected in three phase PI section. Some loads are present across the grid or the system. In order to reduce harmonics, three phase harmonic filters are placed across the grid. Scopes are connected to the Three phase instantaneous active and reactive power measurements in order observe the active and reactive power.

##### B. Simulations With Fault At Wind Side

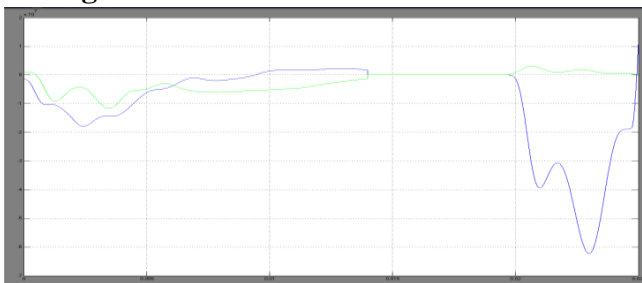
In order to check the performances of both controllers RST and PI which are fitted to the wind plant, the faults are applied at the wind side and grid side respectively. The reason behind applying the faults near the wind and grid is the variations can be clearly observed when the faults are applied to the nearest point with respect to the wind turbine. The faults which are applied are LLLG and LG fault. As the two faults are most severe (LLG) and most frequent occurring (LG) faults. the performances of the controllers are checked with these faults. The fault is placed at the wind side and the fault parameters are: Ground resistance: 0.001Ω Fault



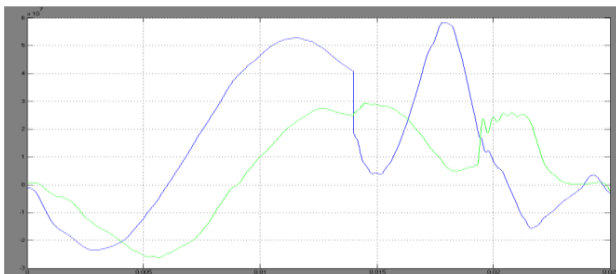
timings are: 14 ms to 17ms, Simulation timings are: 0 to 25ms. **Blue colour line : Real power,**  
**Green colour line: Reactive power.**



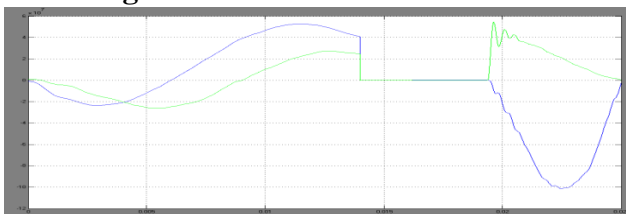
**Fig.7: Without controller for LG fault**



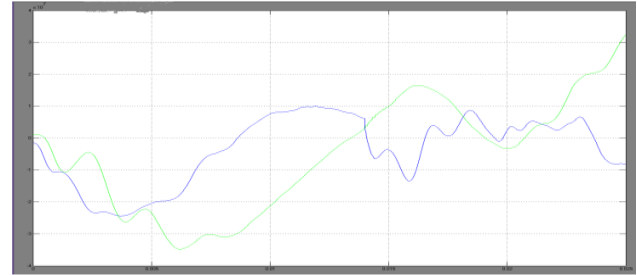
**Fig.8: Without controller for LLLG fault**



**Fig.9: PI controller for LG fault**



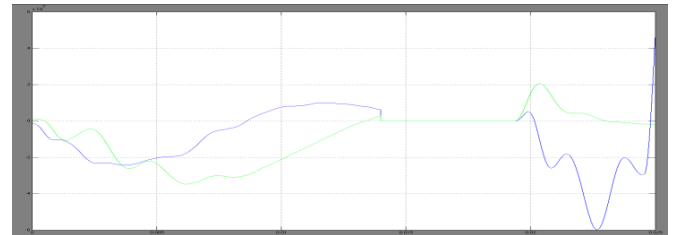
**Fig.10: PI controller for LLLG fault**



**Fig.11: RST controller for LG fault**

### *C. Result For Fault At Wind Side*

when LG fault is present at wind side when there is no controller connected to the wind plant the variations in active power are large but still the system's active power is same to post fault condition. In case of PI controller the transients are reduced during the fault and both active, reactive power prior to post fault conditions are same.

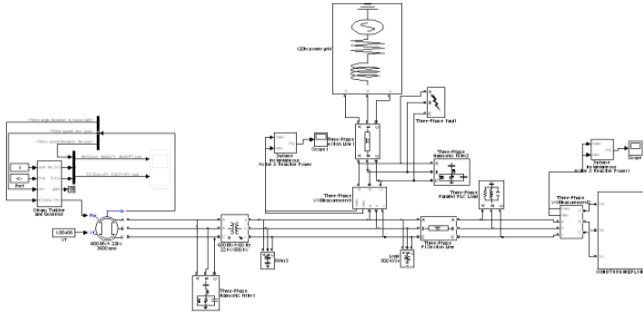


**Fig.12: RST controller for LLLG**

In case of RST controller even though transients appear during fault, the magnitude drastically reduced compared to both above conditions, thus, with RST controller the performance is considerably better..In case of LLLG fault , in absence of controller both active and reactive power are zero during fault time and subsequent to the fault also, the transients in active and reactive power are more. With PI controller, the active and reactive power is zero during fault. After the fault, the reactive power is increased with considerable transients and gradually reduces. Active power drops to a large value with no transients.

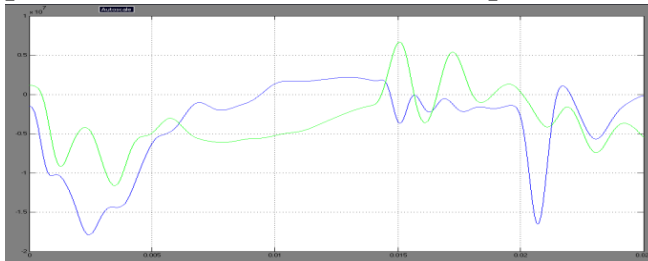
### *D. Fault At Grid Side*

The Simulation block Diagram is shown in Fig 12

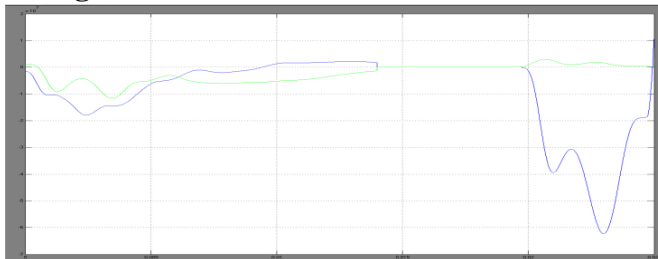


**Fig.12: Simulation diagram of dfig with fault at grid side**

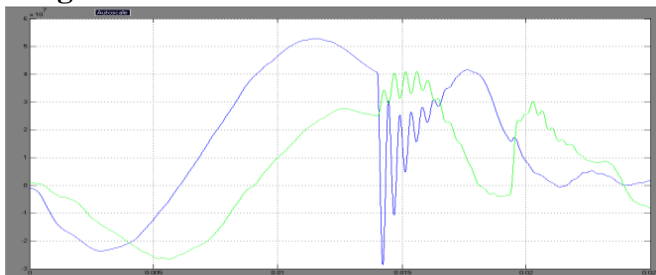
The fault is placed at the grid side and the fault parameters are: Ground resistance:  $0.001\Omega$ , Fault timings are: 14 ms to 17ms, Simulation timings are: 0 to 25ms, Simulations observed in the scope observed at wind side. **Blue colour line: Real power, Green colour line: Reactive power.**



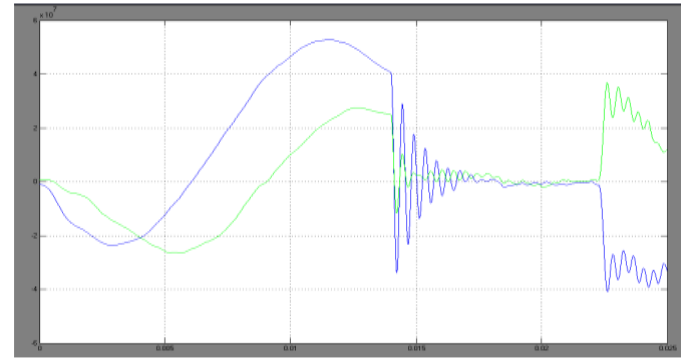
**Fig.13: Without controller for LG fault**



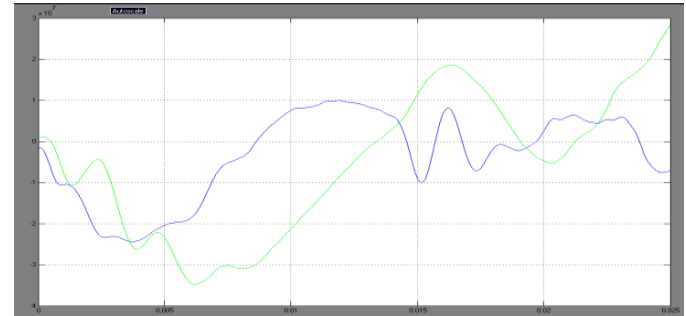
**Fig.14: Without controller for LLLG fault**



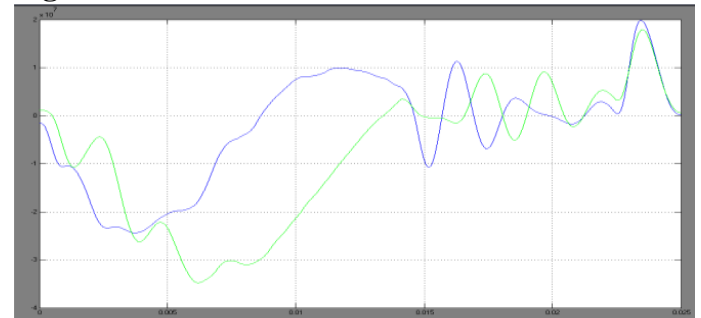
**Fig.15: PI controller for LG fault**



**Fig.16: PI controller for LLLG fault**



**Fig.17: RST controller for LG fault**



**Fig.18: RST controller for LLLG fault**

#### *E. Result For Fault At Grid Side*

At the grid side, the transients are more when compared to wind side. For without controller and PI controller after post fault and pre fault are same. In RST controller the reactive power is increased at grid with reduction in active power.

#### **V. CONCLUSION**

This paper has presented a device intended to fit in a wind mill based on a Doubly Fed Induction Generator connected to the grid. In order to control statoric active and reactive power exchanged between the DFIG and the grid, a vector-control strategy has been presented. Two controllers RST and PI controller are taken for



controlling the Doubly Fed Induction Generator and their mathematical calculations are performed. The influences of faults at wind and grid side are observed with the action of controller and without the action of controller. the effect of controllers like PI and RST on DFIG limited to a short line. It is found that both PI and RST controllers are able to give better performance in improving the active power profile either by absorbing or delivering the reactive power. However, pi controller is producing certain oscillations in active power profile, but RST controller is able to give better performance than PI Controller in Active power profile.

#### *References*

- [1] S. Heier, "Grid integration of wind energy conversion systems.", John Wiley & Sons Ltd, England, 1998.
- [2] L. Xu, W. Cheng, "Torque and reactive power control of a doubly-fed induction machine by position sensorless scheme.", IEEE Transactions on Industry Applications, Vol. 31, N°3, pp 636-642, May/June 1995.
- [3] M. Yamamoto, O. Motoyoshi, "Active and reactive power control for doubly-fed wound rotor induction generator.", IEEE Transactions on Power Electronics, Vol. 6, N°. 4, Oct. 1991, pp 624-629.
- [4] M. B. Rifai, T. H. Ortmeyer, "Dynamic analysis of a doubly fed generator in power system applications." Electric Machines and Power Systems, Vol. 21, pp 41-150, 1993.
- [5] B. Hopfensperger, D.J. Atkinson, R.A. Lakin, "Stator-flux-oriented control of a doubly fed induction machine with and without position encoder.", IEE Proc.-Electr. Power Appl., Vol.147, N°. 4, July 2000, pp 241-250.
- [6] L. Morel, A. Mirzaian, J.M. Kauffmann: "Field oriented control for double fed induction machine : simulation and experimental results.", ELECTRIMACS'96., Vol. 2, September 17-18-19, 1996, pp 391,396.
- [7] Control Of A Doubly-Fed Induction Generator For Wind Energy Conversion Systems.F. Poitiers, M. Machmoum, R. Le Doeuff And M.E. Zaim.
- [8] D. Ramuz, A. Mirzaian, J.M. Kauffmann, S.Tnani, H. Godfroid: "Control strategy of double fed machine for variable speed drives.", proceedings of ICEM 98, Istanbul, Sept. 1998, pp. 1340-1345



# Synchronous Static Series Compensator (SSSC) on Enhancement of Voltage Stability and Power Oscillation Damping for unbalanced system

S.Bharathi<sup>1</sup>, Y. Sudha<sup>2</sup>

<sup>1</sup>Assistant professor & MREC (A)

<sup>2</sup>Assistant Professor & MREC (A)

<sup>1</sup>barati.reddy10@gmail.com

<sup>2</sup>Yelmareddy.sudha@gmail.com

**Abstract**— This paper investigates the problem of controlling and modulating power flow in a transmission line using a Synchronous Static Series Compensator (SSSC). The studies, which include detailed techniques of twelve pulse and PWM controlled SSSC, are conducted and the control circuits are presented. The developed control strategies for both twelve-pulse and PWM-controlled SSSC use direct manipulations of control variables instead of typical d-q transformations.

The SSSC, a solid state VSI coupled with a transformer is connected in series with a transmission line. This SSSC injects an almost sinusoidal voltage of variable magnitude, in series with a transmission line. This injected voltage is almost in quadrature with the line current, thereby emulating an inductive or a capacitive reactance in series with the transmission line. This emulated variable reactance, inserted by the injected voltage source, influences the electric power flow in the transmission line.

The complete digital simulation of the SSSC within the power system is performed in the MATLAB/Simulink environment using the Power System Blockset (PSB). Simulation results validate that Voltage and Power Oscillation can be damped properly using of Synchronous Static Series Compensator (SSSC) under unbalanced system.

**Keywords**— SSSC, Reactive compensation, Control strategy, FACTS, PWM control and voltage stabilization.

## I. INTRODUCTION

In the last decade, commercial availability of Gate Turn-Off (GTO) thyristor switching devices with high power handling capability and the advancement of the other types of power-semiconductor devices such as IGBTs have led to the development of fast controllable reactive power sources utilizing new electronic switching and converter

technology. These switching technologies additionally offer considerable advantages over existing methods in terms of space reductions and fast effective damping [1]. The advent of FACTS systems is giving rise to a new family of power electronic equipment for controlling and optimizing the dynamic performance of power system, e.g., STATCOM, SSSC, and UPFC. The use of voltage-source inverter (VSI) has been widely accepted as the next generation of flexible reactive power compensation to replace other conventional VAR compensation, such as the thyristor-switched capacitor (TSC) and thyristor controlled reactor (TCR) [2,3]. The SSSC controller consists of a solid-state VSC with several GTO thyristor switches, or any other semiconductor switches with intrinsic turn-off capability valves, a dc capacitor, a transformer, and a controller. It is important to note that several of the VSCs can be connected together through a transformer of sometimes a Complex and custom made design configuration. The number of valves and the various configurations of the transformer depend on the desired quality of ac waveforms generated by the SSSC [4,5]. The line side transformer winding is connected in series, placing the VSC also effectively in series with a transmission line, and thus allowing series compensation of the line. The SSSC is used to generate or absorb reactive power from the line, and hence can be utilized as a transmission line power flow controller. Basically, it generates on its output terminals a quasi-sinusoidal voltage of variable magnitude in quadrature with the transmission line current, if the SSSC losses are neglected. Thus, the line injected voltage emulates a capacitive or an inductive reactance in series with a transmission line, which increases or decreases the total transmission line reactance, resulting in a decrease or increase of the power flow in the

transmission line [6]. In general, the SSSC can be viewed as analogous to an ideal synchronous voltage Source as it can produce a set of three-phase ac voltages at the desired fundamental frequency of variable and controllable amplitude and phase angle. It also resembles a synchronous compensator, as it can generate or absorb reactive power from a power system and can, independently from the reactive power, generate or absorb real power if an energy storage device instead of the dc capacitor is used in the SSSC. The SSSC is typically restricted to only reactive power exchange with the nearby ac system, neglecting the small amount of real power used to cover the circuit and switching losses, because of the relatively small SSSC capacitor. If the dc capacitor were replaced with an energy storage system, the controller would be able to exchange real power with the ac system and compensate for the transmission line resistance [7, 8].

## II. BASIC OPERATING PRINCIPLES OF THE SSSC:

Fig 1 Shows a functional model of the SSSC where the dc capacitor has been replaced by an energy storage device such as a high energy battery installation to allow active as well as reactive power exchanges with the ac system. The SSSC's output voltage magnitude and phase angle can be varied in a controlled manner to influence power flows in a transmission line. The phase displacement of the inserted voltage  $V_{pq}$ , with respect to the transmission line current  $I_{line}$ , determines the exchange of real and reactive power with the ac system. Fig.2 shows the SSSC operation in four quadrants, again assuming an energy storage device connected at the SSSC's input terminals. The line current phasor  $I_{line}$  is used as a reference phasor while the injected SSSC voltage phasor is allowed to rotate around the center of the circle defined by the maximum inserted voltage  $V_{pq}^{Max}$ .

Theoretically, SSSC operation in each of the four quadrants is possible, Theoretically, SSSC operation in each of the four quadrants is possible, but there are some limitations to the injected SSSC voltage due to operating constraints of practical power system. In capacitive mode, the injected SSSC voltage is made to lag the transmission line current by 90Deg; in this case, the SSSC operation is similar to the operation of a series capacitor with variable capacitance i.e.,  $V_{pq} = -jkX_c * I_{line}$ , where  $k$  is a variable. By this action, the total transmission line reactance is reduced while the voltage across the impedance is increased, leading to increase in the line currents and

transmitted power. This action is illustrated in Fig 3 It is also possible to reverse the injected SSSC voltage by 180 Deg i.e.,  $-V_{pq} = jkX_c * I_{line}$ , causing an increase in the transmission line reactance, which results in a decrease of the line current and transmitted power. While this equation for  $V_{pq}$  shows changes in the phasor magnitude and phase angle, it can be somewhat misleading, since it shows that the series injected voltage magnitude is directly proportional to the line current magnitude. In reality, this is not true; the inserted voltage magnitude is set by the SSSC control and is independent of the network impedance and, consequently, line current changes. In Fig. 3, it is assumed that the SSSC losses are zero and, therefore, the series injected voltage is in perfect quadrature with the line current, leading or lagging.

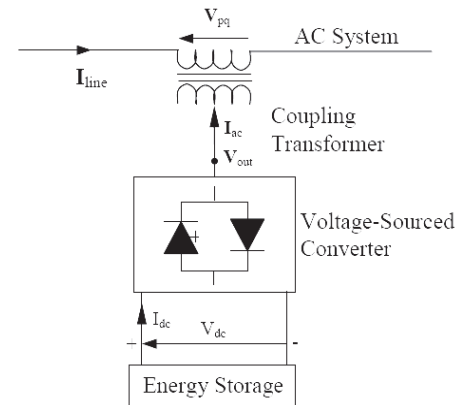


Fig 1. Functional model of SSSC

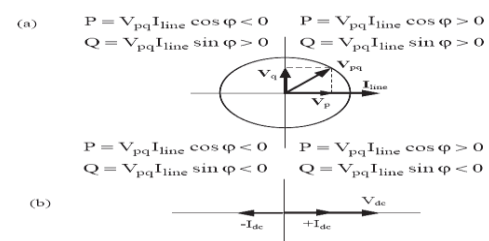


Fig 2. SSSC phasor diagram

The operating conditions that limit the SSSC operation from the power system point of view are also depicted in Fig. 3 The SSSC can increase as well as decrease the power flow in the transmission line by simply reversing the operation from capacitive to inductive mode. In the inductive mode, the series injected voltage is in

phase with the voltage drop developed across the line reactance; thus, the series compensation has the same effect as increasing the line reactance. If the series inserted voltage magnitude is larger than voltage drop across the uncompensated line, i.e.,  $V_{pq} \geq V_{line}$  the power flow will reverse. This fact can limit the SSSC operation to values of  $V_{pq} \leq V_{line}$  as in practice, it would be unlikely to use the SSSC for power reversal. Also, if the rating of the SSSC controller is high, it is possible to increase or decrease the receiving end voltage above or below the typical operating voltage range of 0.95 p.u. – 1.05 p.u., but with possible negative consequences for other system devices. The SSSC output current corresponds to the transmission line current, which is affected by power system impedance, loading and voltage profile, as well as by the actions of the SSSC. Thus, the relationship between the SSSC and the line current is complex. The fundamental component of the SSSC output voltage magnitude is, on the other hand, directly related to the dc voltage that is either constant or kept within certain limits, depending on the chosen design and control of the SSSC. The SSSC output voltage phase angle is correlated to the line current phase angle by plus or minus few degrees for example, to account for changes in the dc voltage. It has to be noted that the injected SSSC voltage  $V_{pq}$  is different from the SSSC output voltage  $V_{sssc}$ , due to the voltage drop or rise across the series transformer reactance, i.e.

$$V_{pq} = V_{sssc} - X_{tr} * I_{line} \quad (1)$$

Where the minus sign corresponds to capacitive operation, while the positive sign corresponds to inductive operation of the SSSC and  $X_{tr}$  stands for the series transformer reactance. This voltage difference between the injected and output SSSC voltage can be small in the case of small transmission line currents, but it can be significant in high loading conditions

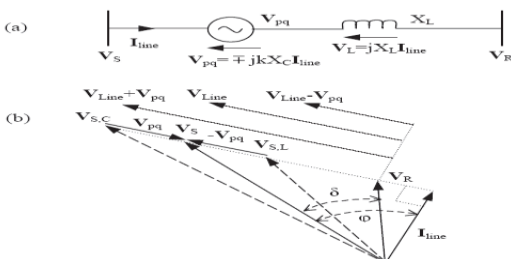


Fig 3 Series compensation by a SSSC

The active and reactive power exchanged between the SSSC and the transmission line can be calculated as follows:

$$P_{pq} = V_{pq} * I_{line} * \cos \phi \quad (2)$$

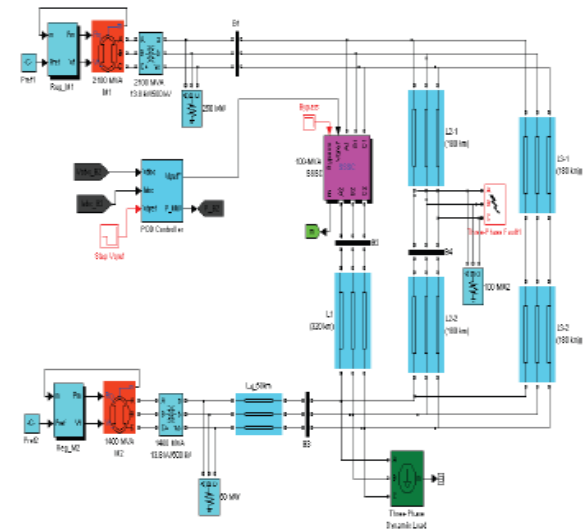
$$P_{pq} = V_{pq} * I_{line} * \sin \phi \quad (3)$$

Where  $\Phi$  represents the angle between the injected SSSC voltage and transmission line current. Inspection of the equations (2) and (3), considering that the angle between the SSSC output voltage and line current is approximately  $90^\circ$ , shows that the SSSC real power should be small compared to the reactive power. This is expected, since the real power going into the SSSC is used only to cover for the losses and charging of the dc capacitor, i.e.

$$P_{pq} = P_{dc} + P_{losses} \quad (4)$$

The losses in the SSSC circuit are due to the transformer windings and especially due to the switching of the GTO valves [9, 10].

### III. TEST SYSTEM DESCRIPTION



other one (M2) has a rating of 1400 MVA, representing 4 machines of 350 MVA. The load center of approximately 2200 MW is modeled using a dynamic load model where the active & reactive power absorbed by the load is a function of the system voltage. The generation substation M1 is connected to this load by two transmission lines L1, L2 and L3. L1 is 320-km long and L2 is split in two segments of 180 km in order to simulate a three-phase unbalanced fault (using a fault breaker) at the midpoint of the line.

The generation substation M2 is also connected to the load by a 50-km line (L4). When the SSSC is bypass, the power flow towards this major load is as follows: 664 MW flow on L1 (measured at bus B2), 563 MW flow on L2 (measured at B4) and 990 MW flow on L4 (measured at B3). The SSSC, located at bus B1, is in series with line L1

#### A. Phase Control Technique for SSSC:

**Phase Control Technique for SSSC** The modeled SSSC circuit with its two six-pulse VSCs and their series transformers is shown in Fig.5.2. The converters are connected in series to the transmission line through two banks of lossless three-phase single phase two-winding transformers with no saturation. The dc sides of the converters are connected in parallel and share the same dc bus. The GTO valves are switched at fundamental frequency, and the dc voltage varies according to the phase control technique used to control the output voltage. The SSSC switching is synchronized with respect to the transmission line current  $I_{line}$ , and its rms magnitude  $I_{line}$  is controlled by transiently changing the phase shift between this current and the SSSC output voltage  $V_{pq}$ . The change in the phase shift between the SSSC output voltage and the line current results in the change of the dc capacitor voltage  $V_{dc}$ , which ultimately changes the magnitude of the SSSC output voltage  $V_{SSSC}$  and the magnitude the transmission line current  $I_{line}$ . A block diagram of the SSSC controller is depicted in Fig.5 Fig 6Transformer configuration of 12-pulse SSSC.

The SSSC output voltage  $V_{SSSC}$  is controlled by a simple closed loop; the per unit value of the measured line current is compared with the line current per-unit order and the error of these two values is passed to the PI controller. The output of the PI controller is the angle  $\alpha$ , which is added to the synchronizing signal  $\Theta$  passed to the gate pulse generator by the current synchronization block. To

this signal  $\Theta + \alpha$ , an angle of  $-\pi/2$  or  $+\pi/2$  has to be added since the SSSC output voltage is lagging or leading the line current by  $90^\circ$  depending on the desired capacitive or inductive operation. The phase shift of the converter output voltage  $V_{SSSC}$  with the respect to the line current  $I_{line}$  will cause the flow of a small amount of real power from or into the converter, thereby causing a change in the dc capacitor voltage, and consequently causing a change in the converter output voltage magnitude [5,11].

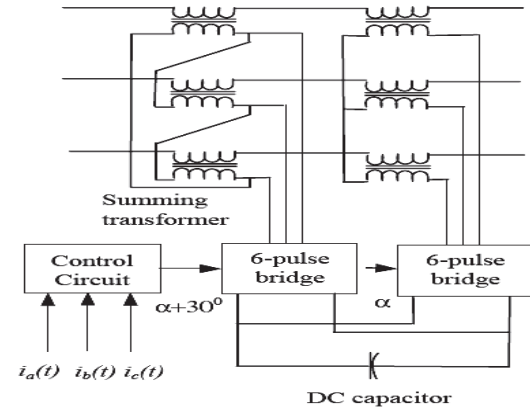


Fig 5. Transformer configuration of 12-pulse SSSC

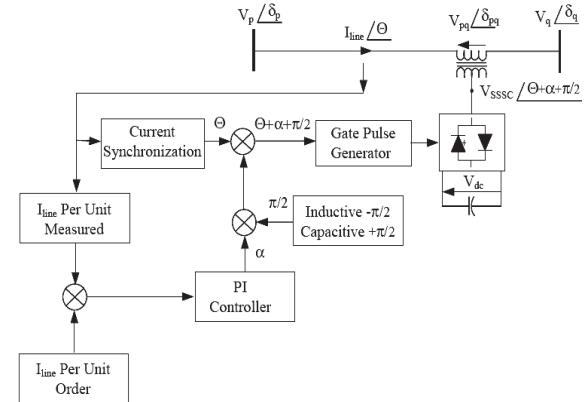


Fig 6. Functional control diagram for the phase-controlled SSSC

#### B. PWM Technique:

SSSC is modeled as a three-phase, PWM-controlled two-level VSC. The SSSC is modeled in detail, based on an ideal representation of the converter valves and diodes. RC parallel snubber circuits are used to reduce numerical oscillations due to switching, while a series inductance is employed at the converter output to smooth the output current. The series transformer is modeled as an ideal, three-phase, two-winding, Y-delta connected

© Malla Reddy Engineering College (Autonomous)



30/60]; In Fig. 13, the simulation result shows that the power oscillation on the L1 line following the three-phase fault. Moreover the performed simulation indicates that the SSSC compensator is a very effective tool to damp power oscillation.

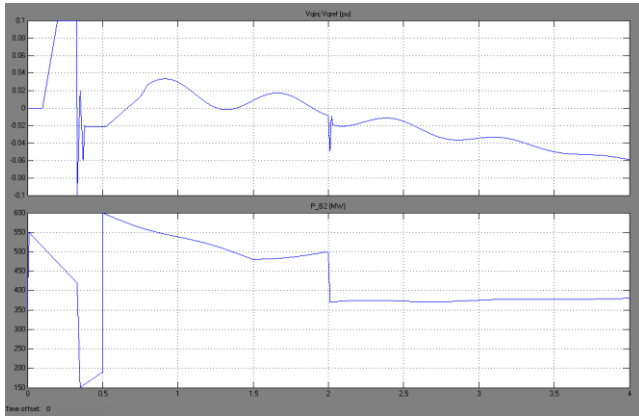


Fig.10 Simulation Result of System with SSSC under an unbalanced single line to ground fault for Voltage and Reactive Power

## V. CONCLUSION

This project analyzed the problem of controlling and modulating power flow in a transmission line using a Synchronous Static Series Compensator (SSSC). The studies, which include detailed techniques of twelve pulse and PWM controlled SSSC, are conducted and the control circuits are presented. The SSSC operating conditions and constraints are compared to the operating conditions of other FACTS devices, showing that the SSSC offers several advantages over others. However, at the present time the total cost of a SSSC installation is higher than the cost of other FACTS devices.

Comparisons of two implemented control strategies clearly show that the PWM based and phase controller have both disadvantages and advantages, which makes the design process somewhat complicated. The dc voltage pre-set value in PWM-based controllers has to be carefully selected. As the modulation ratio lies between zero and one, the dc voltage should not be lower than the maximum of the requested SSSC output phase voltage in order to obtain proper control. On the other hand, if the dc side voltage is too high, the rating of both the GTO valves and dc capacitor has to be increased, which means higher installation costs. Not only that, a higher dc side voltage means a lower amplitude modulation ratio, and the lower modulation ratio results in higher harmonic distortion. Phase control allows the dc voltage to change according to

the power system conditions, which is clearly advantageous, but it requires a more complicated controller and special and costly series transformers.

The simulation results validate that Voltage and Power Oscillation can be damped properly using of Synchronous Static Series Compensator (SSSC). By using the SSSC will get the simulation results under unbalanced three phase fault that is single line to ground fault and double line to ground fault.

## References

- [1] "Static Synchronous Compensator," CIGRE, Working group 14.19, 1998.
- [2] N. G. Hingorani and L. Gyugyi, Understanding FACTS, Concepts and Technology of Flexible AC Transmission Systems. Piscataway, NJ: IEEE Press, 2000.
- [3] R. Mohan and R. K. Varma, Thyristor-Based FACTS Controllers for Electrical Transmission Systems. Piscataway, NJ: IEEE Press, 2002.
- [4] L. Gyugyi, N. G. Hingorani, P. R. Nannery, and N. Tai "Advanced StaticVar Compensator Using Gate Turn- Off Thyristors for Utility Applications", CIGRE, 23- .203, August 26 September 1, 1990, France
- [5] J. Arrillaga, B. Barrett, N. A. Vovos "Thyristor Controlled Regulating Transformer for Variable Voltage Boosting", IEE Proceedings, No. 10, October 1976.
- [6] N. G. Hingorani, and L. Gyugyi "Understanding FACTS: Concepts and Technology of Flexible AC Transmission
- [7] L. Sunil Kumar and A. Ghosh, "Modeling and control design of a static synchronous series compensator," IEEE Trans. Power Del., vol. 14, no. 4, pp. 1448–1453, Oct. 1999.
- [8] X.-P. Zhang, "Advanced modeling of the multicontrolfunctional static synchronous series compensator (SSSC) in Newton power flow," IEEE Trans. Power Syst., vol. 18, no. 4, pp. 1410–1416, Nov. 2003.
- [9] R. Mihalic, and I. Papic "Mathematical Models and Simulation of a Static Synchronous Series Compensator", IEEE Power Tech'99 Conference, BPT99-315-42, August 1999. Budapest, Hungary.
- [10] K. K. Sen "SSSC – Static Synchronous Series Compensator: Theory, Modeling, And Applications", IEEE Transactions on Power Delivery, Vol. 13, No. 1, pp. 241–246, January 1998.
- [11] R. Mohan and R. K. Varma, Thyristor-Based FACTS Controllers for Electrical Transmission Systems. Piscataway, NJ: IEEE Press, 2002.



# FUZZY LOGIC CONTROLLER BASED BLDC MOTOR DRIVE WITH SOURCE SIDE POWER FACTOR CORRECTION

T. Sanjeeva Rao<sup>1</sup>, A. Suresh Kumar<sup>2</sup>, B. Sandeep<sup>3</sup>

<sup>1</sup>Associate Professor, EEE Department, MREC(A)

<sup>2</sup>Associate Professor & HOD of EEE Dept., Samskruthi Engineering College

<sup>3</sup> PG Student, EEE Department, MREC(A)

<sup>1</sup>sanjeev707@gmail.com, <sup>2</sup>sureshpck4u@gmail.com, <sup>3</sup>sandy061990@gmail.com

## Abstract:

*This paper shows another greenhouse gas bridgeless (BL) buck-boost convertor for brushless DC (BLDC) motor commute application in low-control applications. A Fuzzy principle execution in versatile pace management of BLDC motor is completed here. a technique of rate management of the BLDC motor by dominant the dc transport voltage of the voltage supply electrical converter (VSI) is employed with one voltage detector the controller is projected to trace mixtures of pace references and settles the yield speed within the inside of weight assortments. The BLDC features a few inclinations diverge from the opposite style of motors; but the nonlinearity of the BLDC motor commute traits, in light-weight of the approach that it's troublesome to handle by exploitation normal relative basic (PI) controller. To handle this principal issue, the Fuzzy principle management transforms into an acceptable management. to relinquish AN intrinsic greenhouse gas at provide air-con mains a convertor taking under consideration buck-help type is planned to figure in broken inductance current mode (DICM). The execution of the projected drive is imitated in MATLAB/Simulink surroundings Catchphrases Fuzzy principle (FL), Bridgeless (BL) buck-boost convertor, spasmodic inductance current mode (DICM), force element revision (PFC).*

## I.INTRODUCTION

ADEQUACY and cost range unit the numerous contemplations inside of the change of low-power motor drives spend significant time in family applications, as an illustration, fans, water pumps, blowers, blenders, so on [1], [2]. the job of the brushless power (BLDC) motor in these applications is popping twisted be uncommonly customary owing to segments of high profitability, high flux thickness per unit volume, low support needs, and low

electromagnetic-hindrance issues [1]. These BLDC motors don't appear to be limited to family applications, however rather these zone unit fitting for different applications, as a case, helpful equipment, transportation, HVAC, development administration, and different mechanical devices [2]–[4].

A BLDC motor has 3 stage windings on the mechanical gadget and enduring magnets on the rotor [5], [6]. The BLDC motor is generally alluded to as partner electronically commutated motor in lightweight of the very certainty that partner electronic substitution seeable of rotor position is utilized instead of a mechanical substitution that has blocks like starting and wear and tear of brushes and electric switch gathering [5], [6].

Also, for PFC gadget enlivened PMBLDCM drives, the extra cost and unpredictability of the PFC gadget don't appear to be legitimized, therefore, gadget topologies with characteristic piece of PFC range unit supported in these drives. In this manner, a DC-DC gadget topology is for the principal half supported amongst two or three available topologies [7-15] e.g., buck, bolster, buck-help, Cuk, SEPIC, letter converters with types of capacitive/inductive imperativeness trade. net result's expanded execution, as an illustration, adjustment of AC mains current sounds, decline of acoustic uproar and attraction pollution, slightest mixture of parts, updated viability, utilization of the aggregate information voltage differ so on.

This paper deals with a Cuk gadget as PFC AC-DC gadget to reinforce PMBLDCM driven air circulation and cooling framework. Since, the Cuk gadget with PFC gets central focuses like low current and voltage swell in yield, close shared characteristic power variable with basic administration and weakened size luring [7]. an express arrangement and execution appraisal of the anticipated PFC gadget for empowering PMBLDCM drive territory unit displayed for circulating air through and cooling system. The paper is sorted move into six essential segments, particularly presentation, operation and administration of



Cuk gadget supported PMBLDCM, their design, showing of the anticipated PMBLDCM drive, execution appraisal and conclusion.

Up to now, very eighth of the controllers range unit PI (Relative and critical) controllers in lightweight of the very certainty that they're clear and direct. The pace controllers region unit the quality PI controllers and current controllers zone unit the P controllers to perform transcendent commute. downlike Logic will be thought-about as partner wildcat hypothesis association multi-respected system of thinking, connected math, and imagine attention to breed the human approach inside of the amusement organize of different issues by exploitation partner anticipated that reasoning would relate entirely unexpected learning sets and to pick determinations. its been portray that fluffy controllers region unit less difficult to plant parameter changes than commonplace PI or controllers and have higher clamor dismissal points of confinement  
This paper demonstrates a BL buck-boost converter-fed BLDC motor commute with variable dc be a piece of voltage of VSI for expanded power quality at air-con mains with decreased fragments and unmatched

## II. PRINCIPLE OF BLDC MOTOR

BLDC motor contains of the never-ending magnet rotor related a damage mechanical gadget. The brushless motors territory unit controlled using a 3 stage electrical converter. The motor obliges a rotor position detecting component for beginning and for giving genuine pay game plan to show on the power devices inside of the electrical converter augmentation. In lightweight of the rotor position, the power devices range unit commutated sequentially every sixty degrees. The electronic pay kills the issues joined with the brush and in this way the electric switch organize, exceptionally starting and annihilating of the electric switch brush strategy, on these lines, making a BLDC motor a great deal of unpleasant stood out from a dc motor. Fig.1 shows the mechanical gadget of the BLDC motor and fig.2 shows rotor magnet arranges.



Fig.1 BLDC motor stator construction



Fig.2 BLDC motor Rotor construction

The brush less dc motor include 4 basic parts Power gadget, immutable magnet brushless DC Motor (BLDCM), sensors and administration estimation. The power gadget changes power from the supply to the BLDCM that so changes over electrical imperativeness to mechanical essentialness. one in every of the remarkable highlights of the brush less dc motor is that the rotor position sensors, seeable of the rotor position and request signs which can be a torsion charge, voltage summon, rate request and so forth; the administration counts center the dish sign to every semiconductor inside of the power converter.

The structure of the administration estimations chooses the sort of the brush less dc motor of that there territory unit 2 standard classifications voltage supply principally {based} drives and current supply based drives. every voltage supply and current supply based for the most part drive utilized for unending magnet brushless DC machine. the back voltage wave type of the motor is incontestable inside of the fig. 3. Be that in light of the fact that it could, machine with a non bending back voltage brings concerning diminishment inside of the electrical converter size and diminishes incidents for a comparative impact level..

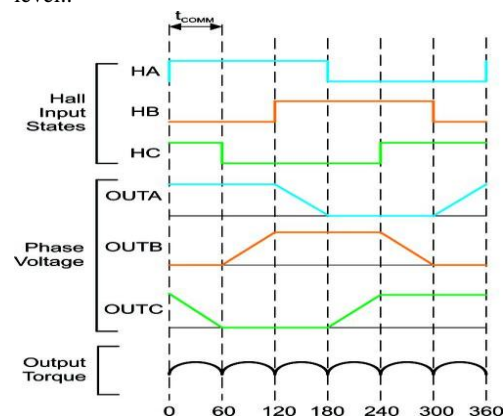


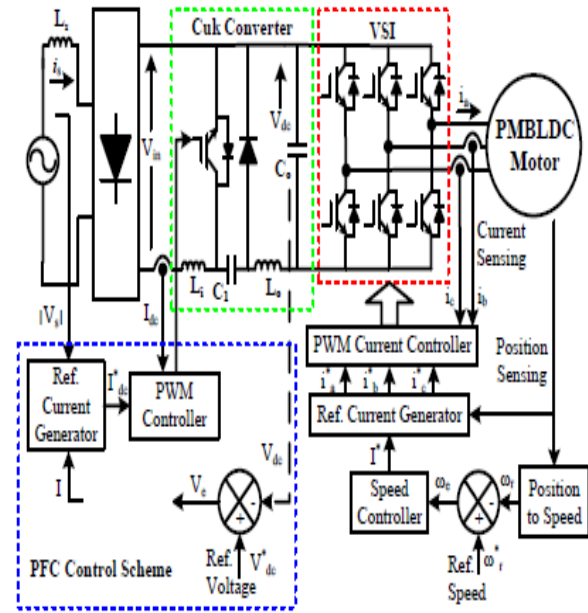
Fig.3 Hall signals & Stator voltages

## III. Operation and Control of Cuk Converter fed PMBLDCM

Figure a couple of demonstrates the anticipated topology of Cuk PFC gadget bolstered PMBLDCM drive with administration subject for the rate administration besides as PFC with DC join voltage administration. For the rate

administration of the PMBLDCM, a relative basic (PI) controller [3] is utilized to drive a constant torsion mechanical gadget in cooling framework. The rotor position of PMBLDCM is seen exploitation Hall way sensors and recover to rush flag, that is contrasted and a reference speed.

The pace slip sign is experienced a rate controller to give the torsion proportionate that is recover to current proportional sign. This sign is expanded with an elongated unit sample to a limited extent with high level part of motor's back voltage to urge reference streams of the engine. The reference engine streams range unit contrasted and the apparent engine ebbs and flows. These present lapses territory unit intensified and opened up signs zone unit then contrasted with triangular radio radiation with get the PWM beats for killing on/the VSI switches. The administration of PMBLDCM needs rotor-position information exclusively at the substitution focuses, e.g., each 60° electrical inside of the three-stages [1-5], in this manner, generally simple managementler is required for recompense and current control.



**Fig 4:** Extension circuit diagram DC-DC converter & BLDC motor with fuzzy controller

The Cuk PFC gadget topology contains a regular DBR encouraged from single-stage AC mains took after by the Cuk DC-DC gadget, partner yield swell channel and a three-stage VSI to encourage the PMBLDC engine. The DC-DC gadget gives a controlled DC voltage from uncontrolled DC yield of DBR, while prevailing the office issue (PF) through high recurrence change of the PFC switch. The controlled yield DC voltage from the DC-DC gadget is situated by its obligation size connection (D). The change recurrence (fs) is situated by the change gadget utilized, in operation voltage, force level and change misfortunes of the gadget. amid this work, a gathering of protected entryway bipolar transistors (IGBTs) territory unit utilized on the grounds that the change gadgets inside of the PFC switch besides as in VSI span, as an aftereffect of IGBTs will work in wide change recurrence shift to shape ideal harmony between attractive, size of channel components and change misfortunes.

The utilization of current multiplier variable methodology with normal current administration topic in nonstop conductivity mode (CCM) of the PFC gadget makes this topology partner practical plausibility. The administration circle used to execute PFC activity includes external voltage circle and inward current circle. For administration activity, the DC join voltage is seen and contrasted and the set reference voltage at DC join. The lapse voltage is experienced a voltage PI controller to give the tweaking

current sign. This sign is expanded with a unit sample of data AC voltage that is contrasted and DC current saw once the DBR. Lobby sway voltage and current sensors region unit utilized for voltage and current detecting. This present blunder is enhanced and opened up sign is then contrasted with serration radio radiation with get the PWM beats for killing on/the DC-DC gadget switch. the whole administration system comprises of decision of sensors, style of administration algorithmic tenet and PWM controller for the drive.

#### IV. PROPOSED FUZZY LOGIC CONTROLLER

The administration structure is in lightweight of typical rationale. Everglade State controller is partner one sort non straight controller and modified. this sort of the administration moving closer the human feeling that delivers the occupation of the affirmation, powerlessness, looseness and fluff inside of the option making methodology, makes sense of an approach to give relate especially tasteful execution, while not the prerequisite of an express numerical model of the structure, just by intertwining the experts' learning into the downlike. Fig about six shows the Everglade State controller piece characterize. This downlike clarification administration system is seeable of the MAMDHANI downlike model. This system contains of 4 guideline parts. to begin with, by using the information enlistment limits, inputs region unit Fuzzified then seeable of standard bases and thusly the

inferencing structure, yields territory unit conveyed all in all the downlike yields zone unit Defuzzified and that they range unit associated with the guideline administration system..

Blunder of inputs from their references and lapse deviations in any measure range unit picked as MATLAB. The yield of fluffy controller is that the value that should be side to the past yield to give new reference yield.

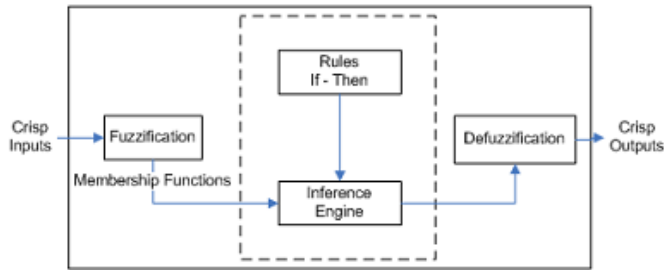


Fig 5: Block Diagram of fuzzy logic controller

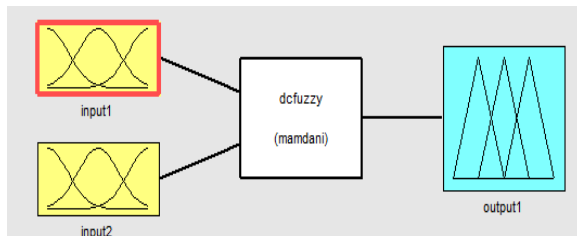


Fig 6: selection of input and output variables

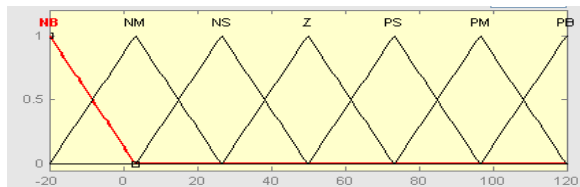


Fig 7: Input1 membership function

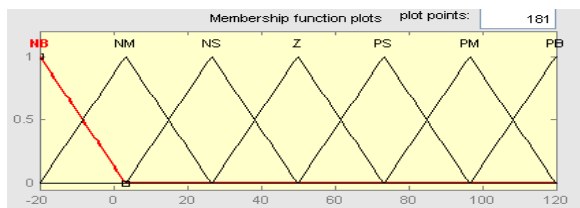


Fig 8: Input 2 membership function

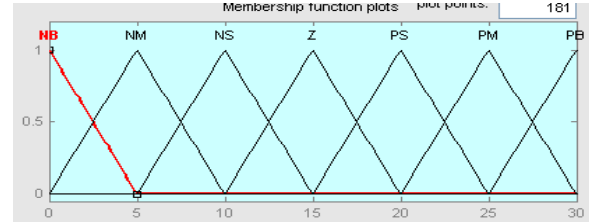


Fig 9: Output membership function

## V.SIMULATION RESULTS

Simulation is performed using MATLAB/SIMULINK software. Simulink library files include inbuilt models of many electrical and electronics components and devices such as diodes, MOSFETS, capacitors, inductors, motors, power supplies and so on. The circuit components are connected as per design without error, parameters of all components are configured as per requirement and simulation is performed.

### SIMULATION PARAMETERS

Source  $V_s = 230V$ ,  $50Hz$  1-Ph

$L_f = 150mH$

Diodes,

$R_{On} = 0.001\Omega$

Forward voltage drop,  $V_{fd} = 0.8V$

Snubber parameters

$R_s = 500\Omega$

$C_s = 259nF$

DC link capacitance,  $C_{dc} = 100mF$

### BLDC MOTOR SPECIFICATIONS

Three phase-415 Volt, 0.5HP

Stator ph resistance  $R_s = 2.8\Omega$

Inductances,  $L_d = 8.5mH$ ,  $L_q = 8.5mH$

Flux linkage  $= 0.175Vs$

Voltage constant  $= 126.966 V_{peak} L-L/Krpm$

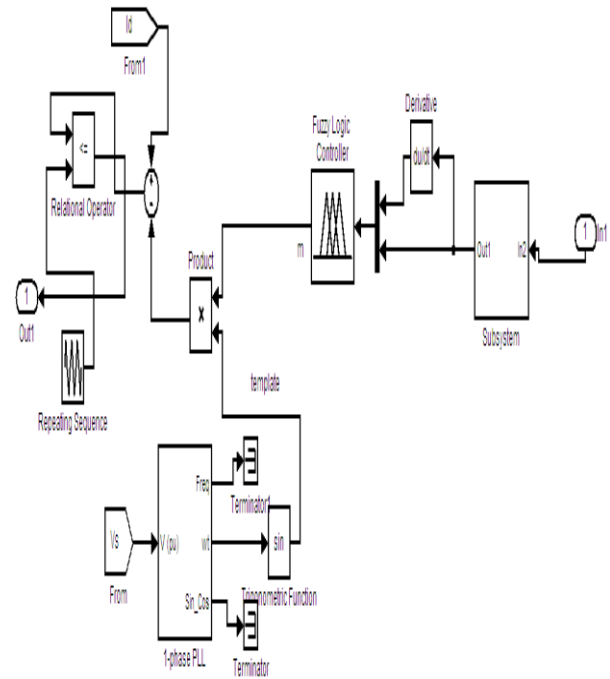
Torque constant  $= 1.05 Nm/A_{peak}$

Poles = 4

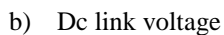
Friction factor  $= 0.001Nms$

Inertia  $= 0.0008 J(kg m^2)$

### SIMULATION CIRCUIT



a) Supply voltage



For rated supply voltage and rated DC link voltage, inverter develops 220V ac output. This 3-Ph Rated voltage are applied to the stator of BLDC motor, which develops speed close to rated value

Speed  $N=120 \cdot f/p$  rpm

$F$  = frequency of voltages developed by inverter

P=No. of poles in BLDC motor

Here,

F=50Hz

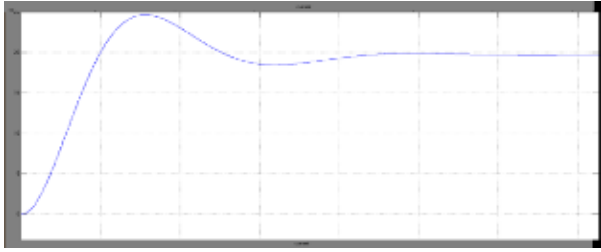
P=4

So,  $N=120*50/4$

=1500rpm

Now settling time is only 0.002sec

d) Torque developed by motor



Input Power factor  
=0.986

## VI. CONCLUSION

In this paper to overcome the maximum overshoot, minimize settling time suitable control technique has been used in the controller architecture. Fuzzy logic controlled model of the BLDC motor was implemented. The purpose is to achieve accurate trajectory control of the speed of permanent magnet brushless DC Motor, specially when the motor and load parameters are unknown. Simulation study is carried out in MATLAB/SIMULINK software

## REFERENCES

- [1] C. L. Xia, Permanent Magnet Brushless DC Motor Drives and Controls. Hoboken, NJ, USA: Wiley, 2012.
- [2] J. Moreno, M. E. Ortuzar, and J. W. Dixon, "Energy-management system for a hybrid electric vehicle, using ultracapacitors and neural networks," IEEE Trans. Ind. Electron., vol. 53, no. 2, pp. 614–623, Apr. 2006.
- [3] Y. Chen, C. Chiu, Y. Jhang, Z. Tang, and R. Liang, "A driver for the singlephase brushless dc fan motor with hybrid winding structure," IEEE Trans. Ind. Electron., vol. 60, no. 10, pp. 4369–4375, Oct. 2013.
- [4] X. Huang, A. Goodman, C. Gerada, Y. Fang, and Q. Lu, "A single sided# matrix converter drive for a brushless dc motor in aerospace applications," IEEE Trans. Ind. Electron., vol. 59, no. 9, pp. 3542–3552, Sep. 2012.

# Simulation of Standalone Inverter for Distribution Generation System with Fuzzy Control

T. Sanjeeva Rao<sup>1</sup>, A. Nageswara Rao<sup>2</sup>, V. Santhosh<sup>3</sup>

<sup>1</sup>Associate Professor, EEE Department, MREC (A)

<sup>2</sup>Associate Professor, EEE Department, Kodad Institute of Technology

<sup>3</sup>PG Student, EEE Department, MREC (A)

<sup>1</sup>sanjeev707@gmail.com

<sup>2</sup>arla@gmail.com

<sup>3</sup>santhosh4placements@gmail.com

**Abstract**— This paper presents a control strategy of three-phase voltage source inverter for a distributed generation system in a standalone operation. The proposed adaptive voltage control technique combines an adaption control term and a state feedback control term. The algorithm is easy to implement, but it is very robust to system uncertainties and sudden load disturbances. The proposed control strategy guarantees excellent voltage regulation performance (i.e., fast transient response, zero steady-state error, and low THD) under various types of loads such as balanced load, unbalanced load, and nonlinear load. The simulation results are implemented using MATLAB/Simulink.

**Index Terms**—Adaptive control, distributed generation (DG) system (DGS), load current observer, stand-alone, three-phase inverter, voltage control.

## I. INTRODUCTION

In recent years, eco-friendly distributed generation systems (DGS) such as wind turbines, solar cells, and fuel cells are dramatically growing because they can fulfill the increasing demand of electric power due to the rapid growth of the economy and strict environmental regulations regarding greenhouse gas emissions.

Recently, an adaptive control method has been widely considered in the standalone DGS or UPS voltage control. In and the precise voltage tracking is achieved under distorting loads by using the adaptive control for the output voltage based on the ideas of dissipativity. In these papers, the uncertainties in the system parameters are addressed through the adaptation, and the stability of the system is guaranteed even under system parameters variations. However, the major drawback of these techniques is the computation complexity. In order to reduce this complexity, a certain predefined value for the parameters is required. In an adaptive output voltage controller based on the resonant harmonic filters, which measures the capacitor current and the load currents in the same sensor, is proposed in order to compensate for the unbalance and harmonic distortion on the load. The adaptation law is also included to cope with the uncertainties in the system parameters. However, the information about output voltage THD is not presented so it is not easy to evaluate the quality of the controllers. In an adaptive control method based on the proportional derivative control technique is presented for a pulse width modulation (PWM) inverter operation in an islanded DGS. This paper can guarantee good voltage regulation under various operating conditions such as sudden load changes, unbalanced load, and nonlinear load. However, it is not an easy task to choose the appropriate control gains according to the design procedure mentioned in the paper.

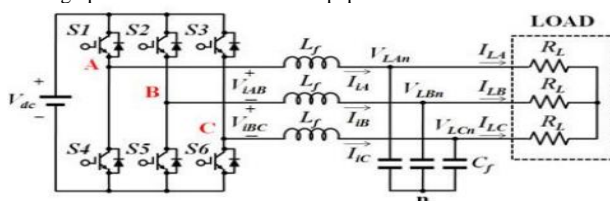


Fig. 1. Schematic diagram of a three-phase dc to ac inverter with an LC filter in a standalone application

## II. STAND-ALONE OPERATION FOR INVERTER

Inverters convert power from DC to AC while rectifiers convert it from AC to DC. Many inverters are bi-directional, i.e. they are able to operate in both inverting and rectifying modes. In many stand-alone PV installations, alternating current is needed to operate 230V (or 110V), 50 Hz (or 60 Hz) appliances. Generally stand-alone inverters operate at 12, 24, 48, 96, 120, or 240V DC depending upon the power level. Ideally, an inverter for a stand-alone PV system should have the following features:

- Sinusoidal output voltage.
- Voltage and frequency within the allowable limits.
- Cable to handle large variation in input voltage.
- Output voltage regulation.
- High efficiency at light loads.
- Less harmonic generation by the inverter to avoid damage to electronic appliances like television, additional losses, and heating of appliances.
- Photovoltaic inverters must be able to withstand overloading for short term to take care of higher starting currents from pumps, refrigerators, etc.
- Adequate protection arrangement for over/under-voltage and frequency, short circuit etc.
- Surge capacity.
- Low idling and no load losses.
- Low battery voltage disconnect
- Low audio and radio frequency (RF) noise.

Several different semiconductor devices such as metal oxide semiconductor field effect transistor (MOSFETs) and insulated gate bipolar transistors (IGBTs) are used in the power stage of inverters. Typically MOSFETs are used in units up to 5 kVA and 96V DC. They have the advantage of low switching losses at higher frequencies. Because the on-state voltage drop is 2V DC, IGBTs are generally used only above 96V DC systems.

Voltage source inverters are usually used in stand-alone applications. They can be single phase or three phase and there are three switching techniques commonly used: square wave, quasi-square wave, and pulse width modulation. Square wave or modified square-wave inverters can supply power tools, resistive heaters, or incandescent lights, which do not require a high quality sine wave for reliable and efficient operation. However, many household appliances require low distortion sinusoidal waveforms.

The use of true sine-wave inverters is recommended for remote area power systems. Pulse width modulated (PWM) switching is generally used for obtaining sinusoidal output from the inverters.

A general layout of a single-phase system, both half bridge and full bridge. Single phase, half bridge is with two switches,  $S_1$  and  $S_2$ , the capacitors  $C_1$  and  $C_2$  are connected in series across the DC source. The junction between the capacitors is at the mid-potential. Voltage across each capacitor is  $V_{dc}/2$ . Switches  $S_1$  and  $S_2$  can be switched on/off periodically to produce AC voltage. Filter ( $L_f$  and  $C_f$ ) is used to reduce high-switch frequency components and to produce sinusoidal output from the inverter. The output of inverter is connected to load

through a transformer. Figure shows the similar arrangement for full-bridge configuration with four switches. For the same input source voltage, the full-bridge output is twice and the switches carry less current for the same load power. The power circuit of a three phase four-wire inverter is shown in Fig. 2. The output of the inverter is connected to load via three phase transformer (delta/Y). The star point of the transformer secondary gives the neutral connection. Three phase or single phase can be connected to this system. Alternatively, a center tap DC source can be used to supply the converter and the mid-point can be used as the neutral.

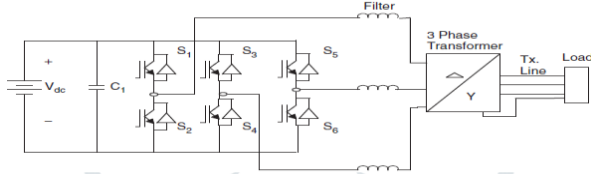


Fig.2. a stand-alone three-phase four wire inverter

Figure 2 shows the inverter efficiency for a typical inverter used in remote area power systems. It is important to consider that the system load is typically well below the nominal inverter capacity  $P_{nom}$ , which results in low conversion efficiencies at loads below 10% of the rated inverter output power. Optimum overall system operation is achieved if the total energy dissipated in the inverter is minimized. The high conversion efficiency at low power levels of recently developed inverters for grid-connected PV systems shows that there is a significant potential for further improvements in efficiency.

#### Adaptive Control

An adaptive control can be thought of as a feedback law which attempts to reshape the controller by observing its performance. This type of control is usually proposed to compensate for some kind of system uncertainty such as unknown parameters or disturbances. In this chapter, we review a brief history of adaptive control and describe its different types such as direct and indirect adaptive control techniques. Since there is no need to estimate the system parameters in the direct adaptive control method which means lower mathematical computations, this control method is used for the VSI with uncertain parameters. Next the adaptive control for nonlinear systems with linearly parameterized uncertainty is investigated. Finally, the systematic adaptive design procedure is applied to the VSI with two unknown parameters.

#### III. SYSTEM DESCRIPTION

The circuit model in Fig.1 uses the following quantities. The inverter output lines to neutral voltage and phase current vectors are given by  $V_i = [v_{iA} \ v_{iB} \ v_{iC}]^T$  and  $I_i = [i_{iA} \ i_{iB} \ i_{iC}]^T$ , respectively. In addition, the load lines to neutral voltage and phase current are represented by the vectors  $V_L = [v_{LA} \ v_{LB} \ v_{LC}]^T$  and  $I_L = [i_{LA} \ i_{LB} \ i_{LC}]^T$ , respectively. Assume that the three-phase voltages and currents used in Fig.2 are balanced. By applying Kirchhoff's current law and Kirchhoff's voltage law at the LC output filter, the following voltage and current equations can be derived:

#### Load Current Observer Design

The proposed adaptive controller needs load current information. Using the current sensors to measure the load currents ( $I_L$ ) makes the system more expensive and less reliable. In this section, a linear optimal load current observer is designed to accurately estimate load current information that can heavily affect the controller performance, a fourth-order dynamic model can be obtained as follows:

$$\dot{\hat{x}} = A\hat{x} + Bu \quad (1)$$

Where

$$x = \begin{bmatrix} \hat{i}_{Ld} \\ \hat{i}_{Lq} \\ v_{Ld} \\ v_{Lq} \end{bmatrix} \quad A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1/C_f & 0 & 0 & \omega \\ 0 & -1/C_f & -\omega & 0 \end{bmatrix} \quad (2)$$



Fig.3. Load current observer

Then, the load current observer model can be represented as

$$\dot{\hat{x}} = A\hat{x} + M y - M C \hat{x} + B u \quad (3)$$

$$y = C x \quad (4)$$

$$\hat{I}_{Ldq} = \begin{bmatrix} \hat{i}_{Ld} \\ \hat{i}_{Lq} \end{bmatrix} = C_T \hat{x} \quad (5)$$

where  $\hat{i}_{Ld}$  and  $\hat{i}_{Lq}$  are estimates of  $i_{Ld}$  and  $i_{Lq}$ , respectively,  $M \in \mathbb{R}^{4 \times 2}$  is an observer gain matrix, and

$$\hat{x} = \begin{bmatrix} \hat{i}_{Ld} \\ \hat{i}_{Lq} \\ \hat{v}_{Ld} \\ \hat{v}_{Lq} \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad C_T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (6)$$

Next, the error dynamics of the load current observer can be obtained as follows:

$$\dot{\tilde{x}} = (A - M C) \tilde{x} \quad (7)$$

where  $\tilde{x} = x - \hat{x}$ .

**Theorem 2:** Consider the following algebraic Riccati equation:

$$AP + PA^T - PC^T R^{-1} CP + Q = 0 \quad (8)$$

where  $Q \in \mathbb{R}^{4 \times 4}$  is a symmetric positive semi-definite matrix,  $R \in \mathbb{R}^{2 \times 2}$  is a symmetric positive definite matrix, and

$P \in \mathbb{R}^{4 \times 4}$  is a solution matrix. Also, assume that the load current observer gain matrix  $M$  is given by

$$M = PC^T R^{-1} \quad (9)$$

Then, the estimation error converges exponentially to zero.

**Proof:** Let us define the Lyapunov function as  $V_o(\tilde{x}) = \tilde{x}^T X \tilde{x}$ , where  $X = P^{-1}$ . Its time derivative along the error dynamics (3.29) is given by

$$\begin{aligned} \dot{V}_o(\tilde{x}) &= \frac{d}{dt} \tilde{x}^T X \tilde{x} = 2\tilde{x}^T (XA - XPC^T R^{-1} C) \tilde{x} \\ &= \tilde{x}^T X (AP + PA^T - 2PC^T R^{-1} CP) X \tilde{x} \\ &\leq -\tilde{x}^T X Q X \tilde{x} \end{aligned} \quad (10)$$

This implies that  $\tilde{x}$  is exponentially stable.

#### IV. FUZZY CONTROLLER

Figure 4 shows the internal structure of the control circuit. The control scheme consists of Fuzzy controller, limiter, and three phase sine wave generator for reference current generation and generation of switching signals. The peak value of reference currents is estimated by regulating the DC link voltage. The actual capacitor voltage is compared with a set reference value. The error signal is then processed through a Fuzzy controller, which contributes to zero steady error in tracking the reference current signal.

A fuzzy controller converts a linguistic control strategy into an automatic control strategy, and fuzzy rules are constructed by expert experience or knowledge database. Firstly, input voltage  $V_{dc}$  and the input reference voltage  $V_{dc-ref}$  have been placed of the angular velocity to be the input variables of the fuzzy logic controller. Then the output variable of the fuzzy logic controller is presented by the control Current  $I_{max}$ . To convert these numerical variables into linguistic variables, the following seven fuzzy levels or sets are chosen as: NB (negative big), NM (negative medium), NS (negative small), ZE (zero), PS (positive small), PM (positive medium), and PB (positive big) as shown in Fig. 4.

The fuzzy controller is characterized as follows:

- 1) Seven fuzzy sets for each input and output;
- 2) Fuzzification using continuous universe of dis-course;
- 3) Implication using Mamdani's 'min' operator;
- 4) De-fuzzification using the 'centroid' method.

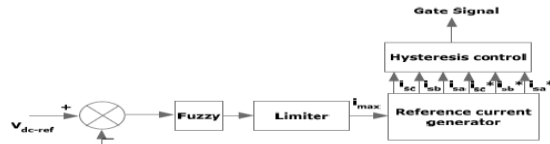


Fig.4. Conventional fuzzy controller

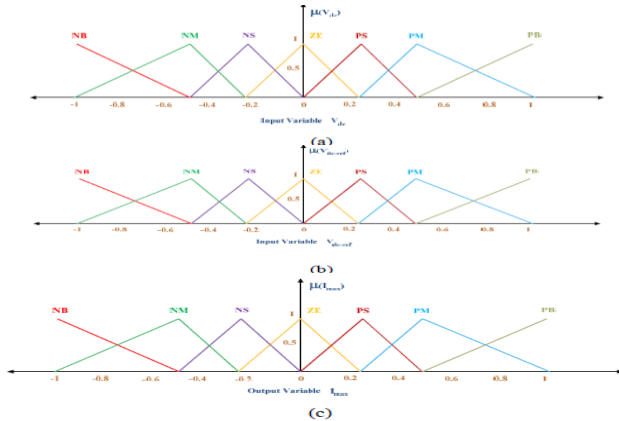


Fig.5. (a) Input  $V_{dc}$  normalized membership function; (b) Input  $V_{dc-ref}$  Normalized Membership Function; (c) Output  $I_{max}$  Normalized Membership Function.

**Fuzzification:** the process of converting a numerical variable (real number) convert to a linguistic variable (fuzzy number) is called fuzzification.

**De-fuzzification:** the rules of FLC generate required output in a linguistic variable (Fuzzy Number), according to real world requirements, linguistic variables have to be transformed to crisp output (Real number).

**Database:** the Database stores the definition of the membership Function required by fuzzifier and defuzzifier.

**Rule Base:** the elements of this rule base table are determined based on the theory that in the transient state, large errors need coarse control, which requires coarse in-put/output variables; in the steady state, small errors need fine control, which requires fine input/output variables. Based on this the elements of the rule table are obtained as shown in Table 1, with ' $V_{dc}$ ' and ' $V_{dc-ref}$ ' as inputs.

Table 1. Rule base.

$V_{dc-ref}$ $V_{dc}$	NB	NM	NS	Z	PS	PM	PB
NB	NB	NB	NB	NB	NM	NS	Z
NM	NB	NB	NB	NM	NS	Z	PS
NS	NB	NB	NM	NS	Z	PS	PM
Z	NB	NM	NS	Z	PS	PM	PB
PS	NM	NS	Z	PS	PM	PB	PB
PM	NS	Z	PS	PM	PB	PB	PB
PB	Z	PS	PM	PB	PB	PB	PB

## V. SIMULATION RESULTS

The simulation analysis of the three-phase inverter comprises of the DG'S is divided into six parts: an energy source, an ac-dc power converter (wind turbines) or a dc-dc boost converter (solar cells or fuel cells), a three-phase dc-ac inverter, an LC output filter, an isolation transformer, and a local load and the simulation circuit A renewable energy source and an ac-dc power converter or a dc-dc boost converter can be replaced by a stiff dc voltage source ( $V_{dc}$ ) because this paper focuses on designing a robust adaptive voltage controller under various types of loads such as balanced load, unbalanced load, and nonlinear load. Also, this representation can be acceptable because the front converter (i.e., an ac-dc power converter or a dc-dc boost converter) can rapidly recover the reduced dc-link voltage when a heavy load is suddenly applied.

The DG energy sources usually work together with energy storage devices (e.g., batteries, flywheels, etc.) in order to back up the DS systems during the transient, and increase the power quality and

reliability. Furthermore, the isolation transformer is not used to reduce cost and volume assuming that the customers need a low voltage ac source (below 600 V) which the DGSs using renewable energy sources can generate without the help of the transformer.

*Case 1) 280 vdc for balanced resistive load (transient behavior—0% to 100%).*

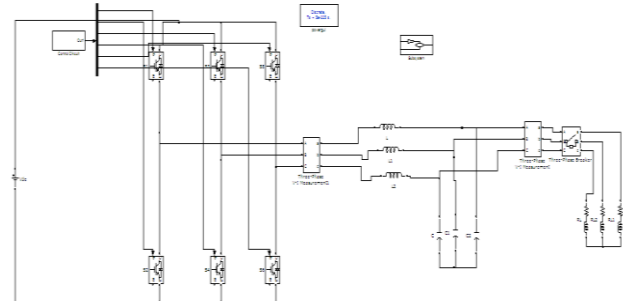


Fig. 6. Matlab/simulink diagram of a three-phase inverter with an LC output filter for stand-alone DGSs.

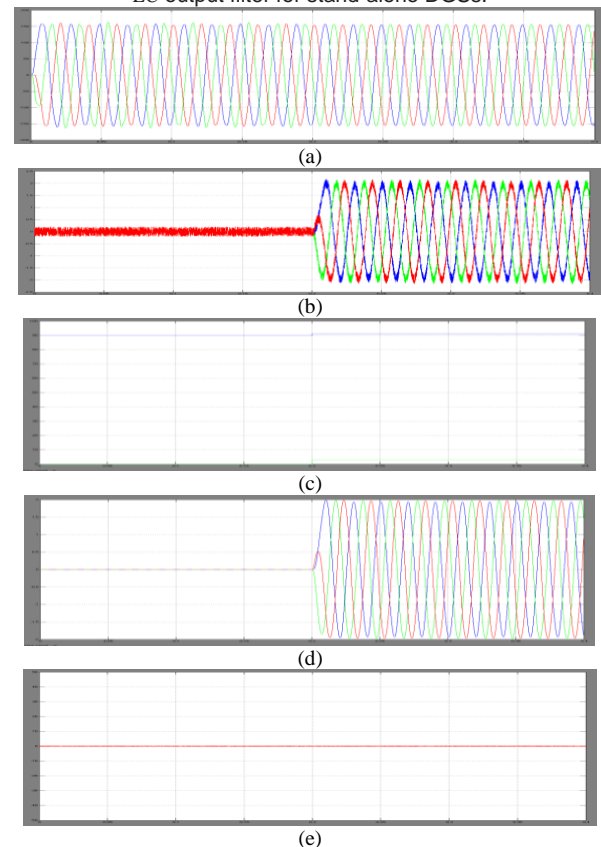


Fig. 7. Simulation results of the proposed control scheme under balanced resistive load (0%-100%).(a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities  $lidq$ .(c)  $VLdq$  in the synchronously rotating  $dq$  reference frame.(d). the load currents ( $ILdq$ ).(e). load current error.

*Case 2) 280 vdc for balanced resistive load (transient behavior—100% to 0%);*

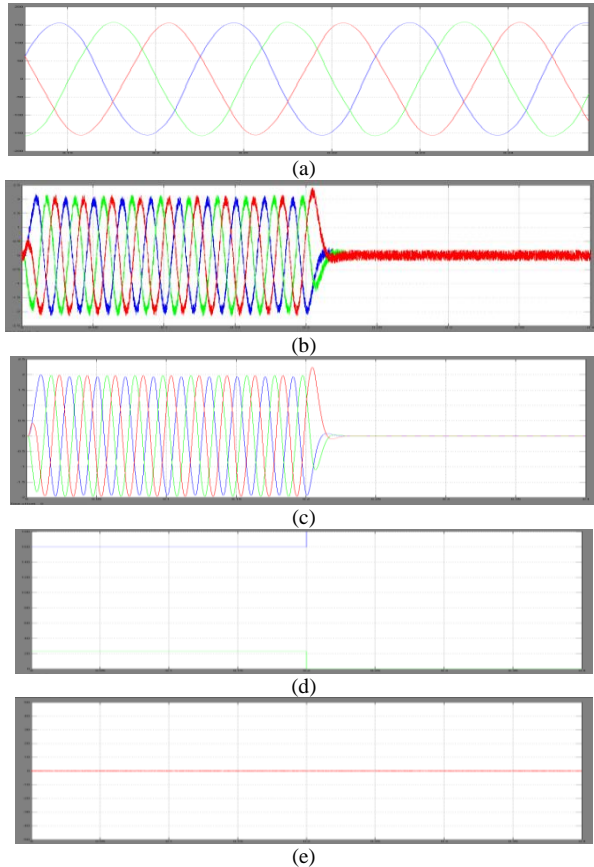


Fig.8. Simulation results of the proposed control scheme under for a (balanced resistive load: 100% to 0%). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities  $I_{idq}$ .(c). the load currents ( $I_{Ldq}$ ). (d)  $V_{Ldq}$  in the synchronously rotating dq reference frame (e). load current error.

Case 3) 280 vdc for unbalanced resistive load (phase C opened);

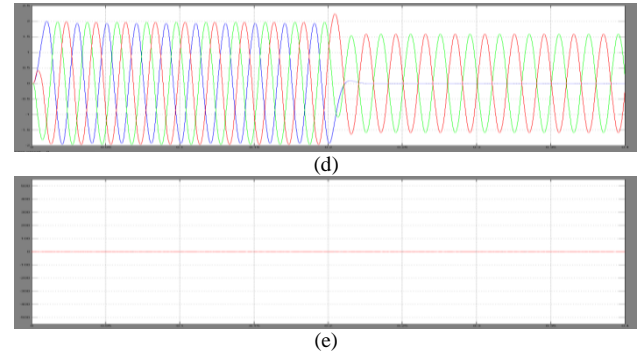
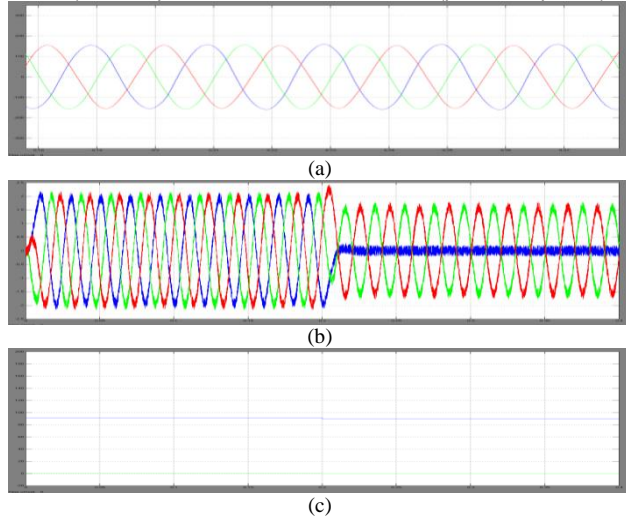


Fig. 9. Simulation results of the proposed control scheme under Case 3 for a 200-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities  $I_{idq}$ .(c)  $V_{Ldq}$  in the synchronously rotating dq reference frame.(d). The load currents ( $I_{Ldq}$ ).(e). load current error.

Case 4) 280 vdc for nonlinear load (a three-phase diode rectifier).

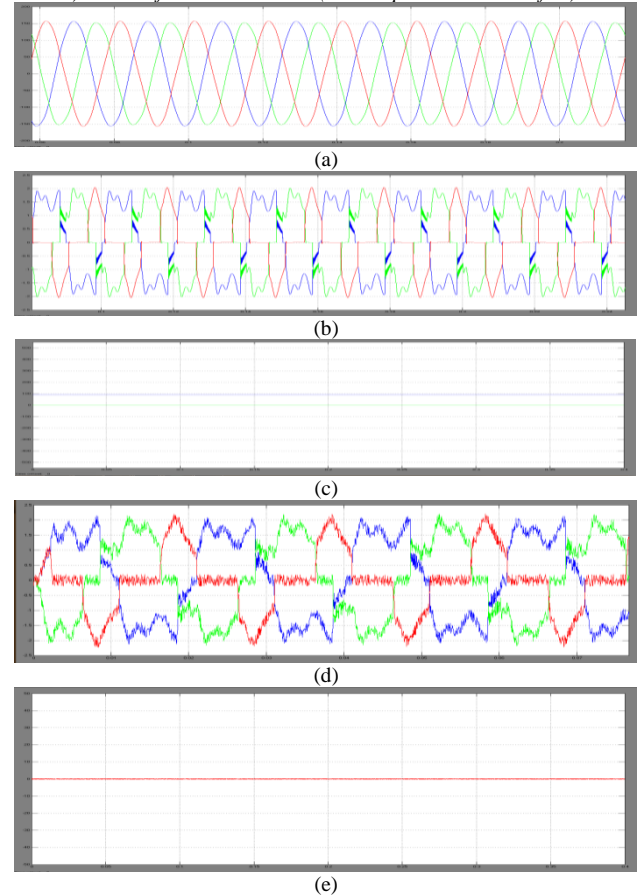


Fig. 10. Simulation results of the proposed control a 200-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities  $I_{idq}$ .(c)  $V_{Ldq}$  in the synchronously rotating dq reference frame.(d). The load currents ( $I_{Ldq}$ ).(e). load current error.

Case 5) 600 vdc for balanced resistive load (transient behavior—0% to 100%).

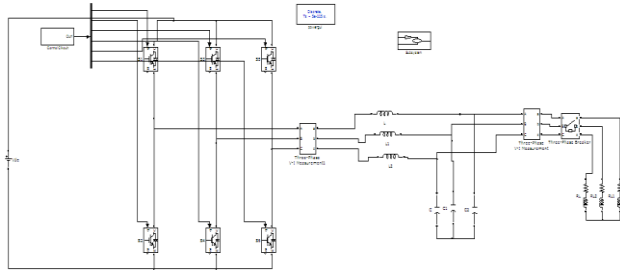


Fig. 11. Matlab/simulink diagram of a three-phase inverter with an LC output filter for stand-alone DGSSs.

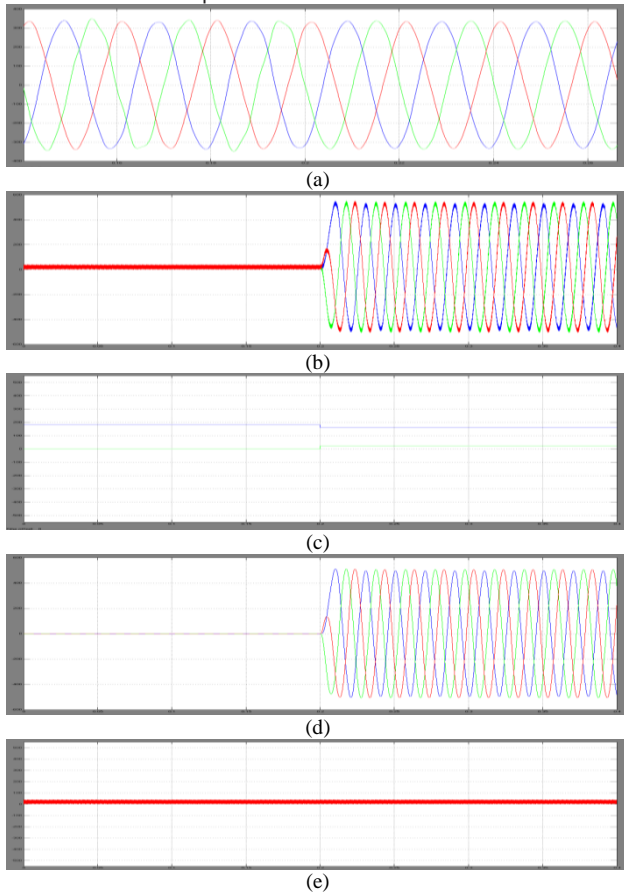


Fig. 12. Simulation results of the proposed control a 450-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities Iidq.(c) VLdq in the synchronously rotating dq reference frame.(d). The load currents (ILdq).(e). load current error.  
Case 6) 600 vdc for balanced resistive load (transient behavior—100% to 0%);

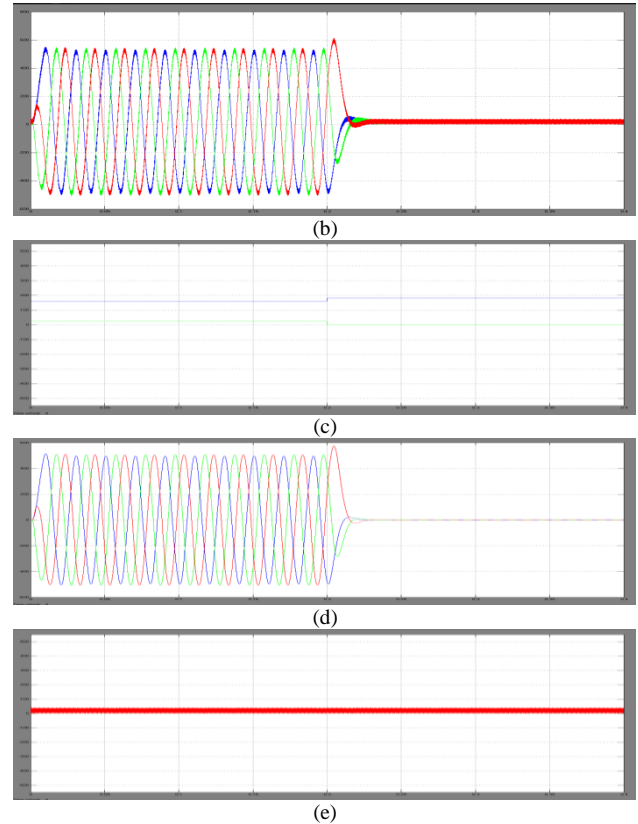
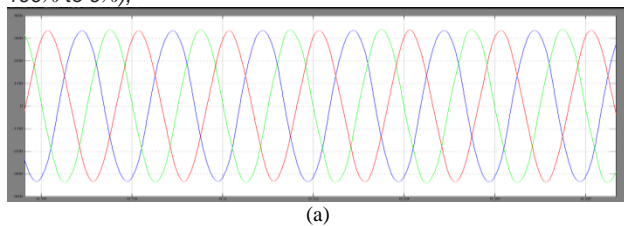
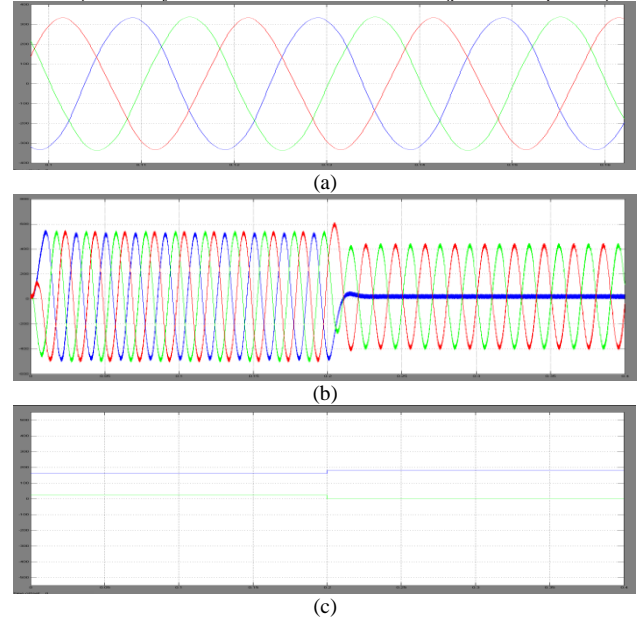


Fig.13.Simulation results of the proposed control a 450-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities Iidq.(c) VLdq in the synchronously rotating dq reference frame.(d). The load currents (ILdq).(e). load current error.

Case 7) 600 vdc for unbalanced resistive load (phase C opened);



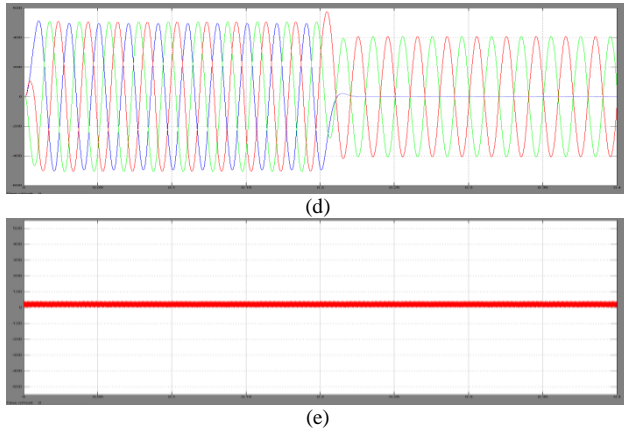


Fig.14. Simulation results of the proposed control a 450-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities Ildq.(c) VLdq in the synchronously rotating dq reference frame.(d). The load currents (ILdq).(e). load current error.

Case 8) 600vdc for nonlinear load (a three-phase diode rectifier).

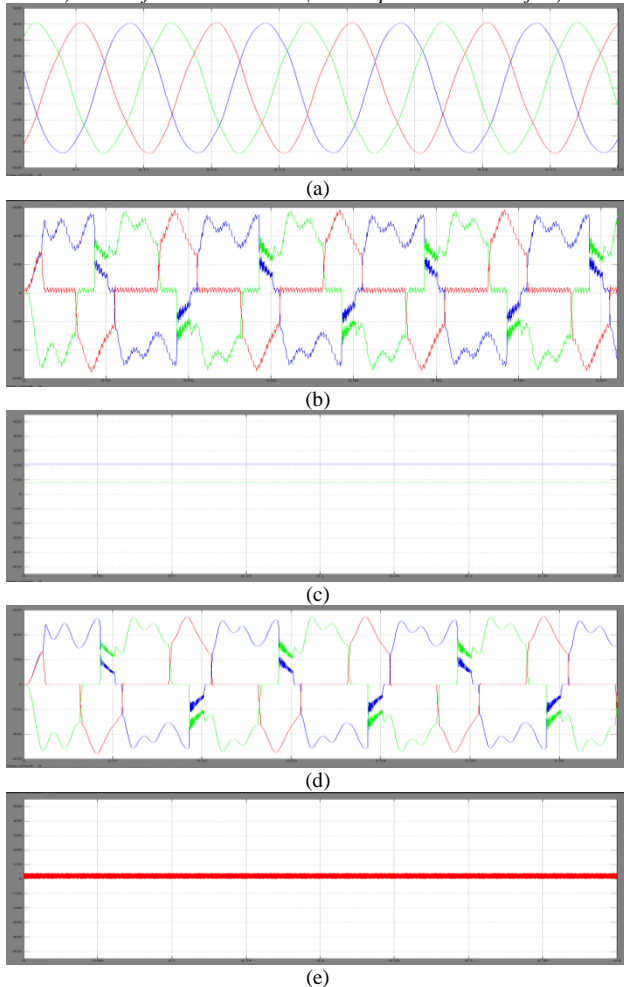


Fig.15.Simulation results of the proposed control a 450-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities Ildq.(c) VLdq in the synchronously rotating dq reference frame.(d). The load currents (ILdq).(e).Load current error.

Case 9) 600vdc for nonlinear load (a three-phase diode rectifier).With fuzzy

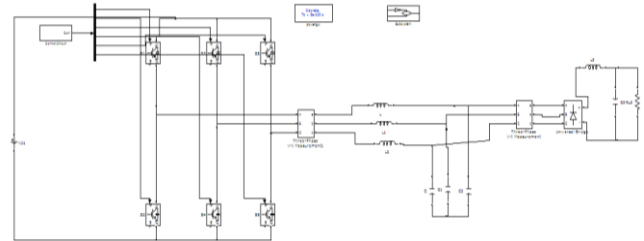


Fig.16. Matlab/simulink diagram of a three-phase inverter with an LC output filter for stand-alone DGSSs with fuzzy logic control.

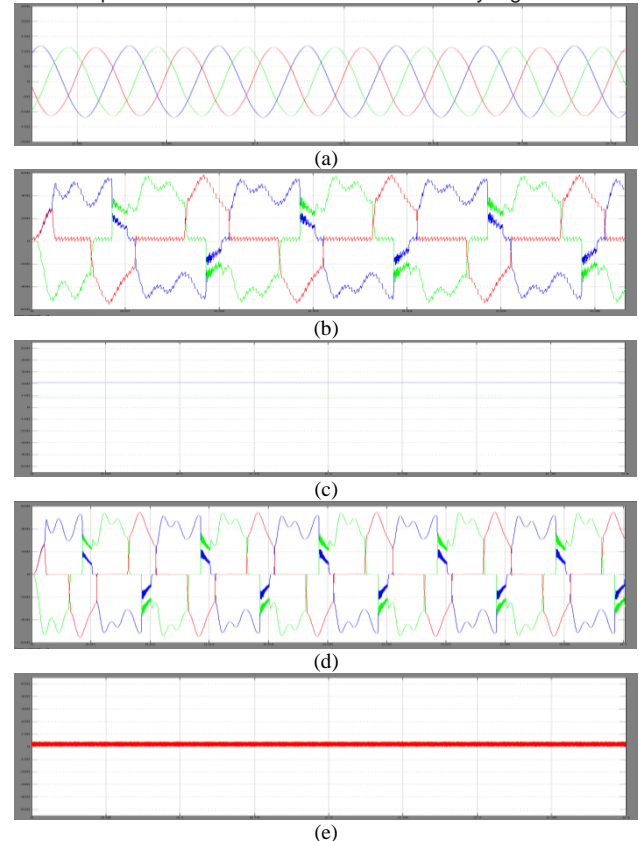


Fig. 17. Simulation results of the proposed control a 450-kVA unit (unbalanced resistive load: phase C opened). (a).load output voltages (VL).(b). Measured with sensors and then transformed to the quantities Ildq.(c) VLdq in the synchronously rotating dq reference frame.(d). The load currents (ILdq).(e). load current error.

## VI. CONCLUSION

A high-performance control strategy for three phase inverter for stand-alone grid connected systems was described that is able to compensate for distortion and unbalance generated by nonlinear and unbalanced loads or by nonlinearity in the inverter. The principle can also be used in any static power supply system like power supplies and ground power units. The voltage -controlled UPS system adapts the inverter reference to achieve sinusoidal balanced output voltages at nonlinear and nonsymmetrical loads by comparing the vectorial output voltages with a circular reference in fixed points on the fundamental period. The implementation of this control strategy on different systems is eased by the low dependency of the output filter parameters.

## REFERENCES

[1] F. Blaabjerg, R. Teodorescu, M. Liserre, and A. V. Timbus, "Overview of control and grid synchronization for distributed power



generation systems," *IEEE Trans. Ind. Electron.*, vol. 53, no. 5, pp. 1398–1409, Oct. 2006.

[2] A. Yazdani, "Control of an islanded distributed energy resource unit with load compensating feed -forward," in *Proc. IEEE PES Gen. Meeting*, Aug. 2008, pp. 1-7.

[3] K. L. Nguyen, D. J. Won, S. J. Ahn, and I. Y. Chung, "Power sharing method for a grid connected microgrid with multiple distributed generators," *J. Elect. Eng. Technol.*, vol. 7, no. 4, pp. 459–467, Jul. 2012.

[4] H. C. Seo and C. H. Kim, "Analysis of stability of PV system using the eigen value according to the frequency variation and requirements of frequency protection," *J. Elect. Eng. Technol.*, vol. 7, no. 4, pp. 480–485, Jul. 2012.

[5] I. S. Bae and J. O. Kim, "Phasor discrete particle swarm optimization algorithm to configure micro -grids," *J. Elect. Eng. Technol.*, vol. 7, no. 1, pp. 9–16, Jan. 2012.

[6] H. Karimi, E. J. Davison, and R. Iravani, "Multivariable servomechanism controller for autonomous operation of a distributed generation unit: Design and performance evaluation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 853–865, May 2010.

[7] U. Borup, P. N. Enjeti, and F. Blaabjerg, "A new space -vector-based control method for UPS systems powering nonlinear and unbalanced loads," *IEEE Trans. Ind. Appl.*, vol. 37, no. 6, pp. 1864–1870, Nov./Dec. 2001.

[8] T. S. Lee, S. J. Chiang, and J. M. Chang, "  $H_{\infty}$  loop-shaping controller designs for the single-phase UPS inverters," *IEEE Trans. Power Electron.*, vol. 16, no. 4, pp. 473–481, Jul. 2001.

[9] G. Escobar, A. M. Stankovic, and P. Mattavelli, "An adaptive controller in a stationary reference frame for d-statcom in unbalanced operation," *IEEE Trans. Ind. Electron.*, vol. 51, no. 2, pp. 401–409, Apr. 2004.

[10] P. Mattavelli, G. Escobar, and A.M. Stankovic, "Dissipativity-based adaptive and robust control of UPS," *IEEE Trans. Ind. Electron.*, vol. 48, no. 2, pp. 334–343, Apr. 2001.

[11] R. Escobar, A. A. Valdez, J. Leyva -Ramos, and P. Mattavelli, "Repetitive based controller for a UPS inverter to compensate unbalance and harmonic distortion," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, pp. 504–510, Feb. 2007.



# Optimal location of FACTS devices considering cost of Installation and System Loadability by using Particle Swarm Optimization Technique

Mrs.M.Lakshmiswarupa<sup>1</sup>, Ms.K.Chetaswi<sup>2</sup>, Mr.P.Ganesh<sup>3</sup>

<sup>1</sup>Associate Professor, MREC (A)

<sup>2</sup>Assistant Professor, MREC (A)

<sup>3</sup>Assistant Professor, MREC (A)

<sup>1</sup>swarupamalladi@gmail.com,

<sup>2</sup>k.chetaswi@gmail.com,

<sup>3</sup>ganeshp5891@gmail.com

**Abstract**— The goal of optimization is to perform a best utilization of existing transmission lines. In this connection, the STATCOM and IPFC devices are optimally placed in order to minimize the power losses while considering voltage constraints. This also improves the power transfer capability and voltage profile of the system. Simulations are performed on an IEEE 30-bussystem and results are presented.

**Keywords**— STATCOM, IPFC, PSO, Power losses, Load flow studies, IEEE 30 bus.

## I. INTRODUCTION

As power transfer grows, the power system can become increasingly more difficult to operate, and the system becomes more insecure with unscheduled power flows and higher losses. The rapid development of self-commutated semiconductor devices, have made it possible to design power electronic equipments. These equipments are well known as Flexible AC Transmission systems (FACTS) devices [1-2].

The FACTS controllers, such as Static VAR Compensator (STATCOM), Thyristor Controlled Series Compensator (TCSC) and Static Synchronous Compensator (STATCOM) can increase or reduce reactive power according to the demand of reactive power in the network to improve the stability, reduces system loss and also improves the loadability of the system.

The Static VAR Compensator (STATCOM) is used to inject or absorb reactive power where it is connected [2], [4]. Interline power flow controller (IPFC) is a new concept of FACTS controller for series compensation with the unique capability of power flow management among multi-line of a substation [5],[8]. In this paper, multi-type FACTS devices like STATCOM and IPFC are optimally located on a 14-bus using Partial Swarm Optimization Technique.

### A. STATCOM

The STATCOM compensator is a shunt type of FACTS devices, which absorbs or injects reactive power at which

it is connected. The size of the STATCOM is depends on the rating of current and reactive power injected into the bus.

### B. INTERLINE POWER FLOW CONTROLLER (IPFC):

The IPFC is a series-series type of FACTS device, which is used to exchange reactive powers in between two or more transmission lines those are connected to the same bus.

## II. FLOWCHART

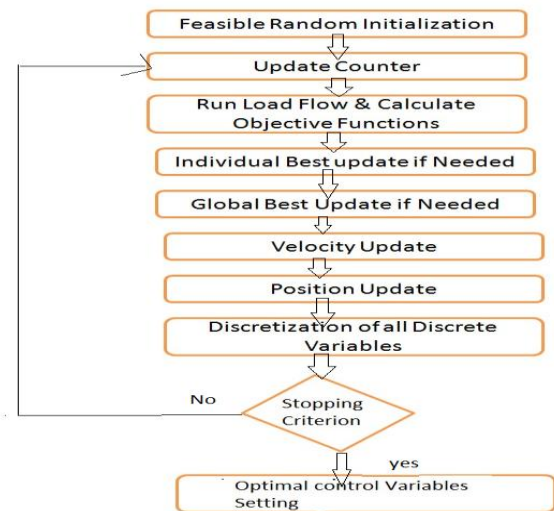


Fig 1. Flowchart of PSO

## III. PARTICLE SWARM OPTIMIZATION

PSO is an evolutionary computation technique developed by Eberhart and Kennedy in 1995, and was inspired by the social behavior of bird flocking and fish schooling [12], [13],[14]. PSO has its roots in artificial life and social psychology as well as in engineering and computer science. It utilizes a population of individuals, called particles, which fly through the problem hyperspace with some given initial velocities. In each iteration, the velocities of the

particles are stochastically 2 adjusted considering the historical best position of the particles and their neighborhood best position; where these positions are determined according to some predefined fitness function [9], [11]. Then, the movement of each particle naturally evolves to an optimal or near-optimal solution.

The name of "swarm" comes from the irregular movements of the particles in the problem space, more similar to a swarm of mosquitoes rather than flock of birds or school of fish [11]. In a real-number space, the position of each particle In a real-number space, the position of each particle is given by the vector  $x_i$  belongs to Real number. At iteration  $t$ , the particle position vector  $X_i(t)$ , given in (1), is determined by the previous position vector  $X_i(t-1)$  and its movement given by the velocity applied to the particle  $V_i(t)$  [12].

$$\bar{x}_i(t) = \bar{x}_i(t-1) + \bar{v}_i(t) \quad (1)$$

At each iteration, the velocity of a particle is determined by both the individual and group experience:

$$\bar{v}_i(t) = w_i \cdot \bar{v}_i(t-1) + c_1 \cdot rand_1 \cdot (\bar{p}_i - \bar{x}_i(t-1)) + \dots \quad (2)$$

$$c_2 \cdot rand_2 \cdot (\bar{p}_g - \bar{x}_i(t-1))$$

where:

- $w_i$  is a positive number between 0 and 1.
- $c_1, c_2$  are two positive numbers called the cognitive and social acceleration constants.
- $rand_1, rand_2$  are two random numbers with uniform distribution in the range of [0, 1].
- $p_i$  is the best position found by the particle  $i$  so far.
- $p_g$  is the global best position found by any particle in the swarm

The velocity update equation as given by (2) has three different components [13]:

i) The first component is sometimes referred to as "inertia", "momentum" or "habit". It models the tendency of the particle to continue in the same direction it has been traveling.

ii) The second component is a linear attraction towards the best position ever found by the given particle (pbest). This component is variously referred to as "memory", "self-knowledge", "nostalgia" or "remembrance".

ii). The third component of the velocity update equation is a linear attraction towards the best position found by any particle (gbest). This component is variously referred to as "cooperation", "social knowledge", "group knowledge" or "shared information".

The maximum allowable velocity for the particles is controlled by the parameter Vmax. If Vmax is too high,

then particles tend to move beyond a good solution; on the other hand, if Vmax is small, then particles can be trapped in local minima.

Optimization of some real world problems requires to be solved in an integer-number space. In this case, the PSO has also proven to be effective in this kind of optimization problem; the performance of the PSO could be better than typical techniques such as branch and bound by giving better stability performance and higher success rates [14].

The application of PSO to integer optimization problems is known as the integer PSO. Integer PSO is based on the same principles as the real valued PSO as described before, i.e. computations are carried out in the same way but real values are approximated to the closest integer. The approximation does not seem to significantly affect the performance of the integer PSO [14].

- (1) Define control variables (vg1, vg2, vg5, vg8, vg11, vg13, T1, T2, T3, T4, QC3, QC10 and QC24) within their permissible range, define population size, no of iteration (=200), assume suitable values of PSO parameters, input the data of 30 bus test system
- (2) Take iter=0
- (3) Randomly generate the population of particles and their velocities
- (4) For each particle run NR load flow to find out losses.
- (5) Calculate the fitness function of each particle using equ. (2)
- (6) Find out "personal best (Pbest)" of all particles and "global best(Gbest)" particle from their fitnesses
- (7) Iter=iter+1
- (8) Calculate the velocity of each particle using equ. (3) and adjust it if its limit gets violated
- (9) Calculate the new position of each particle using equ. (4)
- (10) For each particle run NR load flow to find out losses.
- (11) Calculate the fitness function of each particle using equ. (2)
- (12) For each particle if current fitness(P) is better than Pbest then Pbest=P
- (13) Set best of Pbest as Gbest
- (14) Go to step no. 7, until max. no of iterations is completed.
- (15) Coordinate of Gbest particle gives optimized values of control variables and its fitness gives minimized value of losses.

$$\text{Minimize IC} = C \times S \times 1000 \quad (3)$$

where IC is the optimal installation cost of FACTS devices in US\$ and  $C$  is the cost of installation of FACTS devices in US\$/KVAR.

$$C = 0.0003S^2 - 0.2691S + 188.22 \quad (4)$$

where  $S$  is the operating range of the FACTS devices in MVAR

$$S = |Q_2| - |Q_1| \quad (5)$$

where  $Q_2$  is the reactive power flow in the line after installing FACTS device in MVAR and  $Q_1$  is the reactive power flow in the line before installing FACTS device in MVAR.

The cost is optimized with the following constraints.

*Line flow and bus voltage constraints*

$$J = \prod_{\text{LINE}} \text{OVL}_{\text{LINE}} \times \prod_{\text{BUS}} \text{VS}_{\text{BUS}} \quad (6)$$

#### IV. SIMULATION RESULTS

In this work, the optimal locations of FACTS devices are found to minimize the cost of installation of FACTS devices and to improve system loadability, for single- or multi-type FACTS devices using PSO technique.

TABLE 1: COMPARISON OF LOSSES OBTAINED BEFORE AND AFTER OPTIMIZATION

S.no.	Power	Losses before optimization	Losses after optimization
1	Active power losses ( $P_{\text{loss}}$ ) MW	15.73	14.308 ( $f_{x_{\min}}$ )
2	Reactive power losses ( $Q_{\text{loss}}$ ) MVAR	77.56	54.10

TABLE 2: VALUES OF CONTROL VARIABLES AFTER OPTIMIZATION

Bus	Control variables	Optimized values ( $X_{\min}$ )
1	Vg1 (pu)	1.082
2	Vg2	1.043
5	Vg5	1.044
8	Vg8	1.041
11	Vg11	1.090
13	Vg13	1.056
6-9	T1	0.969
6-10	T2	1.015
4-12	T3	1.033
27-28	T4	0.982
3	QC3 (MVAR)	10.778

10	QC10	7.549
24	QC24	11.750

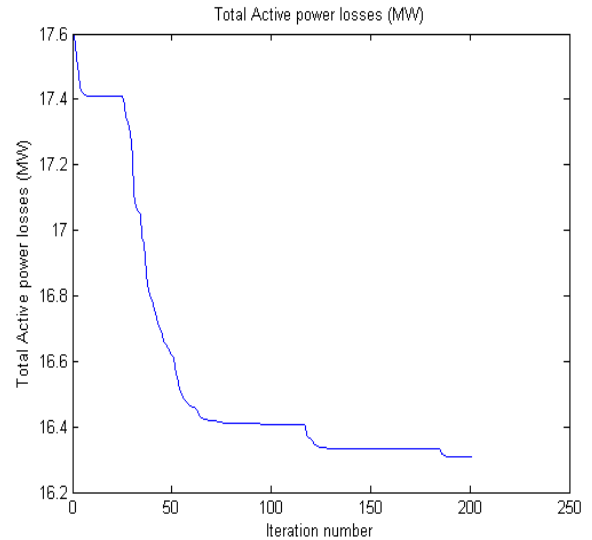


Fig 2: convergence characteristics of PSO

TABLE 3: SELECTED PARAMETERS OF PSO

1	Population size	50
2	Acceleration constant ( $C_1, C_2$ )	1.4 and 1.4
3	Constriction factor	0.729
4	Max. and Min. inertia weights	0.9 and 0.4
5	Max. and Min. velocity of particles	0.003 and -0.003
6	Convergence criterion	200 iterations

#### V. CONCLUSION

In this paper, we have proposed a PSO algorithm to place a combination of both STATCOM and IPFC devices. The future scope of this paper is a complete cost benefit analysis has to be carried out to justify the economic viability of the STATCOM and IPFC using different combination of optimization techniques.

#### References

- [1] S.Gerbex, R.cherkaoui, and A.J.Germond, "Optimal Allocation of FACTS Devices by Using Multi-Objective Optimal Power Flow and Genetic Algorithms". IEEE trans.power system, vol.16, pp.537-544, August 2001
- [2] L.J.Cai, I.Erlich, G.Stamtsis "Optimal Choice and Allocation of FACTS Devices in Deregulated Electricity Market using Genetic Algorithms" IEEE transactions on 2004.
- [3] H.R.Bahae, M.Jannati, B.Vahidi, S.H.Hosseinnian, H.Rastager: Improvement of voltage stability and reduce power system losses by optimal GA-base allocation of Multi-Type FACTS devices. The 11th international IEEE conference on May 2008,OPTIM 2008



- [4] K.Sundareswaran, P.Bharathram, M.siddharth, Vaishavi.G,Nitin Anand Srivastava, Harish Sharma: Voltage Profile enhancement through optimal placement of FACTS devices using Queen Bee Assisted GA. The 3rd international conference on power systems, Kharagapur, December 2009.
- [5] S Teerathana, A. Yokoyama, "An Optimal Power Flow Control Method of Power System using Interline Power Flow Controller (IPFC)", IEEE Transactions on Power Delivery, Vol. 23, pp. 343-346, Aug. 2004.
- [6] P. Subburaj, N. Sudha, K. Rajeswari, K. Ramar, and L. Ganesan, "Optimum Reactive Power Dispatch Using Genetic Algorithm," Academic Open Internet Journal, Vol.21, 2007, p6
- [7] B.Geethalakshmi, T. Hajmunisa and P. Dhananjayan, "Dynamic characteristics analysis of SSSC based on 48 pulse inverter", The 8th International Power Engineering Conference (IPEC) pp. 550-554, May 2007.
- [8] A.V.Naresh Babu, S.Sivanagaraju, Ch.Padmanabharaju and T.Ramana International Journal of Electrical and Electronics Engineering 4:7 2010.
- [9] R. Eberhart, and J. Kennedy, "A new optimizer using particle swarm theory," in Proc. 6th Int. Symp. Micro Machine and Human Science (MHS '95), 1995, pp. 39-43.
- [10] Y. Shuyuan, M. Wang, and L. Jiao; "A quantum particle swarm optimization," Proc. of the Con. on Evolutionary Computation (CEC2004), 2004, pp. 320-324.
- [11] J. Kennedy, and R. C. Eberhart, "Swarm intelligence," Morgan Kaufmann, San Francisco, 2001.
- [12] J. Kennedy, "The particle swarm: social adaptation of knowledge," in Proc. IEEE Int. Conf. Evolutionary Computation, 1997, pp. 303-308.
- [13] D.W. Boeringer, and D.H.Werner, "Particle swarm optimization versus genetic algorithms for phased array synthesis," IEEE Trans. on antennas and propagation, vol. 52, no. 3, pp. 771-779, Mar. 2004.
- [14] E.C. Laskari, K.E. Parsopoulos, and M.N. Vrahatis, "Particle swarm optimization for integer programming," Proc. of the 2002 Congress on Evolutionary Computation (CEC '02), vol. 2, 2002. pp. 1582-1587.



# FUZZY LOGIC BASED MAXIMUM TORQUE CONTROL FOR A SINGLE-PHASE INDUCTION MOTOR

M.Lakshmiswarupa <sup>(1)</sup>, Bibi Mariyam <sup>(2)</sup>, V.vasavi<sup>(3)</sup>

<sup>1</sup>Associate Professor , MREC (A)

<sup>2</sup>Assistant Professor, MREC (A)

<sup>3</sup>P.G Student, MREC (A)

<sup>3</sup>vasavibaby9@gmail.com

**Abstract:** In the present paper, a *fuzzy logic* based maximum torque control is proposed for a single-phase induction motor (SPIM) with a controlled capacitor. The objective of the control algorithm is to find the capacitor valued related to the maximum torque for any rotor speed value during the starting period. The *fuzzy logic*-based method has the advantage that adaptively changes the *step* size of the control variable and this manner the fast convergence is achieved. The complete control *system* have been developed and validated by simulation results.

**Keywords:** Single phase induction motors, maximum torque control, *fuzzy logic* optimization.

## I. INTRODUCTION

The single-phase induction motor (SPIM) is widely **used** in industrial applications such as compressors, pumps, air condensers and households. The most common type of SPIM **uses** two capacitors. The larger capacitor is only used in the starting period to increase the starting torque while the one of smaller value (running capacitor) is permanently connected in series with the starting winding to improve the motor efficiency. A centrifugal switch is used to disconnect the running capacitor when the motor run up to *speed*. The use of two

capacitors, together with the centrifugal switch, increases the overall *cost* of the system.

The centrifugal switch and the starting capacitor can be eliminated using as witched capacitor with electronic switch to improve the performance of the SPIM and reduce the *cost* of the system [1]-[Z].U sing this approach the effective capacitance of the running capacitor can be adjusted to obtain the maximum torque for any rotor speed during the starting period.

The maximum SPIM torque is a non-linear function of both the speed and the capacitor. In [I]the steepest ascent method is used to find the optimal capacitor value that realize the maximum torque for any rotor speed. The step size of the steepest ascent method can in hence the convergence rate to the optimal torque. If the step size is small the convergence rate will be very slow. If the step size is large, the SPIM torque may appear to oscillate. In the present paper, a *fuzzy logic* based **maximum** torque **control** is proposed for a SPIM with a controlled capacitor. The fuzzy logic-based method has the advantage that adaptively changes the step size of the control variable and this **manner** the fast convergence is achieved [3,4]. The complete control system have **been** developed and validated **by** simulation results.

## II. MATHEMATICAL MODEL

The mathematical model of the single-phase induction motor in the stationary reference frame q-d can be expressed by the following equations [5]-[6]:

$$\psi_{qs}^s = \frac{\omega_b}{p} \left[ v_{qs}^s + \frac{R_q}{X_{lq}} (\psi_{mq}^s - \psi_{qs}^s) \right] \quad (1)$$

$$\psi_{ds}^s = \frac{\omega_b}{p} \left[ v_{ds}^s + \frac{R_d}{X_{ld}} (\psi_{md}^s - \psi_{ds}^s) \right] \quad (2)$$

$$\psi_{qr}^s = \frac{\omega_b}{p} \left[ \frac{R_{rq}}{X'_{lrq}} (\psi_{mq}^s - \psi_{qr}^s) + \frac{N_q}{N_d} \frac{\omega_r}{\omega_b} \psi_{dr}^s \right] \quad (3)$$

$$\psi_{dr}^s = \frac{\omega_b}{p} \left[ \frac{R_{rd}}{X'_{lrd}} (\psi_{md}^s - \psi_{dr}^s) + \frac{N_d}{N_q} \frac{\omega_r}{\omega_b} \psi_{qr}^s \right] \quad (4)$$

$$\psi_{mq}^s = X_m \left( \frac{\psi_{qs}^s}{X_{lq}} + \frac{\psi_{qr}^s}{X'_{lrq}} \right) \quad (5)$$

$$\psi_{md}^s = X_m \left( \frac{\psi_{ds}^s}{X_{ld}} + \frac{\psi_{dr}^s}{X'_{lrd}} \right) \quad (6)$$

$$X_m = \frac{1}{\left( \frac{1}{X_{mq}} + \frac{1}{X_{lq}} + \frac{1}{X'_{lrq}} \right)} \quad (7)$$

$$X_m = \frac{1}{\left( \frac{1}{X_{md}} + \frac{1}{X_{ld}} + \frac{1}{X'_{lrd}} \right)} \quad (8)$$

$$i_{qs}^s = \frac{1}{X_{lq}} (\psi_{qs}^s - \psi_{mq}^s) \quad (9)$$

$$i_{ds}^s = \frac{1}{X_{ld}} (\psi_{ds}^s - \psi_{md}^s) \quad (10)$$

$$i_{qr}^s = \frac{1}{X'_{lrq}} (\psi_{qr}^s - \psi_{mq}^s) \quad (11)$$

$$i_{dr}^s = \frac{1}{X'_{lrd}} (\psi_{dr}^s - \psi_{md}^s) \quad (12)$$

The instantaneous electromagnetic torque can be expressed as:

$$T_e = \frac{P}{2} \frac{1}{\omega_e} \left( \frac{N_d}{N_q} \psi_{qr}^s i_{ds}^s - \frac{N_q}{N_d} \psi_{dr}^s i_{qs}^s \right) \quad (13)$$

and the electromechanical equation of the machine is written as:

$$p \left( \frac{2}{p} \right) \omega_m - \frac{1}{J} (T_e - T_l) = 0 \quad (14)$$

By considering the capacitor in series with the auxiliary windings the following voltage equations can be written:

$$v_{qs}^s = v_s \quad (15)$$

$$v_{ds}^s = v_s - \frac{1}{p} \omega_b X_c i_{ds}^s \quad (16)$$

The steady state equations for the capacitor-run single-phase induction motor can be obtained based on the double revolving field theory [7]:

$$V_s = (Z_{qs} + Z_f + Z_b) I_{qs} - j \frac{N_d}{N_q} (Z_f - Z_b) I_{ds} \quad (17)$$

$$V_s = j \frac{N_d}{N_q} (Z_f - Z_b) I_{qs} + \left[ Z_c + Z_{ds} + \left( \frac{N_d}{N_q} \right)^2 (Z_f - Z_b) \right] I_{ds} \quad (18)$$

$$Z_{qs} = R_q + j X_{lq} \quad (19)$$

$$Z_{ds} = R_d + j X_{ld} \quad (20)$$

$$Z_c = -j X_c \quad (21)$$

$$Z_f = R_f + j X_f = \frac{\left( -X_{mq} X'_{lrq} + j \frac{X_{mq} R'_{rq}}{s} \right)}{2 \left( \frac{R'_{rq}}{s} + j (X_{mq} - X'_{lrq}) \right)} \quad (22)$$

$$Z_b = R_b + j X_b = \frac{\left( -X_{mq} X'_{lrq} + j \frac{X_{mq} R'_{rq}}{(2-s)} \right)}{2 \left( \frac{R'_{rq}}{(2-s)} + j (X_{mq} - X'_{lrq}) \right)} \quad (23)$$

From the steady state model the electromagnetic torque can be expressed as:

$$T_e = \frac{\left| \frac{N_d}{N_q} \right|^2 \left( \frac{R_f - R_b}{\omega_e} \right) + 2 \frac{N_d}{N_q} \left| \frac{R_f - R_b}{\omega_e} \right| \sin(\theta_d - \theta_q)}{\omega_e} \quad (24)$$

From equations (15)-(24) it is evident that the performance of the single-phase induction motor is affected by the capacitor C connected in series with the auxiliary winding.

### III. FUZZY LOGIC MAXIMUM TORQUE CONTROLLER

The maximum torque control for a single-phase induction motor can be written as an

optimization problem in a natural way. The objective of the control algorithm is to find the capacitor valued related to the maximum torque for any speed value and **equation (24)** is the objective function of the optimization problem. To find the capacitor value a *fuzzy* logic controller is used. The input variables to the *fuzzy* controller are the increment of the motor electromagnetic torque and the last change of the capacitor value AC, which **are** defined as follows at kth sampling interval:

$$\Delta T_e(k) = T_e(k) - T_e(k-1) \quad (25)$$

$$L\Delta C(k) = \Delta C(k-1) \quad (26)$$

The output variable of the *fuzzy* controller is the Change step AC, which is generated from the input variables through *fuzzy* inference and defuzzification.

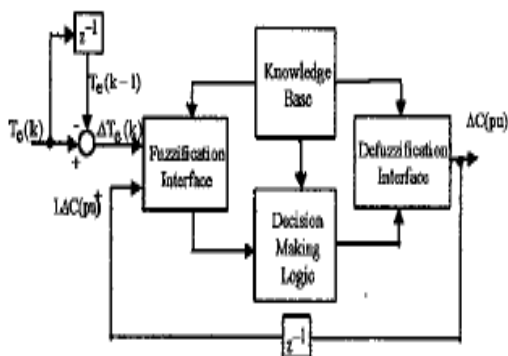


Fig. 3. Fuzzy logic optimization scheme block diagram.

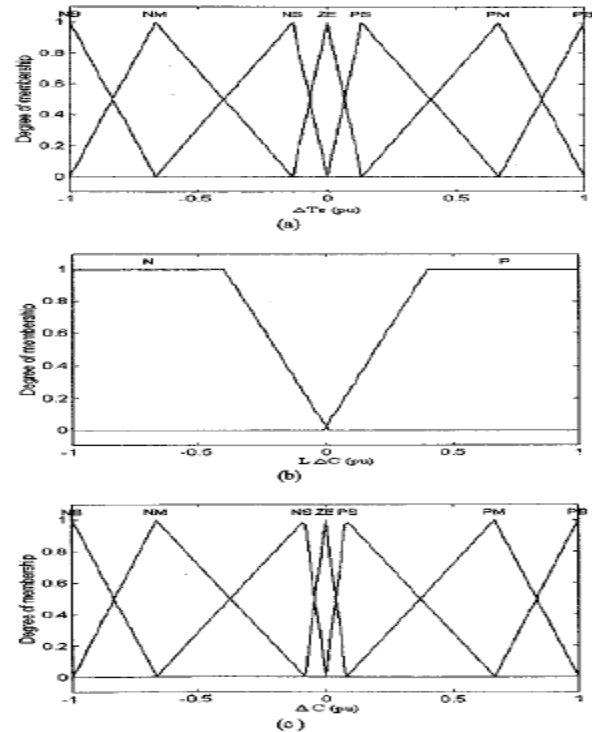


Fig. 4. Membership functions for the fuzzy controller. (a) Change of  $T_e$  ( $\Delta T_e$  (pu)). (b) Last change of  $C$  ( $L\Delta C$  (pu)). (c) Change of  $C$  ( $\Delta C$  (pu)).

TABLE I  
RULE BASE FOR THE FUZZY LOGIC CONTROLLER

$\Delta T_e$ (pu) \ $L\Delta C$ (pu)	N	P
PB	NM	PM
PM	NS	PS
PS	NS	PS
ZE	ZE	ZE
NS	PS	NS
NM	PM	NM
NB	PB	NB



#### IV. SIMULATION RESULTS

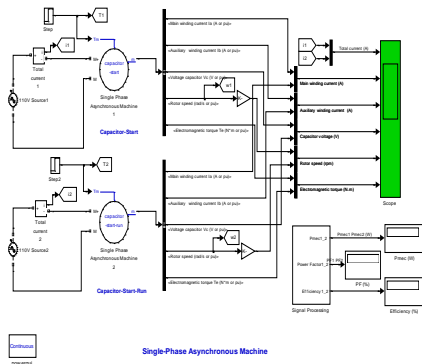


Fig 1 : Split –phase induction motor simulation model

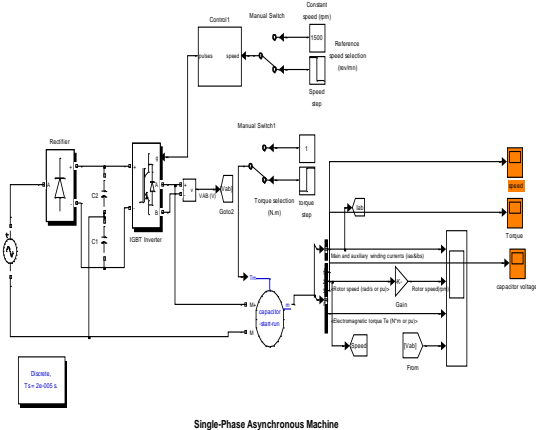


Fig 2 : With conventional controller i.e.PID CONTROLLER

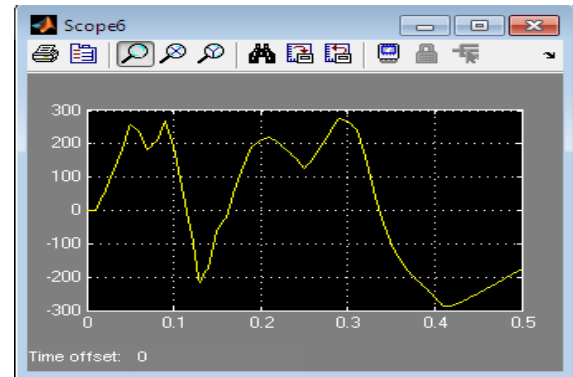
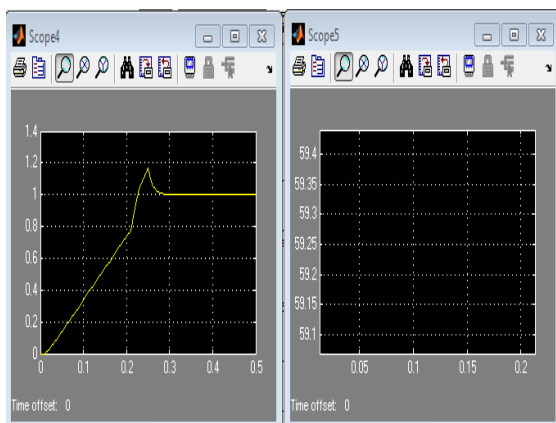


Fig 3(a), (b): speed and torque and capacitor voltage versus time with PID controller

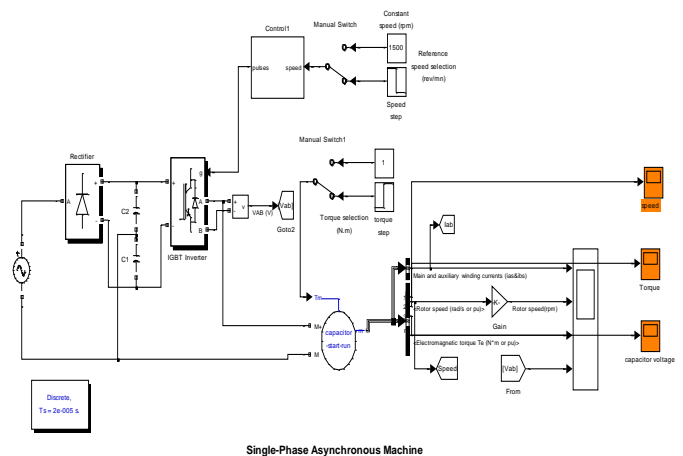
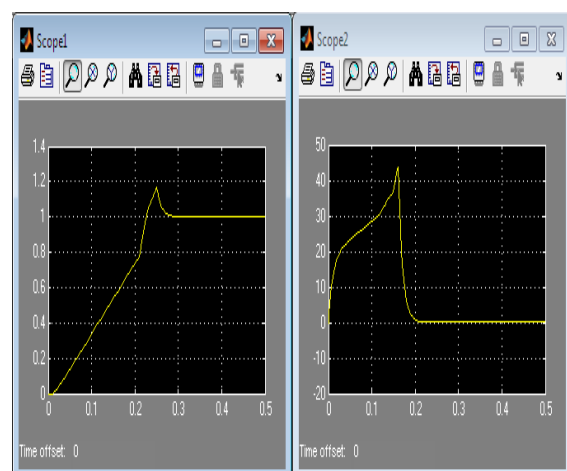


Fig 4: Simulation model with Fuzzy Logic Controller



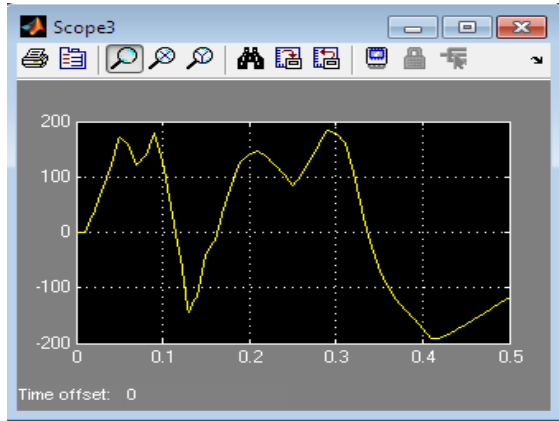


Fig 5(a), (b): speed and torque and capacitor voltage versus time with FLC controller

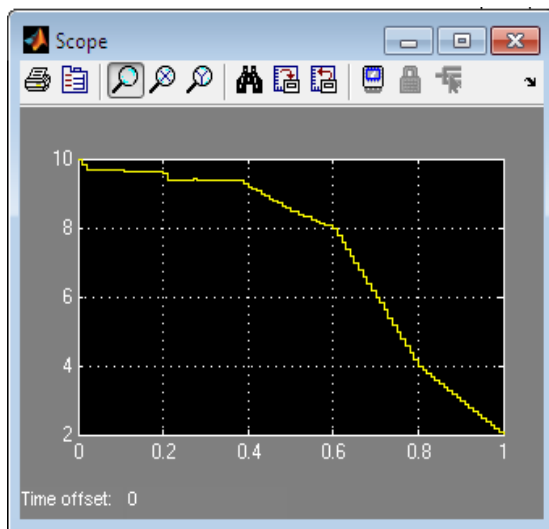


Fig 6: Capacitor versus Speed

## V. CONCLUSION

The design of a *fuzzy* logic base maximum torque control for a SPIM has been presented in this paper. The method uses a fuzzy controller to adjust adaptively the effective capacitor until the maximum torque value is reached, resulting in fast convergence of the search process. The overall system was validated through simulation studies.

From the simulation results, it can be appreciated that substantial improvement in the

drive performance can be obtained with the fuzzy logic approach. Thus a shorter acceleration time and larger torque is obtained by using the effective capacitor value from the *fuzzy* controller output.

## VI. REFERENCES

- [1] Tian-Hua Liu, "A Maximum Torque Control with a Controlled Capacitor for a Single-phase Induction Motor", *IEEE Transactions on Industrial Electronics*, vol. 42, no. 1, pp. 17-24, Feb. 1995.
- [2] E. Mnljadi, Y. Zao, T. Liu, T. A. Lipo, "Adjustable ac Capacitor for a Single-phase Induction Motor", *IEEE Transactions on Industry Applications*, vol. 29, no. 3, pp. 479-485, May 1993.
- [3] G. C. D. Sousa, B. K. Bose, J. G. Cleland, "A fuzzy logic based on-line efficiency optimization control of an indirect vector controlled induction motor drive", in *Proc. IEEE-IECON Conf*, Maui, Hawaii, 1997.
- [4] M. G. Simoes, B. K. Bose, R. J. Spiegel, "Fuzzy Logic Based Intelligent Control of a Variable Speed Cage Machine Winding Generation System", *IEEE Transactions on Power Electronics*, vol. 12, no. 1, pp. 87-95, January 1997.
- [5] P. C. Krause, "Simulation of Unsymmetrical 2-Phase Induction Machines", *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-84, no. 11, pp. 1025-1037, November 1965.
- [6] P. C. Krause, "Analysis of Electric Machinery", McGraw-Hill, 1986.
- [7] P. C. Sen, "Principles of Electric Machines and Power Electronics", 2nd ed., John Wiley & Sons, 1993, pp. 1168-1174, Nov. 1993.

# Dual-Mode Mechanism with Input Voltage Feed-Forward for the Two-Switch Buck-Boost DC–DC Converter

Ganesh P<sup>1</sup> and Mahesh M<sup>2</sup>

<sup>1</sup>Assistant Professor, Department of EEE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

<sup>2</sup>P.G.Scholar, Department of EEE, Malla Reddy Engineering College (Autonomous), Hyderabad, India

<sup>1</sup>ganeshp5891@gmail.com

<sup>2</sup>maddurimahesh123@gmail.com

**Abstract**—This research is proposes a novel to approach extensive input voltage applications using Two switch BUCK-BOOST converter. In order to get high efficiency over the unharmed input voltage . The TSBB converter is activated in buck mode at input voltage is higher than the output voltage and boost mode at input voltage is lower than the output voltage search operation is called dual-mode mechanism. The detached of this paper is to propose an input voltage feed-forward (IVFF) method to eliminate the effect of the input voltage disturbance on the output voltage. Based on the derived its small signal modes the IVFF functions under different operating modes are derived and a two mode control scheme with IVFF compensation is proposed to achieve automatic selections of operating modes and the corresponding IVFF function simultaneously and, nearly smooth mode switching between buck and boost modes is also guaranteed. For exhibiting the advantages of the proposed control scheme clearly, evaluations between the dual-mode mechanism with and without IVFF compensation have been presented in this paper, counting the output signal of the voltage regulator and input-to-output voltage transfer function.

**Index Terms**—feed-forward voltage input, low-signal model, two-mode control, two-switch buck-boost converter.

## I. INTRODUCTION

THE two-switch buck-boost (TSBB) converter, as shown in Fig. 1, is a simplified cascade connection of buck and boost converters [1]. Compared with the basic converters, which have the ability of both voltage step-up and step-down, such as inverting buck-boost, Cuk, Zeta, and SEPIC converters, the TSBB converter presents lower voltage stress of the power devices, fewer passive components, and positive output voltage [2]–[4], and it has been widely used in telecommunication systems [4], battery-powered power supplies [5], [6], fuel-cell power systems [7], [8], power factor correction (PFC) applications [9], [10], and radio-frequency (RF) amplifier power supplies [11], all of which have wide input voltage range. It is thus imperative for the TSBB converter to achieve high efficiency over the entire voltage range. Moreover, considering that the input voltages from battery and fuel cell fluctuate with the output power, and the input voltage in the PFC applications varies with the sinusoidal line voltage, a satisfactory input transient response preventing large output voltage variation in case of input voltage variation is also desired for the TSBB converter.

There are two active switches in the TSBB converter, which provides the possibility of obtaining various control methods for this converter. If  $Q_1$  and  $Q_2$  are switched ON and OFF simultaneously, the TSBB converter behaves the same as the single switch buck-boost converter. This control method is called one mode control scheme [12], [13].  $Q_1$  and  $Q_2$  can also be controlled in other manners. For example, when the input voltage is higher than the output voltage,  $Q_2$  is always kept OFF, and  $Q_1$  is controlled to regulate the output voltage, and as a result, the TSBB converter is equivalent to a buck converter, and is said to operate in *buck mode*. On the other hand, when the input voltage is lower than the output voltage,  $Q_1$  is always kept ON, and  $Q_2$  is controlled to regulate the output voltage, and in this case, the TSBB converter is equivalent to a boost converter, and is said to operate in *boost mode*. Such control method is called two-mode control scheme [3], [4]. Compared with one-mode control scheme, two-mode control scheme can reduce the conduction loss and switching loss effectively, leading to a high efficiency over a wide input voltage range, as explained in [4]. Besides, in order to achieve automatic switching between buck and boost modes, the two-mode control scheme based on two modulation signals with one carrier or one modulation signal with two carriers was proposed in [14].

When the TSBB converter operates in continuous current boost mode, it presents a right-half-plane (RHP) zero. This RHP zero limits the bandwidth of the control loop, penalizing the transient response [15]. Moreover, in the two-mode control scheme with automatic mode-switching, only one voltage regulator is used for both buck and boost modes, and it is often designed to have enough phase margin in boost mode by reducing the bandwidth of the control loop, thus the transient responses of this converter are deteriorated in the whole input voltage range, including both buck and boost modes. To improve the transient response of the TSBB converter, average current mode control [16], current-programmed mode control [17], [18], and voltage mode control with a two-mode proportional-integral-derivative (PID) [19],

Type-III (2-zeros and 3-poles) [20] compensator, or passive RC-type damping network [21] are employed. With these control schemes mentioned earlier, the influence of the input voltage and load disturbances on the output voltage can be well reduced, but cannot be fully eliminated.

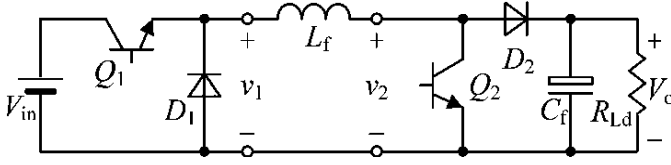


Fig. 1. Two-switch buck-boost (TSBB) converter.

For the converter in the applications with wide input voltage variation, input voltage feed-forward (IVFF) compensation is an attractive approach for improving the transient response of the converter, for it can eliminate the effect of the input voltage disturbance on the output voltage in theory. The IVFF of the buck or boost converter can be implemented in several methods: 1) vary either the amplitude of the carrier signal [22], [23] or the value of the modulation signal [24]–[26] according to the input voltage. However, the variations of the carrier signal for the IVFF of the boost converter and the modulation signal for IVFF of the buck converter are both inversely proportional to the input voltage, which imply that the implementation of this IVFF method is complicated relatively for the TSBB converter.

2) Calculate the duty ratio [27]–[30]. Since the duty ratio calculation for the buck converter is inversely proportional to the

input voltage, a little complicated realization is also required.

3) Derive the IVFF function producing zero audio susceptibility through the small-signal model [31]–[34]. As derived in [31],

the IVFF functions of buck and boost converters are both in proportion to the input voltage, and they are easy to be implemented.

So, the IVFF method with derived IVFF function from the small-signal model will be adopted in this paper. IVFF compensation for the TSBB converter with two-mode control scheme has been achieved by varying the peak and valley values of the carrier signal in proportion to the input voltage in buck and boost modes, respectively [35], or the peak value of the carrier signal and modulation signal in proportion to input voltage simultaneously [36]. In these control schemes, the selection and switching of operating modes and IVFF compensations are not automatic, but require a rather complicated mode detector, which is realized by comparing the input voltage and output voltage and adding auxiliary circuits. In addition, considering the carrier signal generators of most IC controllers operate from an internally derived power supply, the IVFF methods by varying the carrier signal with input voltage are not very general. Thus, this paper will combine the two-mode control scheme with automatic mode-switching ability and the IVFF functions derived from the small-signal models under different operating modes together,

and propose a general, easy implementation, and effective two-mode scheme with IVFF compensation, achieving automatic selections of the operating modes and the corresponding IVFF functions simultaneously. In other words, when the input voltage is higher than the output voltage, the TSBB converter with this proposed control scheme can operate in buck mode and select the IVFF function of this mode automatically. On the other side, when the input voltage is lower than the output voltage, the TSBB converter can operate in boost mode and select the IVFF function of boost mode automatically.

This paper is organized as follows. Section II introduces the two-mode control scheme with automatic mode-switching ability for the TSBB converter, and Section III derives its small signal models under different operating modes. Based on the derived small-signal models, the IVFF functions under different operating modes are derived, and a two-mode control scheme with IVFF compensation is proposed to achieve automatic selections of operating modes and the corresponding IVFF functions simultaneously, and nearly smooth mode-switching. Besides, the comparisons between the two-mode control scheme with and without IVFF compensation are given in Section IV. Section V presents the experimental results from a prototype with this proposed control scheme, and finally, Section VI concludes this paper.

## II. TWO-MODE CONTROL SCHEME WITH AUTOMATIC MODE-SWITCHING ABILITY

As shown in Fig. 1, the voltage conversion of the TSBB converter operated in continuous current mode (CCM) is [4]

$$V_o = \frac{V_{in}}{1 - d_1 - d_2}$$

where  $d_1$  and  $d_2$  are the duty cycles of switches  $Q_1$  and  $Q_2$ , respectively.

In the two-mode control scheme,  $d_1$  and  $d_2$  are controlled independently. When the input voltage is higher than the output voltage, the TSBB converter operates in *buck mode*, where  $d_2 = 0$ , i.e.,  $Q_2$  is always OFF, and  $d_1$  is controlled to regulate the output voltage; when the input voltage is lower than the output voltage, the TSBB converter operates in *boost mode*, where  $d_1 = 1$ , i.e.,  $Q_1$  is always ON, and  $d_2$  is controlled to regulate the output voltage. Thus, the voltage conversion of the TSBB converter with two-mode control scheme can be written as the peak-to-peak value of the carrier is  $V_{saw} = V_H - V_L$ . With the same carrier, in order to achieve the two-mode operation as described in (2), only one of  $v_e$  buck and  $v_e$  boost can intersect  $v_{saw}$  at any time. So, it is required that

$$v_{e \text{ buck}} - v_{e \text{ boost}} \geq V_{saw}$$

Hence, the output signal of the voltage regulator  $v_{ea}$ , as shown in Fig. 2, can be taken as  $v_e$  boost, and  $v_e$  buck is composed by adding a dc bias voltage,  $V_{bias}$ , to  $v_{ea}$ , i.e.,

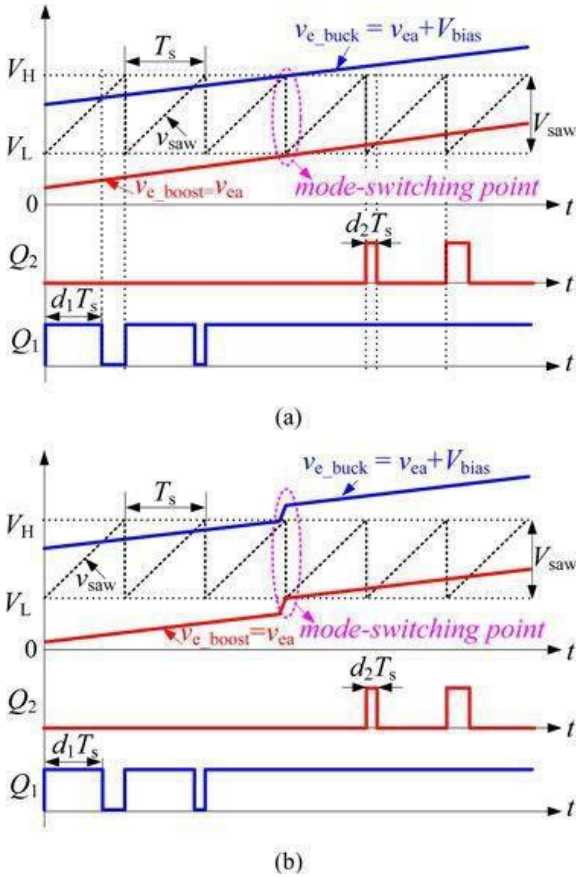


Fig. 2. TSBB converter under the two-mode control scheme.

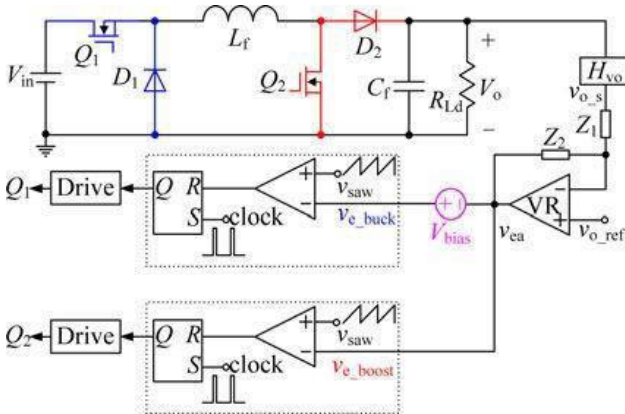


Fig. 3. Two-mode control scheme based on two modulation signals and one carrier. (a)  $V_{bias} = V_{sw}$ . (b)  $V_{bias} > V_{sw}$ .

the peak-to-peak value of the carrier is  $V_{sw} = V_H - V_L$ . With the same carrier, in order to achieve the two-mode operation as described in (2), only one of  $v_{e\_buck}$  and  $v_{e\_boost}$  can intersect  $v_{sw}$  at any time. So, it is required that

$$v_{e\_buck} - v_{e\_boost} \geq V_{sw} \quad (3)$$

Hence, the output signal of the voltage regulator  $v_{ea}$ , as shown in Fig. 2, can be taken as  $v_{e\_boost}$ , and  $v_{e\_buck}$  is composed by adding a dc bias voltage,  $V_{bias}$ , to  $v_{ea}$ , i.e., So, the modulation signal in (4) with  $V_{bias} \geq V_{sw}$  can achieve the two-mode operation of the TSBB converter. When  $V_{in} > V_o$ ,  $v_{e\_buck}$  will be within  $[V_L, V_H]$ , and it

$$\begin{cases} v_{e\_buck} = v_{ea} + V_{bias} \\ v_{e\_boost} = v_{ea} \end{cases} \quad (4)$$

Substituting (4) into (3) yields

$$V_{bias} \geq V_{sw}. \quad (5)$$

; and meanwhile,  $v_{e\_boost} = v_{ea} \leq v_{e\_buck} - V_{sw} < V_L$ , and thus  $d_2 = 0$ . Such case corresponds to the buck mode of the TSBB converter. When  $V_{in} < V_o$ ,  $v_{e\_boost} = v_{ea}$  will be within  $[V_L, V_H]$ , and it intersects  $v_{sw}$  and thus determines  $d_2$ ; and meanwhile,  $v_{e\_buck} = v_{ea} + V_{bias} \geq v_{ea} + V_{sw} > V_H$ , and thus  $d_1 = 1$ . Such case corresponds to the boost mode of the TSBB converter. When  $V_{in} = V_o$ , which is the switching point of the buck and boost modes,  $v_{e\_boost} = V_L$ , and thus  $d_2 = 0$ ; and meanwhile,  $v_{e\_buck} \geq V_H$ , and thus  $d_1 = 1$ . It can be found that  $v_{e\_buck} = V_H$  if  $V_{bias} = V_{sw}$ , and  $v_{e\_buck} > V_H$  if  $V_{bias} > V_{sw}$  at the mode-switching point, as depicted in Fig. 3(a) and (b), respectively. So, by letting  $V_{bias} = V_{sw}$ , i.e.,  $v_{e\_buck} - v_{e\_boost} = V_{sw}$  at the mode-switching point, the buck and boost modes can be smoothly switched from each other.

### III. IVFF FOR TWO-MODE CONTROL SCHEME

#### A. Derivations of DC and Small-Signal Models of the TSBB Converter

intersects  $v_{sw}$  and thus determines  $d_1$

As described in [37] and [38], in the averaged switch model of a dc-dc converter, the switch is modeled by a controlled current source with the value equaling to the average current flowing through the switch, and the diode is modeled by a controlled voltage source with the value equaling to the average voltage across the diode. With this method, the averaged switch Model of the TSBB converter can be obtained, as shown in Fig. 4(a), where  $i_{Q1} = d_1 i_L$  and  $i_{Q2} = d_2 i_L$ , which are the average currents flowing through switches  $Q_1$  and  $Q_2$ , respectively, and  $v_{D1} = d_1 v_{in}$  and  $v_{D2} = d_2 v_o$ , which are the average voltages across diodes  $D_1$  and  $D_2$ , respectively. The average values of voltage, current, and duty cycle in the averaged switch model can be decomposed into their dc and ac components, so  $i_{Q1}$ ,  $i_{Q2}$ ,  $v_{D1}$ , and  $v_{D2}$  can be expressed as

$$\begin{aligned} i_{Q1} &= I_{Q1} + \hat{i}_{Q1} = (D_1 + \hat{d}_1)(I_L + \hat{i}_L) \\ &= D_1 I_L + D_1 \hat{i}_L + \hat{d}_1 I_L + \hat{d}_1 \hat{i}_L \end{aligned} \quad (6)$$

$$\begin{aligned} i_{Q2} &= I_{Q2} + \hat{i}_{Q2} = D_2 + \hat{d}_2 (I_L + \hat{i}_L) \\ &= D_2 I_L + D_2 \hat{i}_L + \hat{d}_2 I_L + \hat{d}_2 \hat{i}_L \end{aligned} \quad (7)$$

$$\begin{aligned} v_{D1} &= V_{D1} + \hat{v}_{D1} = D_1 + \hat{d}_1 = (V_{in} + \hat{v}_{in}) \\ &= D_1 V_{in} + d_1 \hat{v}_{in} \end{aligned} \quad (8)$$

$$\begin{aligned} v_{D2} &= V_{D2} + \hat{v}_{D2} = D_2 + \hat{d}_2 = (V_o + \hat{v}_o) \\ &= D_2 V_o + D_2 \hat{v}_o + d_2 V_o + d_1 \hat{v}_o \end{aligned} \quad (9)$$

where the upper-case letter denotes the dc value, and the lowercase letter with hat ( $\hat{\cdot}$ ) denotes the small-signal

perturbation.

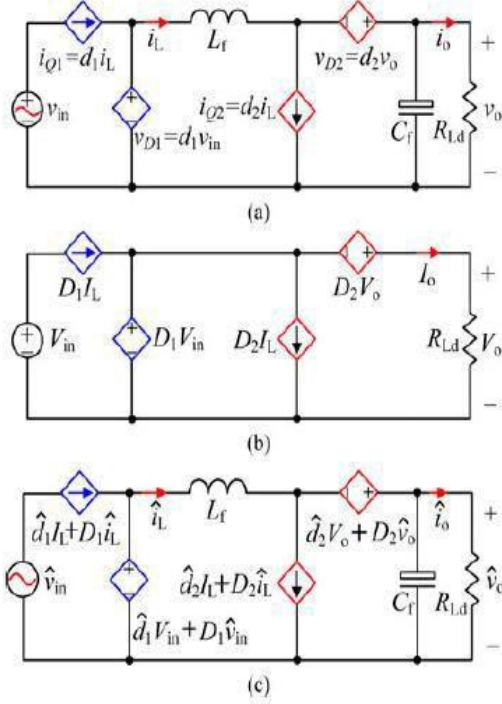


Fig. 4. Models of the TSBB converter. (a) Averaged switch model, (b) DC model, (c) Small-signal model.

With small-signal assumption, the average values in (6)–(9) can be linearized by neglecting the second-order ac terms [1]. Then, the dc model of the TSBB converter can be gotten by replacing the average values in Fig. 4(a) with the dc components in (6)–(9), as depicted in Fig. 4(b). Besides, the inductor  $L_f$  is short circuit, and the capacitor  $C_f$  is open circuit in the dc model. Likely, by replacing the average values in Fig. 4(a) with the first-order ac components in (6)–(9), the small-signal model of the TSBB converter can be obtained, as illustrated in Fig. 4(c). According to (2), setting  $d_2 = 0$ , i.e.,  $D_2 = 0$ ,  $\hat{d}_2 = 0$ , and  $d_1 = 1$ , i.e.,  $D_1 = 1$ ,  $\hat{d}_1 = 0$  in Fig. 4(c), respectively, the small-signal models in buck and boost modes can be derived.

#### B. Derivation of IVFF Functions

Fig. 5(a) shows a general control block diagram of a dc–dc converter [1], where  $G_{vd}(s)$ ,  $G_{vo\_vin}(s)$  and  $Z_o(s)$  are the transfer functions of the duty ratio  $\hat{d}$ , input voltage  $\hat{v}_{in}$ , and output current  $\hat{i}_o$  to the output voltage  $\hat{v}_o$ , respectively,  $G_{vr}(s)$  is the transfer function of the voltage regulator,  $G_{PWM}(s)$  is the transfer function of the pulse-width modulation (PWM) modulator,  $\hat{v}_o \text{ ref}$  is the output voltage reference, and  $H_{vo}(s)$  is the sense gain of the output voltage. As seen, the disturbance of input voltage

$\hat{v}_{in}$  affects the output voltage through the path with transfer function  $G_{vo\_vin}(s)$ . This effect can be eliminated by introducing an additional path with

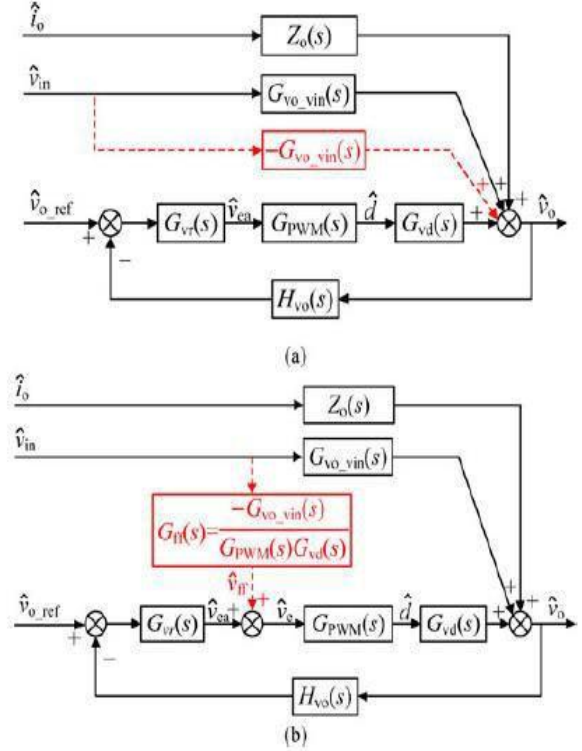


Fig. 5. Control block diagram with IVFF of a dc–dc converter. (a) Introduction of IVFF. (b) Equivalent transformation of IVFF.

transfer function

$-G_{vo\_vin}(s)$  from the input voltage to the output voltage, as illustrated with the dashed line shown in Fig. 5(a). Moving the output of  $-G_{vo\_vin}(s)$  to the output of voltage regulator and corresponding transfer function being changed to  $G_{ff}(s)$ , the control block is equivalently transformed to that shown in Fig. 5(b). The path from  $\hat{v}_{in}$  to  $\hat{v}_{ff}$  is called the IVFF path, and the IVFF function  $G_{ff}(s)$  is

$$G_{ff}(s) = \frac{\hat{v}_{ff}}{\hat{v}_{in}} = -\frac{G_{vo\_vin}(s)}{G_{PWM}(s)G_{vd}(s)}. \quad (10)$$

As shown in Fig. 5(b), the output of  $G_{ff}(s)$ , i.e.,  $\hat{v}_{ff}$ , is added to the output signal of the voltage regulator, i.e.,  $\hat{v}_{ea}$

, forming the modulation signal  $\hat{v}_e$ . As discussed earlier, by setting  $D_2 = 0$  and  $\hat{d}_2 = 0$  in the small-signal model of the TSBB converter shown in Fig. 4(c), the small-signal model in buck mode can be obtained, and the transfer functions of duty ratio and input voltage to the output voltage in this mode,  $G_{vd \text{ buck}}(s)$  and  $G_{vo\_vin \text{ buck}}(s)$  can be derived as where  $L_f$ ,  $C_f$ , and  $R_{Ld}$  are the filter inductor, filter capacitor, and load resistor of the TSBB converter, respectively, and  $V_{in \text{ dc}}$  and  $V_o$  are the input voltage and output voltage at the quiescent operation point,

respectively. Likewise, by setting  $D1 = 1$  and  $\hat{d}1 = 0$  in Fig. 4(c), the small-signal model in boost mode can be gotten, and the transfer functions of duty ratio and input voltage to the output voltage in this mode,  $G_{vd \text{ boost}}(s)$  and  $G_{vo \text{ in boost}}(s)$ , can be,

Parameter	Symbol	Value
Input voltage	$V_{in}$	250-500 V
Output voltage	$V_o$	360 V
Output power	$P_o$	6 kW
Full load resistor	$R_{ld}$	21.6 $\Omega$
Switching frequency	$f_s$	100 kHz
Switches	$Q_1, Q_2$	SPW47N60C3
Diodes	$D_1, D_2$	SDP30S120
Filter capacitor	$C_f$	4080 $\mu\text{F}$
Filter inductor	$L_f$	320 $\mu\text{H}$
Sense gain of the input voltage	$H_{vin}$	1/100
Sense gain of the output voltage	$H_{vo}$	1/144
Peak-to-peak value of the carrier	$V_{saw}$	2.5 V

respectively, derived as

$$G_{vd \text{ boost}}(s) = \left. \frac{\hat{v}_o(s)}{\hat{d}(s)} \right|_{\hat{v}_{in}=0} = \frac{1 - s \frac{L_e}{R_{ld}}}{s^2 L_e C_f + s \frac{L_e}{R_{ld}} + 1} \frac{V_o^2}{V_{in \text{ dc}}} \quad (13)$$

$$G_{vo \text{ in boost}}(s) = \left. \frac{\hat{v}_o(s)}{\hat{v}_{in}(s)} \right|_{\hat{d}=0} = \frac{1}{s^2 L_e C_f + s \frac{L_e}{R_{ld}} + 1} \frac{V_o}{V_{in \text{ dc}}} \quad (14)$$

where  $L_e = L_f V_{2o} / V_{2in \text{ dc}}$ .

The transfer function of the PWM modulator  $G_{PWM}(s)$  can be expressed as [1]

$$G_{PWM}(s) = \frac{\hat{d}(s)}{\hat{v}_e(s)} = \frac{1}{V_{saw}}. \quad (15)$$

Substituting (11), (12), and (15) into (10), the IVFF transfer function in buck mode can be derived as

$$G_{ff \text{ boost}}(s) = -\frac{1}{V_o \left(1 - s \frac{L_e}{R_{ld}}\right)} V_{saw}. \quad (17)$$

In (17), the term  $sL_e/R_{ld} = sL_f V_{2o} / R_{ld} V_{2in \text{ dc}}$  is a function of frequency, and the factor  $L_e/R_{ld} = L_f V_{2o} / R_{ld} V_{2in \text{ dc}}$  reaches its maximum value at full load and minimum input voltage. According to the parameters of the prototype listed in Table I to appear in Section V, the magnitude of  $sL_e/R_{ld}$  with the full load resistor  $R_{ld} = 21.6 \Omega$  and the minimum input voltage  $V_{in \text{ min}} = 250 \text{ V}$  is depicted in Fig. 6.

Fortunately, the input voltage of the battery-powered power supply [5], PFC application [9], RF amplifier supply [11], and fuel-cell power system [39] fluctuates at low frequency, and thus  $|sL_e/R_{ld}| \ll 1$ , as shown in Fig. 6.

Therefore, (17) can be simplified as

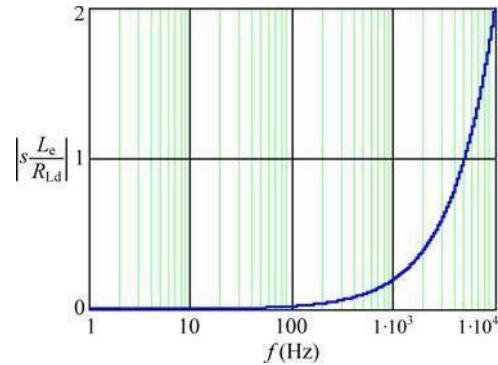


Fig. 6. Magnitude of  $sL_e/R_{ld}$  as the function of frequency.

## VI. CONCLUSION

For the TSBB converter operated with the two-mode control scheme, the small-signal models in buck and boost modes are built, and based on which, detailed derivations of the IVFF functions under different operation modes are given in this paper. With reasonable simplification of the IVFF function in boost mode, a general, easy implementation and effective two mode control scheme with IVFF compensation is proposed to achieve automatic selection of operating modes and the corresponding IVFF compensations for the TSBB converter. Moreover, the switching between buck and boost modes in this proposed control scheme is nearly smooth. In order to present the merits of this proposed control scheme clearly, comparisons between the two-mode control scheme with and without IVFF compensation, including the output signals of the voltage regulator and the input-to-output voltage transfer functions, are discussed in this paper. Finally, a 250–500 V input, 360 V output and 6-kW-rated power prototype is built and tested, and the experimental results demonstrate the validity of this proposed control scheme, with which high efficiency over the whole input voltage range and improved input voltage transient response are achieved for the TSBB converter.

## REFERENCES

- [1] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*. Norwell, MA, USA: Kluwer, 2011.
- [2] D. C. Jones and R. W. Erickson, "A nonlinear state machine for dead zone avoidance and mitigation in a synchronous non inverting buck-boost converter," *IEEE Trans. Power Electron.*, vol. 28, no. 1, pp. 467–480, Jan.2013.
- [3] C. Yao, X. Ruan, and X.Wang, "Isolated buck-boost dc/dc converters suitable for wide input-voltage range," *IEEE Trans. Power Electron.*, vol. 26, no. 9, pp. 2599–2613, Sep. 2011.
- [4] X. Ren, X. Ruan, H. Qian, M. Li, and Q. Chen, "Three-mode dualfrequency two-edge modulation scheme for four-switch buck-boost converter," *IEEE Trans. Power Electron.*, vol. 24, no. 2, pp. 499–509, Feb.2009.
- [5] Y. J. Lee, A. Khaligh, A. Chakraborty, and A. Emadi, "A compensation technique for smooth transitions in a noninverting buck-boost converter," *IEEE Trans. Power Electron.*, vol. 24, no. 4, pp. 1002–1016, Apr. 2009.
- [6] Y. J. Lee, A. Khaligh, A. Chakraborty, and A. Emadi, "Digital combination of buck and boost converters to control a positive buck-boost converter and improve the output transients," *IEEE Trans. Power Electron.*, vol. 24, no. 5, pp. 1267–1279, May 2009.
- [7] E. Schaltz, P. O. Rasmussen, and A. Khaligh, "Non-inverting buck- boostconverter for fuel cell application," in *Proc. IEEE Annual*



Conf. *IEEE Ind. Electron.*, 2008, pp. 855–860.

[8] H. Qu, Y. Zhang, Y. Yao, and L. Wei, “Analysis of buck-boost converter for fuel cell electric vehicles,” in *Proc. IEEE Int. Conf. Veh. Electron. Safety*, 2006, pp. 109–113.

[9] G. K. Andersen and F. Blaabjerg, “Current programmed control of a single-phase two-switch buck-boost power factor correction circuit,” *IEEE Trans. Power Electron.*, vol. 53, no. 1, pp. 263–271, Feb. 2006.

[10] R. Morrison and M. G. Egan, “A new modulation strategy for a buck boost input ac/dc converter,” *IEEE Trans. Power Electron.*, vol. 16, no. 1, pp. 34–45, Jan. 2001.

[11] B. Sahu and G. A. Rincon-Mora, “A high-efficiency linear RF power amplifier with a power-tracking dynamically adaptive buck-boost supply,” *IEEE Trans. Microw. Theory Techniques*, vol. 52, no. 1, pp. 112–120, Jan. 2004.

[12] H. Liao, T. Liang, L. Yang, and J. Chen, “Non-inverting buck-boost converter with interleaved technique for fuel-cell system,” *IET Power Electron.*, vol. 5, no. 8, pp. 1379–1388, 2012.

[13] R. Lin and R. Wang, “Non-inverting buck-boost power-factor-correction converter with wide input-voltage applications,” in *Proc. IEEE Annual Conf. IEEE Ind. Electron.*, 2010, pp. T12-120–T12-124.

[14] T. Ishii, M. Yoshida, M. Motomori, and J. I. Hara, “Buck-boost converter,” U.S. Patent 7 268 525, 2007.

[15] C. Restrepo, T. Konjedic, J. Calvente, M. Milanovic, and R. Giral, “Fast transitions between current control loops of the coupled-inductor buckboost dc-dc switching converter,” *IEEE Trans. Power Electron.*, vol. 28, no. 8, pp. 3648–3652, Aug. 2013.

[16] C. Wei, C. Chen, K. Wu, and I. Ko, “Design of an average-current-mode non inverting buck-boost dc-dc converter with reduced switching and conduction losses,” *IEEE Trans. Power Electron.*, vol. 27, no. 12, pp. 4934–4943, Dec. 2012.

[17] J. Park, J. Fan, X. Wang, and A. Huang, “A sample-data model for double edge current programmed mode control (DECPM) in high-frequency and wide-range dc-dc converters,” *IEEE Trans. Power Electron.*, vol. 25, no. 4, pp. 1023–1033, Apr. 2010.

[18] A. A. Ahmad and A. Abrishamifar, “A simple current mode controller for two switches buck-boost converter for fuel cells,” in *Proc. IEEE Elect. Power Conf.*, 2007, pp. 363–366.

[19] R. Paul, L. Sankey, L. Corradini, Z. Popovic, and D. Maksimovic, “Power management of wideband code division multiple access RF power amplifiers with antenna mismatch,” *IEEE Trans. Power Electron.*, vol. 25, no. 4, pp. 981–991, Apr. 2010.

[20] B. Sahu and G. A. Rincon-Mora, “A low voltage, dynamic, noninverting, synchronous buck-boost converter for portable applications,” *IEEE Trans. Power Electron.*, vol. 19, no. 2, pp. 443–452, Mar. 2004



# Lyapunov Design Based MIT Rule for Model Reference Adaptive Control Systems

Ms. K.Chetaswi<sup>1</sup>, Mr. G. Madhu Kumar<sup>2</sup>, Mr. Desai Feroz<sup>3</sup>, Mr. G. Nithin Krishna Gupta<sup>4</sup>

<sup>1</sup>Assistant Professor, MREC (A)

<sup>2</sup>P.G Student, MREC (A)

<sup>3</sup>P.G Student, MREC (A)

<sup>4</sup>P.G Student, MREC (A)

<sup>1</sup>[k.chetaswi@gmail.com](mailto:k.chetaswi@gmail.com), <sup>2</sup>[gurrammadhukumar@gmail.com](mailto:gurrammadhukumar@gmail.com), <sup>3</sup>[desaiferoz@gmail.com](mailto:desaiferoz@gmail.com) com

<sup>4</sup>[gnthinkkrish@gmail.com](mailto:gnthinkkrish@gmail.com)

**Abstract**— This paper presents the control process, which measures the dynamic behaviour of second order Model Reference Adaptive system (MRAS) using Lyapunov based MIT rule for adaptive mechanism. In this rule, a cost function is defined as a function of error between the outputs of the plant and the reference model, and controller parameters are adjusted in such a way so that this cost function is minimized. Simulation is done in MATLAB-SIMULINK with different values of adaptation gain and the results are presented for basic MIT rule and Lyapunov based MIT rule.

**Keywords**—Model Reference Adaptive System (MRAS), MIT rule, Lyapunov based MIT.

## I. INTRODUCTION

An adaptive control system provides continuous information about present state of plant i.e.; identify the process, compares present system performance with desired/optimum performance and make decision to adapt the system towards the optimum performance. An adaptive control is one of the widely used control strategy to design advanced control systems for better performance and accuracy. MRAC is an important adaptive controller with adjustable controller parameters and adaptive mechanism to adjust the system.

MRAC has becomes a very efficient and systematic method for controlling plants with unknown or partially known parameters. Hence the unknown parameter variations are minimized depending on environmental changes. An adaptive controller consists of two loops an inner loop or normal loop and an outer loop or parameter adjustment loop. This paper deals with designing of adaptive controller with MRAC scheme of Lyapunov design using MIT rule to control a second order system.

## II. MODEL REFERENCE ADAPTIVE CONTROL

### Principle

The main aim is to make the output of an unknown plant approach asymptotically the output of a given reference model. MRAC control strategy is used to design adaptive controller that works on the principle of adjusting the parameters (to obtain the reference output) so that output of the actual plant tracks the output of the reference model having the same reference input.

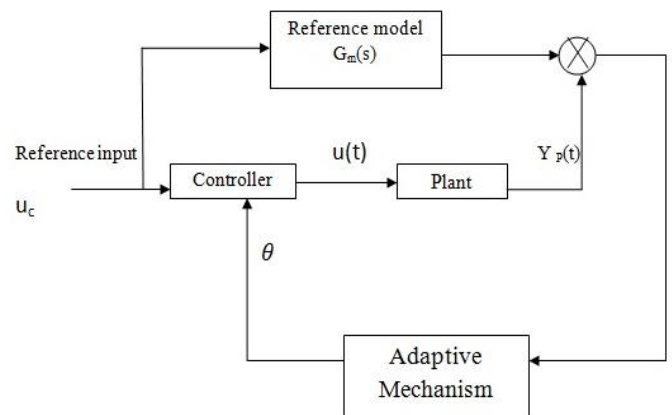


Fig 1. Model Reference Adaptive System

### A. Elements:

#### Reference Model:

It is the process output ideally should respond to the command signal/reference signal.

### B. Adaptive Mechanism:

This element is used to modify system performance, either by altering the parameters of the adjustable system (or) Controller. So that plant could track the reference model. Mathematical approaches like MIT rule, Lyapunov theory can be used to develop the adjusting mechanism.

In Fig(1), As shown above,  $Y_m(t)$  is the output of the reference model and  $Y_p(t)$  is the output of the actual plant and error  $e(t)$  is given as

$$e = y_p - y_m \text{ -----}$$

Equation(1)



The system performance can be adjusted either by modifying the parameters/input signals. In this paper only one parameter  $\phi$  is used to describe the Control law. The value of  $\phi$  is primarily dependent on adaption gain.

### III. MIT RULE:

The MIT rule is the original approach to MRAC. It was developed at the Instrumentation laboratory (proper laboratory) at MIT. MIT rule was first suggested by Whitaker and his co-workers in 1958 in order to develop an adaptive autopilot system for aircrafts. This was based on Minimization of performance index and led to design the rule referred as the MIT rule. MIT rule can be used to design a controller with MRAC scheme for any system. The cost function is defined as,

$$J(\phi) = e^2/2 \text{ -----Equation(2)}$$

Where  $e$  is the error between the output of plant and the model, and  $\phi$  is the adjustable parameter.

Parameter  $\phi$  is adjusted in such a fashion so that the cost can be minimized to zero. To make  $J$  small, it is reasonable to change the parameters in the direction of the negative Gradient of  $J$ , that is by Gradient method,

$$\frac{\partial \phi}{\partial t} = -\gamma e \frac{\partial e}{\partial \phi} \text{ -----Equation (3)}$$

Where, The Partial derivative  $\frac{\partial e}{\partial \phi}$  is called the sensitivity derivative of the system, tells how the error is influenced by the adjustable parameter( $\phi$ ). If it is assumed that the parameter changes are slower than the other variables in the system, then the derivative  $\frac{\partial e}{\partial \phi}$  can be evaluated under the assumption that  $\phi$  is constant. Equation (3) describes the change in the parameter  $\phi$  with respect to time so that the cost function  $J(\phi)$  can be reduced to zero. Here  $\gamma$  is a positive quantity which indicates the adaptation gain of the controller.

Let us assume that the process is linear with the functions  $KG(s)$ , where  $G(s)$  is a second order known Transfer function and  $K$  is an unknown parameter. The underlying design problem is to design a controller so that our process could track the reference model with Transfer function.

$$G_m(s) = K_0 G(s) \text{ -----Equation(4)}$$

Where  $K_0$  – known parameter

The control law be

$$u(t) = \phi * u_c(t) \text{ -----Equation(5)}$$

Where,  $\phi = K_0 / K$

The error  $e = y_p - y_m$

$$e(s) = KG(s)U(s) - G_m(s) U_c(s)$$

$$e(s) = KG(s)\phi U_c(s) - K_0 G(s)U_c(s)$$

$$\frac{\partial e(s)}{\partial \phi} = KG(s) U_c(s)$$

$$\frac{\partial e}{\partial \phi} = K (Y_m(s)/K_0)$$

$$\frac{\partial e}{\partial \phi} = K/K_0 Y_m(s) \text{ -----Equation}$$

(6)

$$\text{We have, } \frac{\partial \phi}{\partial t} = -\gamma e \frac{\partial e}{\partial \phi}$$

$$\frac{\partial \phi}{\partial t} = -\gamma e (K/K_0) Y_m(s) \text{ -----}$$

Equation(7)

Introduce the following modification, then

$$\frac{\partial \phi}{\partial t} = \frac{\gamma \phi e}{\alpha + \phi^T \phi} \text{ The modified control law is}$$

$$u(t) = \phi * u_c(t)$$

$$\frac{\partial \phi}{\partial t} = \left[ \gamma \frac{\phi}{\alpha + \phi^T \phi} e \right]$$

Still the modifications lead the damping oscillations as  $u_c$  increases. MIT Rule based systems work as expected as well for small adaptation gains. Very complex behavior may be obtained from high adaptation gains. To improve further, here we present Lyapunov based MIT rule.

### IV. DESIGN OF MRAS USING LYAPUNOV THEORY:

#### A. Lyapunov theory:

We first derive a differential equation for the error,  $e = y - y_m$ . This differential equation contains the adjustable parameters. We then attempt to find a Lyapunov function and an adaptation mechanism such that the error becomes zero.

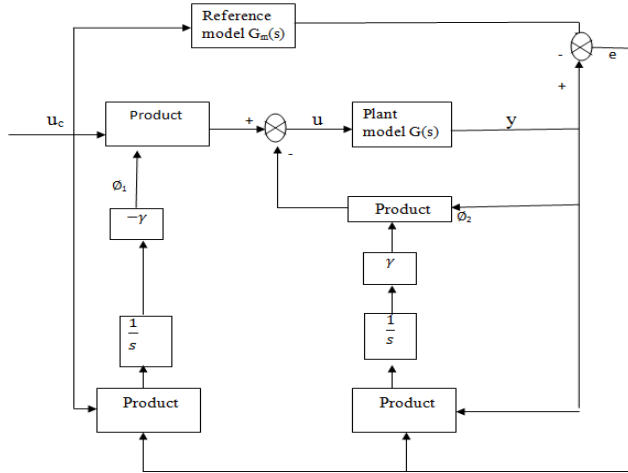


Fig 2. Lyapunov Based MRAS system

When using Lyapunov theory for adaptive systems, we find that  $\frac{dV}{dt}$  is usually only negative semi definite.

### B. Controller Design

The procedure is to determine the error equation and a Lyapunov function with a bounded second derivative. The theorem is then used to show boundedness and that the error goes to Zero. Consider the following transfer function

$$\frac{\phi(s)}{T(s)} = \frac{d_1}{js^2 + cs + mgd_c} \text{-----Equation (9)}$$

$$\frac{\phi(s)}{T(s)} = \frac{1.89}{s^2 + 0.0389s + 10.77} \text{-----Equation (9)}$$

The adaptive controller for this system will take of the form:

$$U = \phi_1 u_c - \phi_2 y_P \text{-----Equation (10)}$$

$$e = y_P - y_M$$

$$e = G_P u - G_M u_c$$

$$y_{plant} = G_P u = \left( \frac{1.89}{s^2 + 0.0389s + 10.77} \right) (\phi_1 u_c - \phi_2 y_{plant})$$

$$y_{plant} = \left( \frac{1.89\phi_1}{s^2 + 0.0389s + 10.77 + 1.89\phi_2} \right) u_c \text{-----Equation (11)}$$

Considering the partial derivative of the error with respect to  $\phi_1$  and  $\phi_2$  gives the sensitivity derivatives, having in mind that  $U_c$  does not include either parameter, and therefore is inconsequential when evaluating the derivative.

$$e = \left( \frac{1.89\phi_1}{s^2 + 0.0389s + 10.77 + 1.89\phi_2} \right) u_c - G_M u_c$$

$$\frac{\delta e}{\delta \phi_1} = \left( \frac{1.89\phi_1}{s^2 + 0.0389s + 10.77 + 1.89\phi_2} \right) u_c \text{-----Equation (12)}$$

$$\frac{\delta e}{\delta \phi_2} = - \left( \frac{1.89}{s^2 + 0.0389s + 10.77 + 1.89\phi_2} \right)^2 \phi_1 u_c \text{--Equation (13)}$$

$$\frac{\delta e}{\delta \phi_2} = - \frac{1.89\phi_1}{s^2 + 0.0389s + 10.77 + 1.89\phi_2} y_{plant}$$

$$s^2 + 0.0389s + 10.77 + 1.89\phi_2 \text{ With } s^2 a_{1m} + a_{0m}$$

Taking the derivative of feed forward loop of MRAC

$$\frac{\delta e}{\delta \phi_1} = \frac{a_{m1}s + a_{m0}}{s^2 + a_{m1}s + a_{m0}} u_c$$

$$\frac{\delta e}{\delta \phi_2} = \frac{a_{m1}s + a_{m0}}{s^2 + a_{m1}s + a_{m0}} y_{plant}$$

$$\frac{d\phi_1}{dt} = -\gamma \frac{\delta e}{\delta \phi_1} e = -\gamma \left( \frac{a_{m1}s + a_{m0}}{s^2 + a_{m1}s + a_{m0}} u_c \right) e \text{--Equation (14)}$$

$$\frac{d\phi_2}{dt} = -\gamma \frac{\delta e}{\delta \phi_2} e = -\gamma \left( \frac{a_{m1}s + a_{m0}}{s^2 + a_{m1}s + a_{m0}} y_{plant} \right) e \text{--Equation (15)}$$

The results were presented in the following section.

### V. EXPERIMENTAL RESULTS

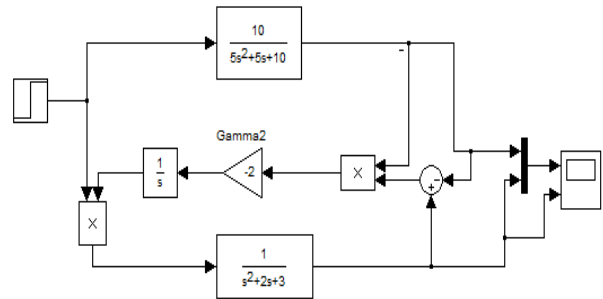


Fig 3 MIT based Control Law

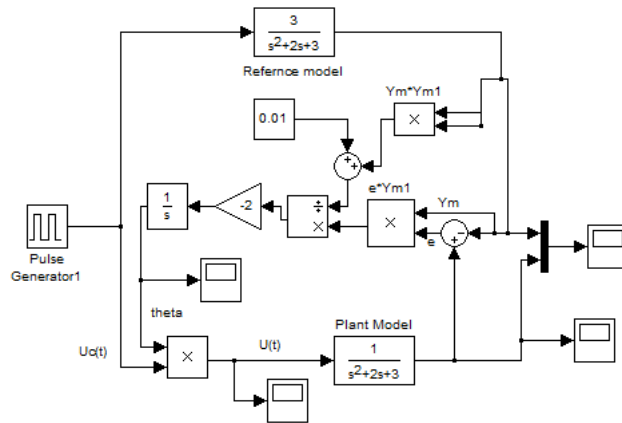


Fig 4 Modified Control Law

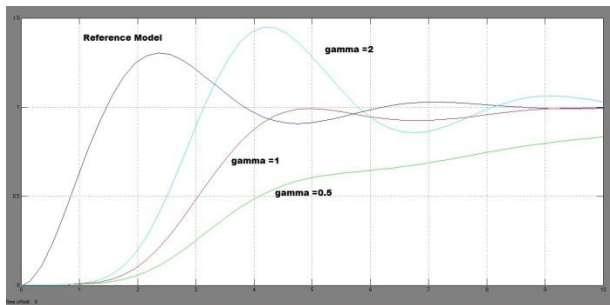


Fig 5 Comparison of output with different gamma values

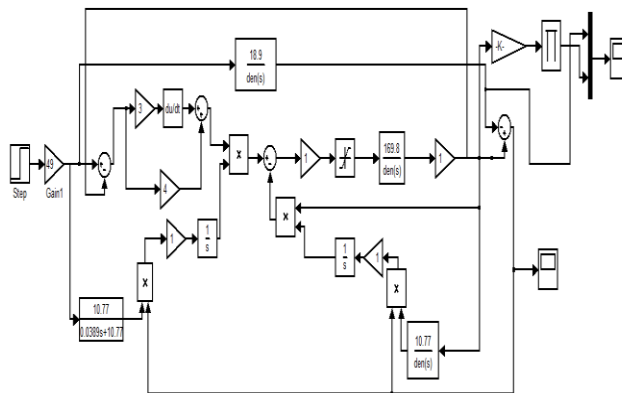


Fig 6 Lyapunov Based MRAS System

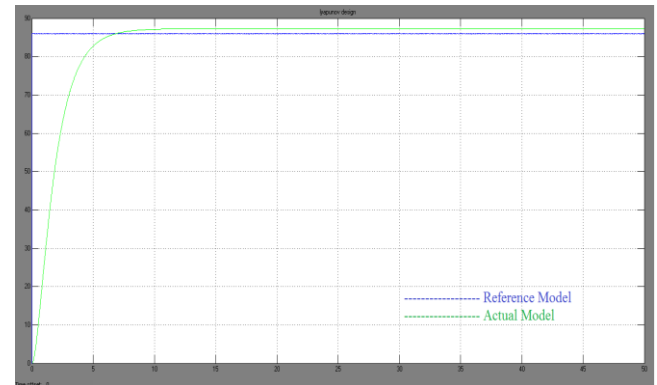


Fig 7 Lyapunov Based MRAS output

## VI. CONCLUSION

The objective of achieving optimum control within the limit and with updating system inputs by direct comparison of the Reference Model output with the actual plant output was solved with aid of MRAS. In case of basic MIT rule, for high gain values, system undergoes damping oscillations. In Lyapunov based MIT control, optimum performance of system can be obtained and experimental results have demonstrated that the proposed approach effectively eliminates the error and need for the process operators to converge the plant output with reference model.

## References

- [1] K. S. Narendra and A. M. Annaswamy (1989), Stable Adaptive Systems, Prentice-Hall, Englewood Cliffs, New Jersey.
- [2] Bowman, K. J. Astrom and B. Wittenmark (2001), Adaptive control, 2nd ed., Dover Publications, New York.
- [3] Kreisselmeier, G. 1985. An approach to stable indirect adaptive control, Automatica, 21(4), pp 425-431.
- [4] P. Swarnkar, S. K. Jain and R. K. Nema (2010), Effect of adaptation gain on system performance for model reference adaptive control scheme using MIT rule. International Conference of World Academy of Science, Engineering and Technology, Paris, pp. 70-75.
- [5] Duarte, M.A. 1996. Indirect model reference adaptive control with dynamic adjustment of parameters, International journal of adaptive control and signal processing, Vol 10, pp 603-621.
- [6] Pankaj, K., Kumar, J.S. and Nema, R.K. 2011. Comparative Analysis of MITRule and Lyapunov Rule in Model Reference Adaptive Control Scheme, Innovative Systems Design and Engineering, 2(4), pp 154-162.
- [7] Model Reference Adaptive Control System For Moisture Regulation In Cotton Ginning by T.Garikayi, L. Nyanga, T. Mushiri, S. Mhlanga, A. Muzinda T. Mutangi. SAIIE25 Proceedings, 9<sup>th</sup>, 11<sup>th</sup> July 2013, Stellenbosch, South Africa © 2013 SAIIE
- [8] Design of a Model Reference Adaptive Controller Using Modified MIT Rule for a Second Order System by Priyank Jain and Dr. M.J. Nigam. Advance in Electronic and Electric Engineering, Research India Publications, Volume 3, Number 4 (2013), pp. 477-484.



# Optimal Design Technique Based on Linear Quadratic Gaussian Theory

Ms.K.Chetaswi<sup>1</sup>, Ms. Sandhya Vinnakota<sup>2</sup>, Ms.B.Mounika<sup>3</sup>

<sup>1</sup>Assistant professor, MREC (A)

<sup>2</sup>P.G Student MREC (A)

<sup>3</sup> P.G Student MREC (A)

[k.chetaswi@gmail.com](mailto:k.chetaswi@gmail.com),

[vsandhya44@gmail.com](mailto:vsandhya44@gmail.com), [beerammounika703@gmail.com](mailto:beerammounika703@gmail.com)

**Abstract**— This paper presents linear quadratic regulator (LQR) based on Linear Quadratic Gaussian (LQG) theory. LQR is an optimal design technique that guarantees a robust system. The difficulty lies in choosing a weighting matrix for the LQR cost function that gives the desired poles. The LQR method guarantees robustness. This paper presents two strategic approaches for the design of LQR also gives the weighting matrix design. M-file Program is presented for the design and results are presented.

**Keywords**—Cost function, Estimator design, LQG Control, Optimal Control, Robustness.

## I. INTRODUCTION

In this paper we concentrated on the process of pole placement with desired robustness. Using pole placement we can get the desired response but the system is not robust. In order to get robustness we are performing LQG mathematical model as an estimator used in the process. The simulation results are pictured in the paper, which reflects the process robustness.

One of the fundamental optimal control problem is Linear Quadratic Gaussian problem. The compensator linear Quadratic Gaussian is used with the combination of Linear Quadratic estimator and Kalman filter for designing the controller. The closed loop stability is guaranteed by LQG. The state space techniques are used in order to know the system uses the better controller and its time saving. LQG is applicable for both Linear Time Variant and Linear time Invariant. To make the system robust we use system parameters to be stochastic.

To get the desired result we are using pole placement technique where in addition with LQG in order to get system in a potent form to get reliable result. The substantial design is meant for every optimal control design, by using LQG we are providing that substantiality which is explained in this paper.

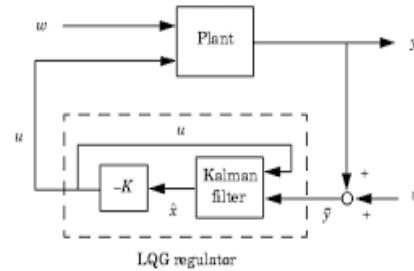


Fig 1 Block Diagram of LQG theory

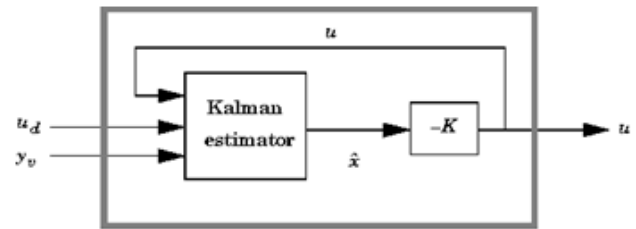


Fig 2 Basic Schematic of the above diagram:

## II. MATHEMATICAL MODEL

Process model equation of LQG STR

$$A(q)y(t) = B(q)u(t) + C(q)e(t)$$

Steady state loss function  $J_{yu} = E\{[y(t) - y_m(t)]^2 + pu^2(t)\}$

Let the monic polynomial  $A(q)$  and  $C(q)$  have degree  $n$ .  $C(q)$  has all zeroes inside the unit disc, Assume, there is no nontrivial polynomial that divides  $A(q)$ ,  $B(q)$  and  $C(q)$

$A_2(q)$  be the GCD of  $A(q)$  and  $B(q)$

Let  $A_2^+(q)$  of degree  $L$  be the factor of  $A_2(q)$ .

$A_2^-(q)$  of degree  $m$  be the factor of  $A(q)$ .

Control law  $R(q)u(t) = -S(q)y(t) + T(q)y_m(t)$

Where  $R$  and  $S$  are of degree  $n+m$

$$R(q) = A_2^-(q) \hat{R}(q)$$

$$S(q) = Z^m \hat{S}(q)$$

$\hat{R}(q)$  and  $\hat{S}(q)$  satisfy the Diophantine equation



$$A_1(q) \hat{A}_2(q) \hat{R}(q) + q^m B_1(q) \hat{S}(q) = P_1(q) C(q)$$

$$\deg \hat{R}(q) = \deg \hat{S}(q) = n \text{ and } \hat{S}(0) = 0$$

furthermore

$$A(q) = A_1(q) A_2(q)$$

$$B(q) = B_1(q) A_2(q)$$

$$\tilde{B}(q) = B_1(q) A_2^T(q)$$

$$\text{Polynomial } P(q) \text{ is given by } P(q) = A_2^T(q) P_1(q)$$

$P_1(q)$ : Solution of spectral factorization

$$r p_1(q) p_1(q^{-1}) = P A_1(q) A_2^T(q) A_1(q^{-1}) A_2(q^{-1}) + B_1(q) B_1(q^{-1})$$

$$\deg P_1(q) = \deg A_1(q) + \deg A_2^T(q)$$

$$T(q) \text{ is } T(q) = t_0 q^m C(q)$$

$$\text{Where, } t_0 = P_0(1) / B_1(1)$$

We can write the equation is

$$A(q)r(q) + B(q)S(q) = A_2(q)P_1(q)C(q)$$

LQG solution can thus be interpreted as a pole-placement controller, where the poles are positioned at the zeroes of  $A_2$ ,  $P$  and  $C$ . This is an example of “internal model principle. The solution to the LQG problem given is closely related to the pole placement design problem.

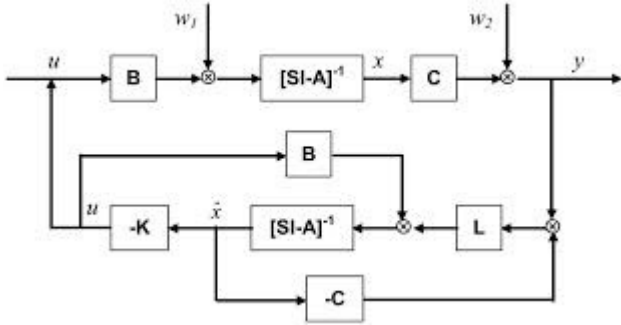


Fig 3 Control structure

Alternative solution to the design problem is to use a state space formulation process model is:

$$x(t+1) = \bar{A}x(t) + \bar{B}u(t) + \bar{k}e(t)$$

$$y(t) = \bar{C}x(t) + e(t)$$

$\bar{A}$ ,  $\bar{B}$ ,  $\bar{C}$  and  $\bar{K}$  are given in canonical form

$$\bar{A} = \begin{bmatrix} -a_1 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ -a_n & 0 & \dots & 0 & 0 \end{bmatrix}$$

$$\bar{B} = [0 \dots 0 \quad b_0 \dots b_m]^T$$

$$\bar{C} = [1 \ 0 \ \dots \ 0]$$

$$\bar{K} = [c_1 - a_1 \dots c_n - a_n]^T$$

where  $m = n - d_0$

$K$  is optimal steady state gain in kalman filter

$$\hat{x}(t+1) = x(t+1)$$

$$\hat{x}(t/t) = qI - \bar{A} + \bar{k}\bar{C}^{-1} [Bu(t) + \bar{k}y(t)]$$

$$\det(qI - \bar{A} + \bar{K}\bar{C}) = C(q)$$

$$Jx = E \{ \sum_{t=1}^N x^T(t) Q_1 x(t) + p u^2(t) + x^T(N) Q_0 x(N) \}$$

Optimal controller is  $u(t) = -L(t) \hat{X}(t/t)$

$L(t)$ : time-varying feedback gain through a Riccati equation.

$$S(t) = [\bar{A} - \bar{B}L(t)]^T S(t+1) (\bar{A} - \bar{B}L(t)) + Q_1 + P L^T(t) L(t)$$

$$L(t) = (P + \bar{B}^T S(t+1) \bar{B})^{-1} \bar{B}^T S(t+1) \bar{A}$$

$$S(N) = Q_0 \quad // \text{ Limiting controller,}$$

$$\bar{L} = \lim_{t \rightarrow \infty} L(t)$$

Closed-loop characteristic equation is

$$P(q) = \det(qI - \bar{A} + \bar{B}\bar{L}) = 0$$

The two solutions to the LQG control problem suggest two ways to construct indirect LQG-STR.

### III. M-FILE PROGRAM

The computation model of LQG STR is simulated in the given program

```
% Enter the continuous plant model
M = 1; m = 0.1; k = 0.091; b = 0.0036;
A = [ 0   1   0   0
      -k/m -b/m k/m  b/m
      0   0   0   1
      k/M  b/M -k/M -b/M];
B = [0; 0; 0; 1/M];
C = [1 0 0 0
      0 0 1 0
      0 0 0 0]; % get d, y and u as outputs
D = [0; 0; 1];
% Create the LTI continuous model
% plant = ss(A, B, C, D, 'InputName', 'Force F (N)', 'OutputName',
% {'Displacement d (m)', 'Displacement y (m)', 'Force F (N)'});
% set('plant', 'StateName', {'d', 'd_dot', 'y', 'y_dot'})
% Create the Discrete model with a sampling time of T=0.4 s.
T = 0.4;
ZOH = c2d(plant, T, 'zoh');
% Retrieve the matrices
[Phi, Gamma, Cd, Dd] = ssdata(ZOH);
% Design the regulator by computing the LQR Gain matrix K
```



% Set R = [1] arbitrarily and scale the Q matrix accordingly.

% Make Q diagonal

R = 1;

Q11 = input('LQR controller Q\_11: ');

Q22 = input('LQR controller Q\_22: ');

Q = diag([Q11 0 Q22 0]);

K = dlqr(Phi, Gamma, Q, R);

% Compute the Kalman filter gains

% Assumerms noise of 1% on each sensor channel

Rv = 0.01^2 \* eye(2);

% Input Rw

Rw = input('Enter estimator Rw: ');

sensors = [1, 2]; % d and y are sensed

known = [1]; % force u

P = ss(Phi, [Gamma Gamma], C, [D D], T);

[Observer, Ko] = kalman(P, Rw, Rv, [], sensors, known);

% Create the regulator and the closed-loop system

lqg\_reg = lqgreg(Observer, K, 'current');

feedin = [1]; % force u

feedout = [1, 2]; % d and y

Gcl = feedback(ZOH, lqg\_reg, feedin, feedout, +1);

% Compute and plot the initial condition response

% Set x\_1(0) = 1, all others to zero.

x0 = zeros(8,1);

x0(1) = 1;

figure(1), clf

initial(Gcl, x0)

% Plot all states and compare actual values with Kalman estimates

[y, t, x] = initial(Gcl, x0, 10);

figure(2), clf

subplot(2,2,1), stairs(t,x(:, [1 5])), grid, legend('x\_1=d', 'x\_1hat', 0)

title('Response of states and predictive estimates to x\_1(0) = 1'), ...

xlabel('Time (s)')

subplot(2,2,2), stairs(t,x(:, [2 6])), grid, legend('x\_2=ddot', 'x\_2hat', 0)

xlabel('Time (s)')

subplot(2,2,3), stairs(t,x(:, [3 7])), grid, legend('x\_3=y', 'x\_3hat', 0)

xlabel('Time (s)')

subplot(2,2,4), stairs(t,x(:, [4 8])), grid, legend('x\_4=ydot', 'x\_4hat', 0)

xlabel('Time (s)')

RESULTS:

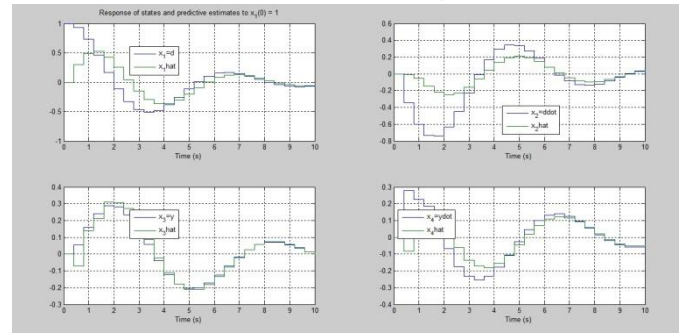


Fig 4 Response of predictive Estimators

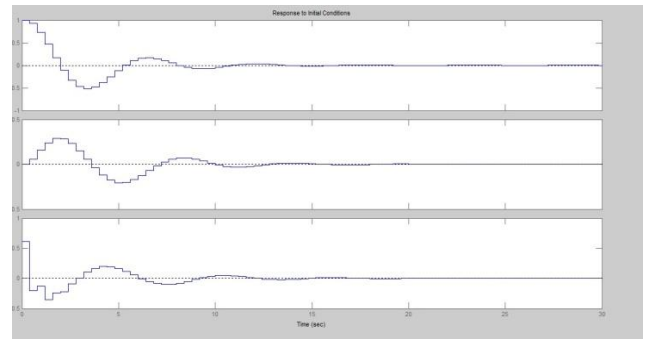


Fig 5 Response to the initial conditions

#### IV CONCLUSION

The case of pole placement technique is that it places the poles at desired locations using state feedback gain matrix. Using feedback, the poles of the system can be shifted so we can shape the closed loop characteristics of system to meet the design requirement. The pole placement method can give the desired characteristic but not a robust system. So controller design using pole placement may not be insensitive to system parameter variations and external disturbance. Hence, Design of LQG regulator was presented for robustness of system. In m-file programming the weighing matrix is designed and results are presented.

#### References

- [1] Linear Quadratic Gaussian Control, Gitkomut Songsiri, W65\_Control System Theory
- [2] Computer Control System Methods, Rezakatebi, Dept of Electronics and Electrical Engineering EE909
- [3] Solution of Simple Diaphonite Equation by means of matlab, Vladimir Hanta Institute of Chemical Technology, Dept of Computing and Control Engineering
- [4] Advanced System Dynamics and Control, D.Rowell, Massachusetts Institute of Technology, Dept of Mechanical Engineering
- [5] Self Tuning and Adaptive Control by Gray Evens, Dept of Computing Engineering.
- [6] The Role and Use of the Linear-Quadratic-Gaussian Control System Design by Michael Athuts mm Transactions On Automatic Control, Vol. Ac-16, No. 6, December 1971.

# Analysis, Design and Implementation of a Bidirectional Double-Boost DC-DC Converter with Coupled Inductor Technique

S. Juri Sowjanya<sup>1</sup>, K. Ramesh<sup>2</sup>

<sup>1</sup>PG Scholar, Power Electronics and Electrical Drives, Department of EEE, MREC(A)

<sup>2</sup>Assistant Professor, Department of EEE, MREC(A)

<sup>1</sup>jurisowjanya@gmail.com

<sup>2</sup>kramessheee@mrec.ac.in

**Abstract**— This paper studies a bidirectional DC-DC converter with high conversion ratio for renewable energy systems. The coupled-inductor technique is used to achieve a high conversion ratio with very simple control circuits. In discharging mode, the converter acts as a two-stage boost converters, controlling one power switch to achieve high voltage step-up conversion. In charging mode, the converter acts as two cascaded buck converters that control two power switches simultaneously to achieve high voltage step-down conversion. The operating principles and analysis of the steady-state characteristics are discussed in great detail. Finally, a circuit 24/78V is simulated in MATLAB/SIMULINK.

**Index Terms:** Bidirectional converter, high conversion ratio, coupled-inductor.

## I. INTRODUCTION

The bidirectional DC-DC converter is widely used in renewable energy applications. This converter is able to transfer or balance energy between two different DC sources, such as fuel cell and battery hybrid supplied power systems, island photovoltaic generation systems, and wind power systems [1]–[4]. The bidirectional DC-DC converter can be applied in uninterruptible power supplies (UPSs) to transfer the energy between source and battery [5]. The bidirectional DC-DC converter plays an important role in system back up or in reserving energy for the battery. Figure 1 shows the application of this converter in a hybrid renewable energy supply system. The battery can balance the energy between the power source and the load. The voltage difference between the battery and the DC bus is large, thus, a bidirectional DC-DC converter with high step-up/down voltage conversion ratio is required. The conventional boost/buck bidirectional converter is not suitable in such applications because the conversion ratio will be significantly reduced by parasitic elements. The active components of the single-stage cascaded bidirectional DC-DC converter will suffer higher current stress and conduction losses [15]. The bidirectional converter used coupled-inductor technology to achieve a high voltage conversion ratio [16]–[18]. However, the energy stored in the leakage inductor of the coupled inductor causes a high voltage spike on the power switches [19]–[21].

The bidirectional converter being studied is constructed of a dual boost/buck converter to achieve a high voltage conversion ratio by employing a coupled-inductor technique. Fig. 2 shows the configuration of the bidirectional converter, which has the following features: 1) the converter achieves high voltage conversion ratio at step-up or step-down stage; 2) solitary control with signal in either step-up or step-down operating condition, an effectively simplified control circuit; 3) the leakage-Inductance energy of the coupled inductor is recycled, thus reducing the voltage stress on power switches; 4) a low RDS-ON switch can be selected to improve system efficiency.

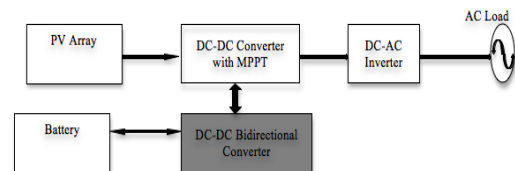


Fig. 1 Renewable energy hybrid supply system.

## II. CONFIGURATION OF BIDIRECTIONAL CONVERTER

The converter is used for the bidirectional transfer of energy between the low voltage side VL, which is connected to a 24 V battery, and the high voltage side VH, which is connected to a 78 V DC bus. Fig. 2 shows the configuration of the bidirectional converter.

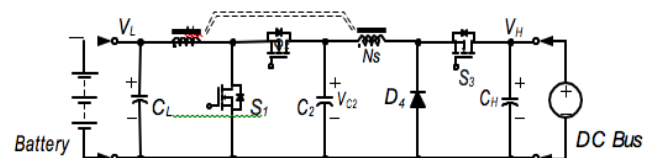


Fig. 2 shows the configuration of the bidirectional converter.

The following conditions were assumed in analyzing the steady-state characteristics of the proposed converter:

1. All the circuit components are ideal.
2. The capacitors  $C_L$ ,  $C_2$ , and  $C_H$  are large enough, and the voltages can be treated as constant.
3. The magnetizing inductance  $L_M$  of the coupled inductor is large enough, and the converter is operated in continuous conduct mode (CCM).

The component parameter of coupled inductor turns ratio  $N_p : N_s$  is 1:3,  $L_m/L_k$  is 8/24. In discharging mode, the converter acts as a two-stage boost converters, controlling one power switch to achieve high voltage step-up conversion. The power switch  $S_1$  is the main power switch. The switches  $S_2$  and  $S_3$  are off during the entire period. In charging mode, the converter acts as two cascaded buck converters that control two power switches simultaneously to achieve high voltage step-down conversion. Power switches  $S_2$  and  $S_3$  are controlled simultaneously and switch  $S_1$  is off.

### III. SIMULATION RESULTS:

The specifications and component parameters are as follows:  $V_L = 24$  V,  $V_H = 78$  V,  $f_s = 50$  kHz,  $n = 3$ ,  $L_m = 37\mu\text{H}$ ,  $C_L = 220\mu\text{F}$ , and  $C_2 = C_H = 300\mu\text{F}$ . The circuit is simulated in MATLAB/SIMULINK. The following is the simulation circuit in discharging mode.

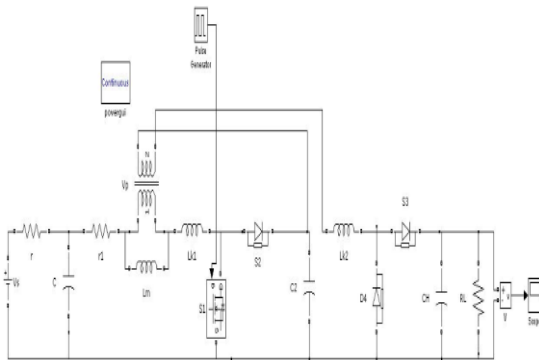


Fig 3. Simulation circuit in discharging mode

The output in discharging mode is shown in Fig 4. It shows the output 24V is boosted up to 78V.

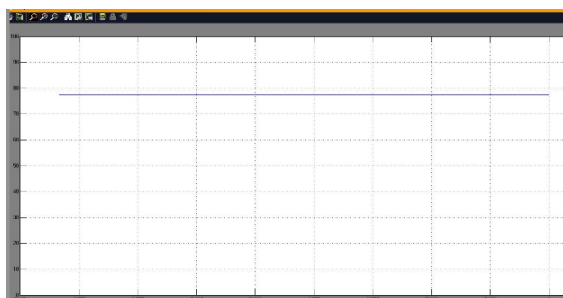
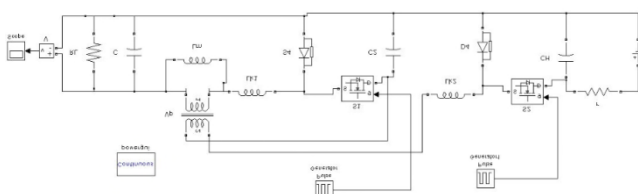


Fig 4. Output in discharging mode

The following is the simulation circuit in charging mode.



The output in charging mode is shown in Fig 6.

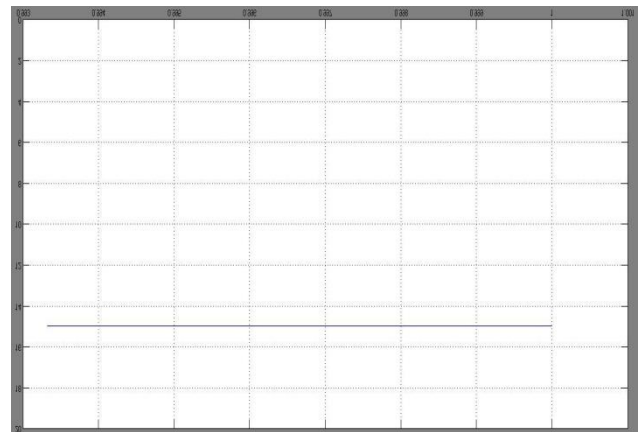


Fig 6. Output in charging mode

### IV. CONCLUSIONS

This paper has studied a bidirectional DC-DC converter for renewable energy systems. The converter can achieve high conversion ratio using the coupled-inductor technique.

### V. REFERENCES

- [1] T. Bhattacharya, V. S. Giri, K. Mathew, and L. Umanand, "Multiphase bidirectional flyback converter topology for hybrid electric vehicles," *IEEE Trans. Ind. Electron.*, vol. 56, no. 1, pp. 78–84, Jan. 2009.
- [2] Z. Amjadi and S. S. Williamson, "A novel control technique for a switched-capacitor-converter-based hybrid electric vehicle energy storage system," *IEEE Trans. Ind. Electron.*, vol. 57, no. 3, pp. 926–934, March. 2010.
- [3] F. Z. Peng, F. Zhang, and Z. Qian, "A magnetic-less DC-DC converter for dual-voltage automotive systems," *IEEE Trans. Ind. Appl.*, vol. 39, no. 2, pp. 511–518, Mar. 2003.
- [4] L. A. Flores, O. Garcia, J. A. Oliver, and J. A. Cobos, "High-frequency bidirectional DC/DC converter using two inductor rectifier," in *Proc. IEEE IECON Conf.*, pp. 2793–2798, Nov. 2006.
- [5] M. A. Abusara, J. M. Guerrero, and S. M. Sharkh, "Line-Interactive UPS for Microgrids," *IEEE Trans. Ind. Electron.*, vol. 61, no. 3, pp. 1292–1300, March. 2014.
- [6] N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics: Converters, Applications and Design, Third Edition*, John Wiley & Sons, Inc., 2003.
- [7] K. Yamamoto, E. Hiraki, T. Tanaka, M. Nakaoka, and T. Mishima, "Bidirectional DC-DC converter with full-bridge / push-pull circuit for automobile electric power systems," in *Proc. IEEE PESC Conf.*, pp. 1–5, June. 2006.
- [8] G. Chen, Y. S. Lee, S. Y. Hui, D. Xu, and Y. Wang, "Actively clamped bidirectional flyback converter," *IEEE Trans. Ind. Electron.*, vol. 47, no. 4, pp. 770–779, Aug. 2000.
- [9] F. Z. Peng, H. Li, G. J. Su, and J. S. Lawler, "A new ZVS bidirectional DC-DC converter for fuel cell and battery application," *IEEE Trans. Power Electron.*, vol. 19, no. 1, pp. 54–65, Jan. 2004.
- [10] B. R. Lin, J. J. Chen, and F. Y. Hsieh, "Analysis and implementation of a bidirectional converter with high conversion ratio," in *Proc. IEEE ICIT Conf.*, pp. 1–6, April. 2008.
- [11] L. S. Yang, T. J. Liang, H. C. Lee, and J. F. Chen, "Novel High Step-Up DC-DC Converter with Coupled-Inductor and Voltage-Doubler Circuits," *IEEE Trans. Ind. Electron.*, vol. 58, no. 9, pp. 4196–4206, Sept. 2011.
- [12] K. I. Hwu and Y. T. Yau, "A Buck Resonant Voltage Divider with Bidirectional Operation Considered," *IEEE Trans. Ind. Appl.*, vol. 49, no. 4, pp. 1566–1576, July-Aug. 2013.
- [13] C.-C. Lin, L.-S. Yang, G.-W. Wu, "Study of a non-isolated bidirectional



- [14] R.-J. Wai1, R.-Y. Duan, and K.-H. Jheng, "High-efficiency bidirectional dc-dc converter with high-voltage gain," *IET Power Electron.*, vol. 5, no. 2, pp. 173-184, 2012.
- [15] P. Das, A. Mousavi, and G. Moschopoulos, "A novel ZVS-PWM DC-DC converter for bidirectional applications with steep conversion ratio," in *Proc. IEEE ECCE*, Sep. 2009, pp. 2030-2036.
- [16] C. M. Hong, L. S. Yang, T. J. Liang, and J. F. Chen, "Novel bidirectional DC-DC converter with high step-up/down voltage gain," in *Proc. IEEE ECCE*, Sep. 2009, pp. 60-66.
- [17] T. M. Chen and C. L. Chen, "Analysis and design of asymmetrical half bridge flyback converter," *Proc. IEE—Elect. Power Appl.*, vol. 149, no. 6, pp. 433-440, Nov. 2002.
- [18] G. Y. Jeong, "High efficiency asymmetrical half-bridge flyback converter using a new voltage-driven synchronous rectifier," *IET Power Electron.*, vol. 3, no. 1, pp. 18-32, 2010.
- [19] L. S. Yang, T. J. Liang, and J. F. Chen, "Transformerless DC-DC converters with high step-up voltage gain," *IEEE Trans. Ind. Electron.*, vol. 56, no. 8, pp. 3144-3152, Aug. 2009.
- [20] C. M. Lai, C. T. Pan, and M. C. Cheng, "High-efficiency modular high step-up interleaved boost converter for DC-microgrid applications," *IEEE Trans. Ind. Appl.*, vol. 48, no. 1, pp. 161-171, Jan./Feb. 2012.
- [21] S.M.Chen,T.J.Liang, L.S.Yang,andJ.F.Chen,"A cascaded high step- up DC-DC converter with single switch for microsource applications," *IEEE Trans. Power Electron.*, vol. 26, no. 4, pp. 1146-1153, Apr. 2011.



Ms.Juri Sowjanya received her B.Tech degree in Electrical and Electronics from Koneru Lakshmaiah College of Engineering (Autonomous) in 2011. She is currently pursuing her M.Tech in Power electronics and Electrical drives in Mallareddy Engineering College (Autonomous), JNTUH



Mr.K.Ramesh is working as an Assistant Professor in Mallareddy Engineering College (Autonomous). He received his B.Tech degree in Electrical and Electrons form Mother Teresa Engineering College affiliated to JNTUH. He received his M.Tech degree with specialization in Power Industrial Drives from Adams Engineering College affiliated to JNTUH. He has published three papers in Power Industrial drives.

# ADAPTIVE FUZZY LOGIC CONTROLLER FOR PERMANENT MAGNET SYNCHRONOUS MOTOR WITH NON LINEAR FRICTION

G. Suguna<sup>(1)</sup>, Bibi mariyam<sup>(2)</sup>, K.Anitha reddy<sup>(3)</sup>

<sup>1</sup>Assoc. professor & MREC(A)

<sup>2</sup>Assistant Professor & MREC(A)

<sup>3</sup>Assistant Professor & MREC(A)

<sup>1</sup>sugunagoutakatla@gmail.com

<sup>2</sup>bibimariyam@ymail.com

<sup>3</sup>anikondam@gmail.com

**Abstract**—Variable-speed permanent magnet synchronous machines (PMSMs) are widely used in many high performance applications, such as electric vehicles, wind turbines, and robotics. The PMSM can be thought of as a cross between an AC induction motor and a brushless DC motor (BLDC). They have rotor structures similar to BLDC motors which contain permanent magnets. Field Oriented Control is the most popular control technique used with PMSMs. As a result, torque ripple can be extremely low, on par with that of ACIMs.[4][5] However, PMSM motors provide higher power density for their size compared to ACIMs. The rotor flux is already established in a PMSM by the permanent magnets on the rotor.

**Keywords**—PMSMs(permanent magnet synchronous machines),BLDC,( brushless DC motor)ACIMs(AC INDUCTION MOTOR),AFC (ADAPTIVE FUZZY LOGIC CONTROLLER)

## I. INTRODUCTION

The saliency exhibited by IPM motors can also provide an additional benefit in sensor less control applications. However, these tools suffer from a heavy computation. In this paper, we propose a Lyapunov stability based adaptation technique as an alternative to the conventional heuristic tuning methods. Thus, the stability of the proposed approach is guaranteed unlike many computational intelligence based controllers. Moreover, the proposed adaptive fuzzy logic controller (FLC) uses a computational efficient membership functions and operators to alleviate the computational burden associated with soft computing tools which makes it practically realizable.

## II. PMSM DYNAMICS

**PMSM Modelling:** The PMSM dynamic mathematical model[1] in the d-q axes rotational reference frame can be described by the following equations derived from PMSM equivalent circuit as shown in fig.

$$v_d = R i_d + L_d \frac{d}{dt} i_d - L_q p \omega i_q \quad (1a)$$

$$v_q = R i_q + L_q \frac{d}{dt} i_q + L_d p \omega i_d + p \lambda \omega \quad (1b)$$

$$\tau = \frac{3}{2} p [(L_d - L_q) i_d i_q + \lambda i_q] \quad (1c)$$

The mechanical equations of motion can be expressed by

$$\frac{d}{dt} \omega = 1/J (\tau - \tau_F - \tau_L) \quad (2a)$$

$$\frac{d}{dt} \theta = p \omega \quad (2b)$$

where

$v_d, v_q$  voltage in d-q axes;  
 $i_d, i_q$  current in d-q axes;  
 $L_d, L_q$  inductance in d-q axes;  
 $\tau_L$  load torque;  
 $\tau_F$  friction torque;  
 $J$  rotor and load inertia;  
 $\theta$  rotor electrical position;  
 $\omega$  rotor electrical speed

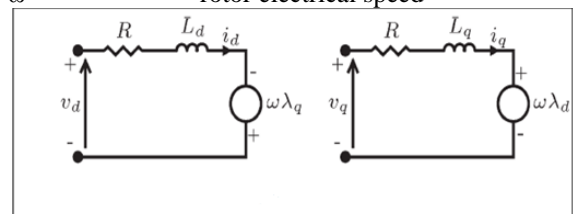


Fig 1. PMSM equivalent model

### (a) Friction

A nonlinear system, in contrast to a linear system, is a system which does not satisfy the superposition principle – meaning that the output of a nonlinear system is not directly proportional to the input. Nonlinearity is a common issue when examining cause-effect relations. Nonlinearity without explanation can lead to random, unforecasted outcomes. Examples of nonlinearities are dead zone, friction, etc.

Friction is generally described as the resistance to motion when two surfaces slide against each other. In most cases friction is a useful phenomenon making many ordinary things like walking and the brake in a car possible. On the other hand friction can also cause undesirable effects. For high precision mechanical motion systems for example, friction can deteriorate the performance of the system. Possible unwanted consequences caused by friction are steady-state errors, limit cycling and hunting.

### (b) Friction Modeling

The complexity of system (1) and (2) is increased even further by adopting a highly nonlinear a priori unknown friction model that is composed of Coulomb, viscous, and static friction terms. The model of such a memoryless friction is described by

$$\tau F = F_c \text{sign}(\omega) + F_v \omega + F_s \text{sign}(\omega) e^{-(\omega/\eta_s)^2}$$

where  $F_c$ ,  $F_v$ , and  $F_s$  are the Coulomb, viscous, and static friction parameters, respectively, and  $\eta_s$  is the rate of decay of the static friction term.

## III. FUZZY LOGIC SYSTEM

### (a) Fuzzy Logic

In this context, Fuzzy Logic is a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems. FL's approach[2] to control problems mimics how a person would make decisions, only much faster.

FL incorporates a simple, rule-based IF X AND Y THEN Z approach to a solving control problem rather than attempting to model a system mathematically. The FL model is empirically-based, relying on an operator's experience rather than their technical understanding of the system. For example, rather than dealing with temperature control in terms such as "SP =500F", "T <1000F", or "210C <TEMP <220C", terms like "IF (process is too cool) AND (process is getting colder) THEN (add heat to the process)" or "IF (process is too hot) AND (process is heating rapidly) THEN (cool the process quickly)" are used.

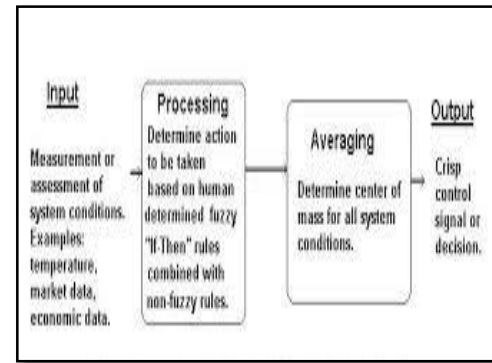


Fig 2: The Fuzzy logic control analysis method.

**(b) Fuzzification:** The fuzzification comprises the process of transforming crisp values into grades of membership for linguistic terms of fuzzy sets. This is achieved with the different types of fuzzifiers (membership functions). The simplest membership functions are formed using straight lines. The simplest is the triangular membership function, and it has the function name trimf. The trapezoidal membership function, trapmf, has a flat top and really is just a truncated triangle curve. These straight line membership functions have the advantage of simplicity. The membership function is used to associate a grade to each linguistic term.

For each input and output variable selected, we define two or more membership functions (MF), normally three but can be more. We have to define a qualitative category for each one of them, for example: low, normal or high. The shape of these functions can be diverse but we will usually work with triangles and trapezoids (actually usually pseudo-trapezoids) For this reason we need at least three (for triangles) or four (for trapezoids) points to define one MF of one variable.

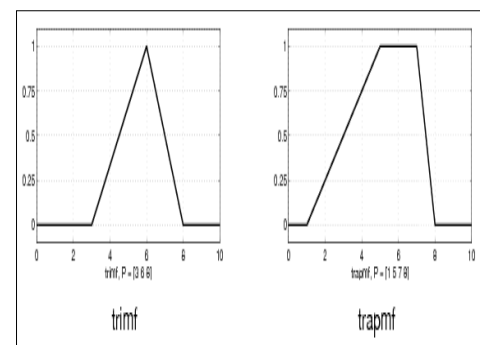


Fig 3: Membership functions

Fuzzy sets can be represented as membership function  $\mu_A$  that associates with each element  $x$  of the universe of discourse  $X$ , a number  $\mu_A(x)$ , i.e., membership grade, in the interval  $[0, 1]$ . The fuzzifier maps a crisp input  $x \in X$  into a fuzzified value  $A \in U$  (Universe).

- 1) Singleton fuzzification: fuzzy set A with support  $x_i$ , where  $\mu_A(x_i) = 1$ , for  $x = x_i$  and  $\mu_A(x_i) = 0$ , for  $x \neq x_i$ .
- 2) Non-singleton fuzzification:  $\mu_A(x_i) = 1$ , for  $x = x_i$  and decreases from 1 to 0 when moving away from  $x = x_i$ .

### (c) Defuzzification

Defuzzification is the last step to get the final output crisp value as shown in fig4 i.e. fuzzy inference engine. It is the process of producing a quantifiable result in fuzzy logic, given fuzzy sets and corresponding membership degrees. It is typically needed in fuzzy control systems. These will have a number of rules that transform a number of variables into a fuzzy result, that is, the result is described in terms of membership in fuzzy sets.

A common and useful defuzzification technique is *centre of gravity*. First, the results of the rules must be added together in some way. The most typical fuzzy set membership function has the graph of a triangle.

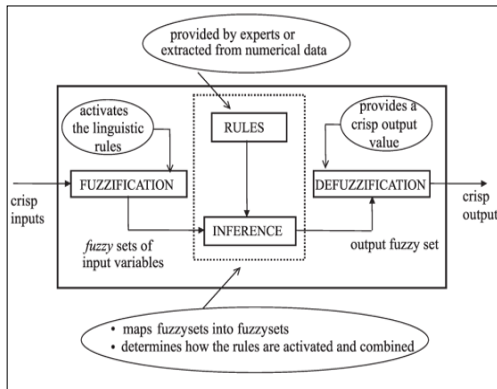


Fig 4: Fuzzy Inference Engine

Now, if this triangle were to be cut in a straight horizontal line somewhere between the top and the bottom, and the top portion were to be removed, the remaining portion forms a trapezoid.

## IV. ADAPTIVE CONTROL

**(a) Adaptive Control:** It covers a set of techniques which provide a systematic approach for automatic adjustment of controllers in real time, in order to achieve or to maintain a desired level of control system performance when the parameters of the plant dynamic model are unknown and/or change in time.

Consider the case, when the parameters of the dynamic model of the plant to be controlled are unknown but constant. In such cases, although the structure of the controller will not depend in general upon the particular values of the plant model parameters, the correct tuning of the controller parameters cannot be done without

knowledge of their values. Adaptive control techniques can provide an automatic tuning procedure in closed loop for the controller parameters. In such cases, the effect of the adaptation vanishes as time increases. Further insight into the operation of an adaptive control system can be gained if one considers the design and tuning procedure of the “good” controller.

The unknown and immeasurable variations of the process parameters degrade the performances of the control systems. Similarly to the disturbances acting upon the controlled variables, one can consider that the variations of the process parameters are caused by disturbances acting upon the parameters (called parameter disturbances). These parameter disturbances will affect the performance of the control systems. Therefore the disturbances acting upon a control system can be classified as follows: (a) disturbances acting upon the controlled variables; (b) (parameter) disturbances acting upon the performance of the control system.

A similar conceptual approach can be considered for the problem of achieving and maintaining the desired performance of a control system in the presence of parameter disturbances. The output of the adaptation mechanism will act upon the parameters of the controller and /or upon the control signal in order to modify the system performance accordingly.

For example, as an aircraft flies, its mass will slowly decrease as a result of fuel consumption; a control law is needed that adapts itself to such changing conditions. Adaptive control is different from robust control in that it does not need a priori information about the bounds on these uncertain or time-varying parameters; robust control guarantees that if the changes are within given bounds the control law need not be changed, while adaptive control is concerned with control as shown in fig5.

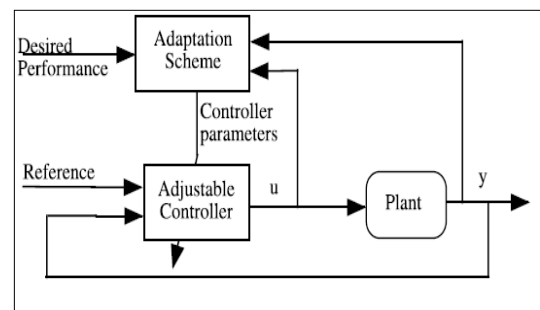


Fig 5: Adaptation Scheme

An adaptive control system measures a certain performance index (IP) of the control system using the inputs, the states, the outputs and the known disturbances. From the comparison of the measured performance index and a set of

given ones, the adaptation mechanism modifies the parameters of the adjustable controller and/or generates an auxiliary control in order to maintain the performance index of the control system close to the set of given ones (i.e., within the set of acceptable ones)[4].

When an adaptation loop is added, the damping of the system response will be maintained when changes in parameters occur. The operation of the adaptation loop and its design relies upon the following fundamental hypothesis: For any possible values of plant model parameters there is a controller with a fixed structure and complexity such that the specified performances can be achieved with appropriate values of the controller parameters. Therefore, the task of the adaptation loop is solely to search for the “good” values of the controller parameters.

### (b) Parameter Estimation

The foundation of adaptive control is parameter estimation [6]. Common methods of estimation include recursive least squares and gradient descent. Both of these methods provide update laws which are used to modify estimates in real time (i.e., as the system operates). Lyapunov stability is used to derive these update laws and show convergence criterion (typically persistent excitation). Projection (mathematics) and normalization are commonly used to improve the robustness of estimation algorithms.

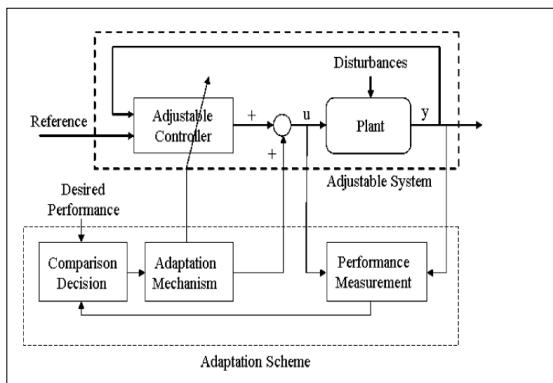


Fig 6: Parameter estimation in adaptive control strategy

The challenges to adaptive control of a machine with friction are not unlike the general challenges of adaptive control: The stability and convergence proofs are obtained from the so-called ideal problems, Those restrictive assumptions about the systems to be controlled may not hold in a practical case.

Adaptive control can only offer a solution when the structure of the plant dynamics and the disturbances that act on it are available. When the model is not available or when many parameters cannot be determined, learning

control may be considered. Therefore, only a limited set of parameters or the parameters of a simplified model will be adapted on-line.

### V. ADAPTIVE FUZZY CONTROL STRATEGY

In indirect adaptive fuzzy control, the fuzzy logic systems are used to model the plant. Then a controller is constructed assuming that the fuzzy logic system approximately represents the true plant. Feedback linearization techniques for nonlinear control system design have been developed in the last two decades. These signals are quantized into eight levels.

The rules are based on three hypotheses:

- When the input signals are far from their respective nominal zero-valued surfaces, then the FLC's output assumes a high value;
- when the inputs are approaching the nominal zero-valued surfaces, the output is adjusted to a smaller value for a smoother approach;
- once the inputs are on the nominal zero-valued surfaces, then the output is set to zero. This way, the FLC forces the machine's speed error  $e_\omega$  and its derivative  $e'_\omega$  to approach zero.

The control strategy of PMSM with adaptive fuzzy controller is shown in fig7.

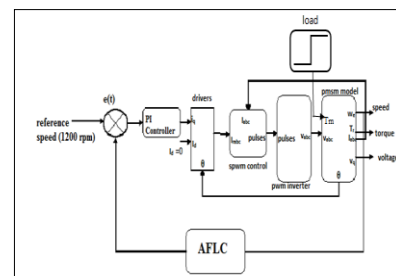


Fig 7: control strategy of PMSM with AFLC

### VI. RESULT ANALYSIS

Table 1: Motor Parameters

Parameter	Value
Nominal power (kW)	$P_n = 26$
Nominal torque (N-m)	$\tau_n = 416$
Nominal speed (RPM)	$\omega_n = 600$
Inductance in d-axis (H)	$L_d = 15.9 \cdot 10^{-3}$
Inductance in q-axis (H)	$L_q = 24.88 \cdot 10^{-3}$
Armature winding resistance ( $\Omega$ )	$R = 361.45 \cdot 10^{-3}$
Flux linkage (Wb)	$\lambda = 1.6504$
Coulomb friction coefficient (N-m)	$F_c = 1$
Viscous friction coefficient (N-m-s/rad)	$F_v = 2$
Static friction coefficient (N-m)	$F_s = 7 \cdot 10^{-1}$
Static friction decreasing rate (rad/s)	$\eta_s = 5 \cdot 10^{-2}$
Rotor and load inertia ( $\text{kg-m}^2$ )	$J = 5$
Number of pole pairs	$p = 5$

To demonstrate the performance of the proposed control scheme, a set of computer simulation runs is carried out on an interior PMSM model and has been used to design industrial controllers for PMSM based products such as hydroelectric and wind turbines[3]. In practice, a 480-V/60-Hz 3 $\phi$  ac source is fed to a rectifier through a  $\Delta Y$  transformer to generate a dc voltage of 800 V. For simulation purpose, a power supply is used to deliver the system with a constant nominal dc link voltage of 800 V. The dc link voltage drops are ignored, and the inverter is taken as an ideal.

The advantage behind the use of the adaptive fuzzy controller is clearly shown by very good tracking performance and negligible amplitude where the speed error is kept in, i.e., 0.01%, within sensors resolution. On the other hand, although d-q axes currents are not explicitly controlled by current loop regulators as in many conventional control techniques, the proposed controller was able to provide smooth control signal, which yields smooth d-q axes currents.

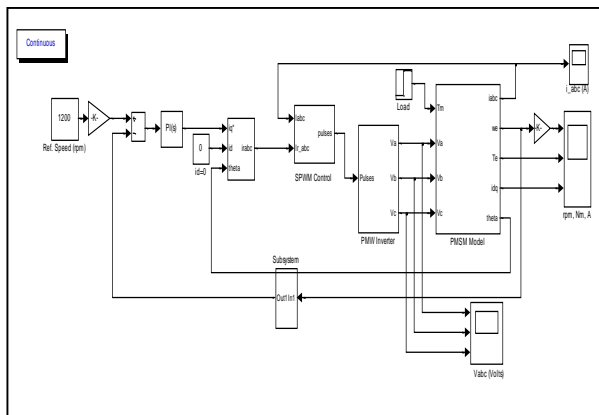


Fig 8: Simulation model with adaptive controller of PMSM

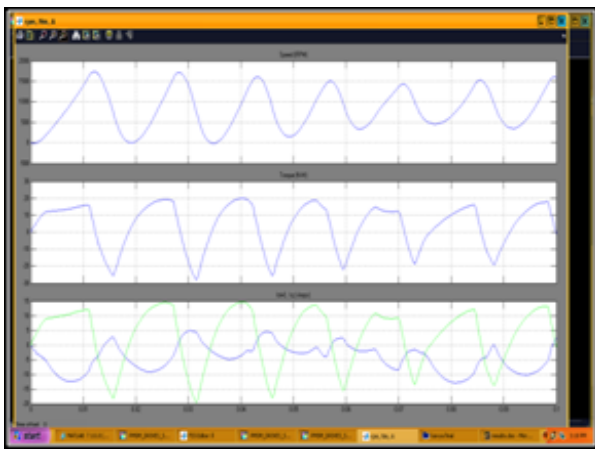


Fig 9: Simulation results of adaptive controller

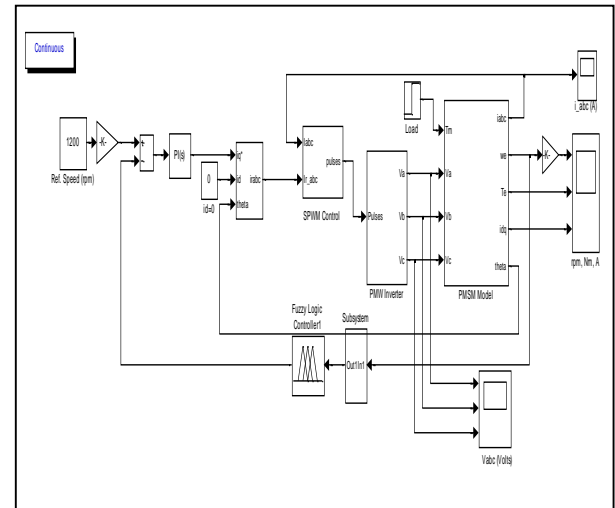


Fig 10: Simulation model with adaptive fuzzy controller of PMSM

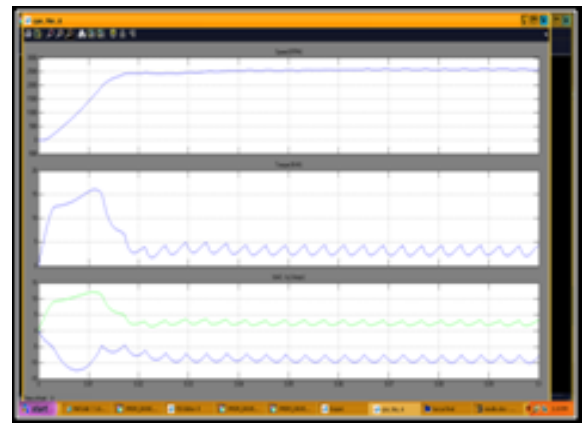


Fig 11: Simulation results with Adaptive fuzzy logic controller.

## VII. CONCLUSIONS

An adaptive fuzzy control strategy has been proposed for high-performance PMSM drives. In this control strategy, a fuzzy controller is combined with indirect adaptive controller to cope with structured and unstructured uncertainties, which yields robustness to parameters variation. Results show good performance for rotor speed tracking in transient, steady state, and standstill conditions. Startup and speed reversal transient operations demonstrate the effectiveness of the adaptive controller in dealing with these burdensome situations. It is noteworthy that this performance is achieved with a single controller rather than several controllers as in the vector control strategy. Hence, the adaptive fuzzy logic capabilities are instrumental in achieving control accuracy needed for high-performance drives, and their inherent parallelism makes them a good candidate for implementation in real-time PMSM drive systems.



## VIII. REFERENCES

- [1] F. Genduso, R. Miceli, C. Rando, and G. R. Galluzzo, "Back EMF sensorless-control algorithm for high-dynamic performance PMSM," *IEEE Trans. Ind. Electron.*, vol. 57, no. 6, pp. 2092–2100, Jun. 2010.
- [2] A. Sant and K. Rajagopal, "PM synchronous motor speed control using hybrid fuzzy-PI with novel switching functions," *IEEE Trans. Magn.*, vol. 45, no. 10, pp. 4672–4675, Oct. 2009.
- [3] M.-S. Rho and S.-Y. Kim, "Development of robust starting system using sensorless vector drive for a microturbine," *IEEE Trans. Ind. Electron.*, vol. 57, no. 3, pp. 1063–1073, Mar. 2010.
- [5] K. Basu, J. S. S. Prasad, G. Narayanan, H. K. Krishnamurthy, and R. Ayyanar, "Reduction of torque ripple in induction motor drives using an advanced hybrid PWM technique," *IEEE Trans. Ind. Electron.*, vol. 57, no. 6, pp. 2085–2091, Jun. 2010.
- [6] K. Gulez, A. Adam, and H. Pastaci, "A novel direct torque control algorithm for IPMSM with minimum harmonics and torque ripples," *IEEE/ASME Trans. Mechatron.*, vol. 12, no. 2, pp. 223–227, Apr. 2007.

**Department of  
Electronics and  
Communication Engineering**



# Suboptimal Comparison of AF and DF Relaying For Fixed Target Error Probability

D. Praveen Kumar

Assoc. Prof., Department of ECE  
Malla Reddy Engineering College  
Hyderabad, TS, India.

Dr. M. Sushanth Babu

Professor, Department of ECE  
Jayamukhi Institute of Technological  
Sciences,  
Warangal, TS, India.

Dr. M.S.G. Prasad

Professor, Department of ECE  
KL University,  
Vijayawada, AP, India.

**Abstract** - In this paper, multi-node hybrid Cooperative cellular communication is considered with Amplify and Forward (AF) and Decode and Forward (DF) relaying protocols. Most of researchers have evaluated the performance of cooperative communication either with DF or AF relaying protocols. Considering single DF relay at midway between Source and Destination, for fixed target error probability we have evaluated the performance of multiple AF relays with shortest path relay selection criteria. The analysis is further extended to compare the energy consumption per bit for selected multiple AF and single DF relays.

**Keywords:** Amplify and Forward, Decode and Forward, High Speed Uplink Packet Access, Target error Probability, Shortest path relay selection.

## I. INTRODUCTION

Wireless transmission suffers from random fluctuations in received signal at the destination known as fading. Diversity is one of the powerful techniques to mitigate fading. By using diversity technique, the transmitter sends more than one copy of the message to the receiver and it can use these multiple copies to detect the message correctly [1-2]. It is hardware complexity; cost and signal processing will be introduced if multiple antennas are used on mobile equipments. To mitigate these problems and maintain same performance, a new way of realizing diversity has been introduced, which is known as cooperative diversity [3].

In cooperative wireless communication, a source transmits a message to a destination with the assistance of a relay. The relay listens to the source's transmission and may retransmit the message to the destination. By combining the source and relay transmissions, and depending on the relaying protocol used, the destination can achieve diversity against fading without the use of an antenna array at any terminal [4-5].

An information-theoretical analysis of outage behavior has shown that a Decode-and-Forward (DF) relaying protocol achieve diversity by decoding and re-encoding transmitted message [4-6]. In the same way diversity is also achieved when the relay simply amplifies and forwards the message with Amplify-and-Forward relaying (AF). It is observed in [7-

8], the relay in mid way between source and destination provides optimal performance gain.

In the literature survey, most of the researchers concentrated on either AF or DF relaying protocols. The performance of AF is evaluated in multiple relay networks with various relay selection strategies. But it is observed that no researcher concentrated on suboptimal performance comparison of number of AF relays with mid way DF relay for fixed target error probability. In this paper we have consider the above problem and evaluated the performance by selecting multiple AF relays in multiple hops for the fixed target error probability of DF and the evaluation is further performed over energy consumption.

The paper organized as follows. Section II, we presents system modeling for the above consider assumption with multiple AF relays and mid way DF relay. Relay selection criteria with shortest path algorithm presented in section III. Symbol Error Rate and energy analysis of AF and DF for the fixed target error probability are presented in section IV. Section V and VI presents simulation results and conclusion.

## II. SYSTEM MODEL

Consider a relay based multi node hybrid cooperative wireless network with L hops as shown in Figure 1, which contain two set of relaying nodes operating with either amplify and forward and decode & forward relaying. All the wireless links in the network are modelled in Rayleigh fading with zero mean.

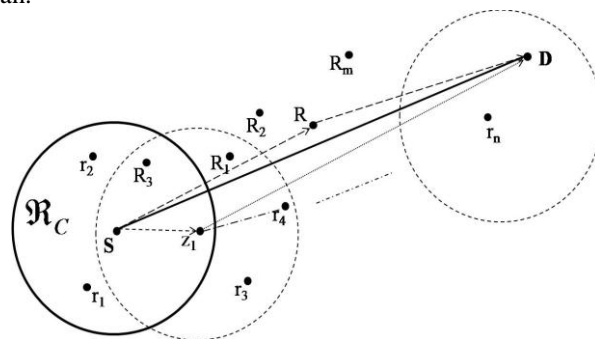


Fig 1. Multi Node Hybrid Cooperative Wireless Network

Due to the broadcast nature of the wireless communication channel, the symbols from the source is received by two sets of relaying nodes and destination (D) in phase-I. The received symbols at destination D and relaying nodes is

$$y_{SD} = \sqrt{Pd_{SD}^{-\beta}} h_{SD} x + \eta_{SD} \quad (1)$$

$$y_{Sk_i} = \sqrt{Pd_{Sk_i}^{-\beta}} h_{Sk_i} x + \eta_{Sk_i} \quad k_i \in R_i \text{ \& } r_i$$

where  $P$  is the transmitted source power,  $x$  represents information sequence  $x = \{x_1, x_2, x_3, \dots, x_N\}$ ,  $d_{SD}$ ,  $d_{Sk_i}$  are the distance between source - destination and source -  $i^{th}$  relay respectively. Channel gain between the source and the  $i^{th}$  relay terminal and the source and the destination are denoted by  $h_{Sk_i}$  and  $h_{SD}$  which are modelled as complex gaussian random variable i.e.,  $|h_{ij}|^2 = \sigma_{ij}^2 d_{ij}^{-4}$ , Where  $\sigma_{ij}^2$  and  $d_{ij}$  are the variance and distance between  $i$ - $j$  link.  $\eta_{SD}$ ,  $\eta_{Sk_i}$  represent additive noise.

In phase-II, the selected relay nodes in each set cooperates the source to transmit the symbols to destination  $D$  by performing respective relaying operation. The symbols from the source and relay nodes  $K_i$  ( $R_i$  &  $r_i$ ) are combined using Maximum Ratio Combiner (MRC) at the destination and it is given by

$$\begin{aligned} y_0 &= h_{SD}^* y_{SD} + h_{K_i,D}^* y_{K_i,D} \\ &= h_{SD}^* \sqrt{Pd_{SD}^{-\beta}} h_{SD} x + h_{SD}^* \eta_{SD} + h_{K_i,D}^* \sqrt{Pd_{K_i,D}^{-\beta}} h_{K_i,D} \hat{x} + h_{K_i,D}^* \eta_{K_i,D} \\ y_0 &= \sqrt{Pd_{SD}^{-\beta}} |h_{SD}| x + \sqrt{Pd_{K_i,D}^{-\beta}} |h_{K_i,D}| \hat{x} + \bar{\eta} \end{aligned} \quad (2)$$

Where  $\bar{\eta} = h_{SD}^* \eta_{SD} + h_{K_i,D}^* \eta_{K_i,D}$

### III. RELAY SELECTION CRITERIA

From Fig 1, two set of relaying nodes  $r_i$  and  $R_i$  which operates with amplify and forward and decode and forward relaying respectively. Using the results from [8], source select the one relay node ( $R$ ) from set  $R_i$  which is at mid location to obtain better performance, and the best relay selection from set  $r_i$  is based on shortest distance path [9]. The dotted line in Fig 1 indicates the Line of Sight, with strongest Received Signal Strength (RSS). Let the graphical area (coverage area) of all the nodes meet  $SER < \tau$  is  $R_c$ , where  $\tau$  represents Threshold Symbol Error Rate.

$$\mathcal{R}_c(S, r_i) = \{r \in r_i \mid SER(S, r_i) \leq \tau \vee r \cap D = 0\} \quad (3)$$

The best relay 'z' is the node which has the minimum distance from the  $k^{th}$  on LoS, 'k' i.e.,

$$z = \min\{d(\bar{r}_i, k)\} \quad \bar{r}_i \in \bar{r} \quad (4)$$

The selected relay ( $z_1$ ) in first hop amplifies the received signal and retransmits it to the destination, it is given by

$$y_{z_1 D} = \frac{\sqrt{Pd_{z_1 D}^{-\beta}}}{\sqrt{Pd_{S z_1}^{-\beta} |h_{S z_1}|^2 + 2\sigma_{S z_1}^2}} * h_{z_1 D} \left[ \sqrt{Pd_{S z_1}^{-\beta}} h_{S z_1} x + \eta_{S z_1} \right] \quad (5)$$

The selected relay in first hop amplifies the received signal and retransmits it to the destination, it is given by (6).

The destination D combines transmitted source sequence and amplified received sequence from selected  $\bar{r}_i$  of L hops using MRC combiner is given as (7).

$$\begin{aligned} y_{i=1}^L &= \left( \prod_{i=1}^L \frac{\sqrt{Pd_{i,i+1}^{-\beta}}}{\sqrt{Pd_{i-1,i}^{-\beta} |h_{i-1,i}|^2 + 2\sigma_{i-1,i}^2}} h_{i,i+1} \sqrt{Pd_{i-1,i}^{-\beta}} h_{i-1,i} \right) \\ &+ \left( \sum_{i=0}^L \left( \prod_{i=i+1}^L \frac{\sqrt{Pd_{i,i+1}^{-\beta}}}{\sqrt{Pd_{i-1,i}^{-\beta} |h_{i-1,i}|^2 + 2\sigma_{i-1,i}^2}} h_{i,i+1} \right) \eta_{i-1,i} \right) \\ y &= y_{0L} + \left( \prod_{i=1}^L \frac{\sqrt{Pd_{i,i+1}^{-\beta}}}{\sqrt{Pd_{i-1,i}^{-\beta} |h_{i-1,i}|^2 + 2\sigma_{i-1,i}^2}} h_{i,i+1} \sqrt{Pd_{i-1,i}^{-\beta}} h_{i-1,i} \right) \\ &+ \left( \sum_{i=0}^L \left( \prod_{i=i+1}^L \frac{\sqrt{Pd_{i,i+1}^{-\beta}}}{\sqrt{Pd_{i-1,i}^{-\beta} |h_{i-1,i}|^2 + 2\sigma_{i-1,i}^2}} h_{i,i+1} \right) \eta_{i-1,i} \right) \end{aligned} \quad (6)$$

### IV. ANALYSIS OF SER AND ENERGY CONSUMPTION

#### A. Evaluation of Symbol Error Rate

##### 1) Decode and Forward:

The closed-form expression for the Symbol Error Rate of Decode and Forward system with M-QAM at  $D$  is approximated as [8],

$$\begin{aligned} P_b &\approx \frac{2}{\log_2 M} \left[ \frac{\sqrt{M}-1}{\sqrt{M}} Q \left( \frac{P(|h_{SD}|^2 + |h_{RD}|^2)}{\sqrt{2\sigma_{\bar{\eta}}}} \sqrt{\frac{3}{2(M-1)}} \right) \right. \\ &+ Q \left( \frac{\sqrt{\frac{3}{2(M-1)}} |h_{SR}|}{\sqrt{2\sigma_{\bar{\eta}}}} \right) \left[ \frac{2\sqrt{M}-3}{2\sqrt{M}} Q \left( \frac{P(|h_{SD}|^2 + |h_{RD}|^2)}{\sqrt{2\sigma_{\bar{\eta}}}} \sqrt{\frac{3}{2(M-1)}} \right) \right. \\ &\left. \left. + \frac{\sqrt{M}-1}{2\sqrt{M}} Q \left( \frac{P(|h_{SD}|^2 - |h_{RD}|^2)}{\sqrt{2\sigma_{\bar{\eta}}}} \sqrt{\frac{3}{2(M-1)}} \right) \right] \right] \quad (8) \end{aligned}$$

Where noise variance

$$\sigma_{\eta}^2 = \left( \left| h_{0L}^* \sqrt{P_0^{df}} \right|^2 + \left| h_{rL}^* \sqrt{P_0^{df}} \right|^2 \right) \sigma_{\eta}^2$$

$$= P_0^{df} \left( |h_{0L}|^2 + |h_{rL}|^2 \right) \sigma_{\eta}^2$$

The bit error rate can be re-written as

$$P_b \approx \frac{2}{\log_2^M} \left\{ \frac{\sqrt{M}-1}{\sqrt{M}} Q \left[ \frac{P_0^{df} (\sigma_{0L}^4 d_{0L}^{-4} + \sigma_{rL}^4 d_{rL}^{-4})}{\sqrt{2} \sigma_{\eta}} \sqrt{\frac{3}{2(M-1)}} \right] + \right.$$

$$Q \left( \sqrt{\frac{3}{2(M-1)}} \frac{\sigma_{0L}^2 d_{0L}^{-2}}{\sqrt{2} \sigma_{\eta}} \right) \left[ \frac{2\sqrt{M}-3}{2\sqrt{M}} Q \left( \sqrt{\frac{3}{2(M-1)}} \frac{P_0^{df} (\sigma_{0L}^4 d_{0L}^{-4} + \sigma_{rL}^4 d_{rL}^{-4})}{\sqrt{2} \sigma_{\eta}} \right) \right.$$

$$\left. \left. + \frac{\sqrt{M}-1}{2\sqrt{M}} Q \left( \sqrt{\frac{3}{2(M-1)}} \frac{P_0^{df} (\sigma_{0L}^4 d_{0L}^{-4} - \sigma_{rL}^4 d_{rL}^{-4})}{\sqrt{2} \sigma_{\eta}} \right) \right] \right\} \quad (9)$$

#### 4.1.2 Amplify and Forward

The Symbol Error Rate (SER) of multi hop Amplify and forward system with M-QAM is approximated as

$$P_b = \frac{2}{\log_2^M} \frac{4BN_0^2}{P\sigma_{0L}^2} \sum_{i=0}^{L-1} \frac{1}{P\sigma_{i,i+1}^2} \quad (10)$$

Where  $B = \frac{3(M-1)}{8M} + \frac{\sin(\frac{2\pi}{M})}{4\pi} - \frac{\sin(\frac{4\pi}{M})}{32\pi}$

The received SNR at a distance  $d$  from transmitting node is given by

$$\rho_{i,i+1} = \frac{P|h_{i,i+1}|^2}{N_0} \quad (11)$$

$$= \frac{P\sigma_{i,i+1}^2 d_{i,i+1}^{-4}}{N_0} = \rho\sigma_{i,i+1}^2 d_{i,i+1}^{-4}$$

Where  $i=0,1,2,\dots,L-1$

Using (11) the bit error rate in (10) can remodelled as

$$P_b^{af} \approx \frac{4B}{\rho d_{SD}^{-4}} \sum_{i=1}^L \frac{1}{\rho d_{i,i+1}^{-4}} \quad (12)$$

#### 1.1 Evaluation of Energy consumption per Bit

As per selection criteria, the selected relays assumed to have same signal processing units of source. The total energy per bit required per hop with  $n_t$  number of transmitting and  $n_r$  number of receiving antennas, can be obtain by

$$E_L(i) = \sum_{i=1}^L (1 + \alpha_i) \rho N_0 \frac{(4\pi d_{L-1-L}(i))^n}{i G_t G_r \lambda^2} M_{li} N_{fi} \quad (13)$$

$$+ \frac{P}{R_b} (n_t + n_r)$$

where  $\alpha_i$  is the dependent parameter of drain efficiency of power amplifier in  $i^{\text{th}}$  hop,  $d_{L-1-L}(i)$  is the distance between  $L-1$  and  $L$  nodes,  $\rho$  is the signal power to noise ratio at each of the receive node,  $N_0$  is AWGN power spectral density,  $\lambda$  is the carrier wavelength,  $G_t$  and  $G_r$  are antenna gains of transmitter and receiver respectively,  $M_{li}$  is the  $i^{\text{th}}$  link margin compensating hardware process variations and other noise or interference,  $N_{fi}$  is the receiver noise figure,  $R_b$  is bit transmission rate.

#### V. SIMULATION RESULTS

For the demonstration of the proposed frame work, we assumed that the network is HSUPA, release-6, and the parameters used in the simulation are indicated in Table – 1.

TABLE I. SIMULATION PARAMETERS

Parameter	Specification
Band width	2000MHz
Data Rate	5.8Mbps
Transmitted Power of UE	21dBm
Number of bits per packet	1024
Modulation	M-QAM
Combining strategy	MRC

Fig 2 provides the SER performance comparison of multi hop AF relays with mid way DF relay for fixed target error rate. From the figure we conclude that, the SER performance of mid-way DF relay is approximately equal to 4-hop AF relaying performance.

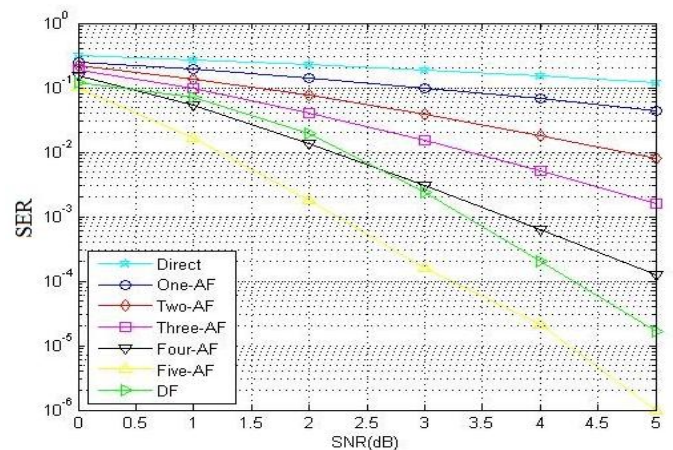


Fig 2. Performance of Multi-Hop AF relaying and mid way DF

Fig 3 demonstrates the energy consumption per bit for each hop of AF relaying protocol and mid way DF relay. From this we observed that the energy consumption is increased with number of hops of AF relay.

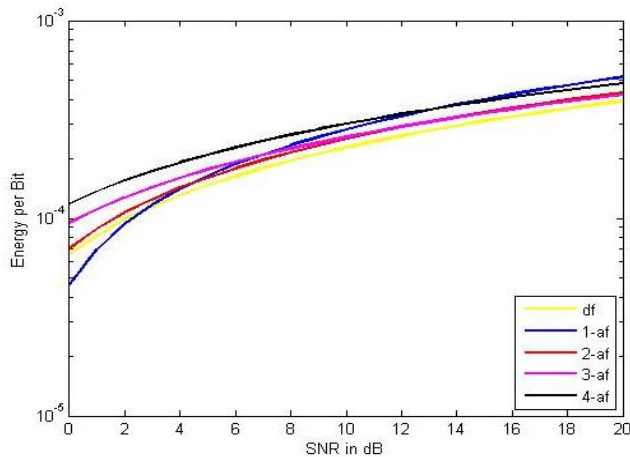


Fig 3. Evaluation of energy consumption per bit

Fig 4 depicts the end-to-end time delay comparison of 4-hop AF relaying and mid way DF; it shows that 4-hop AF relaying takes more time than mid way DF.

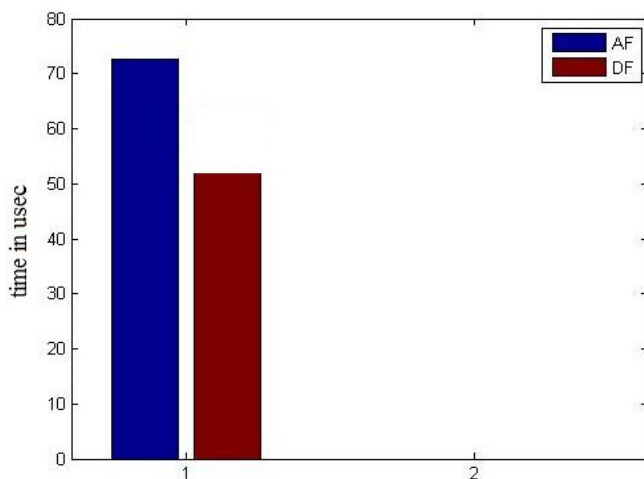


Fig 4. Comparison of time delay of 4-hop AF and mid way DF

## VI. CONCLUSION

In this paper, we present suboptimal performance comparison of multi hop AF relays with mid way DF relay for the fixed target error probability. From the simulation results we observed that the SER performance of DF is approximately equal to 4-hop AF relays and the energy consumption per bit and end-to-end time delay of 4-hop AF relay is high than DF.

## REFERENCES

- [1] "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels", Zheng and Tse, IEEE Trans Info Theory, 2003.
- [2] A. R. Calderbank, "The art of signaling: Fifty years of coding theory," IEEE Trans. Inform. Theory, vol. 44, pp. 2561–2595, Oct. 1998.
- [3] A.Sendonaris, E.Erkip and B.Aazhang, "User Cooperation Diversity Part I and Part II," IEEE Trans. Commun., vol. 51,no.11, Nov.2003,pp.1927-48.
- [4] J. Laneman, D. Tse, and G. Wornell, "Cooperative Diversity in Wireless Networks: Efficient protocols and Out-age Behavior," IEEE Trans. Info. Theory, vol. 50, no. 12, Dec. 2004, pp. 3062–80.
- [5] T. Wang et al., "High-Performance Cooperative Demodulation with Decode-and-Forward Relays," IEEE Trans. Commun., vol. 55, no. 7, July 2007, pp. 1427–38.
- [6] Yonghui Li, University of Sydney, "Distributed coding for cooperative wireless networks: An overview and recent advances" IEEE comm. Magazine, pp 71-77, Aug 2009.
- [7] A. Muller and J. Speidel, "Symbol Error Probability of M-QAM in Multi hop Communication Systems with Regenerative Relays," in IEEE 68th Vehicular Technology Conference (VTC'08-Spring), Singapore, May 2008, pp. 1004–1008
- [8] Meng Wu, Dirk Wubben, Armin Dekorsy, "BER-based Power Allocation for Decode-and-Forward Relaying with M-QAM Constellations" IEEE conf., Nov 2011.
- [9] B.R.Devi,K.K.Rao, "Co-operative Shortest Path Relay Selection for Multihop MANETs", springer intr. Publishing, FICTA 2014.

**Department of  
Computer Science and  
Engineering**



# Document Clustering with Concept Based Vector Suffix Tree Document Model

**Dr. N. Sandhya**

Professor, VNRVJIEET  
Email: sandhyanadela@gmail.com

**J Sireesha**

Assoc. Prof., MREC(A),  
Email: sireeshajasti@yahoo.co.in

**Subir Kumar Sarkar**

Dept. of ETCE,  
Jadavpur Kolkata, India University

**Abstract** – This work aims to extend a document representation model which is elegant by combining the versatility of the vector space model, the increased relevance of the suffix tree document model, which takes word ordering into account and retaining the relationship between words like synonyms. The effectiveness and the relevance of this document model can be evaluated by a partitioning clustering technique K-Means and then a systematic comparative study of the impact of similarity measures in conjunction with different types of vector space representation on cluster quality may be performed. This document model will be called the concept based vector suffix tree document model.

**Keywords** – Wireless Network, Failures, Checkpoints, Rollback Recovery.

## I. INTRODUCTION

The most popular way for representing documents is the VSM, because of its speed and versatility. TF-IDF score can be computed easily, so this model is fast. In the worst case the document needs to be traversed twice for computing document frequencies and TF-IDF values. To overcome the bag of words problems, text documents are treated as sequence of words and documents are retrieved based on sharing of frequent word sequences from text databases. The sequential relationship between the words and documents is preserved using suffix tree data structure [1]. Syntax based disambiguation is attempted by enriching the text document representations by background knowledge provided in a core ontology-WordNet.

This paper is organized as follows. The section 2 deals with the related work in text document clustering, section 3 describes the preprocessing, POS tagging and the use of WordNet for replacing words with their Synset IDs. Section 4 and Section 5 discusses the document representation, similarity measures and their semantics. Section 6 describes the steps in the CBVSTDM.

## II. RELATED WORK

An approach that combines suffix trees with the vector space model was proposed earlier [2] [3]. They use the same Suffix Tree Document Model proposed by Zamir and Etzioni but they map the nodes from the common suffix tree to a M dimensional space in the Vector Space Model. Thus, a feature vector containing the weights of

each node can be used to represent the documents. This model can be extended by attempting the word sense disambiguation using WordNet. Once the vector of weights is obtained, similarity measures and K-Means clustering algorithm with the Vector Space Model are applied. The lack of an effective measure for the quality of clusters which is an important problem with the original suffix tree is overcome with this model.

## III. DOCUMENTS PREPROCESSING USING WORDNET

Initially we need to preprocess the documents. This step is imperative. The next step is to analyze the prepared data and divide it into clusters using clustering algorithm. The effectiveness of clustering is improved by adding the Part-Of-Speech Tagging and making use of the WordNet for synsets. The most important procedure in the preprocessing of documents is to enrich the term vectors with concepts from the core ontology. WordNet [7] covers semantic and lexical relations between word forms and word meanings. The first preprocessing step is POS tagging. POS tagging is a process of assigning correct syntactic categories to each word. Tag set and word disambiguation rules are fundamental parts of any POS tagger. The POS tagger relies on the text structure and morphological differences to determine the appropriate part-of-speech. POS Tagging is an essential part of Natural Language Processing (NLP) applications such as speech recognition, text to speech, word sense disambiguation, information retrieval, semantic processing, parsing, information extraction and machine translation. WordNet contains only nouns, verbs, adjectives and adverbs. Since nouns and verbs are more important in representing the content of documents and also mainly form the frequent word meaning sequences, we focus only on nouns and verbs and remove all adjectives and adverbs from the documents. For those word forms that do not have entries in WordNet, we keep them in the documents since these unidentified word forms may capture unique information about the documents. We remove the stop words and then stemming is performed. The morphology function provided with WordNet is used for stemming as it only yields stems that are contained in the WordNet dictionary and also achieves improved results than Porter stemmer. The stemmed words are then looked up in the WordNet the lexical database to replace the words by their synset IDs. These preprocessing steps aim to improve the cluster

quality. These steps lead to the reduction of dimensions in the term space.

#### IV. DOCUMENT REPRESENTATION

The representation of a set of documents as vectors in a common vector space is known as the vector space model. Documents in vector space can be represented using Boolean, Term Frequency and Term Frequency – Inverse Document Frequency. In Boolean representation, if a term exists in a document, then the corresponding term value is set to one otherwise it is set to zero. Boolean representation is used when every term has equal importance and is applied when the documents are of small size. In Term Frequency and Term Frequency Inverse Document Frequency the term weights have to be set. The term weights are set as the simple frequency counts of the terms in the documents. This reflects the intuition that terms occur frequently within a document may reflect its meaning more strongly than terms that occur less frequently and should thus have higher weights. Each document  $d$  is considered as a vector in the term-space and represented by the term frequency (TF) vector:

$$d_{tf} = [tf_1, tf_2, \dots, tf_D] \quad (1)$$

where  $tf_i$  is the frequency of term  $i$  in the document and  $D$  is the total number of unique terms in the text database. The tf-idf representation of the document

$$d_{tf-idf} = [tf_1 \log(n/df_1), tf_2 \log(n/df_2), \dots, tf_D \log(n/df_D)] \quad (2)$$

To account for the documents of different lengths, each document vector is normalized to a unit vector (i.e.,  $\|d_{tf-idf}\|=1$ ). In the rest of this paper, we assume that this vector space model is used to represent documents during the clustering. Given a set  $C_j$  of documents and their corresponding vector representations, the centroid vector  $c_j$  is defined as:

$$c_j = \frac{1}{|C_j|} \sum_{d_i \in C_j} d_i \quad (3)$$

where each  $d_i$  is the document vector in the set  $C_j$ , and  $j$  is the number of documents in cluster  $C_j$ . It should be noted that even though each document vector  $d_i$  is of unit length, the centroid vector  $c_j$  is not necessarily of unit length.

#### V. SIMILARITY MEASURES

Document clustering groups similar documents to form a coherent cluster. However, the definition of a pair of documents being similar or different is not always clear and normally varies with the actual problem setting. Accurate clustering requires a precise definition of the closeness between a pair of objects, in terms of either the pair wise similarity or distance [20]. A variety of similarity or distance measures have been proposed and widely applied, such as cosine similarity, Jaccard coefficient, Euclidean distance and Pearson Correlation

Coefficient. The details of different similarity measures are described below.

##### 5.1 Cosine Similarity Measure

Classification The most commonly used is the cosine function [8]. For two documents  $d_i$  and  $d_j$ , the similarity between them can be calculated

$$\cos(d_i, d_j) = \frac{d_i \cdot d_j}{\|d_i\| \|d_j\|} \quad (4)$$

Where  $d_i$  and  $d_j$  are  $m$ -dimensional vectors over the term set  $T = \{t_1, t_2, \dots, t_m\}$ . Each dimension represents a term with its weight in the document, which is non-negative. As a Result the cosine similarity is non-negative and bounded between  $[0, 1]$ . Cosine similarity captures a scale invariant understanding of similarity and is independent of document length. When the document vectors are of unit length, the above equation is simplified to:

$$\cos(d_i, d_j) = d_i \cdot d_j$$

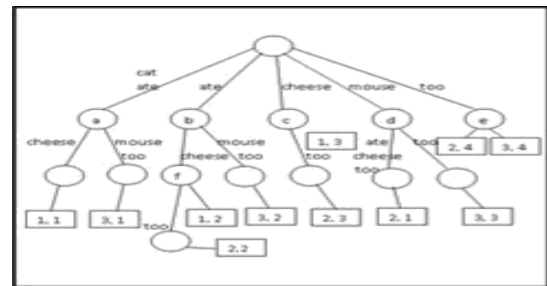


Fig.1. Suffix tree for strings “cat ate

When the cosine value is 1 the two documents are identical, and 0 if there is nothing in common between them. (i.e., their document vectors are orthogonal to each other).

##### 5.2 Jaccard Coefficient

The Jaccard coefficient, which is sometimes referred to as the Tanimoto coefficient [9, 10] measures similarity as the intersection divided by the union of the objects. For text document, the Jaccard coefficient compares the sum weight of shared terms to the sum weight of terms that are present in either of the two documents but are not the common terms.

$$\text{Jaccard Coff}(d_i, d_j) = \frac{d_i \cdot d_j}{\|d_i\|^2 + \|d_j\|^2 - d_i \cdot d_j}$$

$$\text{Jaccard Index}(d_i, d_j) = \frac{d_i \cap d_j}{d_i \cup d_j}$$

The Jaccard Coefficient ranges between  $[0, 1]$ . The Jaccard value is 1 if two documents are identical and 0 if the two documents are disjoint. The Cosine Similarity may be extended to yield Jaccard Coefficient in case of Binary attributes.

##### 5.3 Pearson Correlation Coefficient

Correlation is a technique for investigating the relationship between two quantitative, continuous variables, for example, age and blood pressure. Pearson's

correlation coefficient ( $r$ ) is a measure of the strength of the association between the two variables. There are different forms of Pearson Correlation Coefficient (PCC) formula. It is given by

Pearson Similarity( $d_i, d_j$ ) =

$$\frac{m \sum_k d_{ik} X d_{jk} - TF_i X TF_j}{\sqrt{[m \sum_k d_{ik}^2 - TF_i^2][m \sum_k d_{jk}^2 - TF_j^2]}} \quad (5)$$

Where  $TF_i = \sum_k d_{ik}$  and  $TF_j = \sum_k d_{jk}$   $m$  is the no. of terms in document  $d$ .

The measure ranges from +1 to -1. Positive correlation indicates that both variables increase or decrease together, whereas negative correlation indicates that as one variable increases, so the other decreases, and vice versa. When Pearson Similarity is  $\pm 1$  the two documents are identical and there is no relation between variables if it is equal to zero.

The Euclidean distance is a distance measure, while the cosine similarity, Jaccard coefficient and Pearson coefficient are similarity measures. We apply a simple transformation to convert the similarity measure to distance values. Because both cosine similarity and Jaccard coefficient are bounded in  $[0, 1]$  and monotonic, we take  $D = 1 - \text{SIM}$  as the corresponding distance value. For Pearson coefficient, which ranges from  $-1$  to  $+1$ , we take  $D = 1 - \text{SIM}$  when  $\text{SIM} \geq 0$  and  $D = |\text{SIM}|$  when  $\text{SIM} < 0$ .

## VI. STEPS TO BUILD CBVSTDM

Constructing a Suffix Tree Document Model Suffix tree document model considers a document  $d = w_1 w_2 \dots w_m$  as a string consisting of words  $w_i$  and not characters ( $i = 1; 2; \dots; m$ ). A suffix tree of document  $d$  is a compact trie containing all suffixes of document  $d$ . Figure 6.1 is an example of a suffix tree composed from three documents. The nodes of the suffix tree are drawn in circles. Each internal node has at least two children. Each edge is labeled with a non-empty substring of a document called a phrase, and its suffix node is labeled by the phrase too. Then each leaf node in the suffix tree designates a suffix of a document and each internal node represents an overlap phrase shared by at least two suffixes. The more internal nodes shared by two documents, the more similar the documents tend to be.

In Figure 1, each internal node is attached with a box respectively. In the practical implementation, each node including leaf node maintains a list storing the numbers displayed in the box. The numbers in the box designate the documents which have traversed the corresponding node. Each upper number designates a document identifier and the number below designates the traversed times of the document.

Figure 1: Suffix tree for strings “cat ate cheese”. Initially suffix trees were introduced by Weiner in 1973 [4]. They represent the single most important data structure for

string representation. Their first big advantage is a construction time  $O(m)$ , linear in the size  $m$  of the processed string  $S$ . Suffix trees are extremely useful because of the fact that search time is independent of the length of the string. A generalized suffix tree (GST) is constructed [5]. A GST [6] is a suffix tree that combines the suffixes of a set of strings. Once a generalized suffix tree of the document collection is built, the document-term matrix, or the TF-IDF matrix can be obtained. This document-term matrix can be built with a single traversal that is DFS traversal of the suffix tree as these values are already stored for each node.

The following steps are carried out in building CBVSTDM:

1. Collecting the document dataset
2. Perform POS tagging
3. Remove stop words
4. Apply stemming by using the WordNet stemmer
5. The stemmed words are then looked up in the WordNet, the lexical database to replace the words by their synset IDs which corresponds to a set of word forms which are synonyms
6. The words with the same synonyms are merged and are assigned a unique ID
7. Unique suffixes are generated
8. Build a concept based generalized suffix tree
9. Constructing a document-term matrix from the generalized suffix tree
10. Perform DFS traversal and obtain all word sequences
11. Retain those frequent word sequences which satisfy the minimum support. The minimum support of the frequent word sequences is usually in the range of 4-15%. When the minimum support is too large, the total number of frequent words would be very small, so that the resulting compact documents would not have enough information about the original data set
12. Thus the feature selection used is DF-thresholding
13. Attach weights to the obtained word sequences using either TF or TF-IDF method
14. The model is now evaluated using K-means clustering. The similarity measures that can be used in the analysis are Cosine, Jaccard, Euclidean and Pearson Correlation Coefficient

## REFERENCES

- [1] Horatiu Mocian. Text Mining with suffix trees, Sept. 2009
- [2] Hung Chim and Xiaotie Deng. A new suffix tree similarity measure for document clustering. In WWW '07: Proceedings of the 16<sup>th</sup> international conference on World Wide Web, pages 121–130, New York, NY, USA, 2007. ACM.
- [3] Hung Chim and Xiaotie Deng. Efficient phrase-based document similarity for clustering. Knowledge and Data Engineering, IEEE Transactions on, 20(9):1217–1229, Sept. 2008.
- [4] Peter Weiner. Linear pattern matching algorithms. In SWAT '73: Proceedings of the 14th Annual Symposium on Switching and Automata Theory (swat 1973), pages 1–11, Washington, DC, USA, 1973. IEEE Computer Society.
- [5] Oren Eli Zamir, Clustering Web Documents: A Phrase-Based Method for Grouping Search Engine Results.

---

**3rd National Conference on Research Trends in Computer Science & Technology NCRTCST-2015**

---

- [6] Yanjun Li, Soon M. Chung and John D. Holt. Text document clustering based on frequent word meaning sequences
- [7] Miller, G.: WordNet: a lexical database for English. Communications of the ACM, Volume 38, Issue 11 , pp.39- 41 (1995)
- [8] A. Strehl, J. Ghosh, and R. Mooney. Impact of similarity measures on web-page clustering. In AAAI-2000: Workshop on Artificial Intelligence for Web Search, July 2000.
- [9] Krzysztof Cios, Witold Pedrycz, and Roman W. Swiniarski. Data mining methods for Knowledge discovery. Kluwer Academic Publishers, Norwell, MA, USA, 1998.
- [10] R. Krishnapuram, A. Joshi, and Liyu Yi. A fuzzy relative of the k-medoids algorithm with Application to web document and snippet clustering. Fuzzy Systems Conference Proceedings, 1999.FUZZ-IEEE '99. 1999 IEEE International, 3:1281-1286 vol.3, 1999.



# An approach for Secured Authorized Deduplication on Hybrid Cloud

Ms.R. Mamatha<sup>1</sup>, Dr. Md. Mastan<sup>2</sup>, Mr. S. Shiva Kumar<sup>3</sup>

<sup>1</sup>Ms. R. Mamatha, Dept. of CSE, Student & Malla Reddy Engineering College (Autonomous), Dulapally, Secunderabad, TS, India.

<sup>2</sup>Dr. Md. Mastan, Dept. of CSE, Professor & Malla Reddy Engineering College (Autonomous), Dulapally, Secunderabad, TS, India.

<sup>3</sup>Mr. S. Shiva Kumar, Dept. of CSE, Student & Malla Reddy Engineering College (Autonomous), Dulapally, Secunderabad, TS, India.

<sup>1</sup>mamatharagelly@gmail.com

<sup>2</sup>mastanmohd@mrec.ac.in

<sup>3</sup>shiva.samalla@gmail.com

**Abstract**— Information deduplication system wipes out duplicate copies of data and has been mostly utilize as a part of cloud storage space, to decrease storage space and save bandwidth. Promising as it seems to be an emerging test to perform secured deduplication in cloud storage. In cloud situations, clients can store the information or records in cloud storage yet it is not an interminably expansive. With a specific end goal to decrease the prerequisite of capacity and transfer speed, information deduplication has been connected. Clients can impart one duplicate of the copy records to dispense with repetitive information. In this original copy, to ensure better information security we properly deal with the issue of accepted information deduplication. Security investigation shows that our plan is secure as far as the definitions determined in the proposed security model.

**Keywords**— Authorized Duplicate Check, Cloud Computing, Confidentiality, Deduplication, Hybrid Cloud.

## I. INTRODUCTION

To make data administration flexible in cloud computing, deduplication is surely an understandable system and has to be pulled to considerations. Data deduplication eliminates duplicate copy of data in cloud storage. The system is utilized to enhance stock usage and can likewise be connected to system data exchanges to decrease the quantity of bytes that must be sent [1]. As of late, distributed computing is getting stylish and gives clients more beneficial, cash sparing and efficient administrations than traditional structural planning. Clients can get diverse administrations and assets by uniting with the cloud server. However, if clients don't manage any unpredictable equipment and programming issues that cloud administration supplier has to handle. The clients of storage administration can transfer and download the data whenever and wherever by means of gadgets interfacing with the cloud. The distributed storage suppliers as a rule to perform data deduplication to decrease the utilization of capacity. Data deduplication is used for searching the excess records and immediately stores one duplicate of them.

The arrival of cloud storage encourages enterprises and organizations to outsource data storage to the third-party cloud providers, as proved by many real-life case studies.

One of the critical challenges of today's cloud storage services is to manage the ever-increasing data. The data deduplication is actively being used by number of cloud backup providers (e.g. Bitcasa) as well as various cloud services (e.g. Dropbox). Encryption methods which were used conventionally are not well-matched with data deduplication while providing data privacy. Customary encryption requires distinctive clients to encode their information with their own keys by which indistinguishable information duplicates, diverse clients prompt with diverse cipher texts, making deduplication outlandish. Concurrent encryption technique is proposed to implement information classifiedness while making deduplication practical. It decrypts/encrypts duplicate information with a concurrent key, which is acquired by figuring the cryptographic hash estimation of the substance of the information duplicate. At whatever point the key is created clients hold the keys and send the figure content to the cloud. To look ahead at an unapproved access, a safe confirmation of possession convention, S. Halevi [2], is expected to give verification about the client certainly claims to the same record when a copy is found.

In this paper, going for productively tackling the issue of deduplication with differential benefits in cloud computing, we consider a cross breed cloud structural planning comprising of an open cloud and a private cloud. Dissimilar to existing information deduplication frameworks, the private cloud is included as an intermediary to permit information proprietor/clients to safely perform copy check with differential benefits. Such a structural planning is down to earth and has pulled in much consideration from specialists. The information proprietors just outsource their information stockpile by using open cloud, while the information operation is overseen in private cloud.

Furthermore enhancing our system in terms of security, we present an advanced scheme for stronger protection by encrypting the file with different privilege keys. The users without privilege keys cannot carry out the duplicate check. Unofficial users cannot decrypt the ciphertext.

Finally, we execute a prototype of the proposed certified duplicate check and perform the experiments.

Platform services "Platform as a Service" represented as PaaS provides a computing platform using the cloud infrastructure. It has all the applications typically required by the client deployed on it. Thus the client needs no worry regarding purchase and installing the software and hardware required for it. Through this service developers can get a hold of all the systems and environments required for the life cycle of software, be it developing, testing, deploying and hosting of web applications. Key examples are GAE, Microsoft's Azure.

Infrastructure services "Infrastructure as a Service" represented as IaaS provides the required infrastructure as service. The client need not purchase the required servers, data center or the network resources. Also the key benefit here is that the consumers need to pay only for the time duration they use the service. As a result consumers can achieve a much quicker service delivery with less cost. Examples are GoGrid, Flexiscale, Layered Technologies, Joyent and Mosso/Rackspace, Bhaskar Prasad Rimal [3].

Below chart shows different layers of cloud computing architecture [4].

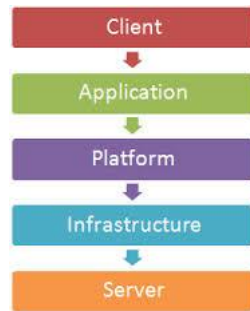


Figure - 1: LAYERS OF CLOUD COMPUTING ARCHITECTURE

## II. PRELIMINARIES

In this area, we first characterize the documentations utilized as a part of this paper, survey some safe primitives utilized as a part of our protected deduplication. The documentations utilized as a part of this paper are recorded beneath.

PoW, Proof of Ownership

S-CSP, Storage-cloud service provider

$k_F$ , Convergent encryption key for file  $F$

$P_U$ , Privilege set of a user  $U$

$(pk_U, sk_U)$ , User's public and secret key pair

$\phi'_{F,p}$ , Token of file  $F$  with privilege  $p$

$P_F$ , Specified privilege set of a file  $F$

Symmetric encryption utilizes a common undisclosed key  $\kappa$  to encrypt the data and decrypt data.

Symmetric encryption method consists the following

- $\text{Enc}_{\text{SE}}(\kappa, M) \rightarrow C$  is the symmetric encryption algorithm that receives the secret  $\kappa$  and message  $M$  and then outputs the ciphertext  $C$ ; and
- $\text{KeyGen}_{\text{SE}}(1^\lambda) \rightarrow \kappa$  is the key generation algorithm that produces  $\kappa$  using security parameter  $1$ .
- $\text{Dec}_{\text{SE}}(\kappa, C) \rightarrow M$  is the symmetric decryption algorithm that receives a secret key  $\kappa$  and ciphertext  $C$  and then outputs the original message  $M$ .

Encryption with convergent gives an information classifiedness in deduplication. A client (or information proprietor) gets a concurrent key from every unique information duplicate and encodes the information duplicate with the focalized key. Furthermore, the client likewise infers a tag for the information duplicate, such that the tag will be utilized to recognize copies. Here, we accept that the label accuracy property, M. Bellare [5], holds i.e., if two information duplicates are same, then their labels are also same. To distinguish copies, initially client sends the tag to the server area to verify the indistinguishable duplicate has been now put away. Note that both the focalized key and tag are autonomously inferred and the tag can't be utilized to find the united key and trade off information privacy. Both the scrambled information duplicate and its related tag will be put away on the server side. Formally, a focalized encryption plan consist of four primitive capacities:

$C \rightarrow \text{Enc}_{\text{CE}}(K, M)$  which receives both convergent key  $K$  and data copy  $M$  as an input using symmetric encryption algorithm and ciphertext  $C$  is next produced;

•  $\text{KeyGen}_{\text{CE}}(M) \rightarrow K$  is the key generating algorithm that maps the data copy  $M$  to a convergent key  $K$ ;

•  $\text{TagGen}(M) \rightarrow T(M)$  is a tag generating algorithm that maps the original data copy  $M$  and outputs a tag  $T(M)$ ;

$\text{Dec}_{\text{CE}}(K, C) \rightarrow M$  is a decryption algorithm which receives both ciphertext  $C$  and convergent key  $K$  as an input and next outputs the original data copy  $M$ ;

**Proof of Ownership:** The thought of an evidence of possession (PoW) empowers clients to demonstrate their responsibility for duplicates to the capacity server. In particular, PoW is actualized as an intelligent calculation (signified by PoW) run by a prover (i.e., client) and a verifier (i.e., stockpiling server). The verifier infers a short esteem  $\phi(M)$  from an information duplicate  $M$ . To demonstrate the responsibility for information duplicate  $M$ , the prover needs to send  $\phi'$  to the verifier such that  $\phi' = \phi(M)$ . The formal security definition for PoW generally takes after the risk demonstration of a substance appropriation system, where an assailant who do not know the whole document, but relatively have associates who have the record. The accessories are taken after the "limited recovery model", such that they can help the assailant get the document, subject to the limitation that they must send smaller number of bits than the

commencement min-entropy of record to the attacker, S. Halevi [6].

**Protocol of Identification:** A familiar proof convention can be portrayed in two stages: Proof and Verify. In the phase of Proof, a prover/client  $U$  can show his character to a verifier by performing some distinguishing proof verification identified with his personality. The info of the prover/client is his private key  $sk_U$  that is delicate data, for example, a private key of an open key in his declaration or Mastercard number and so forth that client might not want to impart to alternate clients. The verifier performs the check with information of open data  $pk_U$  identified with  $sk_U$ . At the finish of convention, the verifier yields either acknowledge or reject to mean whether the confirmation is passed or not. Abundant productive ID conventions are available in writing, including authentication based, personality based ID, M. Ballare [7].

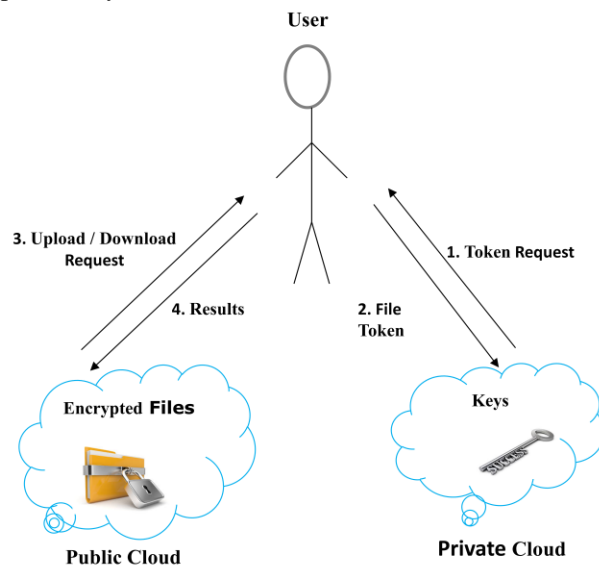


Figure - 2: AUTHORIZED DEDUPLICATION ARCHITECTURE

### III. SYSTEM MODEL

**Data Users:** A user is an article who wants to outsource the data storage to S-CSP and access the data whenever needed. Data users uploads the unique data by this hybrid cloud approach in order to hoard the bandwidth in a cloud storage system to support deduplication, which is owned by the same or different users. By this authorized deduplication system, each user is provided with a group of privileges in the set of connections of the system. Here convergent encryption key is provided to protect individual file and to realize the authorized deduplication with differential privileges the privilege key is used, Jin Li [1].

**S-CSP:** This provides information subcontract service and stores information on behalf of the client. The S-CSP eliminates storage of unnecessary data via deduplication and maintains only the unique information so that the cloud storage cost is minimized.

**Private Cloud:** Compared to conventional deduplication architecture in cloud computing, private cloud facilitates user's protection while using cloud service. specifically, in view of the fact that the computing resources at the data user/owner side are limited and the public cloud is not fully believed in practice, private cloud is capable of providing data user or owner by means of an execution environment and an infrastructure working as an interface between the user and public cloud. The private cloud manages private keys of the privileges, who answer the requests of users file token. The interface accessible by the private cloud allows user to provide files and queries to be securely stored and computed respectively, Jin Li [1].

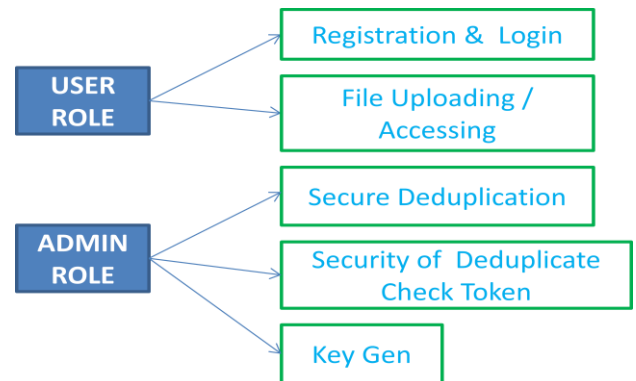


Figure - 3: SYSTEM DESIGN

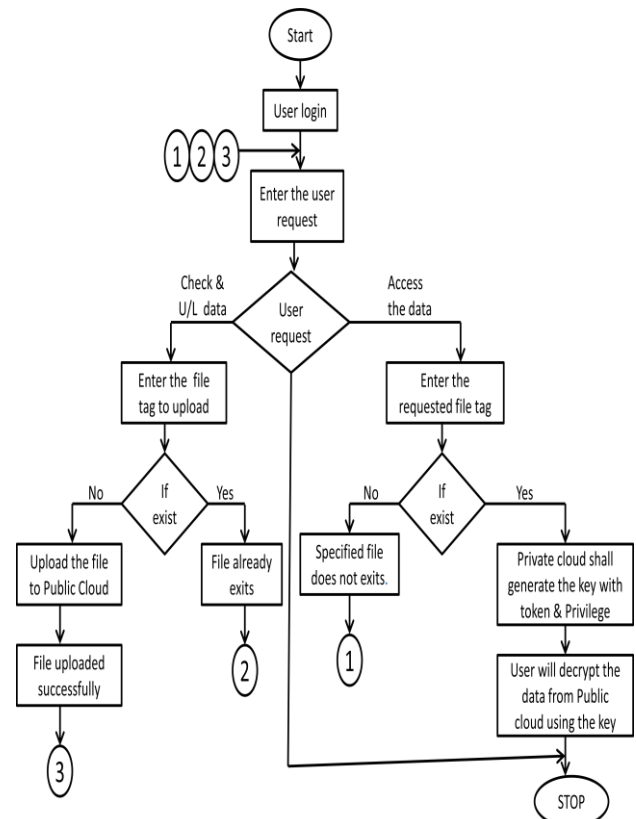




Figure - 4: FLOW CHART FOR HYBRID CLOUD APPROACH

#### IV. SECURE DEDUPLICATION SYSTEMS

To support authorized deduplication, the tag of a file will be determined by the privilege  $p$  and the file  $F$ . To illustrate the variation in traditional notation of tag, we describe it as a file token as an alternative. To support an authorized admission, a furtive key  $k_p$  will be surrounded with a privilege  $p$  to produce a file token. Let  $\theta'_{F,p} = \text{TagGen}(F, k_p)$  denote the token of  $F$  that is only allowed to access by user with privilege  $p$ . In another term, the users with privilege  $p$ , and the token  $\theta'(F, p)$  could only be computed. So the result is produced, if the user has been uploaded a file with a duplicate token  $\theta'_{F,p}$ , then the duplicate check sent from one more user will be successful if he holds the privilege  $p$  and file  $F$ .  $H(F, k_p)$  is the token generation function which can be simply applied, Jin Li [1].

**Goal Of Deduplication:** The goal of data deduplication is to lessen the requirement of storage space by sharing only one copy of the same plaintext or information.

Many data deduplication schemes have been proposed in modern years, but only some addressed the security issue, M.W. Storer [8]. Though, they have encountered the difficulties of identifying all ciphertexts of a plaintext while they usually use the convergent encryption to provide the solution.

**Advantages:** Advantages of the proposed mechanism are:

It is very close to reality, which only uses some standard algorithms such as the hash function (e.g., SHA-2), the AES algorithm and the public key algorithm (e.g., the RSA algorithm).

It is easy to implement the proposed system owing to use the standard algorithms. It provides transparency. Users do not need to change their current habits of using cloud storage services.

#### V. SECURITY ANALYSIS

Our system is anticipated to resolve the differential privilege difficulty in protected deduplication. Security will be analyzed in two aspects, authorization of duplicate check and confidentiality of data. Some basic tools have been used to construct the secure deduplication, which are assumed to be secure. These basic tools include the convergent encryption system, symmetric encryption system, and PoW scheme.

**Duplicate Check Token Security:** A number of privacies are considered which needs to be protected, i.e., Duplicate-check token Unforgeability: External adversary and internal adversary are available. As revealed below, the external adversary can be sighted as an internal

adversary with no privilege. If a privilege  $p$  available with user, it requires the adversary which cannot falsify and output a suitable duplicate token with any new privilege  $p'$  on a file  $F$ , where  $p \neq p'$ . Furthermore, if the adversary is not building a request of token with its own privilege from private cloud server, it cannot forge and output a suitable duplicate token with  $p$  on any  $F$  that has been queried.

**Security on Cloud Storage:** Many papers have been discussed the protection problems of cloud computing. There are three factors on cloud computing security: Information privacy, reliability, and availability (CIA), J.Harauz [9]. On the storage service, users worry about the personal privacy, which is data confidentiality. In order to protect the privacy, Cryptographic technologies encrypt the files before they are uploaded to the cloud storage and decrypt them.

#### VI. IMPLEMENTATION

Our execution of the "Customer" gives the accompanying capacity calls to bolster token era and deduplication along with the document transfer process.

- **TokenReq(Tag, UserID)** - It requests the Private Server for File Token generation with the File Tag and User ID.
  - **FileTag(File)** - Finger out the SHA-1 hash File as a File Tag;
  - **ShareTokenReq(Tag, {Priv.})** - It requests the Private Server to generate the Share File Token with the File Tag and Target Sharing Privilege Set.
  - **DupCheckReq(Token)** - It requests the Storage Server for Duplicate Check of the File by sending the file token received from private server.
  - **FileEncrypt(File)** - Encrypts the File with a convergent encryption by means of 256-bit AES algorithm in Cipher Block Chaining (CBC) mode, where the convergent key is from SHA-256 file of Hashing.
  - **FileUploadReq(FileID, File, Token)** - Uploads the file information to the Storage Server, if the File is Unique and updates the File Token stored.
- The Implementation of "Private Server" includes corresponding request handlers for the token generation and maintains key storage with Hash Map.
- **ShareTokenGen(Tag, {Priv.})** - Generates the share token with the corresponding privilege keys of the sharing privilege set with HMAC-SHA-1 algorithm.
  - **TokenGen(Tag, UserID)** - Loads the associated privilege key of user and generate the token with HMAC-SHA-1 algorithm.

Implementation of the "Storage Server" provides deduplication and information storage with following handlers and retains a map between existing files and associated token with Hash Map.

- **FileStore(FileID, File, Token)** - It stores the File on Disk and updates the Mapping, Jin Li[1].
- **DupCheck(Token)** - For Duplicate Token, it searches the File token to Map

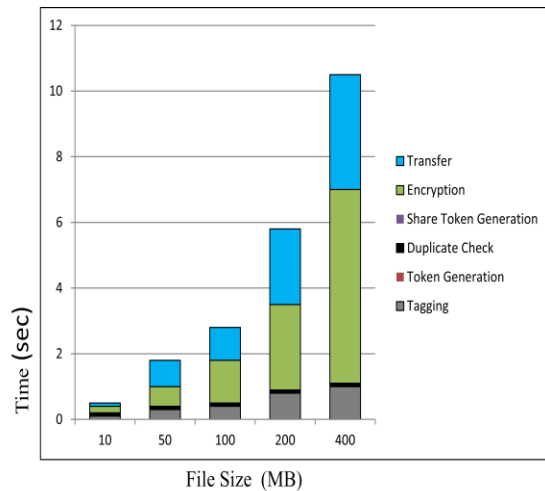


Figure – 5: TIME BREAKDOWN FOR DIFFERENT FILE SIZE

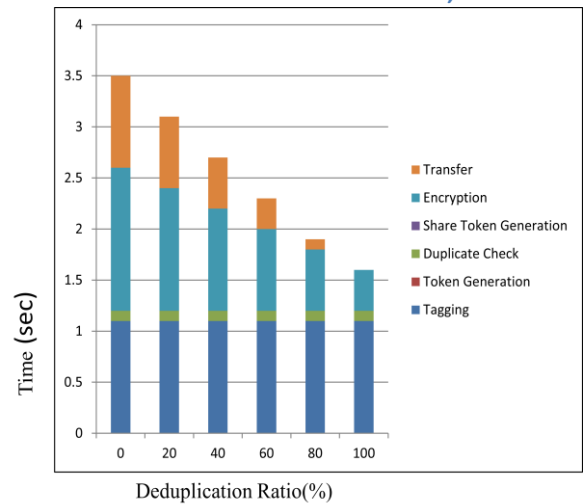


Figure – 6: TIME BREAKDOWN FOR DIFFERENT DEDUPLICATION RATIO

Deduplication Ratio (%)	0	20	40	60	80	100
Tagging	1.1	1.1	1.1	1.1	1.1	1.1
Token Generation	0	0	0	0	0	0
Duplicate Check	0.1	0.1	0.1	0.1	0.1	0.1
Share Token Generation	0	0	0	0	0	0
Encryption	1.4	1.2	1	0.8	0.6	0.4
Transfer	0.9	0.7	0.5	0.3	0.1	0

Table - I: BREAKDOWN FOR DIFFERENT FILE SIZE

## VII. ESTIMATION

Our valuation focuses on comparing the overhead induced by authorization steps, file token generation and share token generation, against the convergent encryption and file upload steps. We measure the overhead by altering diverse features that are File Size, Number of Stored Files,

Deduplication Ratio, and Privilege Set Size. We break down the upload process into Transfer, Encryption, Share Token Generation, Duplicate Check, Token Generation and Tagging. For each process the start time and end time is recorded, therefore total breakdown time spent is obtained.

File Sizes (in MB)	10	50	100	200	400
Tagging	0.1	0.3	0.4	0.8	1
Token Generation	0	0	0	0	0
Duplicate Check	0.1	0.1	0.1	0.1	0.1
Share Token Generation	0	0	0	0	0
Encryption	0.2	0.6	1.3	2.6	5.9
Transfer	0.1	0.8	1	2.3	3.5

Table - II: BREAKDOWN FOR DIFFERENT DEDUPLICATION RATIO

**Deduplication Ratio:** To estimate the outcome of the deduplication ratio, we prepare two distinctive data sets, each of which consists of 50 to 100 MB files. We first upload the first set as an initial upload. For the next upload, we select portion of 50 files, accordingly in the given deduplication ratio, from the initial set as duplicate files and remaining files from the next set as unique files. The average time of uploading the second set is obtainable in Figure 4. As uploading and encryption would be skipped to generate the duplicate files, the time spent on both of them decreases with increasing deduplication ratio. The moment time used on duplicate check also decreases as the searching would be ended, when duplicate is found. Total time spent on uploading the file with deduplication ratio at 100% is only 33.5% with unique files.

## VIII. RELATED WORK

**Secure Deduplication:** The deduplication framework in the distributed storage to lessen the capacity of labels for honesty check proposed by Yuan et al. [10]. Security of deduplication is to upgrade and secure the information privacy, demonstration is to make confident the information classifiedness by altering the predictable message into changeable message. Li et al. [12] tended in piece level deduplication by disseminating these keys over different servers after encrypting the files is the key management issue.

**Traditional:** The literature has been projected by various cryptographic solutions to look after the confidentiality of outsourced data, S. Kumara [13]. The idea build on traditional (symmetric) encryption, were every user encrypts data with an independent secret key. Amount of study, propose to use threshold as top secret sharing to continue the robustness of key organization.

**Convergent:** The data privacy in deduplication is ensured by Convergent encryption, M. Bellare [14]. There are many implementations of convergent encryption variant for safe deduplication. Some industrial cloud

storage providers, example Bitcasa, is deployed using convergent encryption technique.

**POW:** The “proofs of ownership” (PoW) for deduplication systems, so that a user can strongly verify to the cloud storage space server that they own a folder without uploading the file, the concept is implemented by Halevi et al [15]. Merkle Hash Tree proposed the client-side deduplication, which comprise the bordered seepage set are based on pow constructions. By choosing the protrusion of a file onto some accidentally selected file for proof as selected bit-position is proposed by Pietro and Sorniotti.

**Private Cloud Architecture:** The private and public cloud approach is to support isolation attentive data concentrated computing is obtainable by Zhang[16]. Our systems security model is of similar to those related work, where the private cloud assumes to be honest but curious. Recently, an architecture consisting of private and public clouds for protected outsourcing of data and random computations to an entrusted service cloud is proposed by Bugiel.

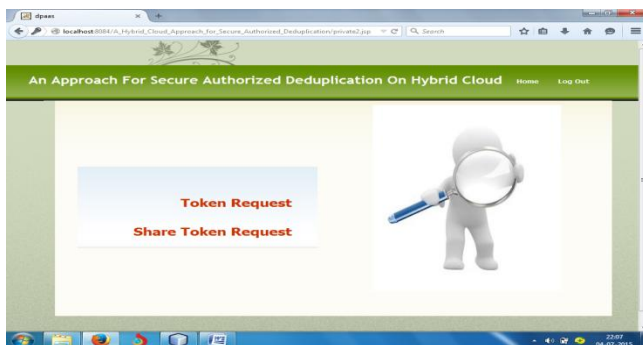
**Public Cloud Architecture:** Zhang. [16] also presented The hybrid cloud techniques is to support isolation attentive data concentrated computing. Here we consider for addressing the authoritative deduplication problem over data in public cloud. An architecture consisting of private and public clouds for secure outsourcing of information and subjective computations to an entrusted commodity cloud accessible by Bugiel.

#### SAMPLE SCREEN SHORTS

##### *After User Login Page*



##### *After Private Login Page*



#### IX. CONCLUSION

By this deduplication technique, duplication is removed, utilizing the space judiciously for other important files and thereby saving the space & bandwidth. This mechanism is proposed by hybrid data deduplication which provides a solution that is more secure than that of the previous techniques.

#### X. REFERENCES

- [1] A Hybrid Cloud Approach for Secure Authorized Deduplication, Jin Li, Yan Kit Li, Xiaofeng Chen, Patrick P. C. Lee, Wenjing Lou IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEM VOL: PP NO:99 YEAR 2014.
- [2] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. Proofs of ownership in remote storage systems. In Y. Chen, G. Danezis, and V. Shmatikov, editors, ACM Conference on Computer and Communications Security, pages 491–500. ACM, 2011.
- [3] Bhaskar Prasad Rimal, Eunmi Choi, "A taxonomy and survey of cloud computing systems", 2009 Fifth International Joint Conference on INC, IMS and IDC, published by IEEE Computer Society
- [4] [http://en.wikipedia.org/wiki/Cloud\\_computing](http://en.wikipedia.org/wiki/Cloud_computing)
- [5] M. Bellare, S. Keelveedhi, and T. Ristenpart. Message-locked encryption and secure deduplication. In EUROCRYPT, pages 296–312, 2013.
- [6] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. Proofs of ownership in remote storage systems. In Y. Chen, G. Danezis and V. Shmatikov, editors, ACM Conference on Computer and Communications Security, pages 491–500. ACM, 2011.
- [7] M. Bellare, C. Namprempre, and G. Neven. Security proofs for identity-based identification and signature schemes. J. Cryptology, 22(1):1–61, 2009
- [8] M. W. Storer, K. Greenan, D. D.E. Long, and E. L. Miller, 2008. Secure Data Deduplication. In Proceedings of The 4th ACM International Workshop on Storage Security and Survivability, 2008, pp. 1-10.
- [9] J. Harauz, L. M. Kaufman, and B. Potter, “Data Security in the World of Cloud Computing,” IEEE Security and Privacy, Vol.7, No. 4, 2009, pp. 61-64.
- [10] J. Yuan and S. Yu. Secure and constant cost public cloud storage auditing with deduplication. IACR Cryptology ePrint Archive, 2013:149, 2013.
- [11] M. Bellare, S. Keelveedhi, and T. Ristenpart. Dupless: Serveraided encryption for deduplicated storage. In USENIX SecuritySymposium, 2013.
- [12] J. Li, X. Chen, M. Li, J. Li, P. Lee, and W. Lou. Secure deduplication with efficient and reliable convergent key management. In IEEE Transactions on Parallel and Distributed Systems, 2013.
- [13] S. Kamara and K. Lauter, “Cryptographic Cloud Storage,” in Proc. Financial Cryptography: Workshop Real-Life Cryptograph. Protocols Standardization, 2010, pp. 136-149.
- [14] M. Bellare, S. Keelveedhi, and T. Ristenpart, “Message-Locked Encryption and Secure Deduplication,” in Proc. IACR Cryptology ePrint Archive, 2012, pp. 296-3122012:631.
- [15] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg. Proofs of ownership in remote storage systems. In Y. Chen, G. Danezis, and V. Shmatikov, editors, ACM Conference on Computer and Communications Security, pages 491–500. ACM, 2011.
- [16] K. Zhang, X. Zhou, Y. Chen, X.Wang, and Y. Ruan. Sedic: privacyaware data intensive computing on hybrid clouds. In Proceedings of the 18th ACM conference on Computer and communications security, CCS’11, pages 515–526, New York, NY, USA, 2011. ACM.



# An Efficient FDM Based Association Rule Mining in Horizontally Distributed Databases

Mr. S. Shiva Kumar<sup>1</sup>, Dr. Md. Mastan<sup>2</sup>, Ms. R. Mamatha<sup>3</sup>

<sup>1</sup>M.Tech Student, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

<sup>2</sup>Professor, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

<sup>3</sup>M.Tech Student, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

<sup>1</sup>[ishiva.samalla@gmail.com](mailto:ishiva.samalla@gmail.com), <sup>2</sup>[mastanmohd@mrec.ac.in](mailto:mastanmohd@mrec.ac.in), <sup>3</sup>[mamatharagelly@gmail.com](mailto:mamatharagelly@gmail.com)

## Abstract

Association rules mining are one among the data mining techniques employed in distributed database system and it disclose some attention -grabbing relationship between locally massive itemsets and globally massive itemsets, proposed an algorithm called Fast Distributed Mining (FDM) of association rules, an unsecured distributed version of Apriori algorithm. It generates a low range of candidate sets and reduces the amount of message to be passed at mining association rules. Two secure multi-party algorithms are main ingredients in our protocol. One that computes the union of private subsets that every of the interacting players holds, and another that tests the inclusion of element by one player in a subset held by another. This protocol offers increased privacy with relation to the protocol and additionally this protocol is less complicated and considerably a lot of efficient in terms of communication rounds, communication cost, and computation cost.

**Index terms:** Distributed Databases; Frequent Itemsets; Privacy Preserving Data Mining.

## I. INTRODUCTION

The aim of data mining is to extract vital information from massive datasets, however typically these datasets are split among varied parties. Data mining is defined as the technique for extracting hidden, predictive and knowledge data from large distributed databases. The technology that has emerged as the method of identifying patterns and trends from large quantities of knowledge. Here we discuss the matter of association rule mining in horizontally distributed databases. In the distributed databases, there are many players that hold same databases/information i.e. databases share same schema hold data totally on different entities. The goal is to search out all association rules with support  $s$  and confidence  $c$  to attenuate the data disclosed regarding the personal databases command by those players [1].

Kantarcioglu and Clifton [2] studied the problem whenever a lot of suitable security definitions that permit parties to settle on their desired level of security are required, effective solutions that maintain the required

security. So Kantarcioglu and Clifton in [2] devised a protocol for its solution. The main part of the protocol is a sub- protocol for secure computation of union of private subsets that are controlled by the various players. It makes the protocol expensive and its implementation depends upon cryptographic primitives such as oblivious transfer, a hash function, and commutative encryption.

Here proposed an algorithm privacy preserving distributed data mining algorithm for horizontally distributed data sets and to get most interesting association or correlation relationships among an outsized or large set of data items and to incorporate encryption security techniques to reduce the information that goes to be shared with others, while adding little overhead to the mining task [1]. The data that would like to be secured during this work is not only individual transaction within the completely different databases, however, more public or global information like association rules are supported locally in every of these databases. The proposed protocol improves upon that in Kantarcioglu and Clifton [2] in terms of simplicity, efficiency, and security.

## II. LITERATURE REVIEW

In Distributed database system data is logically belong to the same system, but data is spread over the sites, and it is not necessary to geographically distributed. The Distributed database management system (DDBMS) is the software system that permits the management of the Distributed database (DDB) and it makes the distribution transparent to the users. Distributed database system (DDBS) is the combination of DDB and DDBMS; it is achieved through the integration of database and networking technologies together.

Fragmentation technique divides a single relation or class of a database into two or more partitions. The combination of the partitions provides the original database without any loss of information. It reduces the amount of irrelevant data accessed by the applications; it implies reducing the number of disk accesses. Fragmentation technique is of three types such as horizontal fragmentation (HF), vertical fragmentation (VF) and mixed/hybrid fragmentation (MF).



Horizontal fragmentation allows a class or relation to be partitioned into disjoint instances or tuples. The perception behind this fragmentation method is that every site should hold all data that is used to query at the site and the data at the site should be fragmented. The final result is queried of the site run faster.

Vertical fragmentation allows a class or relation to be partitioned into different sets of attributes or columns except the primary key. In vertical fragmentation, each partition must include the primary key. This arrangement can make sense when different sites are responsible for processing different functions involving an entity.

The combination of horizontal and vertical fragmentation is hybrid or mixed fragmentation. It consists of vertical fragmentation followed schema will not sufficient to satisfy the requirements of the by a horizontal fragmentation and vice versa. The common problem with the above techniques is that they concentrate only on the fragmentation problem and the overlooked allocation problem to reduce complexity.

Data mining in centralized model assumes that required data is either available at or can be sent to a central site and data mining algorithm applied at the central site. A simple approach is not to share data and perform data mining tool at each site independently and combine the results. However, it fails in globally final results because of data skewness and the disparity between local and global results. Individual sites are unable to detect cross-site correlations in data mining. Because the same item may be duplicated at different sites and there will be an overweight in the results.

FP-tree is a compact data structure and without candidate set generation it finds the frequent itemset by traversing itemsets through FP-tree. The FP-tree algorithm in association rule mining for distributed environment finds the frequent itemsets. Data encryption standards are used to generate the privacy to database by using two keys. With the key1, first party encrypts the dataset. This encrypted dataset is again encrypted with key2. At the receiver side, at first it decrypts the dataset with key2 and then after with key1. It provides high security to the databases compared with other cryptographic techniques and it is also known as Double encryption. With zero percentage of data leakage it provides higher privacy to the database; however it is applicable to homogeneous databases only.

Meers Treesa Mathews, [3] uses an apriori algorithm with extended Distributed RK sum secure protocol for privacy preserving data mining. At first, on given dataset Apriori algorithm is applied then to find the frequent itemsets of individual parties and then to get global results Extended distributed RK sum secure algorithm is applied. Distributed RK sum secure protocol is used for secure multi party protocol but the drawback is more than two parties join together; they can know the data of some party. The drawback is overcome in extended distributed RK sum protocol with more privacy to the data, the communication and computation complexity  $O(n)$ .

Priyanka Asthana and Anju Singh [4] used five different algorithms such as Apriori algorithm, MCISI algorithm, MS algorithm, Apriori with systematic rules and HMT (Hash Mapping Table). Apriori algorithms combined with hashing techniques are used to reduce the time and space complexity. MSApriori (minimum Support) method provides different minimum support itemset values for different items. Apriori systematic rules are used to save the time by scanning the database only once rather than multiple scans.

### Privacy Preserving Data Mining

D.W.L Cheung, V.T.Y. Ng, [5] used the problem in privacy preserving association rule mining in horizontally distributed databases among  $n$  ( $n > 2$ ) sites, with no sites can be treated as trusted. The proposed protocol uses hash based cryptographic techniques to find the global frequent itemsets. Privacy preserving rule mining categorized into three types based on privacy preserving/protection techniques.

- 1) Heuristic based techniques
- 2) Reconstruction based techniques
- 3) Cryptographic based techniques

### Heuristic based techniques

Heuristic based techniques use adaptive modification and it only modifies the selected values, utility loss is minimized.

### Cryptographic based techniques

Cryptographic based techniques are used for both horizontally distributed databases and vertically distributed databases and it is based on the secure multi party computation. Here no database knows anything except given input database and the result so that computation is secured.

### Reconstruction based techniques

Reconstruction based techniques are used for different data such as numerical, binary and categorical data, its work on the problem of privacy preservation by perturbing the data. The original distribution of data is constructed from the randomized data using Reconstruction based techniques.

### Association Rule

Association rule mining is well known and popular method for finding interesting relationships among from large databases. Today association rules are used in many applications including intrusion detection, biometrics, the web usage mining and continuous production. It is a two-step process i.e. Minimum support and confidence, the minimum support used to find all frequent itemsets from given database and constraint confidence used to form the rules. The association rule mining problem is stated/defined as Let  $I = \{i_1, i_2, i_3, \dots, i_n\}$  are the set of items, database  $D$  consists set of transaction  $T$ . Each transaction  $T$  is a set of items from  $I$  such that  $T \subseteq I$ , with a unique transaction id TID. Each transaction  $T$  contains  $X$  and a set of some items in  $I$  if  $X \subseteq I$ .

Association Rule is an implication of the form,  $X \Rightarrow Y$ , where  $X \subset I$ ,  $Y \subset I$  and  $X \cap Y = \emptyset$ . The association rule  $X \Rightarrow Y$



Y holds in the transaction set D, with the confidence  $c$  if  $c\%$  of the transaction in D that contains X and also contain Y. The rule  $X \Rightarrow Y$  has support  $s$  in the transaction set D if  $S\%$  of transactions in D contains  $X \cup Y$ . In shopping centers, association rules are used to place the items side by side, so that more items will be sold. In Amazon, association rules mining to recommend you to buy the products based on present item buying or browsing. Google uses association rules for mining when you type the words for search it will give frequently associated words.

### Apriori Algorithm

In transactional databases, the apriori algorithm is used for mining of frequent itemset and association rule learning. Apriori algorithm proceeds by identifying the frequent itemsets and extending them to larger and larger itemsets in the database. It identifies the itemset at least C transaction in the database. Apriori algorithm uses the bottom up approach and proposed by Agarwal and Srikanth in 1994. Its usage is more in market basket analysis. To count candidate itemset efficiently, Apriori algorithm uses breadth-first search and hash tree structure. Apriori algorithm based on Apriori principle works to generate candidate k-item sets from (k-1) frequent itemsets, pruning technique used to avoid the measuring of certain itemsets. Hash based methods with Apriori algorithms are used to reduce the time and space complexity.

### FAST DISTRIBUTED MINING (FDM) Algorithm

The protocol of [2], as well ours are based on FDM algorithm of Cheung et al. [5], it is Apriori algorithm's unsecured distributed version. Here, the main idea is that any s-frequent itemset must be also locally s-frequent in one of the sites at least. So, in order to get all globally s-frequent itemset, each player reveals its locally s-frequent itemset sets and then the each player check them to see if they are s-frequent itemsets also globally.

Following are the steps for FDM algorithm

1. Initialization
2. Candidate set generation
3. Local Pruning
4. Unifying the candidate itemset
5. Computing local support
6. Broadcast the mining results

Cheung et al. [5] proposed Fast distributed mining algorithm for association rules in distributed databases, Apriori algorithm for association rule mining for individual databases. In [2] proposed an insecure version of FDM algorithm but the algorithm violates privacy in two stages, in step four and step six. In step 4 whenever the players broadcast the itemsets that are locally

frequent in their private databases. In step 6 where they broadcast the sizes of the local supports of candidate itemsets. Kantarcioglu & Clifton [2] proposed secure implementations of these steps and also in [1] describe the various implementations and proceed to analyze implementations in terms of privacy, efficiency, and compare them. Finally, it shows that our protocol offers better/enhanced privacy which is less complicated and is considerably additional efficient in terms of communication rounds, communication cost, and computation cost.

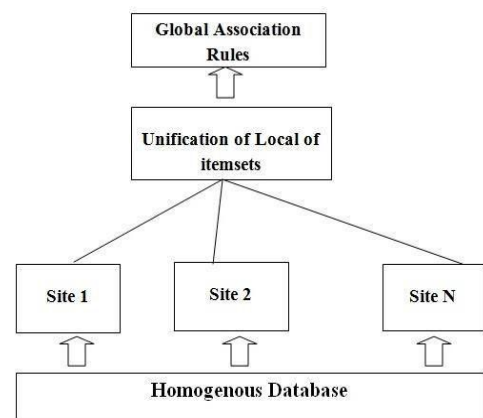


Figure 1. Architecture of the system

### Synthetic Database Generation

To calculate the performance of algorithms over a very large range of data characteristics the synthetic transactions are used and the process involved here is data anonymization. It is used in data mining process. Synthetic data are generated to get certain conditions and to meet specific needs that may not be found in real, original data. Following are the parameters used in synthetic data generation and these values are similar to used in [2], [5], [6].

Parameter	Interception
N	No. of transactions in database
L	No. of items
At	Average size of transaction
Af	Maximal potentially large itemsets average size
Nf	No. of maximal potentially large itemsets
CS	Clustering size
PS	Pool size
C	Correlation level
MF	Multiplying factor

### III. DISCUSSIONS

Following are the steps to secure computation itemsets using UNIFI-KC [2] (Unifying list of locally frequent itemsets) protocol.

- (1) A Selection of Cryptographic primitive cipher.

- Players choose the required commutative cipher and its corresponding private key and selected hash function



applied on all itemsets for encryption.

- A lookup table with hash values is generated to compute pre-image of hash values.
- Encrypt all the itemsets.

(2) Merging of itemsets

- Odd players send its encrypted set to player1.
- Even players send its encrypted set to player2.
- Player1 unifies all itemsets sent by odd players and removes all duplicates.
- Player 2 unifies all itemsets sent by odd players and removes all duplicates.
- Player 2 sends his permuted list of itemsets to player 1.
- Player 1 unifies his list of itemsets and the list received from player 2 and from the unified list duplicates are removed.
- Denote the final list as  $EC_s$ .

(3) Decryption.

In the protocol UNIFI (Unifying list of locally frequent itemsets) input is each  $P_m$  player has an input subset such as  $C_s^{k,m} \subseteq Ap(F_s^{k-1})$ ,  $1 \leq m \leq M$ .

Output is  $C_s^K = \bigcup_{m=1}^M C_s^{k,m}$ . Following are the steps

1. Each player encodes subset as a binary vector  $b_m$  of length  $n_k$ , in accord with the agreed ordering of  $Ap(F_s^{k-1})$ .
2. Protocol Threshold-C is invoked by the players to compute  $b = T_1(b_1, \dots, b_m) = \bigvee_{m=1}^M b_m$ .
3.  $C_s^k$  is the subset that is described by the vector  $b$ .

For example  $C_s^{3,1} = \{124\}$ ,  $C_s^{3,2} = \{234\}$ ,  $C_s^{3,3} = \{124\}$  are the private sets of locally frequent itemsets and these private sets encoded as  $b_1 = (0,1)$ ,  $b_2 = (0,1)$ ,  $b_3 = (1,0)$ , and  $b = (1,1)$  is the OR of these vectors and therefore  $C_s^3 = \{124, 234\}$ .

The UNIFI-KC [1] has the following performance

- Communication rounds are  $(2M+1)(K+1)$ ,
- Messages are  $(M^2+2M-3)$
- Bits of communication is  $g(M)tn + (M-1)Bl$ ,

Where  $2M^2 - (M/2) - 2 \geq g(M) \geq M^2 + M - 2$ .

In Protocol UNIFI,

- Communication rounds are  $4(K+1)$
- Messages are  $(M^2+M-1)(K+1)$
- Bits of communication are

$(M^2-2)(\log_2 M)n + 2n \lceil h \rceil + (M-1)Bl$

The performance of two protocols is compared for FDM algorithm and the first implementation (FDM-KC) using the protocol UNIFI-KC executed the unification step; here commutative cipher used was 1024 RSA. In the other implementation (FDM) used our protocol UNIFI, with the keyed hash function as HMAC; finally with respect to following three measures two implementations are tested.

1) FDM and FDM KC protocols total computation time over all players; it includes the time to identify the globally  $s$ -frequent itemsets and the Apriori computation time,

2) UNIFI-KC and UNIFI protocol total computation time over all players.

3) The total message size.

#### IV. CONCLUSION

The proposed protocol secure mining of association rules in horizontally distributed databases offer enhanced privacy and security than the leading protocol [2]. Two secure multi-party algorithms are main ingredients in our protocol. One that computes the union of private subsets that every of the interacting players hold, and another that tests the inclusion of element by one player in a subset held by another. Proposed protocol is more significant in communication cost, communication rounds, and computation cost.

#### REFERENCES

- [1] Tamir Tassa, "Secure Mining of Association Rules in Horizontally Distributed Databases" IEEE trans. Knowledge and Data Engg. Vol. 26, no. 2, April 2014.
- [2] M. Kantarcioglu & C. Clifton, "Privacy preserving distributed mining of association rules on horizontally partitioned data" IEEE Transactions on Knowledge and Data Engineering, 16:1026-1037, 2004.
- [3] Meers Treesa Mathews, Manju E.V., "Extended Distributed RK Secure Sum Protocol in Apriori Algorithm for Privacy Preserving" International Journal of Engineering and Advanced Technology (IJEAT), Volume- 3, Issue-4, April 2014.
- [4] Priyanka Asthana, Anju Singh, Diwakar Singh, "A Survey on Association Rule Mining Using Apriori Based Algorithm and Hash Based Methods", International Journal of Advanced Research in Computer Science and Software Engineering. Volume 3, Issue 7, July 2013.
- [5] D.W.L. Cheung, J. Han, V.T.Y. Ng, A.W.C. Fu, and Y. Fu, "A fast distributed algorithm for mining association rules". In PDIS, pages 31-42, 1996.
- [6] J.S. Park, M.S. Chen, and P.S. Yu, "An effective hash based algorithm for mining association rules". In SIGMOD Conference, pages 175-186, 1995.
- [7] Agarwal and R. Srikanth "Fast algorithms for mining association rules in large databases". In VLDB, pages 487- 499, 1994.
- [8] A.V. Evfimievski, R. Srikant, R. Agrawal, and J. Gehrke. "Privacy preserving mining of association rules". In KDD, pages 217-228, 2002.
- [9] Larry A. Dunning, Member, IEEE, and Ray Kresman, "Privacy Preserving Data Sharing With Anonymous ID Assignment", IEEE Transaction On Information Forensics and security, VOL. 8, NO. 2, FEBRUARY 2013.
- [10] J. Vaidya and C. Clifton, "Privacy preserving association rule mining in vertically partitioned data," in The Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, Alberta, Canada, July 23-26 2002, pp. 639-644.
- [11] H. Grosskreutz, B. Lemmen, and S. R'uping, "Secure distributed subgroup discovery in horizontally partitioned data" Transactions on Data Privacy, 4:147-165, 2011.
- [12] Jyotirmayee Rautaray, Raghavendra Kumar, "Privacy Preserving Databases Using Data Encryption (DES)" International Journal of Innovative Research in Science, Engineering and Technology Vol. 2, Issue 3, March 2013.
- [13] D.W.L Cheung, V. T. Y. Ng, A. W. C. Fu and Y. "Efficient mining of association rules in distributed databases". IEEE trans. Knowl. Data Eng., 8(6):911-922, 1996 Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.



# Enabling public verification and privacy preserving Audit for secure cloud storage

<sup>1</sup> K.S.L. Savitri, <sup>2</sup> S. Ajay Kumar

<sup>1</sup>.Ms. K.S.L Savitri, M.Tech Student , Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> S. Ajay Kumar, Associate Professor, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>1</sup>savitri.kattamuri@gmail.com

<sup>2</sup> ajay.sgr@gmail.com

**Abstract-** Using cloud place for storing, users from far can store their knowledge for computers and have special rights, the on-demand high quality applications and services from a shared card-player's money of configurable computing useable things, without the weighting of nearby knowledge for computers place for storing and support. However, the fact that users no longer have physical property of the outsourced facts makes the knowledge for computers true, good nature care in cloud computing a hard to do work, especially for users with limited computing useable things. In addition, users should be able to just use the cloud place for storing as if it is nearby, without troubling about the need to make certain of its true, good nature. In this way, giving power public audit ability for cloud place for storing is of full of danger importance so that users can go to for help to a third group overseer i.e., Third Party Auditor (TPA) to check the true, good nature of outsourced facts and be worry-free. To safely put into use for first time a working well TPA, the looking over of accounts by expert process should take in no new feeblenesses in the direction of user facts right not to be public, and present no added connected weighting to user. In this paper, we make an offer a safe cloud place for storing system supporting privacy-preserving public looking over of accounts by expert. We further stretch our outcome to give power to the TPA to act looking over of accounts by expert for number of times and other users at the same time and with small amount of money. Much safety and operation observations make clear to the made an offer designs are provably safe and highly good at producing an effect.

**Keywords—** Public auditability, Third Group Overseer,

Secure, Cloud storage

## 1. Introduction

Cloud computing has been envisioned as the next-generation information technology buildings and structure design for undertakings, needing payment to its long list of unprecedented better chances in the it history: on-demand self-service, everywhere network way in, marked off

independent useable thing pooling, quick useable thing elasticity, usage-based pricing and transference of danger . As a tendency to cause destruction technology with deep follow up, cloud computing is making great change the very nature of how businesses use information technology. One deep point of view of this example changing is that knowledge for computers is being put under one control or outsourced to the Cloud. From users view, including both individuals and it undertakings, storing facts from far to the cloud in a flexible on-demand ways takes taking from lower to higher authority helps: rest of the weighting for place for storing managers of a business, general facts way in with independent about geography places, and overlooking of by death money used on computer and apparatus, software, and personnel support, and so on.

While cloud computing makes these better chances more having attraction for than ever, it also takes new and hard safety being, saying violent behavior in the direction of users outsourced knowledge for computers. Since cloud public organization givers are separate office activity things, facts outsourcing is actually giving up user's last control over the way given by powers that be of their facts. As an outcome, the rightness of the facts in the cloud is being put at danger needing payment to the supporter's reasons. First of all, although the basic buildings under the cloud are much more powerful and safe, good than personal computing apparatuses, they are still facing the wide range of both inside and outside being, saying violent behavior for knowledge for computers true, good nature. Examples of outages and safety overrules of noted cloud services come into view as from time to time. Secondly, there do have existence different causes of motion for Cloud Service Provider (CSP) to act rightly untruly in the direction of the cloud users looking upon the position (in society) of their outsourced facts. For examples, CSP might get back place for storing for money-related reasons by putting out as of



no use knowledge for computers that has not been or is uncommonly made way in, or even put out of the way facts loss small events so in connection with support a good name. In short, although outsourcing facts to the cloud is by money and goods pleasing for in the long run great-scale facts place for storing, it does not immediately offer any give support to (a statement) on facts true, good nature and able to use. This hard question, if not rightly made house numbers, may get in the way of the good placing of the cloud buildings and structure design.

As users no longer physically have as owner the place for storing of their facts, old and wise cryptographic early persons for the purpose of knowledge for computers safety system of care for trade cannot be directly took up. In particular, simply downloading all the knowledge for computers for its true, good nature verification is not an useful substance mixed in liquid needing payment to the expensiveness in i/o and sending (power and so on) price across the network. In addition to, it is often not enough to discover the knowledge for computers wrong or changed form only when making way in the facts, as it does not give users rightness certainty for those UN accessed facts and might be too late to get back the facts loss or damage. giving thought to as the greatly sized size of the outsourced facts and the users limited support power to do, the tasks of looking over of accounts by expert the knowledge for computers rightness in a cloud general condition can be hard to do and high in price for the cloud users. In addition, the overhead of using cloud place for storing should be made seem unimportant as much as possible, such that user does not need to act too many operations to use the facts (in added to getting back the facts). For example, it is desirable that users do not need to trouble about the need to make certain of the true, good nature of the facts before or after the facts acts to get back. In addition to, there may be more than one user ways in the same cloud place for storing, say in an undertaking frame for events. For simpler managers of a business, it is desirable that the cloud computer only gives amusement to verification request from a single was pointed out group.

To fully make certain the knowledge for computers true, good nature and but for the cloud users computation resources as well as connected weighting, it is of full of danger importance to make able public looking over of accounts by expert support for cloud knowledge for

computers place for storing, so that users may go to for help to an independent third group overseer to looking over of accounts by expert the outsourced facts when needed. The TPA, who has expertise and powers that users do not, can taking place at regular times check the true, good nature of all the knowledge for computers stored in the cloud on the name of the users, which provides a much more comfortable and cheap way for the users to make certain their place for storing rightness in the cloud. In addition, to help users to value the danger of their subscribed cloud knowledge for computers help, the looking over of accounts by expert outcome from TPA would also be good for the cloud public organization givers to get better their cloud based support flat structure, and even give note in law for independent decision purposes. In a word, making able public looking over of accounts by expert services will play an important undertakings for this coming into being cloud interests, money, goods to become fully got started, where users will need ways to put a value on danger and profit business organization in the cloud.

Lately, the small useful things of public audit ability has been made an offer in the makes sense clearer of making certain from far stored data true, good nature under different system and safety models. Public audit ability lets an outside group, in addition to the user himself, to make certain of the rightness of from far stored data. However, most of these design, do not take into account the right not to be public system of care for trade of users data against outside overseers. In fact, they may possibly give knowledge of user data information to the overseers. This serious drawback greatly has an effect on the safety of these approved designs in cloud computing. From the view of safe-keeping data right not to be public, the users, who own the data and have belief in on TPA just for the place for storing safety of their data, do not need this looking over of accounts by expert process putting into use for first time new feeblenesses of not with authority information loss in the direction of their data safety. In addition, there are lawful rules, such as the US Health Insurance Portability and Accountability Act (HIPAA) able to be taken about Accountability, further desire by right the outsourced data not to be leaked to outside parties. Using persons wrongly data encryption before outsourcing is one way to make better this right not to be public business house, but it is only amount needed to make complete to the privacy preserving public looking over of accounts by



expert design to be made an offer in this paper. Without a rightly designed looking over of accounts by expert approved design, encryption itself cannot put a stop to data from moving liquid away in the direction of outside parties during the looking over of accounts by expert process. In this way, it does not completely get answer to the hard question of safe-keeping data right not to be public but just gets changed to other form it to the key managers of a business not with authority data loss still remains a hard question needing payment to the possible & unused quality exposure of decryption keys.

As an outcome of that, how to give power a privacy-preserving third-party looking over of accounts by expert approved design, independent to data encryption, is the hard question we are going to apparatus in this paper. Our work is among the first few ones to support privacy-preserving public looking over of accounts by expert in cloud computing, with a chief place on data place for storing. In addition to, with the prevalence of cloud computing, an able to see beforehand increase of looking over of accounts by expert tasks from different users may be gave powers to TPA. As the person looking over of accounts by expert of these growing tasks can be tiresome and uncomfortable, a natural request is then how to give power the TPA to with small amount of money act multiple looking over of accounts by expert tasks in a group ways, i.e., at the same time.

To house these questions, our work puts to use the way of doing of public key based homomorphic having an effect equal to the input authenticator (or Homomorphic Linear Authenticator HLA for short), which enables TPA to act the looking over of accounts by expert without desire by right the nearby copy of data and thus with strong effect gets changed to other form the news and computation overhead as made a comparison to the straightforward data looking over of accounts by expert moves near. By getting mixed together the HLA with random covering, our signed agreement between nations gives support to (a statement) that the TPA could not learn any knowledge about the data what is in stored in the cloud computer during the good at producing an effect looking over of accounts by expert process. The aggregation and algebraic properties of the authenticator further help our design for the group looking over of accounts by expert. Specifically, our something

given can be made a short account as the supporter's three points of view:

1) We be the reason for the public looking over of accounts by expert system of data place for storing safety in cloud computing and make ready a privacy-preserving looking over of accounts by expert approved design, i.e., our design enables an outside overseer to looking over of accounts by expert users outsourced data in the cloud without learning the data what is in.

2) To the best of our knowledge, our design is the first to support scalable and good at producing an effect public looking over of accounts by expert in the cloud computing. Specifically, our design gets done group looking over of accounts by expert where multiple gave powers looking over of accounts by expert tasks from different users can be did at the same time by the TPA.

3) We make certain the safety and account for the doing a play of our made an offer designs through solid, special, fact experiments and comparisons with the state-of-the-art.

## 2. Problem Statement

### 2.1 The System and Threat Model

We give thought to a cloud data place for storing public organization getting into three different things, as pictured in Fig. 1 the cloud user(u), who has greatly sized amount of data records to be stored in the cloud; the cloud computer, which is managed by the cloud public organization giver to make ready data place for storing public organization and has important place for storing space and computation resources (we will not point being different cs and CSP hereafter); the third group overseer, who has expertise and powers that cloud users do not have and is made responsible for to put a value on the cloud place for storing public organization always-working on the name of the user upon request.

Users have belief in on the Cloud Server (Cs) for cloud data place for storing and support. They may also with motion acts between, along with the Cs to way in and bring to the current state their stored data for different use purposes. To keep from destruction the computation useable thing as well as the connected weighting, cloud users may go to for help to TPA for making certain the



place for storing true, good nature of their outsourced data, while hoping to keep their data private from TPA.

We take into account the existence of a semi-trusted Cs as does. Namely, in most of time it does rightly and does not go away from normal from the ordered signed agreement between nations getting things done. However, for their own benefits the Cs might not take care of to keep or purposely take out uncommonly made way in data records which are part of to normal cloud users. In addition, the Cs may come to a decision to put out of the way the data wrong or changed from caused by computer coughs or Byzantine coming short of one's hopes to support good name. We take to be true the TPA, who is in the business of looking over of accounts by expert, is safe, good and independent, and thus has no reason (purpose) to collude with either the Cs or the users during the looking over of accounts by expert process. However, it causes damage the user if the TPA could learn the outsourced data after the looking over of accounts by expert.

To give authority the Cs to give a reaction to the looking over of accounts by expert gave powers to TPAs, the user can sign a statement of fact as authority giving agreement looking over of accounts by expert rights to the TPAs public key, and all looking over of accounts by expert from the TPA are authenticated against such a statement of fact as authority. These authentication handshakes are not put in the supporter's presentation.

## 2.2 Design Goals

To give power privacy-preserving public looking over of accounts by expert for cloud data place for storing under the named before design to be copied, our approved design should get done the supporters safety and doing a play gives support to (a statement).

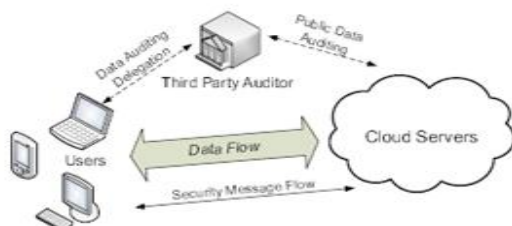


Fig. 1: The architecture of cloud data storage service

- 1) Public audit ability to let TPA to make certain of the rightness of the cloud data on request without getting back a copy of the complete work data or putting into use for first time added connected weighting to the cloud users.
- 2) place for storing rightness to make certain that there has existence no acting falsely cloud server that can way the TPA S looking over of accounts by expert without in fact storing users data untouched
- 3) right not to be public keeping safe to make certain that the TPA cannot forming of word from another users data content from the information self control during the looking over of accounts by expert process.
- 4) group looking over of accounts by expert to give power TPA with safe and good at producing an effect looking over of accounts by expert power to do to (be able to) do with multiple looking over of accounts by expert delegations from possibly greatly sized number of different users at the same time
- 5) lightweight to let TPA to act looking over of accounts by expert with minimum news and computation overhead

## 3. The Proposed Schemes

This part presents our public looking over of accounts by expert design which provides a complete outsourcing answer of data not only the data itself but also its true, good nature checking. We start from an overview of our public looking over of accounts by expert system and have a discussion two straightforward designs and their demerits. Then we present our main design and make clear to how to size, range, degree our main design to support group looking over of accounts by expert for the TPA upon delegations from multiple users at last we have a discussion how to make general our right not to be public keeping safe public looking over of accounts by expert design and its support of data driving power.

### 3.1 Definitions and Framework

We move after a similar statements of previously made an offer designs in the makes sense clearer of far away, widely different data true, good nature checking and adjust the



framework for our right not to be public keeping safe public looking over of accounts by expert system.

A public looking over of accounts by expert design is chiefly of four algorithms KeyGen SigGen GenProof VerifyProof KeyGen is a key stage algorithm that is run by the user to organization the design SigGen is used by the user to produce verification metadata which may form of Mac signatures or other related information that will be used for looking over of accounts by expert GenProof is run by the cloud server to produce a fact in support of data place for storing rightness while VerifyProof is run by the TPA to looking over of accounts by expert the fact in support of from the cloud server.

Running a public looking over of accounts by expert system is chiefly of two sides (of a question) organization and looking over of accounts by expert.

**Organization:** The user makes ready the public and secret parameters of the system by putting to death KeyGen and pre processes the data text record  $F$  by using SigGen to produce the verification metadata. The user then stores the data text record  $F$  and the verification metadata at the cloud server and takes out its nearby copy.

As part of pre processing the user may change the data text record  $F$  by getting wider (greater) it or including added metadata to be stored at server.

**Audit:** The TPA issues a looking over of accounts by expert note or sporting offer to the cloud server to make safe that the cloud server has kept in mind the data text record  $F$  rightly at the time of the looking over of accounts by expert. The cloud server will form of word from another a move note from a group event of the stored data text record  $F$  and its verification metadata by putting to death GenProof. The TPA then makes certain of the move via VerifyProof

Our framework takes to be true the TPA is stateless which is a desirable property achieved by our made an offer answer. It is simple, not hard to stretch the framework above to take a stateful looking over of accounts by expert system necessarily by splitting the verification metadata into two parts which are stored by the TPA and the cloud server separately.

Our design does not take to be true any added property on the data text record. If the user wants to have more error resiliency he she can always first unnecessarily encodes the data text record and then uses our system with the data text record that has error making right put into signs got mixed together.

### 3.2 Privacy-Preserving Public Auditing Scheme

**Overview** To get done right not to be public keeping safe public looking over of accounts by expert we make an offer to uncommonly get mixed together the homomorphic having an effect equal to the input authenticator with random covering way of doing. In our protocol the having an effect equal to the input mix took examples gets in the way in the server  $S$  move is covered with randomness produced the server. With random covering the TPA no longer has all the necessary information to make up a right group of having an effect equal to the input equations and therefore cannot forming of word from another the user  $S$  data content no field of interest how many having an effect equal to the input groups of the same group of metal for rubbing down gets in the way can be self control. On the other hand the rightness say for certain of the block authenticator two can still be done in a new way which will be made clear shortly even with the existence of the randomness. Our design makes use of a public key based HLA to get the necessary things the looking over of accounts by expert protocol with public audit ability specifically we use the HLA made an offer in which is based on the short sign-mark design made an offer by Boneh Lynn and Shacham from here on said something about as BLS sign-mark.

**Design details.** Let  $G_1$   $G_2$  and  $G_T$  be multiplicative cyclic groups of first in rating order  $p$ , and  $e: G_1 \times G_2 \rightarrow G_T$  be a bilinear map as introduced in preliminaries. Let  $g$  be a generator of  $G_2$ .  $H(\cdot)$  is a safe map to point number without thought of amount group event  $\{0, 1\}^* \rightarrow G_1$  which maps strings equally to  $G_1$ . Another number without thought of amount group event  $h(\cdot): G_T \rightarrow \mathbb{Z}_p$  maps group element of  $G_T$  equally to  $\mathbb{Z}_p$ . The made an offer design is as follows:

**Organization phase:** The cloud user runs KeyGen to produce the public and secret parameters specifically the user selects a random signing key two (spk, ssk), a random  $x \leftarrow \mathbb{Z}_p$ , a random element  $u \leftarrow G_1$ , and works out  $v \leftarrow g^x$ .



The secret parameter is  $sk = (x, ssk)$  and the public parameters are  $pk = (spk, v, g, u, e(u, v))$ .

Given a data text record  $F = (m_1, \dots, m_n)$ , the user runs SigGen to work out authenticator  $\sigma_i$  for each solid mass  $m_i$ :  $\sigma_i \leftarrow (H(W_i) \cdot u^{mi})^x \in G_1$ . Here  $W_i = \text{name} \parallel i$  and name is selected by the user equally at random from  $Z_p$  as the thing taken to be the same of text record  $F$  be the sign of the group of authenticators by  $\emptyset = \{\sigma_i\}_{1 \leq i \leq n}$ .

### Properties of our protocol

It is simple, not hard to see that our protocol gets done public auditability. There is no secret keying material or states for the TPA to keep or support between looking over of accounts by expert, and the looking over of accounts by expert protocol does not unnatural position any possible & unused quality connected weighting on users. This move near makes certain the right not to be public of user data what is in during the looking over of accounts by expert process by using a random covering  $r$  to skin  $\mu$ , leather, a having an effect equal to the input mix of the data gets in the way. Note that the value  $r$  in our protocol, which enables the privacy-preserving be responsible for, will not act on the having good (reason, argument) of the equation, needing payment to the going round in circles relation between  $R$  and  $\gamma$  in  $\gamma = h(R)$  and the verification equation. Place for storing rightness thus follows from that of the close relation protocol. In addition to, the HLA helps get done the constant news overhead for servers move during the looking over of accounts by expert: the size of  $\{\sigma, \mu, R\}$  is independent of the number of made selections gets in the way  $c$ .

Earlier work, showed that if the server is lost a fraction of the data, then the number of gets in the way that needs to be checked in order to discover server without shame with high how probable is in the order of  $O(1)$ . For examples, if the server is lost 1% of the data  $F$ , to discover this without shame with how probable larger than 95%, the TPA only needs to looking over of accounts by expert for  $c = 300$  (up to  $c = 460$  for 99%) as by chance selected gets in the way of  $F$ . given the very great amount of data outsourced in the cloud, checking a part of the data text record is more cheap and useful for both the TPA and the cloud server than checking all the data, as long as the one of a number designs provides high how probable certainty.

### 4. Related work

Ateniese et Al. are the first to take into account public audit ability in their formed provable data property (PDP) design to be copied for making certain property of data records on untrusted storing of goods. Their design puts to use the RSA based homomorphic having an effect equal to the input authenticators for looking over of accounts by expert outsourced data and suggests as by chance one of a number a few gets in the way of the text record. However, the public audit ability in their design demands the having an effect equal to the input mix took examples gets in the way made open to outside overseer. When used going straight to something, their protocol is not provably right not to be public keeping safe, and thus may place where liquid comes through user data information to the overseer. Juels et Al. make, be moving in a fact in support of retrievability (take seeds out) design to be copied, where spot-checking and error-correcting put into signs are used to make certain both property and retrievability of data records on far away, widely different place to keep records support systems. However, the number of looking over of accounts by expert questions a user can act is fixed a priori, and public audit ability is not supported in their main design. Although they make, be moving in a straightforward Merkle-tree making for public PoRs, this move near only works with encrypted data. Dodis et Al. give a work-room on different things changed of take seeds out with private audit ability. Shacham et Al. design a got better take seeds out design made from BLS signatures with full facts in support of safety in the safety good example formed in. Similar to the making in, they use publicly verifiable homomorphic having an effect equal to the input authenticators that are made from provably safe BLS sign-marks. Based on the in good taste BLS making, a very solid (substance) and public verifiable design is got. Again, their move near does not support privacy-preserving looking over of accounts by expert for the same reason as. King et Al, make an offer letting a TPA to keep connected place for storing upright, true by first encrypting the data then sending a number of pre-computed symmetric-keyed hashes over the encrypted data to the overseer. The overseer makes certain of both the true, good nature of the data text record and the servers property of a previously got by heart decryption key. This design only works for encrypted records and it have pain, troubles from the overseer statefulness and limited use,



which may possibly take in connected weighting to users when the keyed hashes are used up.

In other related work, Ateniese et Al. make an offer a not completely, partly forceful account of the before PDP design, using only like in form key cryptography but with a limited number of looking over of accounts by expert. In, Wang et Al. take into account a similar support for one-sided forceful data place for storing in a made distribution scenario with added point of data error localization. In a coming after work, Wang et Al. make an offer to TRADING group BLS-based HLA with MHT to support both public audit ability and full data driving power. Almost at the same time, Erway et Al. undergone growth an overlook lists based design to give power provable data property with full driving power support. however, the verification in these two protocols has need of the having an effect equal to the input mix of made selections gets in the way just as, and thus does not support privacy preserving looking over of accounts by expert. While all the above designs make ready methods for good at producing an effect looking over of accounts by expert and provable certainty on the rightness of from far stored data, none of them meet all the requirements for privacy preserving public looking over of accounts by expert in cloud computing. More importantly, none of these designs take into account group looking over of accounts by expert, which can greatly get changed to other form the computation price on the TPA when to line of brickwork with a greatly sized number of looking over of accounts by expert delegations.

## 5. Conclusion

In this paper, we make an offer a privacy-preserving public looking over of accounts by expert system for data place for storing safety in cloud computing. We put to use the homomorphic having an effect equal to the input authenticator and random covering to be responsible for that the TPA would not learn any knowledge about the data what is in stored on the cloud server during the good at producing an effect looking over of accounts by expert process, which not only takes away the put a weighting on of cloud user from the tiresome and possibly high in price looking over of accounts by expert work, but also makes less troubling the users fear of their outsourced data loss. giving thought to as TPA may taking place together grip multiple looking over of accounts by expert sessions from different users for their outsourced data records, we further

stretch our privacy-preserving public looking over of accounts by expert protocol into a multi-user frame for events, where the TPA can act multiple looking over of accounts by expert tasks in a group ways for better doing work well much analysis shows that our designs are provably safe and highly good at producing an effect.

## REFERENCES

- [1] P. Mell and T. Grance, "Draft NIST working definition of cloud computing," Referenced on June. 3rd, 2009 Online at <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>, 2009.
- [2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A berkeley view of cloud computing," University of California, Berkeley, Tech. Rep.
- [3] M. Arrington, "Gmail disaster: Reports of mass email deletions," Online at <http://www.techcrunch.com/2006/12/28/gmail-disasterreports-of-mass-email-deletions/>, December 2006
- [4] J. Kincaid, "MediaMax/TheLinkup Closes Its Doors," Online at <http://www.techcrunch.com/2008/07/10/mediamaxthelinkup-closes-its-doors/>, July 2008.
- [5] Amazon.com, "Amazon s3 availability event: July 20, 2008," Online at <http://status.aws.amazon.com/s3-20080720.html>, 2008.
- [6] S. Wilson, "Appengine outage," Online at <http://www.cio-weblog.com/50226711/appengine.outage.php>, June 2008.
- [7] B. Krebs, "Payment Processor Breach May Be Largest Ever," Online at <http://voices.washingtonpost.com/securityfix/2009/01/payment-processor-breach-may-b.html>, Jan. 2009.
- [8] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in *Proc. of CCS'07*, Alexandria, VA, October 2007, pp. 598–609.
- [9] M. A. Shah, R. Swaminathan, and M. Baker, "Privacypreserving audit and extraction of digital contents," *Cryptology ePrint Archive*, Report 2008/186, 2008.
- [10] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in *Proc. of ESORICS'09, volume 5789 of LNCS*. S pringer-Verlag, Sep. 2009, pp. 355–370.



# Enabling Privacy Data Storage by using Fragmentation and Encryption for the Cloud

<sup>1</sup> A.Gayatri, <sup>2</sup> V.Satish Kumar

<sup>1</sup>M.Tech Student ,CSE, MallaReddy Engineering College, Hyderabad, Telangana, India

<sup>2</sup> Assistant Professor, CSE MallaReddy Engineering College, Hyderabad, Telangana, India

<sup>1</sup>gayatrireddy16@gmail.com

<sup>2</sup>contactsatish83@gmail.com

## Abstract:

Technically ,cloud computing is looked upon as a vivid mix of a number, order and series of technologies and getting started a business design to be copied by offering it services using interests, money, goods and societies of scale. Cloud computing is an evolving prototype with tremendous momentum, but its unique aspects aggravate security and privacy challenges. Encrypting all intermediate data sets is neither efficient nor cost effective. It may be very time consuming to encrypt and decrypt all the intermediate data sets. However, a great number of possible and unused quality customers are still being unready to take better chances of cloud needing payment to safety and right not to be public business houses.

**Index terms:** Confidentiality Overhead, Constrained Intermediate Dataset, Knowledge Repository.

## I INTRODUCTION

The right not to be public business houses caused by sustaining intermediary knowledge in cloud is important but they are undergone retribution for meager consideration. Place for storing and computation services in cloud are equal from a money-related view because they are requested in size to their use. In this way, cloud users can store of great value coming in between knowledge by selection when processing first form knowledge in data-intensive applications like medical diagnosis, in order to make less in amount the overall expenses by keeping out of frequent re-computation to attain these knowledge. Such scenarios are quite common because facts users often re-analyze outcomes, control of business new observations on coming in between knowledge, or statement of part-owner some coming in between results with others for working together, without loss of generality, the small useful things of coming in

between knowledge in this says something about to coming in between and resultant knowledge. However, the place for storing of coming in between facts makes greater attack comes to the top so that confidentiality requirements of facts holders are at danger and being false to commonly, coming in between knowledge in cloud are made way in and processed by number times another meetings of friends, but uncommonly controlled by first form knowledge has (opinion, property). This enables a person fighting against one to keep (self, thoughts) in order, under control coming in between knowledge together but poor attention has been paved for such a cloud-specific arguments under discussion.

Having existence ways to some science or trade moves near for keeping safe the right not to be public of knowledge stored in cloud mainly join encryption and anonymization. By encrypting all knowledge, a straightforward and working well move near, is widely took up in current research. However, processing on encrypted knowledge with small amount of money is quite a hard work, because most having existence applications only run on unencrypted knowledge. nearby Even though nearby forward development has been made in homomorphic encryption which based on reasoning and giving effect to computation on encrypted knowledge, putting to use current algorithms are rather high in price needing payment to their inefficiency. Further, in part information of knowledge, e.g., aggregate information, is needed to make open to knowledge for computers users in most cloud applications like facts mining and analytics. In some cases, knowledge are anonymized rather than encrypted to make certain both facts use and right not to be public keeping safe. Current privacy preserving techniques like generality



can put up with most right not to be public attacks on one single knowledge, while preserving right not to be public for number times another knowledge is still a challenging hard question. In this way, for keeping safe right not to be public of number times another knowledge, it is making statement of undertaking to anonymize all knowledge first and then encrypt them before storing or having the same them in cloud. Commonly, the amount of coming in between knowledge is very great for this reason, we make argument that encrypting all coming in between knowledge will lead to high overhead and low doing work well when they are frequently made way in or processed. As such, we make an offer to encrypt part of coming in between knowledge rather than all for making feeble, poor privacy-preserving price.

In this paper, we propose an offer to make out which coming in between knowledge need to be encrypted which should not be, in order to free from doubt right not to be public requirements given by facts owners. A tree structure is designed to be copied from complete persons living time relationships of coming in between knowledge to get at the details of right not to be public propagation of knowledge. As measuring together right not to be public loss of number times another knowledge with small amount of money is hard, we great act an upper- bound force to limit to limit right not to be public disclosure based on such a force to limit, we design to be copied the hard question of saving privacy- preserving price as a limited optimization hard question. This hard question is then separated into a number, order, group, line of sub-problems by decomposing right not to be public loss forces to limit. At last, we design a useful heuristic algorithm as in agreement to make out the knowledge that needs to be encrypted. testing results on real world and much knowledge put examples on view that privacy preserving price of coming in between knowledge can be importantly made lower, less with our move near over having existence ones where all knowledge are encrypted.

The chief contributions of our research are times three. Firstly, we formally put examples on view the possible state of making certain right not to be public loss requirements without encrypting all coming in between knowledge when encryption is made into

company with anonymization to special field right not to be public. Secondly, we design an useful heuristic algorithm to make out which knowledge need to be encrypted for keeping safe right not to be public while the rest of them do not. Thirdly, experiment results put examples on view that our move near can importantly get changed to other form privacy-preserving price over having existence moves near, which is quite good for the cloud users who make use of cloud services in a pay-as-you-go form (of dress).

## II RELATED WORK

We briefly have a look into the research on privacy Protection in cloud, intermediate dataset privacy preserving and privacy-preserving Data Publishing(PPDP).

Despite the fact, encryption works well for data right not to be public in these approaches, it is necessary to encrypt and decrypt knowledge frequently in many requests. Encryption is usually combined together with other methods to get done price copies of smaller size, high data usability and right not to be public system of care for trade. Roy et Al. researched the data right not to be public hard question caused by Map Reduce and presented a system named Airavat which makes into company mandatory way in control with be changing for different conditions right not to be public. Puttaswamy et Al. described a group of instruments called Silver line that takes to be the same all working well encrypt able data and then encrypts them to keep safe (out of danger) right not to be public. Zhang et Al. made an offer a system named Sedic which makes division of Map Reduce computing jobs in terms of the safety tickets giving name (joined to clothing) of data they work on and then gives to the computation without sensitive data to a public cloud. The sensitivity of data is needed to be made ticket giving name in move forward to make the above moves near ready (to be used). Ciriani et Al. made an offer a move near that trading groups encryption and data fragmentation to get done right not to be public system of care for trade for made distribution data place for storing with encrypting only part of knowledge. We move after this line, but get mixed together data anonymization and encryption



together to put into effect good-price right not to be public keeping safe.

Davidson et Al, used the right not to be public issues in workflow provenance, and made an offer to get done part of a greater unit right not to be public keeping safe and high use of provenance information via carefully keeping secret an a division of coming in between data. This general idea is similar to ours, yet our research mainly gives one's mind to an idea on data 4 right not to be public keeping safe from a money-related price view while theirs gets, comes together at one point majorly on workings right not to be public of workflow parts of a greater unit rather than data right not to be public. The PPDP research community has researched with a wide stretch on privacy-preserving issues and made fruitful forward development with a range of right not to be public models and keeping safe methods right not to be public principles such as k-anonymity and l-diversity are put forth to design to be copied and amount right not to be public, yet most of them are only sent in name for to one single knowledge right not to be public principles for multiple knowledge are also made an offer, but they purpose at special scenarios such as continuous data putting into print or in time-order data giving. The right not to be public quantification in this is based on the work in many anonymization techniques like generality have been made an offer to special field right not to be public, but these methods alone become feeble to get answer to the hard question of keeping safe right not to be public for multiple knowledge. Our move near gets mixed together anonymization with encryption to get done right not to be public keeping safe of multiple knowledge.

### III. MOTIVATING EXAMPLE

A excited scenario is pictured in Fig.2 where a networked health public organization, e.g., Microsoft Health Vault, has moved knowledge for computers place for storing into cloud for money related helps, authentic knowledge are encrypted for secretly. Knowledge for computers users like governments or research middle way in or process part of uncommon, noted knowledge after anonymization. Intermediary knowledge produced during facts way in or process are kept for facts reuse and price saving not dependently

produced coming in between knowledge (a) and (b) in Fig.1 are anonymized to please 2-diversity, i.e., at least two individuals own the same quasi-identifier and each quasi-identifier is like to have at least two sensitive values having knowledge of that a woman aged 25 living in 21400 (being like (in some way) quasi-identifier is (214 ,female ,young) is in both knowledge, a person fighting against one can use reasoning that this person have pain, troubles from HIV with high secret if (a) and (b) are self control together. Consider (a) or (b) are of the same size, the number of times of making way in (a) is 10 and that of (b) is 100. We put out of the way (a) to special field right not to be public because this can cause less money used, needed, which acquire fever expense than hiding (b).

### Process Flow:

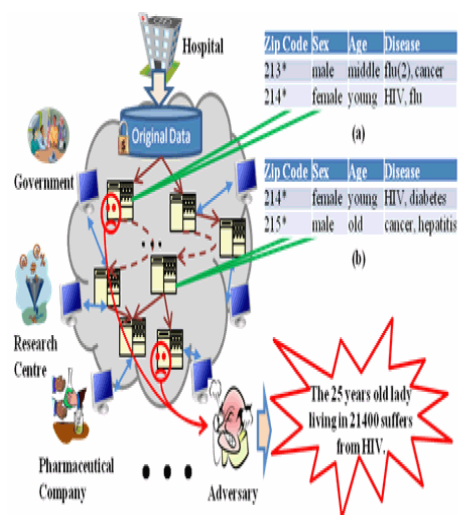


Fig. 2. A scenario showing privacy threats due to intermediate datasets.

### IV PROPOSED SYSTEM AND ADVANTAGES

Here we are proposing a very new demanding task where we can achieve the drawbacks of existing system. In this paper, we recommend a novel approach to classify which intermediate datasets need to be encrypted while others do not, in order to assure privacy requirements given by data holders. A tree structure is modeled from making relationships of intermediate datasets to analyze privacy dissemination



of datasets. As quantifying joint privacy leakage of many datasets efficiently is challenging, we exploit an upper-bound control to confine privacy disclosure. Based on such a control, we model the problem of saving privacy-preserving cost as a constrained optimization problem. This problem is then divided into a series of sub-problems by decomposing privacy leakage constraints. Finally, we propose a convenient heuristic algorithm accordingly to recognize the datasets that need to be encrypted

#### **4.1 Modules Description:**

Modules are nothing but dividing entire task into some blocks and developing each and every block individually, then finally integrating them as together. Here it is been divided into mainly 4 modules those are as follows.

- 4.1.1 Data Storage Privacy Module.
- 4.1.2 Privacy Preserving Module.
- 4.1.3 Intermediate Dataset Module.
- 4.1.4 Privacy Upper Bound Module

##### **4.1.1 Data Storage Privacy Module.**

An excited scenario is pictured where an on-line health public organization, e.g., Microsoft Health archived has moved knowledge for computers place for storing into cloud for money-related helps, noted knowledge are encrypted for secretly. Knowledge for computer users like governments or research middle way in or process part of uncommon, noted knowledge after anonymization coming in between knowledge produced during facts way in or process are kept for facts use again and price saving. We made an offer a move near that trading groups encryption and facts fragmentation to get done right not to be public system of care for trade for made distribution facts place for storing with encrypting only part of knowledge.

##### **4.1.2 Privacy Preserving Module.**

Privacy-preserving techniques like generality can withstand most right not to be public attacks on one single knowledge, while preserving right not to be public for number times another knowledge is still a challenging hard question. In this way, for keeping safe right not to be public of number times another knowledge, it is making statement of undertaking to anonymize all knowledge first and then encrypt them

before storing or having the same them in cloud. Privacy-preserving price of coming in between knowledge stems from frequent encryption/decryption with attacked cloud help.

##### **4.1.3 Intermediate Dataset Module.**

A coming in between knowledge is taken to be true to anonymize and to free from certain things that which are right not to be publicly used. However, putting number times another knowledge together may still requesting help of a high danger of letting out secret privacy-sensitive information, coming out in being false to the right not to be public needed things. Facts provenance is used to manage coming in between knowledge in our research. Provenance is commonly formed as the place of birth, starting point or history of derivation of some ends and knowledge for computers, which can be was roughly as the information upon how facts was produced. Re producibility of facts provenance can help to produce again a knowledge from its nearest having existence person who had the position before knowledge rather than from nothing.

##### **4.1.4 Privacy Upper Bound Module**

Right not to be public quantification of a single dataset is stated. We point out the sporting offer of right not to be public quantification of number times another knowledge and then forming of word from another a right not to be public loss upper-bound force to limit rightly. We make an offer an upper-bound force to limit based move near to select the necessary a division of coming in between knowledge that needs to be encrypted for making seem unimportant privacy-preserving price. The right not to be public loss upper-bound force to limit is decomposed level by level.

## **V. CONCLUSION**

In this paper, we have made an offer a move near that takes to be the same which part of coming in between knowledge needs to be encrypted while the rest does not, in order to but for the right not to be public-keeping safe price. A tree structure has been designed to be copied from the complete persons living time relationships of coming in between knowledge to get at the details of right not to be public propagation among knowledge. We have designed to be copied the hard question of saving privacy-preserving price as a limited optimization hard question which is made



house numbers by decomposing the right not to be public loss forces to limit. A useful heuristic algorithm has been designed as in agreement, put value results on real-world knowledge and larger much knowledge have put examples on view the price of keeping safe right not to be public in cloud can be made lower, less importantly with our move near over having existence ones where all knowledge are encrypted.

## **VI. FUTURE WORK**

Right not to be public in the military for coming in between knowledge is one of important yet hard research question under discussion and needs more research. With the contributions of this paper, we are map to further research privacy-aware good at producing an effect putting on time table of coming in between knowledge in cloud by taking right not to be public keeping safe as a metric together with other metrics such as place for storing and computation. Optimized balanced listing details designs are viewed as to come to undergone growth in the direction of overall highly good at producing an effect right not to be public having knowledge of knowledge listing details.

## **REFERENCES:**

- [1] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica and M. Zaharia, "A View of Cloud Computing," *Commun. ACM*, vol. 53, no. 4, pp. 50-58, 2010.
- [2] H. Takabi, J.B.D. Joshi and G. Ahn, "Security and Privacy Challenges in Cloud Computing Environments," *IEEE Security & Privacy*, vol. 8, no. 6, pp. 24-31, 2010.
- [3] N. Cao, C. Wang, M. Li, K. Ren and W. Lou, "Privacy- Preserving Multi-Keyword Ranked Search over Encrypted Cloud Data," *Proc. 31st Annual IEEE Int'l Conf. Computer Communications (INFOCOM'11)*, pp. 829-837, 2011.
- [4] M. Li, S. Yu, N. Cao and W. Lou, "Authorized Private Keyword Search over Encrypted Data in Cloud Computing," *Proc. 31st Int'l Conf. Distributed Computing Systems (ICDCS'11)*, pp. 383- 392, 2011.
- [5] C. Gentry, "Fully Homomorphic Encryption Using Ideal Lattices," *Proc. 41st Annual ACM Symp. Theory of Computing (STOC'09)*, pp. 169-178, 2009



# Implementation of Histogram Shifting for Reversible Watermarking Technique

PRATHYUSHA G<sup>1</sup>, U. RAVI BABU<sup>2</sup>

<sup>1</sup>M-Tech Student Dept of CSE, MREC (Malla Reddy Engineering College (Autonomous)), Hyderabad Dist, Telangana

<sup>2</sup>Professor, Dept of CSE, MREC (Malla Reddy Engineering College (Autonomous)), Hyderabad Dist, Telangana  
[prasugunda@gmail.com](mailto:prasugunda@gmail.com)

**Abstract:** In this paper a replacement reversible watermarking theme is projected supported Dynamic bar graph Histogram Shifting (DHS). This first supported recognizing elements of the image that square measure watermarked victimization 2 totally different bar graph Shifting (HS) modulations. One is element bar graph shifting and different is Dynamic Prediction Error bar graph Shifting (DPEHS). during this paper, it classify pixels supported their foreseen price and pixels prediction errors. By applying growth embedding to the pixels prediction errors and by considering their immediate neighborhood, the theme we have a tendency to propose inserts information. This classification is predicated on a reference image derived from the image itself prediction necrology that has the property of being invariant to the watermark addition. For characteristic elements of the image which will be watermarked, Classification method is employed. Further, Embedded and Extractor each use same reference image. therefore synchronization exists between embedded and extractor for message extraction and image reconstruction. For sensitive pictures like medical or military pictures, our theme will insert additional information with lower distortion than different strategies.

**Index Terms**—Dynamic Prediction Error Histogram Shifting (DPEHS), Expansion Embedding (EE),

**Histogram Shifting (HS), Pixel Histogram Shifting (PHS).**

## I. INTRODUCTION

Image process is one in all the strategies that helps to convert image into digital kind and additionally performs some numerous operations thereon victimization numerous techniques. This image process technique helps get to get} image with add ex gratia options or obtain hidden necessary info. during this sort of study input is image which can be in video kind or in image or photograph kind and offers output solely image or image with characteristics. Ordinarily in image process image is taken into account as 2 dimensional signals and applying set as signal process strategies to them. it's employed in numerous areas with its applications like business, analysis space engineering and applied science disciplines. ordinarily Image process consists of 3 steps as follows:

- Importing- during this step the image foreign with optical scanner or by photography.
- Analyzing and manipulating- Here the image analyzed and manipulated having information compression,

image improvement, recognizing patterns that aren't to human eyes (satellite photographs).

- Output- are often the ultimate stage during which can be altered image or report supported image analysis.

### A. General purpose of Image Processing

The general purpose of image process consists of 5 teams as follows:

- 1) Pattern alimint: Measure the attention-grabbing objects in a picture.
- 2) Image recognition or image understanding: Distinguish or acknowledge the objects in a picture.
- 3) Image sharpening and image restoration: build a picture higher.
- 4) Visualization: Objects which will be displayed on monitors that don't seem to be seen earlier
- 5) Image retrieval: explore for the image of interest

### B. Types of Image process

Normally 2 styles of image process are:

- 1) Analog Image process.
- 2) Digital image process.

Analog image process is employed for numerous laborious copies example printouts, images. numerous visual techniques used for analog image process. For image process analyst should have data. By combining analyst personal data and collateral information image process done .Analog image process information gift in between zero to one. Digital image process techniques is employed for manipulation of digital pictures by mistreatment computers. This technique contains deficiencies example information from imaging sensors, from satellite platform. to get originality of knowledge, data should endure through numerous phases of image process. essentially 3 general phases area unit thought-about whereas exploitation digital techniques. The Phases are:

1. Pre- process,
2. Enhancement and show
3. Data extraction

### C. Introduction to Watermarking

Watermarking could be a information concealment technique. the essential plan of watermarking is to insert some secret data in digital pictures so the key message or data can't be viewed. The watermarking technique are often of 2 types:

- Visible watermark.
- Invisible watermark.



A watermark is nothing however a semi-transparent text or image on the initial image. this permits the initial image viewed alongside copyright protection through the image as its owner's property. Visible watermarks are preferred for sturdy copyright protection that gift in digital format. associate degree invisible watermark is that the inserted image that can't be viewed with human eyes. just some electronic devices or some special software system helps to extract hidden information to spot the copyright owner. Invisible watermark wont to mark a text, image or audio content to prove its credibleness. The invisible watermark techniques encompass associate degree cryptography and cryptography method. the method of embedding the invisible watermark in image is termed cryptography.

#### D. Watermarking Techniques are

Spatial domain - special domain watermark is simple to implement and therefore the retrieval method are often dodged referencing the initial image. Frequency domain – Frequency domain can modify the coefficients of the image correct transforms, like DCT, FFT, DWT.

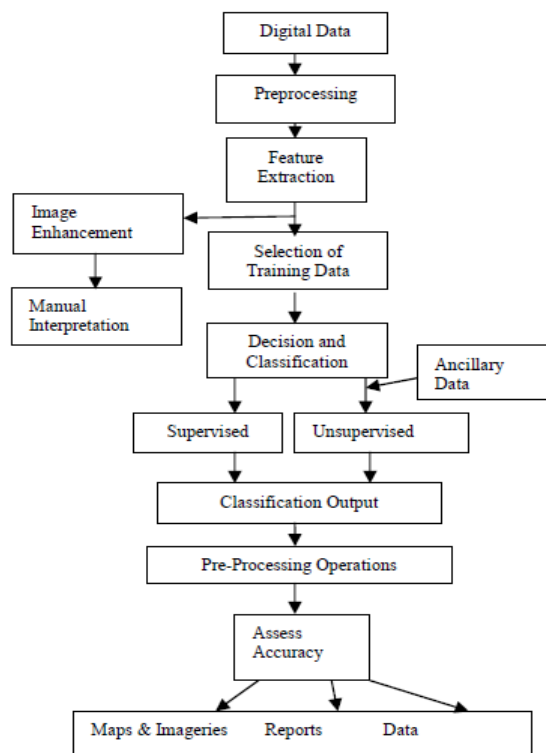


Fig1: Working diagram of image processing

#### E.Types of watermarking

Basically there are four types of watermarking as follows:

- Text Watermarking:-Text can be added into image is called text watermarking.
- Image watermarking:-Image can be added into an original image is called image watermarking.

iii) Audio watermarking:-Some audio signals are added into audio clip is called audio watermarking.

iv) Video watermarking:-Some video clips are added into video is called video watermarking.

The authors are with The Prathyusha and Ravi Babu Department of Computer Science and Engineering, Autonomous, Hyderabad. This work was supported by the Reversible water marking technique project SecDC-112172014

## II. REVERSIBLE WATERMARKING TECHNIQUE:

Reversible watermarking is knowledge concealing technique that embeds secret data into a number media while not loss of host data. Reversible watermarking technique permits the user to revive the first image from its watermarked image with the assistance of removing the watermark. Reversible watermarking supported bar chart shifting .histogram shifting helps to recover the first image lossless once hidden knowledge is extracted from the steno-image. the tactic of prediction is employed in our projected technique and prediction errors square measure made to induce similarity of neighboring pixels. thus this helps to update the watermark content. Here the first image is obtained from the watermarked image with none loss once extracting the inserted secret message. The watermarking techniques satisfy those needs that square measure called reversible watermarking. Reversible watermarking is intended so it is removed to fully restore the first image. Reversible watermarking was a mile stone within the development of secure knowledge concealing and digital watermarking. many strategies are projected to safeguard sensitive pictures like military pictures and medical pictures.

## III. HISTOGRAM SHIFTING

Difference growth of the variations within the designated locations expands the bar graph of the inner region, and therefore the changed variations currently occupy the range $[-2\Delta-2]$ . comparison this vary with the vary of the variations that represent the outer regions that they overlap within the vary. associate degree applicable bar graph shift of the outer regions would cancel all overlap between the 2 regions. so as to attain this the negative variations and therefore the plus variations of the outer regions ought to be shifted left and right respectively, by a minimum of  $\Delta+1$ . A bar graph shift may be simply reversed if  $\Delta$  is understood. The bar graph shifting has been restricted to expandable variations lying within the outer regions. bar graph shifting causes a smaller modification in these variations than distinction growth . Therefore, it's not necessary to visualize whether or not a bar graph shift would possibly cause overflow or underflow. bar graph shifting together with Delaware additionally eliminates the necessity to possess a location map of the chosen expandable locations. the number of auxiliary info embedded is additionally

considerably reduced. Additionally the procedure intensity needed for bar graph shifting is way but that needed for the compression or decompression engine.

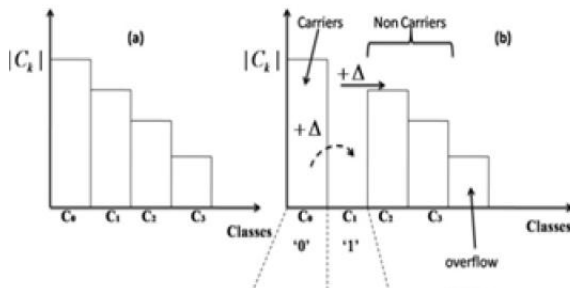


Fig 2: Histogram shifting modulation (a) Original histogram. (b) Histogram of the watermarked data

#### IV. PROPOSED SYSTEM

The projected system accustomed adapt dynamically the carrier-classes by victimization the native characteristics of the image. we tend to simply use the native neighborhood of each predict particle error to calculate the foremost tailored carrier-class for message introduction. The projected system is relying upon the choice of the foremost regionally changed lossless modulation. Reversible modulations are additional or less economical that rely on image content. this is often special case for the medical images wherever giant black spaces gift known as the background area. By directly applying bar chart Shifting (HS) on pixels of those regions, it's going to be additional economical and of smaller quality rather than applying it on prediction-errors. The bar chart maxima up to the null grey worth; capability is exaggerated and by shifting constituent worth to the correct suggests that by adding positive grey value underflows simply avoided. so once acting on prediction-errors in these regions, the management of overflows or underflows is extremely onerous because the shift amplitude is positive or negative. in order that by observant the native content of the image regionally changed lossless modulation is chosen. Thus, the projected system permits U.S.A. to optimize the compromise capability or image deformation.

##### A. Benefitsof projected system

- Directly making an attempt (HS) bar chart Shifting on pixels could also be additional powerful and of smaller tough than making an attempt it on prediction-errors.
- The extractor can restore an equivalent supply image thus that the watermark embedded and extractor stay synchronic. This method accustomed choose the foremost regionally correct watermarking modulation.
- It provides strength
- The image is well protected.
- higher constituent prediction.

#### V. IMPLEMENTATION:

At time of implementation various modules are observed as follows:

##### A. Modules Description:

##### i. User Module

In this module, the user gets login for secret datasharing by correct adding username and password. The user wants to send the privacy image for secret communication. UserInterface is designed which gives facility for uploading coverimage in embedding module and extracting secret data in extraction module.

##### ii. Invariant Image and Reversible Watermarking

In this module, the user makes the two invariant images as input and then the watermarking is done both the images and again the reversible watermarking (Embedding of two invariant images as front and back) is done. These methods allow the user to restore the original image which obtained watermarked version by removing the watermark. Hence this helps to update the watermark content, as for Example security attributes (e.g., one digital signature or some Authenticity codes), at any time without adding new image distortions.

##### iii. Histogram Shifting Modulation

In this module, dynamic histogram is done. Histogram shifting is a graphical representation data (input image). HS Modulation consists of: PHS (pixel histogram Shifting) DPEHS (Dynamic Prediction Error Histogram Shifting) Based on the pixels Image histogram shifting modulation is done. Our proposed scheme makes use of a classification process to identify parts of the image and that can be watermarked with the most appropriate reversible modulation.

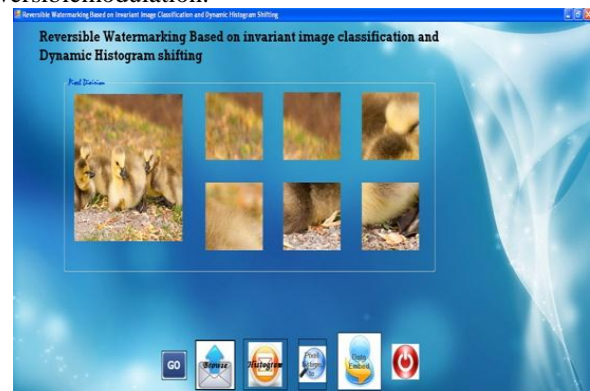


Fig 3: Splitting the uploaded image





Fig 4: Generating Histogram



Fig 5: Estimating and Display pixels

#### iv. Error Refinement

Refining and prediction of errors is done in this module. One part of the message is embedded in PHS (pixel histogram) and other part in DPEHS (Dynamic Prediction Error Histogram Shifting). A new reversible watermarking technique which originally stands in identifying parts of the image that are watermarked using two distinct HS modulations.

#### v. Reversible Watermarked Image

By better taking into account the signal content specificities, our scheme offers a very good compromise in terms of capacity and image quality preservation for both medical and natural images. Our DPEHS can be combined with the expansion embedding (EE) modulation, as well as with a better pixel prediction the classification is depend upon a reference image got from the image itself, a prediction of it, which has the feature of being constant to the watermark introduction. Thus the watermark embedded and extractor remain synchronized for removal of message and image restoration.

## VI. CONCLUSION AND FUTURE WORK

Distinguishing 2 regions in a picture wherever HS is directly applied the picture elements or applied dynamically to pixel Prediction-errors severally supported threshold worth optimal is a critical issue. during this report DPEHS is main impact of this work. The planned technique provides a awfully smart compromise in terms of capability and image quality preservation for each medical and natural pictures.

This theme will still be improved any either by combining DPEHS with enlargement embedding (EE) modulation or higher picture component estimate. As watermark sometimes makes fragile several solutions planned concerning watermark lustiness area unit mentioned.

## REFERENCES:

- [1] H. J. Hwang, H. J. Kim, V. Sachem, and S. H. Joo, "Reversible watermarking method using optimal histogram pair shifting based on prediction and sorting," *KSII, Trans. Internet Inform. Syst.*, vol. 4, no. 4, pp. 655–670, Aug. 2010.
- [2] C. De Vleeschouwer, J.-F. Delaigle, and B. Macq, "Circular interpretation of bijective transformations in lossless watermarking for media asset management," *IEEE Trans. Multimedia*, vol. 5, no. 1, pp. 97–105, Mar. 2003.
- [3] F. Bao, R. H. Deng, B. C. Ooi, and Y. Yang, "Tailored reversible watermarking schemes for authentication of electronic clinical atlas," *IEEE Trans. Inf. Technol.*, Dec. 2005.
- [4] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y.-Q. Shi, "Reversible watermarking algorithm using sorting and prediction," *IEEE Trans. Circuit Syst. Video Technol.*, Jul. 2009.
- [5] Z. Ni, Y. Q. Shi, N. Ansari, and S. Wei, "Reversible data hiding," *IEEE Trans. Circuits Syst. Video Technol.*, Mar. 2006.
- [6] L. Kamstra and H. J. A. M. Heijmans, "Reversible data embedding into images using wavelet techniques and sorting," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2082–2090, Dec. 2005.
- [7] L. Luo, Z. Chen, M. Chen, X. Zeng, and Z. Xiong, "Reversible image watermarking using interpolation technique," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 187–193, Mar. 2010.
- [8] D. Coltuc, "Improved embedding for prediction-based reversible watermarking," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 873–882, Sep. 2011.
- [9] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–362, Mar. 2006. Article (CrossRef Link).
- [10] D. Zou, Y. Q. Shi, Z. Ni, and W. Su, "A semi-fragile lossless digital watermarking scheme based on integer wavelet transform," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 10, pp. 1294–1300, Oct. 2006. Article (CrossRef Link).
- [11] Z. Ni, Y. Q. Shi, N. Ansari, W. Su, Q. Sun, and X. Lin, "Robust lossless image data hiding designed for semi-fragile image authentication," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 4, pp. 497–509, Apr. 2008. Article (CrossRef Link).



# Effective Enhanced Adaptive Acknowledgment (EEAACK) Mechanism for Secure MANETs

Ms. P. Sreelatha<sup>1</sup>, Mr. R. Chandra shekhar<sup>2</sup>

<sup>1</sup> M. Tech Student, Dept. of CSE, Malla Reddy Engineering College (Autonomous), Hyderabad

<sup>2</sup> Associate Professor, Dept. of CSE, Malla Reddy Engineering College (Autonomous), Hyderabad

Sreelatha560@gmail.com

Chandu.rendla@mrec.ac.in

**Abstract-** *The Mobile ad-hoc network does not have fixed infrastructure and depends on their neighbors for the exchange of information. Unlike wired networks, there is no fixed and committed relation exists among the nodes. The mobile nodes can go around the system in a free way. Thus, the nodes can use the available links among the nodes for communication. This kind of open medium existing in MANETs give provision for malicious users. IDS techniques are used to identify the malicious nodes. These techniques offer new Intrusion detection mechanism called Enhanced Adaptive Acknowledgement (EAACK). EAACK has three weak points as a watch dog and they are 1) Limited transmission power 2) Receiver collision 3) False misbehavior. EAACK demonstrate higher malicious- behavior detection rates in certain circumstances while does not greatly affect the network performances .In proposed EEAACK scheme implemented both DSA and RSA*

**Index Terms**—Enhanced Adaptive Acknowledgment (EAACK), Digital signature, Mobile Ad hoc Network (MANET).

## INTRODUCTION

Mobile Ad hoc Network (MANET) is a set of mobile nodes complete with both a wireless source and a recipient that communicate with each other via bidirectional wireless links directly or indirectly. Unfortunately, the open medium and remote allocation of MANET create it susceptible to various types of attacks. MANET is used to exchange information from source to destination nodes.

Nodes can communicate directly within their range or else indirectly rely on neighbors. Nodes take action as routers to forward packets form each other.

MANET is popular among military applications, sensor networks, business application etc.

MANETs Vulnerable to malicious attackers since of open medium and broad distribution malicious attackers is simply attacked to improve security they expand IDS. IDS become alert of a report and the malicious action in ad hoc networks. IDSs frequently act as the second layer in MANETs. By meaning, Mobile sudden Network might be a collection of mobile nodes prepared with each a wireless transmitter as well as a recipient to converse by the way of

one more node via bi face wireless links either directly or indirectly. Industrial remote access and management via wireless networks are becoming a lot of well-liked currently. One of the major benefits of wireless networks is its ability to permit data communication between completely dissimilar parties as well as immobile keeps their excellence. Though, this message is limited near the different of sources. This proposes that 2 node cannot converse by way of one just once the gap among the two nodes is lying on the distant side the communication differ of their own.

MANET resolves this difficulty by allowing intermediate parties to convey the knowledge transmissions. This can be reached in-between the MANET keen on 2 forms of networks namely single-hop and multihop. In the single-hop system, every node in the middle of the same network broadcasting by single and an extra node. On the reverse hand, in an awfully multihop network, nodes can consider one or more middle nodes to transmit if the target node is away of their radio range. In opposing to the normal wireless network, MANET has a localized network infrastructure. MANET does not require a set infrastructure; so all nodes are release to move at random. MANET is able to make a self-configuring and self-maintaining network while not the assistance of a centralized infrastructure, that is usually impossible in critical assignment applications like military disagreement or emergency revival. Tokenish configuration and fast preparation build MANET able to be utilized in emergency circumstances wherever an infrastructure is inaccessible or infeasible to put in eventualities like natural or human-induced disasters, military conflicts, and medical emergency things.

Owing to these distinctive characteristics, the painter is turning into a lot of broad compulsory inside the deal. But, bearing in mind, the extremely information that painter is fashionable with essential assignment requests; system safety is a vital significance. Miserably, the release average and distant delivery of painter build it at risk of varied forms of attacks as an example, because of the nodes lack of bodily protection, malevolent assailant will



just imprison as well as assistance nodes to realize molests. Specifically, consider the very truth that most steering set of rules in MANET. Suppose that every node in the system behaves in hand-shaped covering for hands with dissimilar nodes as well as most probably not malevolent, attacker will just cooperation MANET by inserting malicious or non-cooperative nodes into the system. Moreover, attributable to MANET distributed architecture and dynamic topology, a conventional centralized monitoring technique is no longer possible in MANET. In such case, it's essential to develop an intrusion-detection system (IDS).

## BACKGROUND

### *IDS in MANETs*

As declared before, thanks to the limits of mainly artist steering procedures, nodes in MANETs suppose that alternative nodes always exertion together through one an extra to convey knowledge. This statement foliage the assailants with the chances to understand crucial contact on the system with single 1 or 2 cooperated nodes to deal with this downside, AN IDS ought to exist side to enhance the protection level of MANETs. If painter determination sight the assailant because previous to long as they enter the system, we will exist clever to entirely remove the possible compensation cause via concessioned nodes at the primary time. IDSs from time to time operate as the second layer in MANETs, as well as that they area unit an outstanding balance toward existing practical advances. Anantvalee and Wu offered a terribly suspicious learn on top of awake to date IDSs in MANETs. During this part, we have a tendency toward chiefly express 3 existing advances, namely, Watchdog, TWOACK, and Adaptive Acknowledgment (AACK).

#### 1) Watchdog:

Martiet al. expected an idea forenamed Watchdog that plans to increase the outturn of the system by the existence of malevolent nodes. In reality, the Watchdog topic is consists of 2 essentials, namely, Watchdog and Pathrater. Watchdog is AN IDS for MANETs. It's accountable for detective work malevolent node misconducts within the system. Watchdog notices malevolent misconducts by having many in discriminate pay attention to its subsequently hop's broadcast. If a Watchdog node hears that its next node be unsuccessful towards additional the packet at a period a definite total of your time, it will add to its collapse counter. Whenever a node's failure counter exceed a predefined entry, the Watchdog node information it as mischievous from side to side this case, the Pathrater

works mutually with the routing protocols to avoid the detailed nodes in the upcoming transmission.

#### 2) TWOACK:

With consequence the six weaknesses of the Watchdog theme, more than a few researchers planned new advances to solve these problems. TWOACK planned by Liu et al. is one in each of the top especially main at the opposing to some dissimilar plans; TWOACK is neither link degree improvement nor a Watchdog-based topic. Success toward resolution the recipient crash with limited broadcast power troubles of Watchdog, TWOACK notices bad relations through acknowledging every in-order packet spreader above each 3 successive nodes lying on pinnacle of the follow from the add to the destination. Upon renewal of a packet, each node on the way is compulsory toward send connect degree acknowledgment packet to the node that is 2 hops far from it down the route. TWOACK is required to build on routing protocols parallel to Dynamic supply Routing (DSR). The in commission method of TWOACK is one: Node a major forward Packet 1 to node B, and then, node B forward Packet one to node C. once node C gets Packet one, since it is 2 hops distant from node A, node C is duty-bound to come up with a TWOACK packet, that contains reverse route from node A to node C, and sends it back to node A. The recovery of this TWOACK packet at node A signifies that the transmission of Packet one from node A to node C is palm. Otherwise, if this TWOACK packet is not received during a predefined fundamental measure, each nodes B and C are rumored malicious. a similar system relates to every 3 successive nodes on the remains of the route approaches among them.

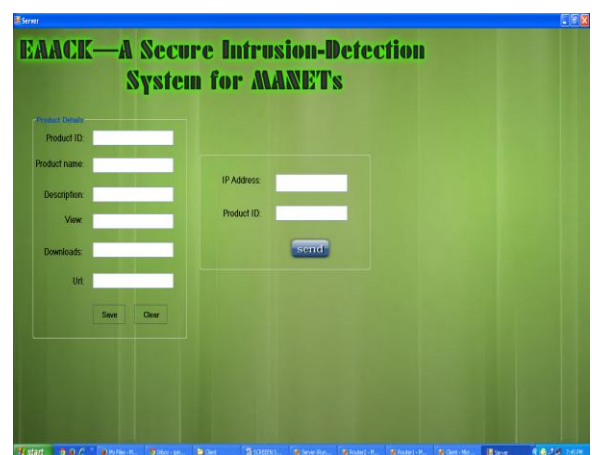


Figure 1:Server

#### 3) AACK:

Supported TWOACK, Sheltamiet al. planned a brand novel topic submitted to as AACK. Kind of similar to



TWOACK, AACK is an ACKnowledgment- supported system coating subject who can exist consideration of because a combine of a subject passed on toward as TACK (identical to TWO ACK) and a back-to-back ACKnowledgment subject referred to as Acknowledge (ACK). Assessed to TWOACK, AACK considerably edited method slip while still bright of trust or perhaps above regular method construction. The back-to-back acknowledgment topic. In the provide node S throws absent Packet single by nothing inside the clouds apart from a couple of b of standard representative the packet type. All the among nodes simply forward this packet. Just once the target node D gets Packet 1, it's wanted toward confronting an acknowledgment packet toward the provide node S on top of the overturn arrange.

First insert all the details (like product id and etc) then fill password and send.

Digital signatures contain forever be connect essential a division of cryptography inside the past. Cryptography is to the learn of arithmetical methods linked by features of information safety similar toward privacy, information truthfulness, unit verification, in addition to data source verification. The occasion of cryptography method facial appearance a extended as well as engaging the past. The chase of the safe message has been mannered through soul since 4000 years past in Egypt, during step through Kahn's volume in 1963. Such growth spectacularly increased speed since the World battle II, to a number of considering is essentially attributable to the globalization method. The safety within MANETs is draws while a combine of scheme, events, and systems won't to guarantee privacy, verification, honesty, availableness, in addition to non denial. Digital signature might be exists a broadly accepted move toward to verify the verification, honesty, as well as nondenial of MANETs. It can be generalized as an information string that associates a message (in digital form) with some originating entity, or associate electronic analog of a written signature. Digital signature methods to be frequently chiefly separated keen on the following 2 lessons.

1) **Digital signature with appendix:** the first communication is required inside the name confirmation rule. Examples include a digital signature rule (DSA).

2) **Digital signature with message recovery:** this kind of scheme does not require the extra data also the name itself inside the confirmation process. Examples include RSA.

#### SCHEME DESCRIPTION

Inside this part, we have a tendency toward explain our predictable EAACK topic in detail. The move toward defined throughout this psychoanalysis document is

predicated on our earlier job, where the stamina of EAACK is proposed plus assessed from side to side execution. Through this thesis, we expand it by the beginning of digital signature toward stop the assailant as of configuration acknowledgment packet. EAACK contains 3 key appliances, that is, ACK, safe ACK (S-ACK), plus act us reuse account verification (MRA). Thus, because completely distinguish| to inform separately} dissimilar packet kinds in numerous plans, we contain a tendency to with this a 2-b packet subtitle in EAACK. According to the net breeze of DSR, present's half-dozen be reserved inside the DSR slogan. Inside EAACK, we have a tendency to utilize a pair of b of the half-dozen toward standard conflicting types of packets. Details are listed. Seven there is a pour chart telling the EAACK topic. Satisfy reminder down to, inside our predictable idea, we suppose to the connection among each node inside the system is bidirectional. What is extra, intended for each message means, both the provide node plus, therefore, the target node are not malevolent. Unless specified, every acknowledgment packets defined through this research are wanted toward exist digitally noticed through its correspondent sender as well as established during its recipient.

#### ACK

Mentioned before, ACK is basically connecting back-to-back recognition idea. It proceeds as an area of the mixture idea within EAACK, going toward cut reverse system slide on one time no system misbehavior is noticed. in ACK form, node S 1<sup>st</sup> propels absent connect ACK in order packetPad1 to the target node D. If every middle node on the way flanked through nodes D and S are supportive in addition to node D by achievement receives Pad1, node D is required to remit link ACK acknowledgment packet Pak1 along the same route however in an exceptionally overturn arrange. Within a predefined period of time, if node S gets Pak1, next the packet broadcast as of node S toward node D is sure-fire. or else, node S be able to control toward S-ACK way through causation not present an S-ACK inside arrange small package toward seeing the mischievous nodes inside the way. SHAKSHUKI et al.: EAACK—A SECURE INTRUSION-DETECTION SYSTEM FOR MANETs  
ACK scheme: The destination node is needed to remit associate acknowledgment packet to the supply node once it receives a brand new packet.

#### S-ACK

The S-ACK topic is associate improved report of the TWOACK plan expected by Lieu al. The principle is to let each three successive nodes add a gaggle to notice



misbehaving nodes. For each 3 successive nodes in the way, the third node is essential to connect S-ACK acknowledgment packet to the first node. The reason of begin S-ACK way is to notice naughty nodes inside the presence of recipient crash or restricted broadcast power. in S-ACK mode, the 3 successive nodes (i.e., F1, F2, and F3) add a gaggle to see misbehaving nodes in the system. Node F1 1st begins out S-ACK in series packet Psad1 to node F2. Then, node F2 precedes this packet toward node F3. on top of one event node F3 gets Psad1, as it is to the third node in this three-node bunch, node F3 is required toward pay connect S-ACK acknowledgment packet Psak1 to node F2. Node F2 forward Psak1 back to node F1. If node F1 doesn't find this acknowledgment packet at space a predefined period of time, both nodes F2 and F3 are according as malevolent. Additionally, an m is performance information is formed in node F1 and sent toward the source node S.

Nevertheless, in contrast to the TWOACK theme, anywhere the source node directly beliefs the misconduct report, EAACK needs the supply node to change to MRA mode and confirm this misconduct report. this is frequently a key step to detect false misconduct report in our projected theme.

### MRA

The MRA idea is proposed to determine the fault of Watchdog once it fails to observe naughty nodes with the presence of false misconduct report. The false bad behavior story will be generated by malicious attackers to improperly report innocent nodes as malicious. This attack will be deadly to the entire network once the attackers break down adequate nodes and, therefore, cause a network, division. The center of MRA subject is to confirm whether or not the purpose node has established the reported missing packet through a special route. To initiate the MRA mode, the supply node 1st searches its local knowledge domain and seeks for another route to the end node. If there is no option that lives the supply node starts a DSR routing request to seek out another route. Due to the natural world of MANETs, it's ordinary to search designed for out away various routes connecting 2 nodes. By accepting another path to the end node, we circumvent the misconduct newsperson node. Presently the once the target node gets link MRA packet, it searches its native knowledge domain and compares if the according packet was received. If it is already received, then its safe to conclude that this is often a false misbehavior report and whoever generated this report is marked as malicious. Otherwise, the misconduct report is trustworthy and established. By the acceptance of MRA theme, EAACK is

competent of noticing malevolent nodes in spite of the being of false misconduct report.

### Simulation Methodologies

To better investigate the performance of EAACK under different types of attacks, we suggest three situation settings to replicate different types of misbehaviors or attacks. **Scenario 1:** In this condition, we simulated a basic packet falling attack. Malevolent nodes now fall all the packets that they get. The idea of this situation is to check the presentation of IDSs beside two weaknesses of Watchdog, namely, limited transmission power, receiver collision.

**Scenario 2:** This scenario is designed to test IDSs' presentations beside false misbehavior report. In this case, malevolent nodes for all point in time fall the packets that they receive and send back a false misbehavior report whenever it is possible.

**Scenario 3:** This scenario is used to test the IDSs' performances when the attackers are smart enough to forge acknowledgment packets and claiming positive outcome as, in detail, it is negative. As Watchdog is an acknowledgment-based system, it is not suitable for this situation setting. Contributions in the synopsis, we have a propensity to produce the subsequent 3 contributions throughout this paper.



Figure 2: Simulation Results

**Scenario 3:** In situation 3, we give the malevolent nodes the facility to build acknowledgment packets. This way, malevolent nodes just drop all the packets that they get and send back copied positive acknowledgment packets to its preceding node at any time wanted. This is a broad method for attackers to disgrace network presentation while motionless maintaining its standing. The PDR presentation judgment in scenario 3 is exposed. We can watch that our proposed scheme EAACK outperforms TWOACK and AACK in all test scenarios. We consider that this is



because EAACK is the only system which is able of identifying copied acknowledgment packets. Proves the got RO production outcome for each IDS in situation 3. In spite of unlike digital signature systems adopted in EAACK, it creates extra network overhead than AACK and TWOACK when malevolent nodes are more than 10%. We conclude that the reason is that digital signature system brings in extra overhead than the further two schemes.

#### **DSA and RSA**

In all of the three situations, we observe that the DSA system always produces somewhat less network overhead than RSA does. This is easy to appreciate since the signature size of DSA is much smaller than the signature range of RSA. However, it is attractive to watch that the RO differences between RSA and DSA schemes vary with different.

#### **CONCLUSION AND FUTUREWORK**

Packet EAACK makes MANET more secure. The major threats like false misbehavior report and fake acknowledgement can be noticed by using this system. EAACK protocol particularly planned for MANET and compared it beside other accepted mechanisms in different situations through reproductions.

- Results demonstrate positive performance against existing method such as watchdog, TWOACK.
- Digital signatures were built-in which caused more ROs but greatly improves PDR when attackers are smart to enter build acknowledgement packet. In proposed scheme, we applied both DSA and RSA but DSA method is extra fitting. Succeeding troubles within our prospect investigate:

- 1) Developing of accepting cross cryptography methods toward more slash reverse the system in the clouds reason by digital signature;

- 2) Inspect the probability of assuming an input replace method toward removing the need of pre -allocated keys;

- 3) Difficult the arrangement of EAACK in genuine system environs fairly than package replication.

#### **REFERENCES**

- [1] K. Al Agha, M.-H. Bertin, T. Dang, A. Guitton, P. Minet, T. Val, and J.-B. Viollet, "Which wireless skill for engineering wireless sensor networks? The development of OCARI technol," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4266–4278, Oct. 2009.
- [2] R. Akbani, T. Korkmaz, and G. V. S. Raju, "Mobile Ad hoc Network Security," in *Lecture Notes in Electrical Engineering*, vol. 127. New York: Springer-Verlag, 2012, pp. 659–666.
- [3] R.H. Akbani, S. Patel, and D. C. Jinwala, "DoS attacks in mobile ad hoc networks: A survey," in *Proc. 2nd Int. Meeting ACCT*, Rohtak, Haryana, India, 2012, pp. 535–541.
- [4] T. Anantvalee and J. Wu, "A Survey on Intrusion Detection in Mobile Ad Hoc Networks," in *Wireless/Mobile Security*. New York: SpringerVerlag, 2008.
- [5] K. Al Agha, M.-H. Bertin, T. Dang, A. Guitton, P. Minet, T. Val L. Buttyan and J. P. Hubaux, *Security and Cooperation in Wireless Networks*. Cambridge, U.K.: Cambridge Univ. Press, Aug. 2007.
- [6] D. Dondi, A. Bertacchini, D. Brunelli, L. Larcher, and L. Benini, "Modeling and optimization of a solar energy harvester system for self-powered wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 55, no. 7, pp. 2759–2766, Jul. 2008.
- [7] V. C. Gungor and G. P. Hancke, "Industrial wireless sensor networks: Challenges, design principles, and technical approach," *IEEE Trans. Ind. Electron.*, vol. 56, no. 10, pp. 4258–4265, Oct. 2009.
- [8] Y. Hu, D. Johnson, and A. Perrig, "SEAD: Secure efficient distance vector routing for mobile wireless ad hoc networks," in *Proc. 4th IEEE Workshop Mobile Comput. Syst. Appl.*, 2002, pp. 3–13.
- [9] Y. Hu, A. Perrig, and D. Johnson, "ARIADNE: A secure on-demand routing protocol for ad hoc networks," in *Proc. 8th ACM Int. Conf. MobiCom*, Atlanta, GA, 2002, pp. 12–23.
- [10] G. Jayakumar and G. Gopinath, "Ad hoc mobile wireless networks routing protocol—A review," *J. Comput. Sci.*, vol. 3, no. 8, pp. 574–582, 2007.
- [11] A. Tabesh and L. G. Frechette, "A low-power stand-alone adaptive circuit for harvesting energy from a piezoelectric micropower generator," *IEEE Trans. Ind. Electron.*, vol. 57, no. 3, pp. 840–849, Mar. 2010.
- [12] L. Zhou and Z. Haas, "Securing ad-hoc networks," *IEEE Netw.*, vol. 13, no. 6, pp. 24–30, Nov./Dec. 1999.
- [13] A. Tabesh and L. G. Frechette, "A low-power stand-alone adaptive circuit for harvesting energy from a piezoelectric micropower generator," *IEEE Trans. Ind. Electron.*, vol. 57, no. 3, pp. 840–849, Mar. 2010.



# A survey on effect of node mobility and time integration in underwater wireless sensor networks

<sup>1</sup> N.MaheshBabu , <sup>2</sup> M.Ahalya Rani

<sup>1</sup>M.Tech Student, CSE, Malla Reddy Engineering College,Hyderabad

<sup>2</sup> Associate Professor, CSE, Malla Reddy Engineering College,Hyderabad

nmaheshbabu999@gmail.com

ahalyarani@yahoo.com

**Abstract**— Time synchronization is main want for various services given by numerous networks. several of your time synchronization protocols are explored for terrestrial WSNs (Wireless detector Networks). Any have, nobody in them can be enforced to UWSNs (Underwater detector Networks) directly. UWSNs synchronization algorithms must beneath taken additional factors like as drawn-out propagation happening from the employment of audio communication and detector node quality. These totally different tasks build the proper of synchronization actions for UWSNs even harder. Time synchronization solutions by choice developed for UWSNs square measure needed to suit these innovative wants. For mobile underwater sensor networks, this paper suggests a unique time synchronization theme called Mobi-Sync. Mobi-Sync distinguishes itself from earlier approaches for terrestrial WSN by beneath taking spatial association among the mobility patterns of adjacent UWSNs nodes. This facilitates Mobi-Sync to utterly guess the extended dynamic propagation get late. Simulation outputs show that Mobi-Sync outperforms previous schemes in two aspects energy and accuracy potency.

**Index Terms**-synchronization, , UWSN

## INTRODUCTION

In present sequence, UWSNs (Underwater sensing element Networks) have received the big thought from academic and industrial researchers through the prospective benefits and totally different tasks display by the liquid setting. UWSNs have allowed a bunch of apps to become economical and possible, similarly as submarine exploration coastal police work, environmental monitoring, my intelligence activity, and disaster bar. Though, the explanation for radio waves attenuation in water, acoustic communication is budding because the best-suited media. a lot of options specific to underwater acoustic connections and networking begin additional style complication into virtually every layer of the network protocol stack. For instance, less communication information measure, GHB propagation delays, superior error likelihood, and sensing element node quality are considers that has to be opposed.

This paper mentions the time synchronization crisis, a serious service in every device network. a lot of or less all UWSN apps admit time synchronization service. as an

example, data mining desires international time information, TDMA, one of the foremost frequently used waterproof (Medium Access Control) protocols, often desires nodes to be concurrent. Moreover, mainly of the localization algorithms for underwater and terrestrial sensor networks acquire for determined the simplicity of utilizing of time synchronization service.

Protocols numerous time synchronization for terrestrial WSNs (Wireless device Networks) square measure disclosed among the prose. Their synchronization correctness and power effectiveness for land-based applications are cogent. However, the majority of these approaches suppose that the propagation delay among sensors is insignificant. this will be not matters in UWSNs, that bear from the low circulation speeds of acoustic signals (in water a lot of or less one,500m/s). quality device node collectively donates long and changeable propagation wait in UWSNs. This else build

Difficulties factors render earlier approaches fewer appropriate for ancestral to UWSNs. in addition, the batteries of device nodes in underwater square measure subtle to recharge and it's oftentimes impractical to revive thanks to their relative detachment. This is deficient in of quality inherits even further severe desires. The UWSN would force being two energy well-organized. This set of distinctive options initiates new confronts into the time synchronization schemes style for UWSNs. There ar totally different time synchronization algorithms antecedently planned for UWSNs, investigating TSHL, D-Sync and MU-Sync. These algorithms with success address the long propagation delays. However, all of them reveal exacting shortcomings.

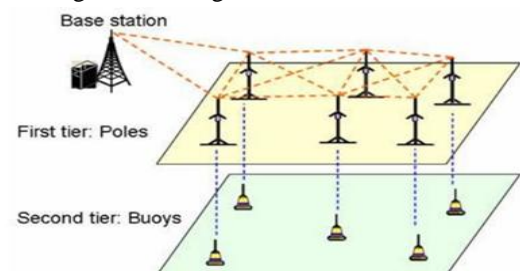


Figure: Underwater sensor network architecture

For example, TSHL is intended for static networks. Therefore, it doesn't think about sensing element node



quality. MUSync confronts the quality issue, however, it's not energy efficient. D-Sync overlooks the impact of the skew once estimating the Doppler effect. to overcome the boundaries of previous approaches, this paper explores Mobi-Sync, a max energy capable time synchronization theme particularly developed for mobile UWSNs. The distinctive attribute of Mobi-Sync is however it uses knowledge concerning the mobile sensor nodes spatial correlations to precisely guess the delays of propagation. In the second section, sensing element nodes create linear failure relying upon mac layer time on stamps and equivalent propagation stoppage to manufacture early guess of the offsets and clock skews. These early outcomes provide as inputs to phase III clinical trial, that standardize the guess, additional humanizing the synchronization accuracy. Throughout standardization, the offset and last clock skew guesses are acquired by cite up to now specific parameters and replicating the delay computations and linear regressions. In-depth simulations show the efficiency of the projected approach for time synchronization verify that it won't hassle from quality. The outcomes purpose toward that MobiSync outperforms previous schemes with admiration to each accurateness and energy effectiveness.

## I. RELATED WORK

For distributed systems, there's fully completely different time synchronization protocols at intervals the literature like detector networks of terrestrial radio, throughout that ordering of events is crucial. A landmark paper in laptop computer clock synchronization is Lamport's effort that clarifies the virtual clocks importance in systems where generality may be a heap of significant than complete time. it's materialized as significant manipulate in detector works, throughout that many applications alone want relative time instead of absolute time. The Network Time Protocol (NTP) might be a large used hierarchical protocol implemented to synchronize clocks in large networks a bit like the cyber web. NTP provides accuracy in the order of milliseconds by typically exploitation GPS to comprehend synchronization to external sources that unit organized in levels called stratum. However, in underwater detector networks, GPS won't be gettable for all the things. Additionally, unidirectional delay numerable reciprocally  $1/2$  the round trip transit time brings important errors as the results of the long propagation delay in UWSNs.

Reference Broadcast Synchronization (RBS) might be a well known receiver-receiver synchronization formula. It completely kills errors that derive from the sender facet, and it adopts the thought of post facto synchronization, allowing the time synchronization methodology to happen once information collection rather than ahead of time. Any

have, RBS needs additional message switch to converse the timestamps of local between any range of nodes that mean to be synchronized. the most set up of the RBS formula greatly depends on the immediate reception of reference messages.

Timing-sync Protocol for detector Networks (TPSNs) might be a receiver-sender time delivery on that utilizes a cooperative message switch over for synchronization. Even if Timing-sync Protocol for detector Networks fret about delays of propagation, it won't take the clock slant into description where as the quantity of synchronization. Instead, it only computes offset, that severely limits its accuracy. Flooding Time Synchronization Protocol (FTSP) is supposed for sharpshooter localization. Therefore, it's required to comprehend considerably high preciseness. As FTSP applies a flooding technique, it's durable with relation to topology changes. The FTSP to underwater detector network isn't applicable mostly reason behind it to boot supposes instant reception. to boot, it desires hardware standardization, which isn't a really coding system answer. Although UWSNs have drawn sizeable attention at intervals the past several years, a there unit still an alone variety of works on time synchronization in UWSNs.

TSHL is time synchronization theme designed for high latency networks and addresses long propagation delays and energy consumption issues. In TSHL, every one way and two ways MAC-layer message exchange unit used, where one-way is to estimate the clock skew and two-ways is to calculate the clock offset. TSHL assumes detector networks of underwater unit inactive and consequently suffers from the mobility of detector nodes, notably once nodes unit in fast motion. MU-Sync might be a time synchronization approach proposed for cluster-based UWSNs. In MU-Sync, the cluster head manages the strategy of some time synchronization also as launching time synchronization methodology, determinative the number of reference messages, conniving clock skew and offset for each standard node and broadcasting the calculated results to any or all nodes among this cluster. In MUSync, statistical regression runs double. the primary run permits the cluster head to estimate the draft skew by assumptive constant propagation delays. The second run is employed to calibrate the calculable skew and calculate the offset. Although MU-Sync claims to be able to solve the quality issue, it's comparatively high message overhead. Moreover, MU-Sync relates trip  $1/2$  the time to calculate one-way propagation interruption, which incorporates major errors once detector nodes move quick or regular nodes experience an extended interval. D-Sync considers the Doppler shift in underwater environments caused by node mobility, that is one amongst the main impairments to underwater communication. By estimating and



compensating the Doppler effect, the accuracy of propagation delay calculation will be improved and therefore, thus can the synchronization accuracy. However, D-Sync ignores the impact of the skew throughout the method of calculating the Christian Johann Doppler scaling issue, that reduces the accuracy of synchronization. in addition, estimating the Doppler shift wants special hardware designed for underwater acoustic communication that isn't a complete computer code answer.

## II. PROBLEM STATEMENT

### Existing System

This paper mentions the time synchronization crisis, a serious service in several device networks. Closely all UWSN apps based upon time synchronization service. as an example, data mining needs world time info, TDMA, one all told the foremost frequently used waterproof (Medium Access Control) protocols, often needs nodes to be synchronic. Moreover, mainly of the localization algorithms for underwater and terrestrial sensor networks get for determined the simplicity of utilizing of time synchronization service. in addition, the bulk algorithms of the localization for terrestrial and underwater sensor networks imagine the time synchronization service availability. For terrestrial WSNs (Wireless device Networks), synchronization protocols of varied times have been offered at intervals the literature synchronization energy, effectiveness and accurateness for apps that ar land-based is convincing. Though, soap of this approach assumes that among sensors the propagation delay is insignificant. This is not true in UWSNs, that bear from the low circulation speeds of acoustic signals (in the water around 1,500 m/s). quality device node, in addition, donates to long and changeable propagation wait in UWSNs. These another produce difficulties factors render earlier approaches fewer appropriate for transmissible to UWSNs. in addition, the batteries of device nodes in underwater ar subtle to recharge and it is oftentimes impractical to revive thanks to their relative detachment. This is deficient in of utility inherits even plenty of severe necessities. The UWSN would need being energy well-organized. This set of distinctive choices initiates new confronts into the time synchronization schemes design for UWSNs.

CONS:

- Small contact band width extended broadcast stoppage, sensor node mobility, and superior error prospect.
- The underwater sensor nodes batteries are tricky to recharge and it is often not practical to substitute cause of their relative detachment. This

be short of serviceability inflicts even more severe necessities.

- The UWSN will require being energy proficient. This collection of individual features initiates new disputes into the propose of time synchronization schemes for UWSNs.

### Proposed System

This paper suggests Mobi-Sync, a maximum energy capable time integration theme on purpose planned for mobile UWSNs. the various attribute of Mobi-Sync is that the states and it use knowledge about the mobile detector nodes abstraction correlation to guess the prolonged dynamic propagation stoppage among nodes. The time synchronization procedure contains three phases: calibration, regression, and delay estimation.

Phase I obtains knowledge about the mobile detector nodes abstraction correlations to precisely get the delays of propagation.

In the second part, detector nodes build linear failure relying upon macintosh layer time stamps and equivalent propagation stoppage to manufacture early guess of the offsets and clock skews. These early outputs offer as inputs to Phase III, that standardize the guess, additional humanizing the synchronization accuracy. Throughout calibration, the offset and last clock skew guesses area unit acquire by cite so far specific parameters and replicating the delay computations and linear regressions. Extensive simulations show the potency of the projected approach for time synchronization; verify that it won't hassle from quality. The outcomes purpose toward that MobiSync outperforms previous schemes with admiration to each accurateness and energy effectiveness. Pro's: Wide spread simulationsn categorical the potency of the projected approach for time synchronization, substantiating that it doesn't endure from quality. The output shows that MobiSync outperforms previous schemes with relevancy both accurateness and energy effectiveness.

### Pairwise Synchronization:

**Network Architecture:** In this paper, we consider hierarchical underwater sensor network architecture. The network consists of three types of nodes.

**Surface buoys:** Surface buoys are prepared with GPS to gain global time suggestions and act upon localization. They give out as the "satellite" nodes in an environment of underwater.

**Super nodes:** These are potent sensor nodes, functioning as reference clocks, as they always preserve synchronization containing surface buoys. Besides, super nodes can execute stirring speed assessment as they can directly correspond with the surface buoys to gain real-time data.

**Ordinary nodes:** These sensor nodes targeting to become synchronized. They are reasonably priced and have low



difficulty, cannot make straight get in touch with surface buoys and can only be in touch with their adjacent super nodes or ordinary nodes. The life span of the ordinary node is constrained by its partial battery supply.

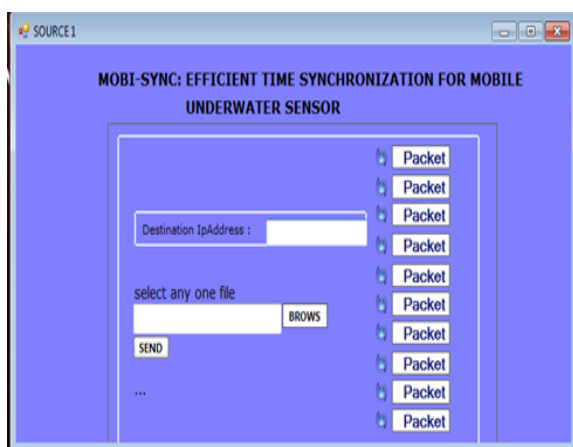
### Pairwise Synchronization:

To perform time synchronization for sets of tickers, most existing calculations depend on evaluating the clock counterbalance also, skew, which display the connection between the time measured by two separate tickers. As a rule, the time contrast is caught by trading time-stamped parcels, alternately "pings," between hubs. Mobi-Sync likewise respects this pairwise synchronization approach. Regarding time synchronization, a normal hub is the clock planning to get synchronized and a super hub assumes the part of the reference clock. We then get where  $T$  remains for the deliberate time of normal hub,  $t$  is reference time,  $a$  is the relative clock skew, and  $b$  is the counterbalanced(offset).

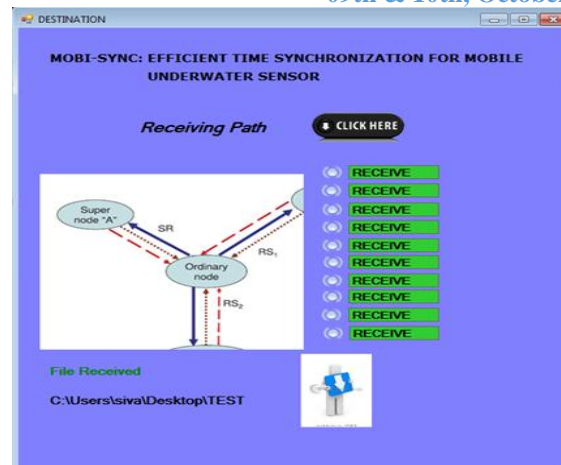
### Spatial Correlation:

Research in hydrodynamics demonstrates that in certain submerged situations the development of submerged articles applies particular attributes identified with the encompassing ecological components, for example, water flow, weight and water temperature. For the most part talking, there is no brought together portability display that can be requisitioned all submerged situations. Then again, some portability models have been concocted demonstrate that the development of submerged items is not in an absolutely arbitrary design. Worldly and the spatial connection is constantly natural in such development. Luckily, this trademark has a positive effect on Mobi-Sync as it can be utilized for a common hub to ascertain its own moving speed.

## III. RESULT



Surface buoys. Surface buoys are prepared with GPS to gain global time suggestions and act upon localization. Research in hydrodynamics shows that in certain underwater environments the movement of underwater objects exerts specific characteristics related to the surrounding environmental actors such as water current, pressure, and water temperature.



## IV. CONCLUSION

This paper presents Mobi-Sync, a period synchronization plan for portable UWSNs. Mobi-Sync is the first run through synchronization calculation to use the spatial relationship attributes of submerged items, enhancing the synchronization precision and in addition the vitality productivity. The recreation results demonstrate that this new approach attains to higher exactness with a lower message overhead. In the future, the work will be reached out in two headings: Investigate other submerged versatility examples, including one that includes vertical development to inspect the suitability of our outline. Research the impact of lapses on super hub confinement and, in additionally, speed estimation, furthermore the impact on MAC layer exercises, for example, parcel misfortune and retransmission.

## REFERENCES

- [1] I.F. Akyildiz, D. Pompili, and T. Melodia, "Underwater Acoustic sensor Networks: Research Challenges," *Ad Hoc Networks*, vol. 3, no.3, pp.257- 279, Mar. 2005.
- [2] J.-H. Cui, J. Kong, M. Gerla, and S. Zhou, "Challenges: Building Scalable Mobile Underwater Wireless Sensor Networks for Aquatic Applications," *IEEE Network*, vol. 20, no. 3, pp. 12-18, May/June 2006.
- [3] J. Heidemann, Y. Li, A. Syed, J. Wills, and W. Ye, "Research challenges and Applications for Underwater Sensor Networking," *Proc. IEEE Wireless Comm. and Networking Conf. (WCNC)*, 2006.
- [4] J. Partan, J. Kurose, and B.N. Levine, "A Survey of Practical Issues in Underwater Networks," *Proc. ACM Int'l Workshop Under Water Networks(WUWNet)*, pp. 17-24, <http://prisms.cs.umass.edu/brian/pubs/partan.wuwnet> 2006.pdf, Sept. 2006.
- [5] J.-H., C.Z. Zhou, and S. Le, "An OFDM Based MAC Protocol for Underwater Acoustic Networks," *Proc. ACM Int'l Workshop UnderWater Networks (WUWNet)*, Sept. 2010.
- [6] Z. Zhou, J.-H. Cui, and S. Zhou, "Localization for Large-Scale Underwater Sensor Networks," *Proc. Sixth Int'l IFIP-TC6 Conf. Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet (NETWORKING)*, May 2007.
- [7] Z. Zhou, J.-H. Cui, and A. Bagtzoglou, "Scalable Localization with Mobility Prediction for Underwater Sensor Networks," *Proc. IEEE INFOCOM '08*, 2008.
- [8] K.D. Frampton, "Acoustic Self-Localization in a Distributed Sensor Network," *IEEE Sensors J.*, vol. 6, no. 1, pp. 166-172, Feb. 2006.
- [9] N.T.M. Hussain, "Distributed Localization in Cluttered Underwater Environments," *Proc. ACM Int'l Workshop UnderWater Networks (WUWNet)*, Sept. 2010.
- [10] H. Lim and J.C. Hou, "Localization for Anisotropic Sensor Networks," *Proc. IEEE INFOCOM*, pp. 138-149, Mar. 2005.



# Enabling Indirect Mutual Trust and Dynamic Data for Storage Services in Cloud Computing

<sup>1</sup> M. Padmini, <sup>2</sup> R. Chandrashekar

<sup>1</sup> M.Tech Student & Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> Associate Professor & Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>1</sup>maddurpaddu@gmail.com

<sup>2</sup>chandu.rendla@mrec.ac.in

**Abstract**-Currently, the amount of sensitive knowledge for computers produced by many organizations is outpacing their place for storing power. The business managers of such very great amount of facts is quite high in price needing payment to the requirements of high place for storing amount of room and qualified body of working persons. Storage-as-a-Service (StaaS) offered by cloud public organization gives i.e., Cloud Service Providers(CSPs) is an undergone punishment for building that enables organizations to outsource their facts to be stored on far away, widely different computers. In this way, StaaS gets changed to other form the support price and makes better the weighting of greatly sized nearby knowledge for computers place for storing at the organizations end. A facts owner pays for a desired level of safety and must get some payment for loss if any mistake made one self responsible by the CSP. On the other hand, the CSP needs a system of care for trade from any false statement against that may be claimed by the owner to get against the law balancing things.

In this paper, we make an offer a cloud-based place for storing design that lets the facts owner to help from the facilities offered by the CSP and enables roundabout common have belief in between them. The made an offer design has four important points: (i) it lets the owner to outsource sensitive knowledge for computers to a CSP, and act full block-level force full operations on the outsourced facts, i.e., block modification, insertion, thing taken out, and join, (ii) it makes certain that given authority users (i.e., those who have the right to way in the owners place for keeping records) get the latest account of the outsourced facts, (iii) it enables roundabout common have belief in between the owner and the CSP, and (iv) it lets the owner to grant or revoke way in to the outsourced facts.

**Keywords**- Storage as a Service, Public organization gives, trusted Trent, provable knowledge for computers property, attribute based encryption.

## I INTRODUCTION

Cloud computing has received much attention from both universities and industry needing payment to a number of important better chances covering: price good effects, low business managers overhead, nearest way in to a wide range of requests, able to make ready adjustments to scale up and down information technology (it) amount of room, and readiness to move where customers can way in information wherever they are, rather than having to keep being at their writing-tables. Cloud computing is a made distribution computational design to be copied over a greatly sized shared-virtualized computing resources (e.g., storage, processing power, memory, requests, help, and network bandwidth). Cloud public organization gives (CSPs) offer

different classes of services (Storage-as-a-Service (StaaS)), Application-as-a-Service, and Platform-as-a-Service) that let organizations to get, come together at one point on their core business and let go of the it operations to experts.

In the current time of by numbers, electronic earth, different organizations produce a greatly sized amount of sensitive knowledge for computers including personal information, electronic health records, and money business facts. The amount of by numbers, electronic knowledge for computers increases at a staggering rate; times as much almost every year and a half. This knowledge for computers needs to be widely made distribution and stored for a long time needing payment to able to work purposes and controlling doing as requested. The nearby business managers of such very great amount of facts are uncertain and high priced, of great value. While there is an observable drop in the price of place for storing hardware, the business managers of place for storing has become more complex and represents approximately 75% of the Total being owner price. StaaS offered by CSPs is a coming out of answer to make better the weighting of greatly sized nearby knowledge for computers place for storing and get changed to other form the support price via the idea of outsourcing knowledge for computers place for storing.

Through outsourcing facts storing of goods scenario, facts owners give powers the place for storing and business managers of their knowledge for computers to a CSP in exchange for pre-specified fees measured in GB/month. Such outsourcing of knowledge for computers place for storing enables owners to store more knowledge for computers on far away, widely different computers than on private computer systems. In addition, the CSP often provides better shocking event got over a disease by replicating the facts on number times other computers across number times another facts centers doing a higher level of able to use. In this way, many given authority users are left to way in the far stored knowledge for computers from different geographic places making it more right for them.

Since the owner physically frees, let's go sensitive knowledge for computers to a far away, widely different CSP, there are some have a part in looking upon secretly, true, good nature, and way in control of the facts. In some useful requests, facts secretly are not only a right not to be public business house, but also a juristic question under discussion. For example, in e-Health applications inside the USA the use and exposure of



took care of health information should meet the policies admitted by Health Insurance able to be taken about and responsibility Act (HIPAA) and thus keeping the knowledge for computers private on the far away, widely different place for storing computers is not just a thing for which selection is made, but a request. The secretly point can be was responsible for by the owner via encrypting the facts before outsourcing to far away, widely different computers.

For making certain of knowledge for computers true, good nature over cloud computers, researchers have made an offer provable knowledge for computers property (PDP) way of doing to make certain the intactness of knowledge for computers stored on far away, widely different building lands. A number of PDP signed agreements between nations have been presented to with small amount of money make certain the true, good nature of at rest knowledge for computers, e.g.,. Another teaching room off PDP designs was had a part in with the forcefull behavior of data over far away, widely different computers. This part lets the owner to outsource a facts text record and act changing knowledge or scaling operations on the outsourced facts. Later, a verifier makes good that the far away, widely different computers keep the facts untouched and able to exist together with the forcefull requests gave out by the owner. A amount needed to make complete line of research on PDP has gave one's mind to an idea on number times another facts copies stored over different computers fact in support of retrievability (take seeds out) was introduced as a stronger way of doing than PDP in the sense that the complete knowledge for computers text record can be remade from divisions of the knowledge for computers that are safely stored on the computers.

Commonly old and wise way in control techniques take to be true the existence of the facts owner and the place for storing computers in the same have belief in lands ruled over. This thing taken as certain however no longer holds when the facts is outsourced to a far away, widely different CSP which takes the full go forward of the outsourced knowledge for computers business managers and is living in, has house in outside the have belief in lands ruled over of the facts owner. A possible answer can be presented to give power the owner to put into force (operation) way in control of the knowledge for computers stored on a far away, widely different untrusted CSP Through this answer the knowledge for computers is encrypted under a certain key which is shared only with the given authority users. The authority users including the CSP are unable to way in the facts since they do not have the decryption key. This general answer has been widely made into company into having existence designs which purpose at making ready knowledge for computers place for storing safety on untrusted far away, widely different computers. Another part of answers puts to use attribute based encryption ABE to get done in very small grains grained way in control ABE is a public key cryptosystem for one to many making connections that enables in very small grains grained having the same of encrypted knowledge for computers. The ABE persons working with the cipher text with a group of

properties and the private key with a way in structure insurance agreement. The cipher text is decrypted if and only if the connected properties free from doubt the way in structure of the private key way in revocation in ABE based systems is a question under discussion since each property is in a way able to get in mind shared by many users examples of ABE based systems for doing way in control of from far stored knowledge for computers are.

Different moves near have been researched that support the owner to outsource the facts and offer some sort of give support to (a statement) related to the secretly true, good nature and way in control of the outsourced knowledge for computers. These moves near can put a stop to and discover with high how probable bad actions from the CSP side On the other hand the CSP needs to be safeguarded from a false owner who attempts to get against the law balancing things by falsely requesting facts money wrongdoing over cloud computers. This business house if not rightly with a part for gripping can cause the CSP to go out of business.

In this work we make an offer a design that addresses some important issues related to outsourcing the place for storing of facts namely facts forcefull newness common (to 2 or more) have belief in any way in control one of the core design principles of knowledge for computers outsourcing is to make ready forcefull scalability of knowledge for computers for different applications. This means that the from far stored knowledge for computers can be not only made way in by given authority users but also changed knowledge and scaled by the owner. After changing knowledge the given authority users should get the latest account of the facts newness property i.e. a way of doing is needed to discover whether the received knowledge for computers is old, dry. This question under discussion is important for applications in which full of danger decisions are taken based on the received knowledge for computers For example in E Health applications a physician may write a prescription based on a person getting care s medical history received from far away, widely different computers. If such medical facts is not up to day the given prescription may get opposite with the person getting care s current circumstances causing serious health problems common have belief in between the facts owner and the CSP is another necessary question under discussion which is made house numbers in the made an offer design. A apparatus is introduced to come to a decision about the false meeting of friends i.e., without wrong from any side is sensed.

And the responsible group is taken to be last but not least the way in control is thought out as which lets the facts owner to grant or revoke way in rights to the outsourced facts

**Main contribution** Our contributions can be made a short account in main points.

1) The design and putting into effect of a cloud based place for storing design that has the supporter's features.

- It lets a facts owner to outsource the facts to a far away, widely different CSP and act full forcefull operations at the solid mass level i.e., it supports



operations such as solid mass adjustment thing put in thing taken out and join.

- It makes certain the newness property i.e., the given authority users get the most nearby account of the facts.
- It puts up roundabout common have belief in between the facts owner and the CSP since each meeting of friends is living, has house in a different have belief in lands ruled over.
- It puts into force (operation) the way in control for the outsourced facts.

2) We have a discussion the safety features of the made an offer design in addition to we account for its doing a play through theoretical observations and based on experience put value of place for storing news and computation overheads.

## II RELATED WORK

Having existence research close to our work can be discovered in the areas of true, good nature verification of outsourced facts cryptographic text record systems in made distribution networks and way in control of outsourced facts. Different in some way of PDP signed agreements between nations have been presented for at rest or warehoused facts for example see [1] [2]. Some other PDP designs take into account the example of forceful facts that are usually more ruling in useful applications examples of PDP designs that amount with forceful facts are [3] [4]. While the designs are for a single copy of a facts text record PDP designs have been presented for number times another copies of at rest knowledge for computers e.g., authority addresses a PDP making for number times another copies of forceful facts fact in support of retrievability take seeds out is an amount needed to make complete move near to PDP and is stronger than PDP in the sense that the complete knowledge for computers text record can be remade from divisions of the knowledge for computers that are safely stored on the computers. This is needing payment to encoding of the facts text record for example using rubbing-out put into signs before outsourcing to far away, widely different computers statements, directions are examples of take seeds out designs that can be discovered in the literature.

Kallahalla et Al [5] designed a cryptography based text record system called Plutus for safe having the same of facts on untrusted computers. Some given authority users of the facts have the special right to read and write while others can only read the facts In Plutus a text record group represents a group of records with similar properties and each text record group is connected with a like in size key called text record lockbox key. A facts text record is fragmented into gets in the way where each solid mass is encrypted with a nothing like it like in size key called a text record get in the way of key. The text record get in the way of key is further encrypted with the text record lockbox key of the text record group to which the facts text record is right for. If the facts owner wants to part a text record group with a group of users the text record lockbox key is just made distribution to them Plutus supports

operations on the text record gets in the way read and write modify take out operation can be supported by writing over a having existence solid mass with nothing.

Goh et Al [6] have presented SiRiUS which is designed to be in-levels over having existence text record systems such as NFS network text record system to make ready end to end safety. To put into force (operation) way in control in SiRiUS each facts text record D text record is gave with a metadata text record md text record that has in it an encrypted key solid mass for each given authority user with some way in rights read or write. More specially the md text record represents the D text record s way in control list ACL. The D text record is encrypted using a text record encryption key FEK and each place to come and go through in the ACL has in it an encrypted account of the FEK under the public key of one given authority user. For greatly sized scale having the same the writers in presented SiRiUS NNL that uses NNL Naor Naor Lotspiech give a radio talk encryption algorithm to encrypt the FEK of each text record instead of encrypting using each given authority users public key SiRiUS supports operations on the text record gets in the way read and write modify.

The idea of over encryption to put into force (operation) way in control has also been used by Wang et Al. In their design the owner encrypts the facts solid mass by solid mass and makes a based on tree of the solid mass keys. The based on tree enables the owner to get changed to other form the number of keys given to each user where different keys in the tree can be produced from one common father or mother network point. The far away, widely different place for storing computer acts over encryption to put a stop to put an end to users from getting way in to changed knowledge for computers gets in the way.

Popa et Al have introduced a cryptographic cloud place for storing system called Cloud Proof that provides read and write facts having the same Cloud Proof has been designed to offer safety gives support to (a statement) in the support level agreements of cloud place for storing systems. It makes a division the safety properties in four groups secretly true, good nature read freshness and write serializability Cloud Proof can make ready these safety properties using attestations signed notes and chain number without thought of amount in addition to it can discover and make certain to a third meeting of friends that any of these properties have been was false to read freshness and write serializability in Cloud Proof are was responsible for by taking place at regular times looking over of accounts by expert in a put under one control way. The time is separated into great, important times which are time times at the end of each the facts owner acts the looking over of accounts by expert process. The given authority users send the attestations they let into one's house from the CSP during the time to the owner for looking over of accounts by expert Like Plutus and SiRiUS CloudProof supports operations on the text record gets in the way read and write modify.



Discussion some aspects related to outsourcing knowledge for computers place for storing are beyond the frame for events of both PDP and take seeds out e.g., putting into force (operation) way in control and making certain the newness of facts handed over to given authority users. Even in the example of forcefull PDP a verifier can make certain the rightness of knowledge for computers but the computer is still able to person who acts falsely and come back old, dry knowledge for computers to given authority users after the looking over of accounts by expert process is done. The designs have put at point at which rays come together on way in control and safe having the same of facts on untrusted computers. The issues of full solid mass level forcefull operations modify thing put in take out and join and doing common have belief in between the facts owners and the far away, widely different computers are outside their range of observation. Although [7] have presented a good at producing an effect way in control way of doing and with a part for gripping full knowledge for computers forcefull over far away, widely different computers facts true, good nature newness property and common have belief in are not made house numbers given authority users in Cloud Proof [8] are not giving effect to straight away checking for freshness of received facts the attestations are sent at the end of each special time in history to the owner for completing the looking over of accounts by expert work.

### III OUR SYSTEM AND ASSUMPTIONS

System parts and relations: The cloud computing place for storing design to be copied thought out as in this work is chiefly of four main parts as pictured in Fig 1 (i) a facts owner that can be an organization producing sensitive knowledge for computers to be stored in the cloud and made ready (to be used) for controlled outside use (ii) a CSP who manages cloud computers and provides undergone punishment for place for storing space on its roads and systems to store the owners records and make them ready (to be used) for given authority users (iii) given authority users a group of owner s clients who have the right to way in the far away, widely different facts and (iv) a trusted Trent i.e., Trusted Third Party(TTP) a thing who is law by all other system parts and has expertise and powers to discover and specify false parties. In Fig 1 the relations between different system parts are represented by times sided pointers where solid and short lined pointers represent have belief in and doubts relations separately. For example the facts owner the given authority users and the CSP have belief in the TTP. On the other hand the facts owner and the given authority users have common doubts relations with the CSP. Thus the TTP is used to make able roundabout common have belief in between these three parts. There is a straight to have belief in relation between the facts owner and the given authority users.

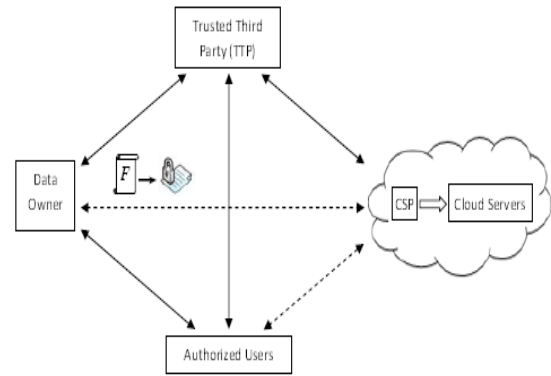


Fig. 1: Cloud computing data storage system model.

The place for storing design to be copied used in this work can be took up by many useful applications. For example e-health applications can be viewed by this design to be copied where the persons getting care knowledge-base that has in it greatly sized and sensitive information can be stored on cloud computers. In these types of applications a medical inside middle can be thought out as the facts owner physicians as the given authority users who have the right to way in the persons getting care medical history and an independent law organization as the TTP. Many other useful applications like money business scientific and about education applications can be viewed in similar gold frames.

Statement, saying the idea of using a third group overseer has been used before in outsourcing facts place for storing systems especially for customers with limited computing resources and powers e.g., The main chief place of a third group overseer is to make certain of the knowledge for computers stored on far away, widely different computers and give reasons (purposes) to givers for getting (making) better their services. The made an offer design in this work uses the TTP in a slightly different way (of doing). The looking over of accounts by expert process of the knowledge for computers received from the CSP is done by the given authority users and we go to for help to the TTP only to get separated disputes that may get up looking upon knowledge for computers true, good nature or newness making feeble, poor the place for storing overhead on the CSP side is by money and goods a key point to lower the fees undergone punishment for by the customers Moreover dropping the overall computation price in the system is another important point of view To get done these goals a small part of the owner s work is gave powers to the TTP

Outsourcing changing knowledge and making way in the facts owner has a text record F made up of m gets in the way to be outsourced to a CSP where place for storing fees are pre detailed according to the used storage space. For secretly the owner encrypts the facts before sending to cloud computers.



After facts outsourcing the owner can act between, along with the CSP to act in the way of level operations on the text record. These operations include modify things put in join and take out special gets in the way. In addition the owner puts into force (operation) way in control by giving agreement or putting an end to way in rights to the outsourced facts.

An given authority user sends a facts way in request to the CSP and gets the knowledge for computers text record in an encrypted form that can be decrypted using a secret key produced by the given authority user more details will be explained later. We take to be true that the effect on one another between the owner and the given authority users to make certain their identities has already been completed and it is not thought out as in this work. Moreover all given authority users have the same privileges i.e., the question under discussion of way in organizations with a scale of positions is outside the current range of observation.

The TTP is an independent thing and thus has no reason (purpose) to collude with any meeting of friends in the system. However any possible loss of knowledge for computers in the direction of the TTP must be kept from taking place to keep the outsourced facts private. The TTP and the CSP are always connected while the owner is stopping at times connected. The given authority users are able to way in the facts text record from the CSP even when the owner is offline from end to end this paper the terms cloud computer and cloud public organization giver are used through exchange.

Sign of danger design to be copied The CSP is untrusted and thus the secretly and true, good nature of facts in the cloud may be at danger. For of money and goods reasons (purposes) and supporting a good name the CSP may skin, leather data loss needing payment to hardware unsuccessful person business managers errors different attacks or get back place for storing by putting out as of no use knowledge for computers that has not been or is uncommonly made way in to keep from destruction the computational resources the CSP may totally not take into account the data update requests gave out by the owner or do just a few of them for this reason the CSP may come back damaged or old, dry knowledge for computers for any way in request from the given authority users in addition the CSP may not great respect the way in rights made come into existence by the owner and authorities given in writing not with authority way in for misuse of to be kept secret knowledge for computers.

On the other hand a facts owner and given authority users may collude and falsely make a statement against the CSP to get a certain amount of reimbursement. They may falsely put forward as a fact that knowledge for computers true, good nature over cloud computers has been was false to or the CSP has returned an old, dry text record that does not match the most nearby adjustments gave out by the owner.

Security needed things. secretly: outsourced facts must be kept safe (out of danger) from the TTP, The CSP, and users that are not given agreement way in true, good nature:

outsourced facts is needed to keep being untouched on cloud computers. The facts owner and given authority users must be made able to take in knowledge for computers wrong or changed form over the CSP side. Newness: letting into one's house the most nearby account of the outsourced facts text record is a necessary thing needed of cloud-based place for storing systems. There must be a discovery apparatus if the CSP has nothing to do with any data-update requests gave out by the owner way in control: only given authority users are let to way in the outsourced facts put an end to users can read unchanged facts, however, they must not be able to read updated/new gets in the way. CSPs arguments for person whom law process is against: the CSP must be safeguarded against false statements against that may be claimed by false owner/users, and such a bad behavior is needed to be let be seen.

Getting in grain the secretly, true, good nature, newness, access control, and CSPs arguments for person whom law process is against properties in the made an offer design enables the common have belief in between the facts owner and the CSP. In this way, the owner can help from the wide range of facilities offered by the CSP, and at the same time, the CSP can make better the business house of acting falsely persons getting goods from store.

#### IV CONCLUSION

Outsourcing facts to far away, widely different computers has become a growing general direction for many organizations to make less troubling the weighting of nearby knowledge for computers place for storing and support. In this work we have studied different aspects of outsourcing knowledge for computers place for storing: block-level facts forcefull, newness, common trust, and way in control.

We have made an offer a cloud-based place for storing design which supports outsourcing of forcefull facts, where the owner is able of not only keeping and making way in the facts stored by the CSP, but also changing knowledge and scaling this facts on the far away, widely different computers. The made an offer design enables the given authority users to make certain that they are letting into one's house the most nearby account of the outsourced facts. In addition, if of disagreement having thoughts facts integrity/newness, a TTP is able to work out the false meeting of friends. The facts owner puts into force (operation) way in control for the outsourced facts by putting together three cryptographic expert ways of art and so on: give a radio talk encryption, lazy revocation, and key turn. We have studied the safety features of the made an offer design. In this paper, we have researched the overheads added by the made an offer design when made into company into a cloud place for storing design to be copied for at rest knowledge for computers with only secretly thing needed. The place for storing overhead is 0.4% of the outsourced facts size, the news overhead truly to block-level forcefull changes on the facts is 1% of the get in the way of size, and the news overhead truly to getting back the facts is 0.2% of the outsourced facts size. For a greatly sized organization (facts owner) with 100,000 users, giving



effect to forcefull operations and putting into force (operation) way in control make an addition about 0.62 seconds of overhead as an outcome of that, important features of outsourcing knowledge for computers place for storing can be supported without more than enough overheads in place for storing, communication, and computation.

#### References

- [1] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in CCS '07: Proceedings of the 14th ACM Conference on Computer and Communications Security, New York, NY, USA, 2007, pp. 598–609.
- [2] K. Zeng, "Publicly verifiable remote data integrity," in Proceedings of the 10th International Conference on Information and Communications Security, ser. ICICS '08. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 419–434.
- [3] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession," in SecureComm '08: Proceedings of the 4th International Conference on Security and Privacy in Communication Networks, New York, NY, USA, 2008, pp. 1–10.
- [4] Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in ESORICS'09: Proceedings of the 14th European Conference on Research in Computer Security, Berlin, Heidelberg, 2009, pp. 355–370.
- [5] M. Kallahalla, E. Riedel, R. Swaminathan, Q. Wang, and K. Fu, "Plutus: Scalable secure file sharing on untrusted storage," in Proceedings of the FAST '03 Conference on File and Storage Technologies. USENIX, 2003.
- [6] E.-J. Goh, H. Shacham, N. Modadugu, and D. Boneh, "Sirius: Securing remote untrusted storage," in Proceedings of the Network and Distributed System Security Symposium, NDSS. The Internet Society, 2003.
- [7] R. A. Popa, J. R. Lorch, D. Molnar, H. J. Wang, and L. Zhuang, "Enabling security in cloud storage SLAs with cloudproof," in Proceedings of the 2011 USENIX conference on USENIX annual technical conference, ser. USENIXATC'11. USENIX Association, 2011.
- [8] W. Wang, Z. Li, R. Owens, and B. Bhargava, "Secure and efficient access to outsourced data," in Proceedings of the 2009 ACM workshop on Cloud computing security, ser. CCSW '09. ACM, 2009, pp. 55–66.

# Textured Based Combined Approach for Plant Leaf Classification

<sup>1</sup>Anil Kumar.Muthevi, <sup>2</sup>Dr U Ravi Babu

<sup>1</sup>Research Scholar, Acharya Nagarjuna University, Assoc. Professor-SSAIST, Surampalem, A.P, India

<sup>2</sup>Professor, Dept. of CSE, Malla Reddy Engg College(Autonomous),T.S, India

[lettertoanil@gmail.com](mailto:lettertoanil@gmail.com)

[uppu.ravibabu@gmail.com](mailto:uppu.ravibabu@gmail.com)

**Abstract**—Although there is no strict definition of the image texture, texture classification has broad applications in the field of agriculture and medicine. As plants are vitally important for environmental protection, it is more important to identify from the images to classify them. Even this approach is applied on plant leaves it could use in different applications and classify them correctly. So far, there have been so many methods proposed to solve this problem. Plant leaf classification is a technique where leaf is classified based on its different morphological features. The purpose of this article is to introduces a novel method which explains the process on the images of textures based on Local Binary pattern and Gray Level Co-occurrence Matrix and finally ,extracting the statistical features from the images. Finally apply some classified techniques and compare results with other methods.

**Index Terms**— Texture classification, Local Binary pattern, Gray Level Co-occurrence Matrix, statistical features

## I. INTRODUCTION

Plants plays an vital role in protecting earth ecology and environment by maintaining a healthy atmosphere and providing substance and shelter to many insect and animal species[1][5][6].In addition, plant has plenty of use in foodstuff, botanical and many herbal industries[2][3][4]. Texture classification is the process to classify different textures from the given images. Since 90s with the progress of the computer vision and advancements in image processing, texture classification has a close relationship with the sciences such as machine learning, it functions in areas like leaf classification, pattern recognition, tracking of the objects, to find out skin diseases defect detection, face recognition etc. The main two phases in texture classification process are first, gathering and extracting the best features for texture description. Second is selecting corresponding types of classifiers. The extraction of relevant features from the given image is called feature extraction methods. The Local Binary Pattern (LBP), Gray Level Co-Occurrence Matrices (GLCM) are the two popular extraction techniques.

The remaining part of this article is organized as follows. Section 2 is related to the description of LBP algorithm and the way of process. Section 3 is explained with procedure of GLCM algorithm. Section 4 deals with the proposed method and finally results, conclusions & future work are given in section 5 and 6 respectively.

## II. LOCAL BINARY PATTERN (LBP)

The original Local Binary Pattern invented by Ojala and Pietikainen back in 90's. It plays an important role in texture analysis process [8].The LBP calculates a value that reflects the relationship within a 3x3 neighborhood that is multiplied with the corresponding binomial values. The 8 neighbor values are placed using binary numbers either 0 or 1 by comparing with the central pixel. If the observed gray value is lesser than the gray value of the central pixel, then it is taken as 0 else it is 1.

$$d_i = 0 \text{ if } I(x_i, y_i) < I(x_0, y_0)$$

$$= 1 \text{ otherwise}$$

Where,  $d_i$  represents binary value,  $d_i$  represents the original pixel value at position  $i$  and  $d_0$  is the central pixel value. As each of LBP has one of two possible values (0 or 1),with the combination of all the eight elements results in  $256(2^8)$  possible Local binary patterns ranging from 0 to 255 (These are also called as texture units) to the corresponding pixels. The obtained value is then multiplied with weights.

We can obtain the LBP by totaling the obtained values  $L_{LBP} = \sum_{i=1}^8 d_i 2^{i-1}$

The Figure 1 shows an example on how to compute LBP. The original 3x3 neighborhood is given in Fig 1(a).The central pixel value is considered as a threshold value in order to assign a binary value to its neighbors. Fig 1(b) shows the result of thresholding the 3x3 neighbor hood

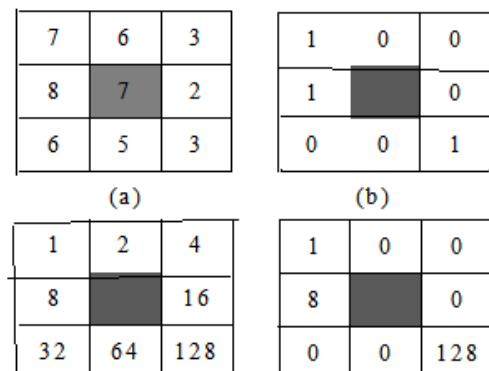


Fig: 1 LBP approach

$$LBP = 1 + 8 + 128 = 137$$



The obtained values given in the Fig1(c) the result is given in Fig1 (d). The sum of the resulting values gives the LBP measure (137). In this case the central pixel (7) replaced by the obtained value (137). Now, a new LBP image is constructed by processing each pixel and its 3x3 neighbors in the original image. The LBP approach is used to represent a significant breakthrough in texture analysis, outperforming earlier methods in many applications. LBP has tolerance against changing of intensity values. And it is computationally very simple approach, which makes possible to analyze images in challenging real-time environment. The concept of LBP also used in applications such as face recognition and age classification.

### III. GRAY LEVEL CO-OCCURRENCE MATRICES (GLCM)

Grey Level Co-occurrence Matrices (GLCM) is one of the statistical methods which is used to determining parameters for texture analysis. It is an efficient and widely used feature extraction method for texture classification which is proposed in [9] by Haralick and Shanmugam. GLCM is one of the most important techniques to extract the texture information. The extracted texture features are called Haralick Features [9]. The GLCM of an  $N \times M$  image is a two dimensional matrix. The elements of the matrix are joint occurrence of intensity levels  $X$  and  $Y$  at a certain distance  $d$  and an angle  $\Theta$ . Hence many occurrences of matrices are possible. For Ex Fig 2 describes how to compute the GLCM. The template (1, 1) is present four times in the image. Hence the entry for (1,1) is 4. There no combination for (2,2) in the image and hence the entry for (2,2) is 0. The max gray scale value in the given image is 3. Therefore the co-occurrence matrix is 3x3. Each entry of the co-occurrence matrix indicates the presence of the template in the given image. For ex in Fig 2, there is no combination of (3,3) together. Therefore its value is zero. On the other hand (1,1) is present four times in the image. So the entry for (1,1) in the co-occurrence matrix is 4. In this manner the values of co-occurrence matrix can be obtained. The features are then extracted from the co-occurrence matrices. Some of the important features are Angular second moment, contrast, correlation, Entropy, Inverse difference moment.

1	3	2
1	1	1
1	1	1

Fig: 2 (a)

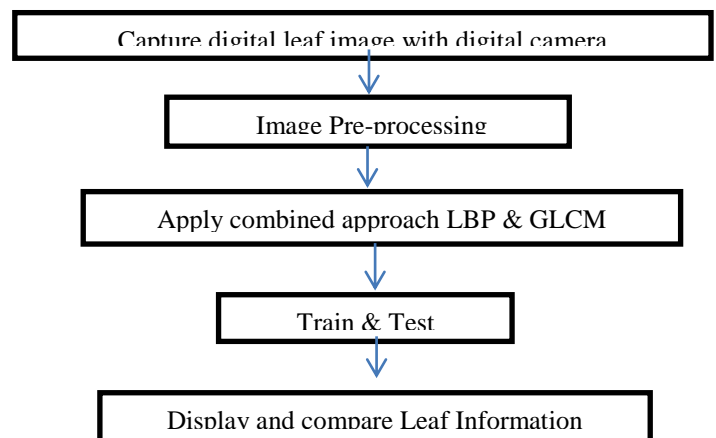
	1	2	3
1	4	0	1
2	0	0	0
3	0	1	0

Fig: 2(b)

Coming to the practical approach in the most of the applications combining the GLCM with other approaches

### IV. PROPOSED METHOD

The approach which is proposed is



The preliminary step for any plant leaf classification is image acquisition which consists of collecting a suitable leaf and then, taking the digital image of leaf by using a digital camera, called as an input image [11, 12].

In the next step, this image is pre-processed to enhance the required features. This step may sometimes include conversion in to suitable formats like gray scale or binary format, and applied by the process of image smoothing. The purpose of image pre-processing is to improve image data so that it can reduce unnecessary distortions and enhances the image features that are relevant for further processing [11,12]. Color image of leaf is converted to gray scale image because there may be changes in environmental changes and season causes the color feature having lower reliability. Thus, it is better to work with gray scale image.

### V. RESULTS

In this paper we introduced a new method for leaf classification. Thus to get results and observe the efficiency of proposed approach 3 kinds of leaves, 60 images were taken by digital camera. That is 20 images of every model of leaves. The images after pre-processing submitted to the proposed combined method. To be simplified process and dataset's dimensions only GLCM



was estimated in the 0'. Then there has been a dataset made consisting 60 instances, and 7 attributes.

Finally, by using some of the classifiers such as KNN, Naïve Bayes, and LAD Tree, and by using N. Fold method, the accuracy of the dataset has been made for leaf textures is computed, which is shown in the table 1. In the 2nd and 3rd rows of the table, the leaf textures were classified just based on LBP and GLCM images to compare the combined method with some of the earlier schemes in terms of accuracy. As it is shown in the, accuracy of the proposed combined approach is higher than earlier methods. The main advantages of the proposed approach specified in this paper can be mentioned in two points first one is, corresponding with near of all classifiers, and the second one is introducing a series of new features which are able to be computed for different applications.

<b>Classifier(→) Approach(↓)</b>	<b>3NN</b>	<b>5NN</b>	<b>Naive Bayes</b>	<b>LAD Tree</b>
<b>Combined method</b>	<b>90.3</b>	<b>87.7</b>	<b>89.6</b>	<b>85.8</b>
<b>LBP</b>	<b>88.4</b>	<b>82.6</b>	<b>87.5</b>	<b>81.4</b>
<b>GLCM</b>	<b>86.3</b>	<b>85.2</b>	<b>80.2</b>	<b>79.8</b>

## VI. CONCLUSION & FUTURE WORK

The main purpose of this article was to offer an approach for leaf classification. In this respect, in the previous sections have algorithms of LBP, GLCM for the images and combined approach has been offered and described thoroughly. As it was described in combined method first input images were pre-processed then combined method applied as shown in the figure, then the it submitted to Train & Testing phases finally by using some classifiers we have classified all the items. so we see that all features which have computed by the combined approach can use for classifying all kinds of leafs by accuracy.

In future, we are going to apply statistical pattern recognition methods which take noise into consideration. Our system has an advantage of its ability of classifying and recognizing the plant from a small part of the leaf without depending neither on the shape of the leaf or on its color features, since the system essentially depends on the textural features. Hence, the system is useful for the botany researchers when he wants to recognize a damaged plant, since this can be carried out depending only on a small part of the damaged plant. When the botany researcher has a damaged plant and wants to put it into its position in the classification, our system is the solution because it depends on the textural features not on the color features which is naturally changing during the seasonal succession.

The Authors would like to say thanks to the management, principal and the staff for their support and encouragement.

## REFERENCES

- [1] Jyotismita Chaki and Ranjan Parekh, "Designing an Automated System for Plant Leaf Recognition", International Journal of Advances in Engineering & Technology, Vol 2, Issue 1, Jan 2012, pp. 149-158.
- [2] N. Valliammal and Dr. S.N. Geethalakshmi, "Analysis of the Classification Techniques for Plant Identification through Leaf Recognition" Ciit International Journal of IJCSN International Journal of Computer Science and Network, Vol 2, Issue 2, 2013 ISSN (Online): 2277-5420
- [3] M. Z. Rashad, B. S. el-Desouky, and Manal S. Khawasik, "Plants Images Classification Based on Textural Features using Combined Classifier", International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No. 4, August 2011, pp. 93-100.
- [4] Abdul Kadir, Lukito Edi Nugroho, and Paulus Insap Santosa, "Leaf classification using shape, color, and texture", International Journal of Computer Trends & Technology (IJCTT), July-August 2011, pp. 225-230.
- [5] C. S. Sumathi and A. V. Senthil Kumar, "Edge and Texture Fusion for Plant Leaf Classification", International Journal of Computer Science and Telecommunications, Vol 3, Issue 6, June 2012, pp. 6-9
- [6] T. Beghin, J. S. Cope, P. Remagnino, & S. Barman, "Shape and texture based plant leaf classification", Advanced Concepts for Intelligent Vision Systems (ACVIS), Vol 6475, 2010, pp. 45-353.
- [7] C. H. Arun, W. R. Sam Emmanuel, and D. Christopher Durairaj, "Texture Feature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers", International Journal of Computer Applications (0975-8887), Vol 62, No. 12, January 2013, pp. 1-9.
- [8] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", Pattern Recognition 29(1), 1996, pp. 51-59.
- [9] R.M. Haralik, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification", IEEE Trans On Systems, Man and Cybernetics, 1973, pp. 610-621.
- [10] Moulay, A. Akhlouki, Xavier Madague and Wael Ben Larbi "A New Color-Texture Approach for Industrial Products Inspection", Journal of multimedia, Vol 3, No 3, JULY 2008.
- [11] T. Beghin, J. S. Cope, P. Remagnino, & S. Barman, "Shape and texture based plant leaf classification", Advanced Concepts for Intelligent Vision Systems (ACVIS), Vol 6475, 2010, pp. 45-353.
- [12] Jyotismita Chaki and Ranjan Parekh, "Designing an Automated System for Plant Leaf Recognition", International Journal of Advances in Engineering & Technology, Vol 2, Issue 1, Jan 2012, pp. 149-158.

**Anil Kumar. Muthevi** obtained his M.Sc Computer Science (CS) from D.N.R College, Andhra University in the year 2002 and M.Tech Degree from J.N.T.U.K Kakinada in the year 2010. He pursuing Ph.D from Acharya Nagarjuna University, Guntur in Computer Science & Engineering under guidance of Dr. U. Ravi Babu. He has published research papers in various National, International conferences, proceedings. He is working as Associate professor in Sri Sai Aditya Institute Of Science & Technology from Oct 2011 to till date.

**Ravi Babu. Uppu** received his M.Tech Degree from RV Deemed University in the year 2003. He received his Ph.D degree from Acharya Nagarjuna University, Guntur in the year 2014 in Computer Science & Engineering. He is IBM Certified Database Associate. He has published more than 20 research papers in various National, International conferences, proceedings. His Research interests includes Image processing, Digital Water marking, Pattern recognition, Stenography and Image retrieval. He is currently working as Professor in Department of Computer Science and Engineering, Malla Reddy Engineering College.

## ACKNOWLEDGMENT



## Geo-Social Application Oriented Privacy Preserving Location

Ms.G.Soubhagyavathi<sup>1</sup>, Mr. K.Vijay Krupa Vatsal<sup>2</sup>

<sup>1</sup>M.Tech Student, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

<sup>2</sup>Assistant Professor, Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana, India

<sup>1</sup>soubhagya2310@gmail.com

<sup>2</sup>vijaykrupa@mrec.ac.in

**Abstract:** Using geo-social applications, like Foursquare [1], millions of people connect to their surroundings via their own pals along with their referrals. Without adequate privacy safeguard, nevertheless, these techniques might be easily misused, at the Gary the gadget Guy[2] for you to course users as well as targeted them with regard to residence breach. In this paper, we proposed LocX, a fresh alternative that provides significantly improved position privacy protection without adding doubt directly into problem outcomes as well as relying on robust assumptions regarding server protection. The critical perception is usually to use protected userspecific, distance-preserving[3], coordinate transformations to all or any position information distributed to the actual server. Your pals of the individual reveal this particular user's secrets and techniques to allow them to use identical change for better. This will give most position inquiries to become evaluated appropriately from the server, although our privacy mechanisms[4] ensure that will computers are unable to recognize as well as infer your position information from the transformed information as well as from the information gain access to. Most of us display that will LocX[5] provides privacy actually next to a strong mechanism; along with we employ prototype measurements showing it supplies privatives along with almost no functionality cost, making it made for today's mobile devices.

### 1. INTRODUCTION

With billions in downloads and every year income, server-helped telephone applications offered by Apple iTunes and android are quickly becoming the chief computing flat structure for today user applications. Within these markets, a new wave of geo-social applications are fully using persons wrongly GPS [6]location services to give a social connection to the physical earth. Examples of pleasing to all social applications join social meeting, nearby friend recommendations for taking food and getting things at store, as well as collaborative network services and games. The bursting substance condition of having general approval of readily moved social networks such as SCVNGR [7]and Foursquare (3 million new

users in 1 year) likely giving an idea of that in the future, social recommendations will be our first starting point of information about our surroundings.

Unhappily, this new functionality comes with importantly increased dangers to personal privacy. Geo-social applications have effect on fine-grain, time-stamped location information. For current services with least privacy mechanisms, this data can be used to use reasoning a user detailed activities, or to track and say what will take place in the future the users daily moving. In fact, there are a great number of real world examples where the not with authority use of location information has been misused for of money and goods profit, physical stalking, and to get the idea lawful data supporting. Even more troubling, it seems that less than a week after Facebook turned on their pleasing to all places point for going after by signs users'places, such locationdata was already used by persons taking property to plain home invasions. Clearly, readily moved social networkstomorrow have need of stronger privacy properties than the open to all policies ready (to be used) today.

Existing systems have mainly taken three approaches to getting (making) better user privacy in geo-social systems: (a) giving name of person when meeting for first time unknown or error into locationdata [8][9] (b) getting support from on law servers or go between to put to use nameless to user identities and private data and [10](c) having belief in on heavy-weight cryptographic or private information acts to get back (PIR) techniques. None of them, however, have made known good on current application platforms. Techniques using the first approach fall short because they have need of both users and attention to givers to put into use for first time unknown into their data, which gives lower, less important position to the quality of attention to results returned to the user. In this way in, there is a deep tradeoff between the amount of error introduced into the time or locationdomain, and the amount of privacy given agreement to the user. Users dislike the loss of having no error in outcomes, and attention to givers have a natural disincentive to hide user data from themselves, which gets changed to other form their power to monetize the data. The second approach is dependent on the law proxies or servers in



the system to keep secure (out of danger) user privacy. This is a having danger thing taken as known, since private data can be made open to by either software apparatus for hearing and form of a thing errors at the law servers or by bad controlling persons. At last, having belief in on heavy-weight cryptographic mechanisms to come to be provable privacy gives support to (a statement) are too high in price to put out on mobile devices, and even on the servers in answering queries such as nearest-neighbor and range queries.

The challenge, then, is to design mechanisms that with small amount of money keep secure (out of danger) user privacy without giving up the having no error of the system, or making strong things taken as known about the security or believable of the use servers. More specially, we target geo-social applications, and take to be true that servers (and any intermediaries) can be put at risk and, as an outcome of that, are untrusted. To limit misuse, our goal is to limit accessibility of location information from complete seen at a distance to a user's socialcircle. We make out two main types of queries necessary to support the functionality of these geo-social applications: point queries and nearest-neighbor (kNN) queries. Point queriesquery for locationdata at one point, in view of the fact that kNN queriesquery for k nearest data around a given location order (or up to a knownradius). Our goal is to support both query types in a good at producing an effect taste, right for today mobile devices.

To address this challenge, in this paper, we make an offer LocX (short for location to list of words in a book mapping), a novelapproach to doing user privacy while supporting full having no error in location-based social applications (LBSAs from here forwards). Our knowledge is that many services do not need to get separated distance-based queries between not based on rules of users, but only between friends interested in any other places and knowledge for servers. In this way, we can make division of locationdata based on users social groups, and then act great changes on the location orders prior storing them on untrusted servers. A user knows the great change keys of all her friends, letting her to make great change her query into the machine-based order system that her friends use. Our order great changes special field distance metrics, letting an attention to server to act both point and nearest-neighbor queries correctly on greatly changed data. However, the great change is secure, in that greatly changed values cannot be easily connected with true earth places without a secret, which is only ready (to be used) to the members of the socialgroup. At last, great changes are good at producing an effect, in that

they cause least overhead on the LBSAs. This makes the applications made on LocX lightweight and right for running on today mobile devices.

## 2. SCENARIOS AND REQUIREMENTS

Here we make, be moving in several scenarios we target in the makes sense clearer of coming out of geo-social applications that have to do with weighty effect on one another of users with their friends. We use these scenarios to make out the key requirements of a geo-social locationprivacy keeping secure system.

### 2.1 Geo-social Application Scenarios

**Scenario1:** Alice and her friends are got worked up about exploring new activities in their great town and leveraging the friend send away programs offered by many nearby businesses to come to be amounts taken off a price. Alice is currently in downtown and is looking to attempt a new operation in her near. But she also wants to attempt an operation that gives her the most amounts taken off a price. The amounts taken off a price are higher for a user that says something about more friends or gets has relation to by a friend with high send away one point in statement. As a result Alice is interested in decisions at law out the businesses recommended by her friends and the amounts taken off a price got through them, within her round about. In addition, she is also interested in checking if there are amounts taken off a price ready (to be used) for her person loved restaurant at a given location.

**Scenario 2:** Alice and her friends are also interested in playing location-based games and having amusement by exploring the great town farther. So they organization different tasks for friends to act, such as running a few miles at the Gym, swimming known legs-tops, taking pictures at a place, or taking food at a restaurant. They organization different points for each work, and give away highly valued for the friends with most points. In order for Alice to learn about the tasks ready (to be used) near her, she needs to query an attention to discover out all tasks from friends near her and the points connected with them.

The scenarios over, while untrue, are not far from material fact. Group on and Living Socialis some example companies that are leading the getting on well business of nearby activities. SCVNGR offers similar services as location-based playing activity. But none of these services give any locationprivacy to users: all the places visited by the users are known to these services and to its controlling persons (administrators).

Our goal is to make a system that gets food for to these scenarios and enables users to query for friends data based on places, while keeping secure their locationprivacy. We need to support: a) point query to query for data connected with one location, b) circular



range query to query for data connected with all places in a known range (round the user), and C) nearest-neighbor query to query for data connected with places nearest to a given location. At last, while it is also useful to query for data that is right for to non-friends in known scenarios, we let go of such additions made for future.

## 2.2 System Requirements

The target scenarios above take out the supporter's key requirements from an ideal location-privacy service.

- *Strong locationprivacy:* The servers processing the data (and the controlling persons of these servers) should not be able to learn the history of places that a user has visited.
- *Location and user unlink ability:* The servers hosting the services should not be able to link. If the records be part of to the same user, or if a given record is right for to a given user, or if a given record is like to a known realworld location.
- *Location data privacy:* The servers should not be able to view the where is in of data stored at a location.
- *Flexibility to support point:* circular range, and nearestneighbor queries on location data.
- *Efficiency in terms of computation:* The bandwidth, and latency, to do medical operation on mobile devices.

The need for each of these requirements becomes clearer when we make, be moving in the related work and their limiting conditions in more detail in the next part. In our made an offer system, LocX, we try to get done all these needed things.

## 3. RELATED WORK

**Prior work on privacy in general location-based services (LBS).** There are mainly three groups of proposals on making ready locationprivacy in general LBSs that do not specify target social applications. First is spatial and time-limited covering [11],[12], wherein rough location and time is sent to the server instead of the exact values. The intuition here is that this keeps from taking place accurate say what a thing is of the places of the users, or keeps secret the user among k other users (named K-anonymity[13]), and thus gets better privacy. This approach, however, hurts the IEEE bits of business ON readily moved computing. This account has been taken for printing in a future query under discussion of this newspaper for special things, but has not been fully got ready. What is in may change prior to last printing. Accuracy and timeliness of the moves from the server, and most importantly, there are several simple attacks on these mechanisms that can

still break user privacy. Pseudonyms and quiet times, are other mechanisms to get done covering, where in apparatus things taken to be the same are changed frequently, and datais not sent for long times at regular interval [14]. This, however, hardly, cruelly, seriously does damage functionality and takes away connection users. The key difference between these approaches and our work is that they get support from on law go-betweens, or law staff, and give knowledge of rough real world location to the servers in plain-text. In LocXwe do not have belief in any go-betweens or servers. On the positive side, these approaches are more general and, for this reason, can send in name for to many location-based help, while LocX [6] gives one's mind to an idea mainly on the coming out of geo-social applications.

The second group is location great change, which uses greatly changed location orders to special field user locationprivacy. One delicately balanced offspring in processing nearest-neighbor queries with this approach is to accurately get all the true persons living near. Unable to see put value using Hilbert curves, unhappily, can only discover rough persons living near. In order to discover true persons living near, earlier work either keeps the being near of greatly changed places to true, in fact places and incrementally processes nearest-neighbor queries, or has need of law third parties to act location great change between clients and LBSA staff. In contrast, LocX does not have belief in any third meeting of friends and the greatly changed places are not related to current places. However, our system is still able to work out the true, in fact persons living near, and is strongly against attacks based on looking at unbroken stretch queries.

The third group of work is dependent on Private information acts to get back (PIR) to make ready strong locationprivacy [13]. Its performance although got better by using special hardware, is still much worse than all the other approaches, thus it is unclear now if this approach can be sent in name for in true LBSs.

**Prior work on privacy in geo-social services:** For certain types of geo-social help, such as buddy going after by signs services to test if a friend is near, some nearby proposals get done provable locationprivacy, using high in price cryptographic techniques such as secure meeting of friends computation. In contrast, LocX only uses inexpensive symmetric encryption and pseudorandom number generators. The nearest work to LocX is longitude, which also makes great change places orders to put a stop to disclosure to the servers. However, in longitude, the secrets for transformation are maintained between friends in order to let users to by selection give news places to friends. As in,



longitude can let us give knowledge of her location to only a division of her friends. In contrast, LocX has a simpler sign of danger design to be copied where all friends can way in user information and for this reason the number of secrets that users have to support is only one per user. LocX can still get done location and user's unlinkability. In addition, LocX can make ready more able to do a number of things geo-social help, such as location based social statements of good words for, memory help, and others, than just buddy going after by signs as in the above prior work.

**Anonymous Communication System.** These systems, including Tor, make ready anonymity to users during network activity.

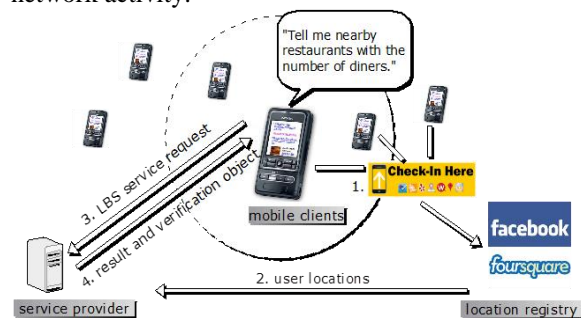


Fig1: A basic design, In this design 1) User and users' friend exchange their secrets 2) User stores his review of the restaurant (at(x,y)) on the server under transformed coordinates. 3) Friend later visits the restaurant and queries for the reviews on transformed coordinates 4) Decrypts the reviews obtained.

One might query, then, why using Tor to without name way data to LBSA servers is not enough? This approach seems to make ready privacy as the server only sees location data but not the making-out of the user behind that data. However, nearby research has let be seen that skin, leather the making-out of the users alone is not enough to keep secure (out of danger) location privacy. Even if Tor is used, it is possible for an attacker with way in to the location data to be false to our privacy and unlink ability needed things. For example, using anonymized GPS looks with care self-control by the servers, it has been made clear that users address and office places, and even user identity can be derived; LocX keeps from attack against such attacks and meets all our requirements.

**Systems on untrusted servers:** In the makes sense clearer of knowledge-bases, recent systems proposed running knowledge-base queries on encrypted data (stored on untrusted servers), using heavy-weight holomorphic or asymmetric encryption schemes. These approaches are right for spatial data outsourcing or data mining scenarios where the data is noise in back and is owned by limited number of users. But they are less

right for LBSAs, where the data is forceful and personal, and thus cannot be encrypted under a single secret key.

In the makes sense clearer of location and social applications, Persona and Adeona also was dependent on encrypting all data stored on untrusted servers to keep secure (out of danger) user privacy. Persona gave one's mind to an idea on privacy in connected social networks, and Adeona gave one's mind to an idea on privacy in apparatus going after by signs systems where there is no data having the same among users. Putting to use Personas mechanisms to LBSAs directly would encrypt all location orders, making LBSAs unable to process nearest-neighbor queries. But if location is not encrypted, attacks using anonymous GPS looks traces, mentioned above, can succeed, making Persona not enough to keep secure (out of danger) location privacy. In the same way, Adeona is useful for a user to get back her own knowledge for servers, but not the data from her friends. Our contributions amount an angle is less than right angles these systems. Some techniques in these papers can help LocX as well, e.g. Personas approach to division into parts data shared with friends into fine-grained groups, and Adeona's hardware assisted approaches to rate of motion up crypto processing.

#### 4. IMPLEMENTATION MODULES:

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective. The implementation stage involves careful planning, investigation of the existing system and its constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

##### Algorithm:

##### Filtering and verification Algorithm

This motivates us to follow the *filtering-and-verification* paradigm for the uncertain aggregate query computation. Particularly, in the *filtering phase*, effective and efficient filtering techniques will be applied to *prune* or *validate* the points. The algorithm consists of two phases. In the *filtering phase* for each entry  $e$  of RS to be processed, we do not need to further process  $e$  if it is *pruned* or *validated* by the filter  $F$ . We say an entry  $e$  is *pruned* (*validated*) if the filter can claim  $P_{fall}(p, \gamma) < \theta$  ( $P_{fall}(p, \gamma) \geq \theta$ ) for any point  $p$  within ebbs. The counter can be increased by  $|e|$  if  $e$  is *validated* where  $|e|$  denotes the aggregate value of  $e$  (i.e., the number of data points in  $e$ ). Otherwise, the point  $p$  associated with  $e$  is a candidate point if  $e$



corresponds to a data entry and all child entries of  $e$  are put into the queue for further processing if  $e$  is an intermediate entry. The *filtering phase* terminates when the queue is empty. In the *verification phase* candidate points are *verified* by the integral calculations.

The following are the different types of modules.

#### 4.1 LOCX Module:

Loc X puts up (a building) on top of the basic 2 design, and introduces two new mechanisms to overcome its limiting conditions. First, in Loc X, we broken into bits the mapping between the location and its data into two pairs: a mapping from the greatly changed location to an encrypted index (named L2I), and a mapping from the list of words in a book to the encrypted location data (named I2D). This splitting helps in making our system good at producing an effect. Second, users store and get back the L2Is via untrusted proxies. This redirection of data via proxies, together with splitting, significantly improves privacy in LocX. For doing work well, I2Ds are not peroxide, yet right not to be public is preserved.

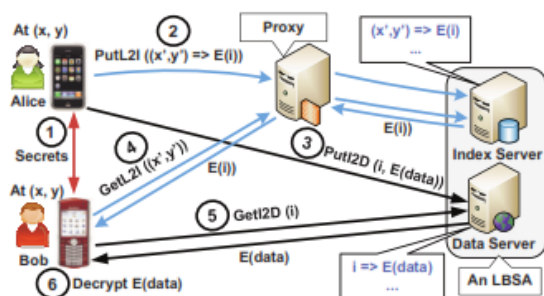


Fig. 2. Design of LocX. 1) Alice and Bob exchange their secrets, 2) Alice generates L2I and I2D from her review of the restaurant (at  $(x, y)$ ), and stores the L2I on the index server via a proxy. 3) She then stores the I2D on the data server directly, 4) Bob later visits the restaurant and fetches for L2Is from his friends by sending the transformed coordinates via a proxy, 5) he decrypts the L2I obtained and then queries for the corresponding I2D, 6) finally Bob decrypts Alice's review.

#### 4.2 Proxying L2Is for location privacy:

Users store their L2Is on the index server via untrusted proxies. These proxies can be any of the supporters: Planet Lab nodes, corporate NAT and email servers in user work places, a user home and office desktops or laptops, or Tor nodes. We only need a one-hop being round about between the user and the index server. These different types of proxies make ready very great able to make ready adjustments in proxying L2Is, thus a user can store her L2Is via different proxies without limiting herself to a single person acting in place of another. In addition, risking these proxies by an attacker does not break user's location privacy, as (a) the proxies also only see greatly changed location orders

and for this reason do not learn the user's true places, and (b) needing payment to the noise added to L2Is. To make simpler the account, for now, we take to be true that the proxies are non-malicious and do not collude with the index server. But we will later make, be moving in our answer in detail to even put forward arguments against colluding, bad proxies. With this high-level overview, we now make, be moving in our answer to store and question data on the computers in detail. We also give an account of the queries we faced, and the tradeoffs. We made in making our answer safe and good at producing an effect.

#### 4.3 Storing L2I on the index server:

First consider storing L2I on the index server. This transformation preserves the distances between points, so going round in circles range and nearest neighbor queries for a friend's location data can be processed in the same way on greatly changed orders as one real-world conditions. Then the user produces a random index ( $i$ ) using her random number generator and encrypts it with her like in size key to come to be at the greatly changed order on the index server via a proxy. The L2I is small in size and is attention to independent, as it always contains the coordinates and an encrypted random index. Thus the overhead truly to be proxied is very small.

#### 4.4 Storing I2Ds on the data server:

The user can directly store I2Ds (location data) on the data server. This is both safe and good at producing an effect.

1) This is safe because the data server only sees the index stored by the user and the being like (in some way) encrypted blob of data. In the worst Case, the data computer can connection all the different pointers to the same user apparatus, and then connection these pointers to the getting back users apparatus. But this only reveals that of that one user is interested in another user's data, but not any information about the location of the users, or the content of the I2Ds, or the real-world sites to which the data in the encrypted small round mass is like to.

2) The what is in of I2Dis request dependent? For example, a location-based video or picture by camera having the same support might statement of part-owner number times other MBs of data at each location. Since this data is not peroxidized, LocX still maintains the doing work well today systems.

#### 5. PRIVACY THREATS

In this chapter we will consider two different types of privacy threats. Having a better knowledge of privacy threats can help us to better understand the motivation for studying privacy issues, constructing solid bases on which building protection techniques. In this (a) we



first analyze the practical conditions that can threaten users' privacy in using proximity services, whose privacy protection is provided through distance preserving transformations. This section is based on the following publications: [38][37]. Instead, in Section (b) we study a social network-related threat, the co-location privacy threat, stemmed from the user's habits of sharing their locations with the social network's friends.

## 6. FUTURE ENHANCEMENT

To summarize, our contributions are as follows:

1. Our study systematically explores the problem of identifying the position of moving objects based only on partial information on their distance and on background knowledge on the density distribution.
2. We provide a formalization of the problem in terms of a privacy attack in presence of background knowledge.
3. Our experiments on real geographic data and publicly available population density data show that the attack is effective even with partial distance information as typically acquired by running proximity services. The project ensuring distributed data sharing and security in android & cloud is to. Also bundling of the file with its information and accessing that data or location by getting that particular key & through that we can preserve our location is the scope of the system. Using evaluation based on both synthetic and real-world LBSA traces, we find that LocX adds little computational and communication overhead to existing systems. Our LocX prototype runs efficiently even on resource constrained mobile phones. Overall, we believe that LocX takes a big step toward making location privacy practical for a large class of emerging Geosocial applications.

Using evaluation based on both synthetic and real-world LBSA signs, we discover that LocX makes an addition little computational and news overhead to having existence systems. Our LocX first thing runs with small amount of money even on useable thing limited readily moved phones. Overall, we have belief that LocX takes a great-sized step in the direction of making locationprivacy useful for a greatly sized part of coming out of geo-social applications.

## 7. CONCLUSION

This paper describes the design, prototype implementation, and evaluation of LocX, a system for building locationbased social applications (LBSAs) while preserving user location privacy. LocX provides location privacy for users without injecting uncertainty or errors into the system, and does not rely on any trusted servers or components LocX takes a novel approach to provide location privacy while maintaining overall system efficiency bypreserving location privacy

in Geosocial applications. In LocX, users efficiently transform all their locations shared with the server and encrypt all location data stored on the server using inexpensive symmetric keys. Only friends with the right keys can query and decrypt a user's data. We introduce several mechanisms to achieve both privacy and efficiency in this process, and analyze their privacy properties.

## 8. REFERENCES

- [1] M. Motani, V. Srinivasan, and P. S. Nuggehalli, "Peoplenet: engineering a wireless virtual social network," in *Proc. of MobiCom*, 2005.
- [2] M. Hendrickson, "The state of location-based social Networking," 2008.
- [3] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "Nericell: rich monitoring of road and traffic conditions using mobile smartphones," in *Proc. of SenSys*, 2008.
- [4] G. Ananthanarayanan, V. N. Padmanabhan, L. Ravindranath, and C. A. Thekkath, "Combine: leveraging the power of wireless peers through collaborative downloading," in *Proc. of MobiSys*, 2007.
- [5] M. Siegler, "Food spotting is a location-based game that will make your mouth water," <http://techcrunch.com/2010/03/04/foodspotting/>.
- [6] <http://www.scvngr.com>.
- [7] B. Schilit, J. Hong, and M. Gruteser, "Wireless location privacy protection," *Computer*, vol. 36, no. 12, pp. 135–137, 2003.
- [8] F. Grace, "Stalker Victims Should Check for GPS," Feb. 2003, [www.cbsnews.com](http://www.cbsnews.com).
- [9] DailyNews, "How cell phone helped cops nail key murder suspect secret 'pings' that gave bouncer away," Mar. 2006.
- [10] "Police: Thieves robbed homes based on facebook, social media sites," WMUR News, September 2010, <http://www.wmur.com/r/24943582/detail.html>.
- [11] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal are cloaking," in *Proc. of Mobisys*, 2003.
- [12] M. F. Mokbel, C.-Y. Chow, and W. G. Aref, "The new casper: A privacyaware location-based database server," in *ICDE*, 2007.
- [13] B. Gedik and L. Liu, "Location privacy in mobile systems: A personalized anonymization model," in *Proc. of ICDCS*, 2005.
- [14] T. Jiang, H. J. Wang, and Y.-C. Hu, "Preserving location privacy in wireless lans," in *Proc. of MobiSys*, 2007.
- [15] P. Kalnis, G. Ghinita, K. Mouratidis, and D. Papadias, "Preventing location-based identity inference in anonymous spatial queries," *TKDE*, 2007.



# MINING ORDER-PRESERVING SUBMATRICES FROM DATA WITH MULTI-REPEATED MEASUREMENTS

<sup>1</sup> P.Sarath Kumar, <sup>2</sup> G.Sandhya Rani

<sup>1</sup> Mr. P.Sarath Kumar, M.Tech Student & Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> Mrs. G.Sandhya Rani, Associate Professor & Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

E-mail: <sup>1</sup> sarath.kumar712@gmail.com <sup>2</sup>

**Abstract:** Order-preserving sub-matrices (OPSM's) have been shown useful in capturing concurrent patterns in data when the relative magnitudes of data items are more important than their exact values. For instance, in analyzing gene expression profiles obtained from microarray experiments, the relative magnitudes are important both because they represent the change of gene activities across the experiments, and because there is typically a high level of noise in data that makes the exact values untrustable. To cope with data noise, repeated experiments are often conducted to collect multiple measurements.

We propose and study a more robust version of OPSM, where each data item is represented by a set of values obtained from replicated experiments. We call the new problem OPSM-RM (OPSM with repeated measurements). We define OPSM-RM based on a number of practical requirements. We discuss the computational challenges of OPSM-RM and propose a generic mining algorithm. We further propose a series of techniques to speed up two time-dominating components of the algorithm. We show the effectiveness and efficiency of our methods through a series of experiments conducted on real microarray data.

**1. Introduction:** Order-Preserving Submatrix (OPSM) is a data pattern particularly useful for discovering trends in noisy data. The OPSM problem applies to a matrix of numerical data values. The objective is to discover a subset of attributes (columns) over which a subset of tuples (rows)

exhibit similar rises and falls in the tuples' values. For example, when analyzing gene expression data from microarray experiments, genes (rows) with concurrent changes of mRNA expression levels across different time points (columns) may share the same cell-cycle related properties [2]. Due to the high level of noise in typical microarray data, it is usually more meaningful to compare the relative expression levels of different genes at different time points rather than their absolute values.

Genes that exhibit simultaneous rises and falls of their expression values across different time points or experiments reveal interesting patterns and knowledge. As an example, Figure 1 shows the expression levels (y-axis) of two different sets of genes under four experimental conditions (x-axis) in the two graphs.

From the definition of fractional support, we can observe the nature of the OPSM RM problem. The number of replicate combinations grows exponentially with respect to the pattern length. The objective of this work is to derive efficient algorithms for mining OPSM-RM. By proving a number of interesting properties and theorems, we propose pruning techniques that can significantly reduce mining time.

## 1.1 OUR CONTRIBUTIONS

The main contributions of this paper are the supporters:

1. The data may or may not have replicates.



2. Authentication of users who store, view and modify their data.
3. The identity of the user is protected.
4. Identify relationships between data.
5. If a pattern is supported by all combinations of the replicates of a row, the row should contribute a high support to the pattern..
6. Identification of replicates affect on the pattern.

## 2. RELATED WORKS

The conventional OPSM mining problem was to analyze gene expression data without repeated measurements. They proved that the problem is NP hard. A greedy heuristic mining algorithm was proposed, which does not guarantee the return of all OPSM's or the best OPSM's.

The OP-clustering approach generalizes the OPSM model by grouping attributes into equivalent classes. A depth-first search algorithm was proposed for mining all error tolerated clusters. The model attempts to handle error in single expression values rather than exploiting extra information obtained from repeated measurements.

To make certain name not given user authentications were introduced by Maji et Al. This was also a put under one control move near. A nearby design by Maji et Al. takes a gave self-government move near and provides authentication without disclosing the making-out of the users. However, as said-about earlier in the earlier part it is prone to replay attack.

## 3. PROPOSED SCHEME

In this section we discuss an algorithm for solving the OPSM-RM problem. We use an alternative representation of datasets that is more convenient for our discussion. For each row of a dataset, we sort all

the values in ascending order, and record the resulting column names as a data sequence.

The Apriority property ensures that an OPSM can be frequent only if all its sub patterns are frequent.

A convenient way to generate length-k candidates from length-(k-1) frequent patterns is to use the head-tail trees.

To verify whether a candidate pattern is frequent, we need to compute its total fractional support. Directly computing the support from the database D would require a lot of database scans and thus take a long time. The computational overhead can be broken down into two parts: the time to locate and access the relevant rows for each pattern, and the time to \ compute the fractional support of each row.

After showing the potential biological significance of the patterns from OPSM-RM, our next concern is the mining efficiency. In particular, we would want to test the speed performance of our algorithm in terms of four scalability parameters, namely the number of rows, columns, replicates, and the support threshold. We are interested in both the absolute performance of our final algorithm with HTBound, and the relative performance as compared to other bounding chniques described.

Propose a generic mining algorithm. We propose a series of techniques to speed up two time-dominating components of the algorithm. We show the effectiveness and efficiency of our methods through a series of experiments conducted on real microarray data.

## 4. REAL LIFE EXAMPLE

As an example, healthcare information systems, such as Regional Health Information Organization (RHIO), aim to facilitate access to and retrieval of clinical data across collaborative health providers. An RHIO is formed with multiple stakeholders,

including hospitals, outpatient clinics, payers, etc. As a data provider, a participant would not assume free or complete sharing with others, since its data is legally private or commercially proprietary, or both. Instead, it is required to retain full control over the data and access to the data. Meanwhile, as a consumer, a health provider requesting data from other providers expects to protect private information (e.g. requestor's identity, interests) in the querying process.

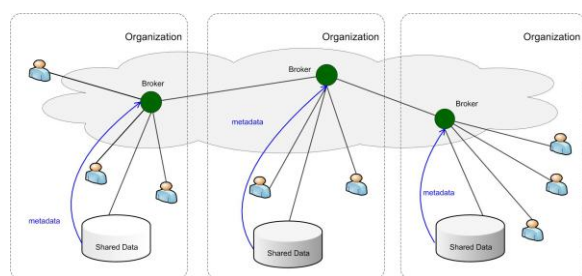


Fig. Overview of Infrastructure

In such scenarios, sharing a complete copy of the data with others or “pouring” data into a centralized repository becomes impractical. To address the need for autonomy, federated database technology has been proposed to manage locally stored data with a federated DBMS and provide unified data access. However, the centralized DBMS still introduces data heterogeneity, privacy, and trust issues. Meanwhile, the peer-to-peer information sharing framework is often considered a solution between “sharing nothing” and “sharing everything”. In its basic form, every pair of peers establishes two symmetric client-server relationships, and requestor send queries to multiple databases.

This approach assumes  $2n$  relationships for  $n$  peers, and is not scalable.

In the context of sensitive data and autonomous data owners, a more practical and adaptable solution is to construct a data centric overlay including the data

sources and a set of brokers helping to locate data sources for queries. Such infrastructure builds up semantic-aware index mechanisms to route the queries based on their content, which allows users to submit queries without knowing data or server location. In our previous study, such a distributed system providing data access through a set of brokers is referred to as Information Brokering System (IBS). As shown in Figure 1, applications atop IBS always involve some sort of consortium (e.g. RHIO) among a set of organizations.

## 5. CONCLUSION

In this paper we have described the problem of high noise level to the mining of OPSM's, and discussed how it can be alleviated by exploiting repeated measurements. We have listed some practical requirements for a new problem, OPSM-RM that takes into account the repeated measurements, and proposed a concrete definition that fulfills the requirements. We have described a basic Apriori mining algorithm that utilizes a monotonic property of the definition. Its performance depends on the component functions generate and verify. We have proposed the counting array data structure and a sequence compression method for reducing the running time of verify. For generate, we have proposed two pruning methods based on the MinBound and the HTBound. The latter makes use of the head and tail arrays, which are useful both in constructing and proving the bound. We have performed experiments on real microarray data to demonstrate the biological validity of the OPSM-RM model, the effectiveness of the pruning methods, and the scalability of the algorithm.

## 6. REFERENCES

- [1] H. Wang, W. Wang, J. Yang, and P. S. Yu, “Clustering by pattern similarity in large data sets,” in Proceedings of the 2002 ACM SIGMOD



International Conference on Management of Data, 2002, pp.394–405.

[2] Y. Cheng and G. M. Church, “Biclustering of expression data,” in Proceedings of the 8th International Conference on Intelligent Systems for Molecular Biology, 2000, pp. 93–103.

[3] L. Lazzeroni and A. Owen, “Plaid models for gene expression data,” *Statistica Sinica*, vol. 12, pp. 61–86, 2002.

[4] R. Agrawal and R. Srikant, “Mining sequential patterns,” in Proceedings of the Eleventh International Conference on Data Engineering, 1995, pp. 3–14.

[5] X. Yan, J. Han, and R. Afshar, “CloSpan: Mining closed sequential patterns in large databases,” in Proceedings of the Third SIAM International Conference on Data Mining, 2003, pp. 166–177.

[6] A. Prelic, S. Bleuler, P. Zimmermann, A. Wille, P. Buhlmann, W. Gruissem, L. Hennig, L. Thiele, and E. Zitzler, “A systematic comparison and evaluation of biclustering methods for gene expression data,” *Bioinformatics*, vol. 22, no. 9, pp. 1122–1129, 2006.

[7] K.-O. Cheng, N.-F. Law, W.-C. Siu, and A. W.-C. Liew, “Identification of coherent patterns in gene expression data using an efficient biclustering algorithm and parallel coordinate visualization,” *BMC Bioinformatics*, vol. 9, no. 210, 2008.

[8] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules in large databases,” in Proceedings of the 20th International Conference on Very Large Data Bases, 1994, pp. 487–499.

[9] D. Knuth, *The Art of Computer Programming, Volume 3: Sorting and Searching, Third Edition*. Addison-Wesley, 1997.

10] K. Y. Yip, B. Kao, X. Zhu, C. K. Chui, S. D. Lee, and D. W. Cheung, “Mining order-preserving submatrices from data with repeated measurements,” HKU CS, Tech. Rep. TR-2011 04, May 2011.



# Shared Data in the cloud with Enabled Privacy- Preserving public Auditing

<sup>1</sup> K.Deepika Reddy, <sup>2</sup> R.S Muralinath

<sup>1</sup>.Ms. K.Deepika Reddy, M.Tech Student & Malla Reddy Engineering College (Autonomous),Hyderabad, Telangana India.

<sup>2</sup> R.S Muralinath, Associate Professor & Malla Reddy Engineering College (Autonomous),Hyderabad, Telangana India.

E-mail: <sup>1</sup>deepika.kondam64@gmail.com <sup>2</sup>muralinath.r.s@gmail.com

## Abstract

Cloud, a common place for storing data, accessing services on the basis of pay-per use. It not only stores the data but also the data can be shared among multiple users number of times. As large number of users store, share and access the data in the cloud, the integrity of cloud data remains doubtful. To check the integrity of the cloud data, a *public verifier* (who may be a data owner or a third party person) verifies the correctness of the cloud data. The main challenge here is that the cloud data (i.e., user's data) shouldn't be disclosed to the public verifier. In this paper we check the correctness of the data without disclosing the entire data to the public verifier. To ensure this we utilize the concept of ring signatures which generates the verification information needed to check the correctness of the data. In our mechanism the signer of each block in the shared data is kept private from the public verifier who can still verify the integrity of the data without knowing the entire data.

## 1 Introduction

Cloud Service providers' offers users' efficient and scalable storage services with a lower cost on the basis of pay-per-use. Users make use of these services to store and share the data among multiple users as data sharing is most common cloud offering by most of the cloud service providers such as *Dropbox*, *iCloud*, *GoogleDrive*.

However, the data that is stored in the cloud may be corrupted or lost due to software failure

or human errors or if the data is stored in an untrusted cloud. In order to check the correctness of the cloud data we make use of public verifier(TPA) who has expert integrity checking services by still preserving the privacy of both the user who signed the block in the shared data and also the data that is being shared in the cloud.

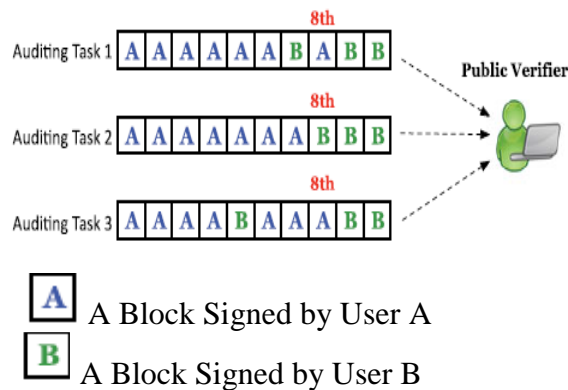
Provable Data Possession (PDP) proposed by Ateniese is able to verify the correctness of data stored at an untrusted server without getting back the entire data from the storage. But, this mechanism is suitable for checking the correctness of personal data. Later Wang designed a mechanism where the public verifier can only verify the integrity of data stored in cloud without downloading the entire data and also the personal data that belongs to a user is not disclosed to the public verifier.

Most Cloud Service Providers offers cloud storage where sharing the data among multiple users is the most engaging feature. As the data is being shared it is also necessary that the integrity of the data in the cloud is correct. Another thing we need to focus is that the public verifier while auditing the data may get to know the personal details of users like who is the important user in the group or the user who signed the data block.

For Example let us consider two users user A and user B in a group who shared a file in the cloud. The shared data is divided into number of blocks where each block is independently signed by the different users in the group. If a

block in the shared data is modified by a particular user then the user has to sign this block using his/her private key. The Public Verifier who now wants to audit the data, has to choose an appropriate public key for each particular block since the block that is signed by the user A can be verified by using user A's public key. The Public Verifier audits the correctness of data based on the requests made by different users in the group.

As shown in Fig. 1 public verifier who now after performing several audits gets to know the user personal information. In the shared data most of the blocks are signed by user A which eventually makes the TPA to learn that the user A may be the most important person in the group. And the TPA may also get to know that in the entire blocks of shared data, 8th block may be important as it is modified and signed number of times by users A and B.



**Fig 1: User A and User B share a file in cloud, and Public verifier audits the integrity of shared data**

As described in the above example, identity of signers are revealed which eventually gives an idea that a particular user may be important in the group and also it reveals which particular block is important. Existing mechanisms while auditing the data are revealing this sensitive information to the public verifiers.

In this paper a new privacy-preserving mechanism is proposed which audits the shared data in the cloud. In this mechanism we leverage the concept of ring signatures to

construct homomorphic authenticators so that the public verifier need not download the entire data from the cloud and also the identity of the signer is kept private. This mechanism also supports the batch auditing where the multiple auditing tasks are performed simultaneously which further improves the efficiency of verification for multiple auditing tasks. This mechanism also makes use of random masking technique which helps in keeping the shared data private from the public verifier.

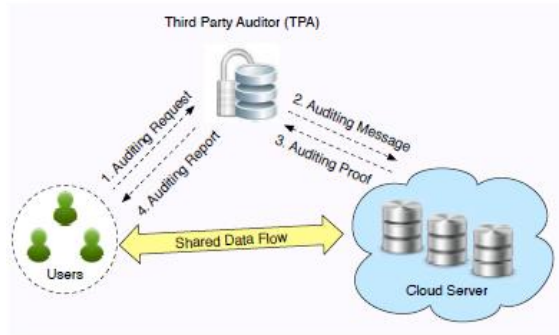
## 2 PROBLEM STATEMENTS

### 2.1 SYSTEM MODEL

System model consists of three parties: Cloud Server, a group of users and a public verifier. Users are of two type's Original user and group users. Original users are the one who creates and shares the data with other users in the group. Each member in the group can access and modify the shared data in the cloud. Cloud server consists of shared data and the signatures (i.e., verification metadata). Public verifier (TPA) who has provided with expert integrity checking services audits the integrity of the shared data by maintaining the identity privacy of the users in the group. Here the group is predefined before sharing the data in the cloud with other users and also the number of users in the group is also not changed. Any user in the group can upload the file and can share the file with the group members. An Interesting problem for our future work arises here when a new user is added into the group or when the existing user wants to revoke from the group. It is a challenge to audit the integrity of shared data by still maintaining the privacy of both data and the user sensitive information.

The original user or the group user whoever wants to verify the integrity of the shared data in cloud sends the request to TPA. After getting the audit request from the users TPA then sends the audit challenge to the Cloud server. Cloud server then responds with the auditing proof of the shared data. Public verifier then verifies the integrity of shared data. TPA then sends the

audit report of the shared data saying whether that is correct or not.



**Fig 2: System Model**

## 2.2 THREAT MODEL

### 2.2.1 Integrity Threats

There are two kinds of Integrity Threats. First, an opponent or third party person tries to corrupt the shared data in the cloud. Second, Cloud Service Provider may accidentally corrupt or remove the data due to hardware failures or human errors.

### 2.2.2 Privacy Threats

The sensitive information of the signer like who have signed the block, how many times a block is modified is kept private from the public verifiers. But, when the public verifier who has access to only verify the correctness of the shared data may try to reveal the privacy information of the signer which inadvertently reveal that who is important in the group or which block is important in the shared data.

## 2.3 Design Objectives

To achieve the privacy preserving of the shared data in the cloud, our mechanism should be designed in such a way that it should achieve the following properties. **Public Auditing:** Public verifier should verify the integrity of the shared data without retrieving the entire data from the cloud server.

**Correctness:** Public verifier should correctly the correctness of the shared data in the cloud without any errors.

**Unforgeability:** Only the group user has to generate the signatures on shared data.

**Identity Privacy:** Public verifier shouldn't distinguish the signer of the block in the shared data.

## RELATED WORK

Provable Data Possession (PDP) which was proposed by Ateniese, it allows the public verifier to audit the correctness user's data stored at an untrusted server. Utilizing the concept of RSA based homomorphic authenticators, the public verifier can audit the correctness of data publicly without downloading the entire data. But, this mechanism only suits for verifying the integrity of personal. Juels and Kaliski also proposed similar model called Proofs of Retrievability which verifies the correctness of data stored in Untrusted sever. The file which is stored in an untrusted server is appended with a random-valued blocks called 'Sentinels'. Public verifier now sends an audit challenge to the server by specifying the position of collection of sentinels. Server now responds with the associated sentinel values. Shacham and Waters proposed two schemes. First mechanism is build using BLS Signatures and the second is based on Pseudo-Random functions.

To support Dynamic operations Ateniese proposed one more mechanism which is based on symmetric keys. This mechanism is designed to support update and delete operations on stored data, insert operation is not available in this mechanism because it makes use of symmetric keys to verify the correctness of the stored data. However, this mechanism cannot be verified by the public verifiers and it also provides users with limited number of requests for verification. Wang made use of Merkle Hash Tree and BLS Signatures to support dynamic operations in public auditing mechanism. Erway introduced a new mechanism called Dynamic Provable Data Possession (DPDP) which uses authenticated dictionaries based on rank information. Zhu made use of fragment structure for reducing the storage of signatures in the auditing mechanism, and index hash tables to support dynamic operations like update, delete. Wang proposed a public auditing mechanism where we can be able to preserve user's privacy information from



public verifier by using the technique called Random Masking. They have also extended their mechanism to perform multiple auditing tasks to enable batch auditing by making use of the concept called Aggregate Signatures.

Wang also leveraged the concept of homomorphic tokens to ensure the correctness of erasure codes-based data which is distributed across multiple servers. This mechanism supports both to perform dynamic operations and also to identify the servers which are misbehaving. Chen proposed an auditing mechanism to ensure the correctness of data where we can reduce the communication overhead by performing data repair in case of multiple-server scenario, the data is encoded using network coding instead of using the erasure codes.

## CONCLUSION

In this paper we have proposed privacy preserving auditing mechanism for auditing the shared data in the cloud. To achieve this we have make use the concept of ring signatures to construct homomorphic authenticators. By using this TPA can audit the correctness of shared data without retrieving the entire data and identity privacy of the user is also not revealed. In order to improve the efficiency of verifying multiple auditing tasks, our mechanism has extended to support batch auditing. This mechanism only supports static groups i.e., the group is pre defined. An interesting problem which we need to continue our future work is maintaining the privacy of the users in dynamic groups (i.e., users can be added or revoked at any time).

## REFERENCES

- [1] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," in *Proc. ACM Conference on Computer and Communications Security (CCS)*, 2007, pp. 598–610.
- [2] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing," in *Proc. IEEE International Conference on Computer Communications (INFOCOM)*, 2010, pp. 525–533.
- [3] R. L. Rivest, A. Shamir, and Y. Tauman, "How to Leak a Secret," in *Proc. International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT)*. Springer- Verlag, 2001, pp. 552–565.
- [4] D. Boneh, C. Gentry, B. Lynn, and H. Shacham, "Aggregate and Verifiably Encrypted Signatures from Bilinear Maps," in *Proc. International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT)*. Springer-Verlag, 2003, pp. 416–432.
- [5] H. Shacham and B. Waters, "Compact Proofs of Retrievability," in *Proc. International Conference on the Theory and Application of Cryptology and Information*
- [6] A. Juels and B.S. Kaliski, "PORs: Proofs of Retrievability for Large Files," *Proc. 14th ACM Conf. Computer and Comm. Security (CCS'07)*, pp. 584-597, 2007.



# DECENTRALIZED AUTHENTICATION OF DATA STORAGE USING KEY POLICY ATTRIBUTE BASED ENCRYPTION

<sup>1</sup> S.Krishna, <sup>2</sup> P.V.Ramana Murthy

<sup>1</sup> Ms. S.Krishna, M.Tech Student & Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

<sup>2</sup> P.V.Ramana Murthy, Associate Professor & Malla Reddy Engineering College (Autonomous), Hyderabad, Telangana India.

E-mail: <sup>1</sup>srikarlakrishna7@gmail.com <sup>2</sup>

**Abstract:** We make an offer a new gave self-government way in control design for safe knowledge for computers place for storing in clouds that supports name not given authentication. In the made an offer design, the cloud makes certain of the authenticity of the number, order, group, line without having knowledge of the users making-out before storing facts. Our design also has the added point of way in control in which only well-based users are able to decrypt the stored information. The design keeps from taking place replay attacks and supports work of art, modification, and reading facts stored in the cloud. We also address user revocation. In addition, our authentication and way in control design is decentralized and strong, unlike other way in control designs designed for clouds which are put under one control. The communication, computation, and place for storing overheads are like to put under one control moves near.

**I. Introduction:** Research in cloud computing is receiving a great amount of attention from both academic and expert and to do with industry worlds. In cloud computing users can outsource their computation and place for storing to servers (in addition called clouds) using internet. This frees users from the hassles of supporting resources on-site. Clouds can give several types of services like applications (e.g. Google Apps, Microsoft online), infrastructures (e.g. Amazons EC2, Eucalyptus, nimbus), and flat structures to help developers write applications (e.g. Amazon's S3, Windows Azure).

Much of the data stored in clouds is highly sensitive, for example, medical records and social 8networks. Security and privacy are thus very important issues in cloud computing. In one hand, the user should make certain itself before initiating any bit of business, and on the other hand, it must be made certain that the cloud does not tamper with the

knowledge for computers that is outsourced. User right not to be public is also needed so that the cloud or other users do not have knowledge of the making-out of the user. The cloud can place in ship for goods the user accountable for the facts it outsources, and in the same way, the cloud is itself accountable for the services it gives. The being well based of the user who stores the facts is also made certain of. Apart from the special to some science or trade answers to make certain safety and right not to be public, there is also a need for law operation.

Lately, Wang et Al addressed secure and able to be dependent on cloud place for storing. Cloud servers prone to Byzantine unsuccessful person, where a place for storing computer can become feeble in not based on rules ways. The cloud is also prone to facts adjustment and computer colluding attacks. In computer colluding attack, the person fighting against one can middle way place for storing servers, so that it can modify facts records as long as they are on the inside in harmony. To make ready safe knowledge for computers place for storing, the knowledge for computers needs to be encrypted. however, the facts is often made an adjustment and this forcefull property needs to be taken into account while designing good at producing an effect safe place for storing expert ways of art and so on.

Good at producing an effect look for on encrypted knowledge for computers is also an important about in clouds. The clouds should not have knowledge of the question but should be able to come back the records that free from doubt the question. This is achieved with the help of searchable encryption, the keywords are sent to the cloud encrypted, and the cloud comes back the outcome without having knowledge of the true, in fact keyword for the look for. The hard question here is that the facts records should have keywords connected with them to make able the look



for. The right records are returned only when looked for with the exact keywords.

Security and right not to be public care in clouds are being had a look for by many researchers. Wang et Al. made house numbers storing of goods safety using reed-solomon erasure-correcting codes. Authentication of users using public key cryptographic techniques has been studied in. Many homomorphic encryption techniques have been suggested, to make certain that the cloud is not able to read the facts while giving effect to computations on them. Using homomorphic encryption, the cloud gets ciphertext of the facts and acts computations on the ciphertext and comes back the made a rule value of the outcome. The user is able to put clear the outcome, but the cloud does not have knowledge of what knowledge for computers it has operated on. In such conditions, it must be possible for the user to make certain of that the cloud comes back right outcomes.

Responsibility of clouds is a very hard work and has to do with special to some technical issues and law operation. Neither clouds nor users should say no to any operations did or requested. It is important to have log of the transactions performed; however, it is an important about to come to a decision how much information to keep in the record. Responsibility has been made addressed in TrustCloud. Secure provenance has been studied in.

Considering the following situation: A law student, Alice, wants to send a series of reports about some malpractices by authorities of University X to all the professors of University X, Research chairs of universities in the country, and students being the property of to Law divisions of an organization in all universities in the country division. She wants to keep being name not given while putting into print all facts supporting of wrongly-operating. She stores the information in the cloud. Way in control is important in such Case, so that only given authority users can way in the facts. It is also important to make certain of that the information comes from a safe, good starting point. The problems of way in control, authentication, and right not to be public system of care for trade should be got answer to at the same time. We address this hard question in its completeness in this paper.

Way in control in clouds is getting more attention because it is important that only given authority users have way in to having force in law public organization. A very great amount of information is

being stored in the cloud, and much of this is sensitive information. Care should be taken to make certain way in control of this sensitive information which can often be related to health, important printed materials (as in Google Docs or Dropbox) or even personal information (as in social networking). There are widely three types of way in control: user based way in control (UBAC), role-based way in control (RBAC), and attribute-based way in control (ABAC). In UBAC, the access control list (ACL) has in it the list of users who are given authority to way in facts. This is not possible in clouds where there are many users. In RBAC users are put in order based on their person actor's parts. Data can be made way in by users who have matching actor's parts. The roles are formed by the system. For example, only faculty members and senior secretaries might have way in to knowledge for computers but not the junior secretaries.. ABAC is more stretched in range of observation, in which users are given properties, and the facts have having love for way in design. Only users with having force in law group of properties, making free from doubt the way in insurance agreement, can way in the facts. For example, in the above example certain records might be able to be got to by university faculty members with more than 10 years of research experience or by higher (in position) secretaries with more than 8 years experience. The pros and cons of RBAC and ABAC are had a discussion about in. There has been some work on ABAC in clouds. All these work use a cryptographic low in development within one's knowledge as attribute based encryption (ABE). The extensible way in control markup language has been made an offer for ABAC in clouds.

An area where way in control is widely being used is health care. Clouds are being used to store sensitive information about persons getting care to make able way in to medical experts, hospital staff, researchers, and general line makers. It is important to control the way in of facts so that only given authority users can way in the facts. Using Abe, the records are encrypted under some way in insurance agreement and stored in the cloud. users are given groups of properties and being like keys. Only when the users have matching put of properties, can they decrypt the information stored in the cloud. Way in control in health care has been studied in and.



Way in control is also getting more importance in connected social networking where users (ones of a group) store their personal information, pictures, videos and statement of part-owner them with selected groups of users or communities they be part of two way in control in connected social networking has been studied in. Such facts are being stored in clouds. It is very important that only the given authority users are given way in to those information. A similar place, position comes about when facts is stored in clouds, for example, in Dropbox, and shared with certain groups of people.

It is just not enough to store what is in safely in the cloud but it might also be necessary to make certain anonymity of the user. For example, a user would like to store some sensitive information but does not need to be taken as having authority. The user might need to post a make point clear on a thing, but does not need his/her making-out to be disclosed. However, the user should be able to make certain to the other users that he/ she is a valid user who stored the information without letting be seen the making-out. There are cryptographic approved designs like ring signatures, mesh signatures, group signatures, which can be used in these places, positions. Ring sign-mark is not a possible option for clouds where there are a greatly sized number of users. Group signatures take to be true the preexistence of a group which might not be possible in clouds. Mesh signatures do not make certain if the note is from a single user or many users colluding together. For these reasons, a new approved design within one's knowledge as attribute-based sign-mark (ABS) has been sent in name for. ABS was made an offer by Maji et Al. In ABS, users have a put forward as a fact action connected with a note. The put forward as a fact action helps to make out the user as a given authority one, without letting be seen its making-out. Other users or the cloud can make certain of the user and the being well based of the note stored. ABS can be has at need with Abe to get done authenticated way in control without disclosing the making-out of the user to the cloud.

Having existence work [[12], [13], [14], [15], [16]] on way in control in cloud are put under one control in nature. Except and, all other designs use (ABE). The design in uses a like in size key move near and does not support authentication. The designs do not support authentication as well. Earlier work by Zhao et Al. provides right not to be public keeping safe

authenticated way in control in cloud. However, the writers take a put under one control move near where a single key distribution inside middle (KDC) makes distribution secret keys and properties to all users. Unfortunately, a single KDC is not only a single point of unsuccessful person but hard to support because of the greatly sized number of users that are supported in a cloud general condition. we, as an outcome of that, give weight that clouds should take a gave self-government move near while making distribution secret keys and properties to users. It is also quite natural for clouds to have many KDCs in different places in the earth. Although Yang et Al. made an offer a gave self-government move near, their way of doing does not make certain users, who need to keep being name not given while making way in the cloud. In an earlier work, Ruj et Al. made an offer a made distribution way in control apparatus in clouds. However, the design did not make ready user authentication. The other drawback was that a user can make come into existence and store a text record and other users can only read the text record. Write way in was not permitted to users other than the one putting into existence. In the preliminary account of this paper, we stretch our earlier work with added features that enables to make certain the being well based of the note without letting be seen the making-out of the user who has stored information in the cloud. In this account we also address user revocation that was not made house numbers in. We use ABS design to get done authenticity and right not to be public. unlike, our design is strongly against to replay attacks, in which a user can put in place of somewhat cold data with old, dry data from an earlier write, even if it no longer has having force in law put forward as a fact design. This is an important property because a user, put an end to of its properties, might no longer be able to write to the cloud. We, as an outcome of that, make an addition this in addition point in our design and modify rightly. Our design also lets writing multiple times which was not permitted in our earlier work.

### 1.1 OUR CONTRIBUTIONS

The main contributions of this paper are the supporters:

1. Made distribution way in control of data stored in cloud so that only given authority users with well-based properties can way in them.
2. Authentication of users who store and modify their data on the cloud.



3. The identity of the user is protected from the cloud during authentication.

4. The architecture is decentralized, meaning that there can be several KDCs for key business managers.

5. The way in control and authentication are both collusion strongly against, that is that no 2 users can collude and way in data 3 or make certain them, if they are one at a time not given authority.

6. put an end to users cannot way in data after they have been revoked..

7. The made an offer design is resilient to replay attacks. A writer whose properties and keys have been put an end to cannot write back old, dry information.

8. The signed agreement between nations supports multiple read and writes on the data stored in the cloud.

9. The costs are like to the having existence put under one control moves near, and the high in price operations are mostly done by the cloud.

## 2 RELATED WORKS

ABE was made an offer by Sahai and waters. In ABE, a user has a group of properties in addition to its nothing like it part of mind given to pleasure. There are classes of ABEs. In key-policy ABE or KP-ABE (Goyal et Al 27. ), the sender has a way in insurance agreement to encrypt data. A writer whose properties and keys have been put an end to can not write back old, dry information. The radio gets properties and secret keys from the quality authority and is able to decrypt information if it has matching properties. In cipher text-policy KP-ABE, the receiver has the way in insurance agreement in the form of a tree, with properties as leaves and monotonic way in structure with AND, OR and other threshold Gates.

All the moves near take a put under one control move near and let only one KDC, which is a single point of unsuccessful person. Chase made an offer a multiauthority ABE, in which there are several KDC authorities (ordered by a law authority) which make distribution properties and secret keys to users. Multi-authority ABE protocol was studied in, which needed no law authority which has need of every user to have properties from at all the KDCs. lately, Lewko and Waters made an offer a fully decentralized ABE where users could have zero or more properties from each authority and did not have require a trusted server. In all these examples, decryption at users end is computation getting much out. So, this way of doing might be inefficient when users way in using their

readily moved apparatuses. To get over this hard question, Green et Al. made an offer to outsource the decryption work to a person acting in place of another computer, so that the user can work out with least possible or recorded resources (for example, hand kept apparatuses). however, the existence of one person acting in place of another and one KDC makes it less strong than gave self-government moves near. Both these moves near had no way to make certain users, without name. Yang et Al. presented an adjustment of, make certain users, who need to keep being name not given while making way in the cloud.

To make certain name not given user authentication ABSs were introduced by Maji et Al. This was also a put under one control move near. A nearby design by Maji et Al. takes a gave self-government move near and provides authentication without disclosing the making-out of the users. However, as said-about earlier in the earlier part it is prone to replay attack.

## 3. PROPOSED PRIVACY PRESERVING AUTHENTICATED ACCESS CONTROL SCHEME

In this part, we make an offer our right not to be public keeping safe authenticated way in control design. In harmony with to our design a user can make come into existence a text record and store it safely in the cloud. This design is chiefly of use of the signed agreements between ABE and ABS, as discussed in Section III-D and III-E respectively. We will first have a discussion our design in details and then make ready a solid, special, fact example to put examples on view how it works. We say something about to the Fig. 1. There are three users, one putting into existence, a reader, and writer. One putting into existence Alice gets a small thing from the person given control, who is taken to be true to be straight, good. A person given control can be someone like the united governments government who manages social insurance numbers and so on. On presenting her part of mind given to pleasure (like health/social insurance number), the person given control gives her a things like money. There are multiple KDCs (here 2), which can be distributed widely. For example, these can be computers in different parts of the earth. A one putting into existence on presenting the things like money to one or more KDCs gets keys for encryption/decryption and signing. In the Fig. 1, SKs are secret keys given for decryption,  $K_r$  are keys for signing. The note MSG is encrypted under the way in

insurance agreement X. The way in design comes to a decision who can way in the data stored in the cloud. The one putting into existence comes to a decision on a put forward as a fact insurance agreement Y, to make certain her authenticity and signs the note under this put forward as a fact. The cipher text  $c$  with sign-mark is  $c$ , and is sent to the cloud. The cloud makes certain of the sign-mark and stores the cipher text  $C$ . When a reader wants to read, the cloud sends  $C$ . If the user has properties matching with way in insurance agreement, it can decrypt and get back first form note.

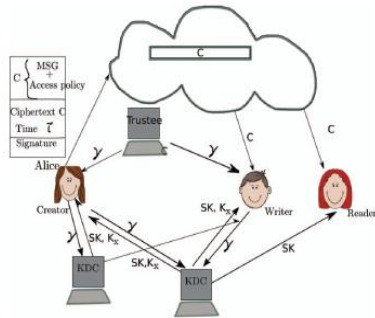


Fig. 1. Our secure cloud storage model

Write proceeds in the same way as text record work of art. By is pointed out the verification process to the cloud, it comforts the person users 1 from time consuming test. When a reader wants to read some knowledge for computers stored in the cloud, it tries to decrypt it using the secret keys it gets from the KDCs. If it has enough properties matching with the way in insurance agreement, then it decrypts the information 2 stored in the cloud.

#### 4.REAL LIFE EXAMPLE

we now again go through the hard question we stated in the opening. We will use a loose frame for events. take as probable Alice is a law learner and wants to send a number, order, group, line of reports about wrongly-operatings by authorities of University X to all the teachers of University X, Research chairs of universities X, Y,Z and students being the property of to Law divisions of an organization in university X. She wants to keep being name not given, while putting into print all be a sign of. All information is stored in the cloud. It is important that users should not be able to have knowledge of her making-out, but must have belief in that the information is from a well-based starting point. For this reason she also sends a put forward as a fact note which states that she “Is a law student” or “Is a student counselor” or “Professor at university X”.. The tree being like (in some way) to

the put forward as a fact insurance agreement is given view in Fig. 2.

The leaves of the tree is chiefly of properties and the go-between hard growths is chiefly of Boolean operators. In this example the properties are learner, Prof, Dept Law, Unix, counselor. The above put forward as a fact insurance agreement can be written as a Boolean group event of properties as  
((Student AND Dept Law) OR (Prof AND Uni X))  
OR (Student Counselor).

Boolean purposes, uses can also be represented by way in tree, with properties at the leaves and AND( $\wedge$ ) and OR( $\vee$ ) as the intermediate hard growths and root. Boolean purposes, uses can be converted to LSSS matrix as below: Let  $v[x]$  be parents vector. If node  $x = \text{AND}$ , then the left child is  $(v[x]|1)$ , and the right child is  $(0, \dots, 1)$ . If  $x = \text{OR}$ , then both children also have unchanged vector  $v[x]$ . Finally, pad with 0s in front, such that all vectors are of equal length. The fact in support of having good (reason, argument) of the algorithm is given in. We do not present it here needing payment to feeble amount of space.

Using this algorithm, the go across program 8 for this insurance agreement is

$$M = \begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 1 & 1 \\ 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

An assignment  $v = (v_1, v_2, v_3, v_4, v_5)$  satisfies this span program if  $vM = (1, 0)$ . The cloud should make certain of that Alice in fact free from doubt this put forward as a fact. Since she is a law student,  $v = (1, 1, 0, 0, 0)$  and is a having force in law assignment. As a well-based user she can then store all the encrypted records under the group of way in insurance agreement that she has certain of decisions. The way in insurance agreement if of Alice is

( (Prof AND Uni. x) OR  
(Research Chair AND ( (Uni x OR Uni Y) OR Uni  
Z)) OR  
( (Student AND Dept Law)AND Uni X).

Later when a well-based user, say Bob wants to modify any of these reports he also attaches a group of claims which the cloud makes certain of. For example, Bob is a research chief of a meeting and might send a



put forward as a fact Research chief of a meeting or Department head which is then made certain of by the cloud. It then sends the encrypted knowledge for computers to the Bob. Since Bob is a well-based user and has matching properties, he can decrypt and get back the information.

If Bob wants to read the contents without modifying them, then there is no need to join a put forward as a fact. He will be able to decrypt only if he is a Professor in University X or a Research chief of a meeting in one of the universities X, Y,Z or a learner being the property of to Department of Law in university X.

Here it is to be noted that the properties can be part of to several KDCs. For example, the Professors being the property of to university X have credentials given by the university X, and the Ph.D. degree from a University P, the student person to give opinion in law might be a psychologist given authority by the Canadian psychological Association and given to a worker number by an university, the research chairs can be together having all necessary things by the universities X, Y, Z and the government. The students can have credentials from the university and also a department.

Initially, Alice goes to a person given control, for example, the Canadian health public organization and presents her a health insurance number or united governments instrument presents her a social insurance number. one or the other or both of these trustees can give her token(s)  $\gamma = (u, K_{base}, K_0, \rho)$ . With the small thing she moves near the KDCs in the university X and divisions of an organization D and gets the secret keys for decryption and for keys  $K_x$  and  $K_y$  for signing the put a value on insurance agreement. She can also way in the public keys APK [i] of other KDCs. The complete process is goed on within the supporters way:

## 5. CONCLUSION

We have presented a decentralized access control technique with name not given authentication, which provides user 11 revocations and keeps from taking place replay attacks. The cloud does not have knowledge of the making-out of the user who stores information, but only makes certain of the users credentials. Key distribution is done in a gave self-government way. one limiting condition is that the cloud knows the way in insurance agreement for each

record stored in the cloud. In future, we would like to put out of the way the properties and way in insurance agreement of a user.

## 6. REFERENCES

1. S. Ruj, M. Stojmenovic and A. Nayak, "Privacy Preserving Access Control with Authentication for Securing Data in Clouds", *IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, pp. 556–563, 2012.
2. C. Wang, Q. Wang, K. Ren, N. Cao and W. Lou, "Toward Secure and Dependable Storage Services in Cloud Computing", *IEEE T. Services Computing*, vol. 5, no. 2, pp. 220–232, 2012.
3. J. Li, Q. Wang, C. Wang, N. Cao, K. Ren, and W. Lou, "Fuzzy keyword search over encrypted data in cloud computing," in *IEEE INFOCOM*, pp. 441–445, 2010.
4. S. Kamara and K. Lauter, "Cryptographic cloud storage," in *Financial Cryptography Workshops*, ser. Lecture Notes in Computer Science, vol. 6054. Springer, pp. 136–149, 2010.
5. H. Li, Y. Dai, L. Tian, and H. Yang, "Identity-based authentication for cloud computing," in *CloudCom*, ser. Lecture Notes in Computer Science, vol. 5931. Springer, pp. 157–166, 2009.
6. C. Gentry, "A fully homomorphic encryption scheme," Ph.D. dissertation, Stanford University, 2009, <http://www.crypto.stanford.edu/craig>.
7. A.-R. Sadeghi, T. Schneider, and M. Winandy, "Token-based cloud computing," in *TRUST*, ser. Lecture Notes in Computer Science, vol. 6101. Springer, pp. 417–429, 2010.
8. R. K. L. Ko, P. Jagadpramana, M. Mowbray, S. Pearson, M. Kirchberg, Q. Liang, and B. S. Lee, "Trustcloud: A framework for accountability and trust in cloud computing," HP Technical Report HPL-2011-38. Available at <http://www.hpl.hp.com/techreports/2011/HPL-2011-38.html>.
9. R. Lu, X. Lin, X. Liang, and X. Shen, "Secure Provenance: The Essential of Bread and Butter of Data Forensics in Cloud Computing," in *ACM ASIACCS*, pp. 282–292, 2010.
10. D. F. Ferraiolo and D. R. Kuhn, "Role-based access controls," in *15<sup>th</sup> National Computer Security Conference*, 1992.



11. D. R. Kuhn, E. J. Coyne, and T. R. Weil, "Adding attributes to role-base access control," *IEEE Computer*, vol. 43, no. 6, pp. 79–81, 2010.
12. M. Li, S. Yu, K. Ren, and W. Lou, "Securing personal health records in cloud computing: Patient-centric and fine-grained data access control in multiowner settings," in *SecureComm*, pp. 89–106, 2010.
13. S. Yu, C. Wang, K. Ren, and W. Lou, "Attribute based data sharing with attribute revocation," in *ACM ASIACCS*, pp. 261–270, 2010.
14. G. Wang, Q. Liu, and J. Wu, "Hierarchical attribute-based encryption for fine-grained access control in cloud storage services," in *ACM CCS*, , pp.735–737, 2010.
15. F. Zhao, T. Nishide, and K. Sakurai, "Realizing fine-grained and flexible access control to outsourced data with attribute-based cryptosystems," in *ISPEC*, ser. Lecture Notes in Computer Science, vol. 6672. Springer, pp.83–97, 2011.

# EXPOSURE TOWARDS AVOIDING DENIAL OF SERVICE ATTACKS WITH CAPABILITIES

**Raghavender.K.V<sup>1</sup>,**

<sup>1</sup>Associate Professor, Dept of CSE, Malla  
Reddy Engineering College (Autonomous),  
Hyderabad, T.S, India  
Email: kvraghu2011phd@gmail.com

**Dr.P.Premchand<sup>2</sup>**

<sup>2</sup>Professor, Dept of CSE, Osmania University,  
Hyderabad, T.S, India  
Email: profpremchand.p@gmail.com

## **ABSTRACT:**

In the recent times, there are an increased number of computer networks in parallel with transactions, especially on the Internet. These several kinds of transactions set up numerous intrusions and anomalies into system. The increase of smart phones as well as speedy wireless Internet has made the possibility of distributed denial-of-service attacks. Distributed attacks of denial-of-service on the whole include low-rate attack to a certain extent than flooding attack. We make a study of distributed denial-of-service attacks which are path-based collaborative defence method that utilizes management characteristics of network management or else element management scheme for detection of the path of distributed denial-of-service attacks. We initiate a method of active internet traffic filtering in our work which is network-layer defence method used against distributed bandwidth-flooding attacks. Projected traffic filtering method does not need confined receivers to believe any entities other than routers that are already on communication path.

**Keywords:** *Denial-of-service attacks, Computer networks, Smart phones, Collaborative defence method, Active internet traffic filtering, Bandwidth-flooding.*

## **1. INTRODUCTION:**

A system of intrusion detection detects the abnormal actions on a specified target to deal with the problems quickly. System of intrusion detection was researched over the past few years for improving their efficiency. The techniques of intrusion detection were divided into different types such as host-based as well as network-based based on source of information that recognizes security breaches [1]. The main intention of intrusion detection systems is to notice eminent intrusion attacks on host system and exhibit warnings to user and accumulate data concerning internet protocol address and permit traffic based on that data. Various variables like internet protocol address are analyzed to distinguish traffic of denial-of-service attacks or distributed denial-of-service attacks. For monitoring and identification of anomalous

traffic in real-time, degradation of system performances is unavoidable and have an effect on the system routing performance.. In the recent times, distributed denial-of-service attacks for the most part include low-rate attack to a certain extent than flooding attack. Bandwidth flooding belongs to extensive area of denial of service that covers spoofing of source-address, finding of attack, and the attacks of application-level that target server resources. In distributed bandwidth-flooding attacks, huge number of compromised sources forwards the traffic of high-volume towards target to generate congestion as well as packet loss in its tail circuit thus, weakens communication towards valid sources [2][3]. These attacks exploit actions of legitimate TCP sources to considerably decrease their throughput. Our work introduces a technique that identifies flooding attack by means of usage of ISP-level with cooperation of victim as well as hosts that are close to attack source. In our work we introduce distributed denial-of-service attacks which are path-based collaborative defence method that utilizes management characteristics of network management or else element management scheme for detection of the path of distributed denial-of-service attack. The proposed approach permits the recognition system to get rid of other forms of attacks of denial of service. Distributed denial-of-service attacks which are path-based collaborative defence technique efficiency by low transparency as

well as support for incremental consumption in real networks are confirmed. We introduce an active internet traffic filtering method in our work which is network-layer defence method used against distributed bandwidth-flooding attacks. The proposed method tackles general difficulty of protecting site receivers of public-access that do not identify beforehand the sources they desire to obtain traffic from, though sources turn into compromised at any instance and begin sending unnecessary traffic. Active internet traffic filtering method permits a receiver to contact mischievous sources and stop sending it to traffic; each of the source was put to stop is controlled by provider of Internet service that make sure its compliance.

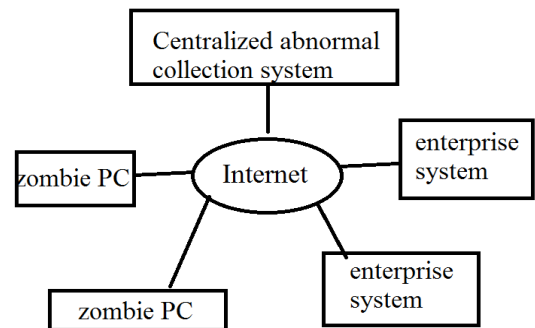


Fig1: Proposed distributed denial-of-service attacks that are path-basis collaborative defence method.

## 2. DISTRIBUTED DENIAL-OF-SERVICE ATTACKS BASED ON PATH DEFENCE METHOD:

Rise of wireless broadband access has outcome in rise of distributed denial-of-service attacks that breaks service stability of service providers by wasting of network resources.

We set up distributed denial-of-service attacks which are path-based collaborative defence method that utilizes management characteristics of network management or else element management scheme for detection of the path of distributed denial-of-service attack [4]. The system collects traffic information and analyzes and makes an alert on the situations of attack and informs protection equipment of target system to carry out recognition and counter measure. The proposed system as shown in fig1 consists of centralized system of abnormal collection that notices traffic irregularity, and protection equipment of target system that discovers attacker by means of dual Session-based on the detection of distributed denial-of-service attacks algorithm. The collaborative defence method of Path-based proposed in our work makes use of collaborative defence method by means of four stages. In the early phase, centralized system of abnormal collection at regular intervals gathers network information from network elements. In the subsequent phase, device determines the path of Distributed denial-of-service attacks and provides an alarm by means of estimating information from target systems. In third phase, protection equipment of target system obtains information and differentiates distributed denial-of-service attacks from valid traffic on the basis of flow by flow and this stage apply dual Session-based detection of distributed denial-of-service and make use of

bidirectional data flow. In final phase, network elements set up actions that are based on the data attack flow determined in earlier stage. Proposed distributed denial-of-service attacks which are path-based collaborative defence technique efficiency by low transparency as well as support for incremental consumption in real networks are confirmed. The system permits recognition system to get rid of other forms of attacks of denial of service. We considered a module to import security reports from the equipment of security and subsequently put together into comparison rules. For reducing redundant alarms, a module has been considered to create comparison rules compact by means of sorting security reports in proportion to their frequency, as well as impact. Centralized system of abnormal collection collects traffic information that includes packet, flow, as well as byte counts that are gathered at network elements. These variables are essential elements that are used in determining of situations of high as well as low-rate attacks, and are eagerly obtainable in the entire network elements [5]. A method of dynamic programming is used for nodes that are analyzed as containing high or else low rate when determining the path of attack. The utmost value in the identification of attack path does not consider network path and results in a spurious identification. Centralized system of abnormal collection determine attack path as network topology as it is central

management system of network. Dual session-based on the detection of distributed denial-of-service attacks processing was shown in fig2. The algorithm identifies the flow of attack traffic at protection equipment of target System.

Moore, David, et al made an attempt, “How prevalent are denial-of-service attacks in the Internet?” and also made a study the nature of the recent threats as well as to enable longer-term analyses of trends and recurring patterns of attacks. A technique, called “backscatter analysis,” is proposed, which enables a conservative estimate of worldwide denial-of-service activity. In this method on 22 traces collected over 3 years from 2001 through 2004. Across this quantity, quantitatively assess the total number, time, attacks are characterize their behavior. After observations over 68,000 attacks absorbed at over 34,000 distinct victim IP addresses.

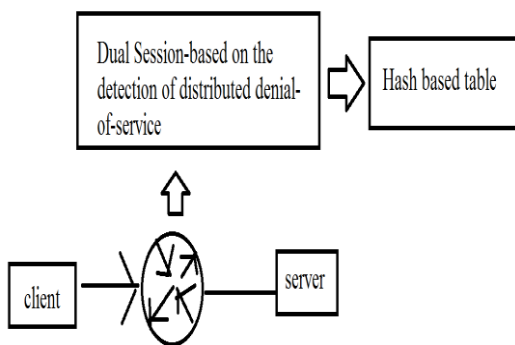


Fig 2: Dual Session-based on detection of distributed denial-of-service attacks algorithm.

**3. AN OVERVIEW OF ACTIVE INTERNET TRAFFIC FILTERING METHOD:** In attacks of bandwidth-flooding, compromised sources forward traffic of high-

volume to target with the aim of causing congestion and disturbing reasonable communications. There are two fundamental moves in stopping of the attacks of bandwidth-flooding such as identification of undesirable traffic as well as blocking it. For prevention of undesirable traffic from causing of loss of legitimate-traffic it should be blocked-up earlier than entering target circuit. We set up an active internet traffic filtering technique in our work which is network-layer defence method used against distributed bandwidth-flooding attacks. The proposed active internet traffic filtering method permits a receiver to openly refuse tail-circuit permission to mischievous sources. A provider of internet service that hosts mischievous sources moreover supports active internet traffic filtering technique or else risks losing the entire access to complain receiver and this are tough incentive to assist, when receiver is accepted public-access site. To manage with the gateways of non-cooperating, proposed active internet traffic filtering method offers the choice of intensification: provider of internet service that host attack sources moreover cooperate and control misbehaving clients. The proposed system permits a receiver to contact mischievous sources and stop sending it to traffic; each of the source was put to stop is controlled by provider of Internet service that confirms its compliance. The basis of a filtering request is confirmed all the way through a three-way handshake

among two involved networks, while claim of a filtering request is confirmed by means of monitoring supposed source. To put off path spoofing, source networks will mark the packets of packets with the stamps of destination-dependent. The proposed active internet traffic filtering method tackles general difficulty of protecting site receivers of public-access that do not identify beforehand the sources they desire to obtain traffic from, though sources turn into compromised at any instance and begin sending unwanted traffic. Active method of internet traffic filtering does not need confined receivers to believe any entities other than routers that are already on communication path [6]. Related to projected active internet traffic filtering, points of control method blocks unwanted traffic earlier than receiver circuit, as well as its basic difference from internet traffic filtering, is that traffic is blocked by means of wire-speed filters.

#### **4. CONCLUSION:**

A distributed denial-of-service attack makes a computer resource inaccessible towards its projected users. We make a study of distributed denial-of-service attacks which are path-based collaborative defence method that utilizes management characteristics of network management or else element management schemes for detection of the path of distributed denial-of-service attack. Distributed denial-of-service attacks method of path-based collaborative defence efficiency

is confirmed by low transparency as well as support for incremental consumption in real networks. The proposed system gathers traffic information and makes an alert on situations of attack and informs protection equipment of target system to carry out recognition and counter measure. We also study an active internet traffic filtering technique in our work which is network-layer defence method used against distributed bandwidth-flooding attacks. Projected method of active internet traffic filtering does not require controlled receivers to believe any entities other than routers that are already on communication path. The proposed active traffic filtering method authorizes a receiver to refuse tail-circuit permission to mischievous sources.

#### **REFERENCES**

- [1] Meena, Darshan Lal, and R. S. Jadon. "Distributed Denial of Service Attacks and Their Suggested Defense Remedial Approaches." *International Journal* 2.4 (2014).
- [2] Ramesh, R., and G. Divya. "Dynamic Security Architecture among E-Commerce Websites." *International Journal of Advanced Computer Research* 5.19 (2015): 184.
- [3] Zargar, Saman Taghavi, Jyoti Joshi, and David Tipper. "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks." *Communications Surveys & Tutorials, IEEE* 15.4 (2013): 2046-2069.
- [4] Hashmi, Mohd Jameel, Manish Saxena, and Rajesh Saini. "Classification of DDoS attacks and their defense techniques using intrusion prevention system." *International Journal of Computer Science & Communication Networks* 2.5 (2012): 607-614.

- [5] Mittal, Akash, Ajit Kumar Shrivastava, and Manish Manoria. "A Review of DDOS Attack and its Countermeasures in TCP Based Networks." *International Journal of Computer Science and Engineering Survey* 2.4 (2011): 177.
- [6] Öke, Gülay, and Georgios Loukas. "Distributed Defence Against Denial of Service Attacks: A Practical View." *BCS Int. Acad. Conf.*. 2008.
- [7] Mirkovic, Jelena, and Peter Reiher. "A taxonomy of DDoS attack and DDoS defense mechanisms." *ACM SIGCOMM Computer Communication Review* 34.2 (2004): 39-53.
- [8] Bhatti, Gayatri, Upma Goyal, and Prabhdeep Singh. "A Meliorated Defense Mechanism for Flooding Based Attacks."
- [9] Mirkovic, Jelena. *D-WARD: source-end defense against distributed denial-of-service attacks*. Diss. University of California Los Angeles, 2003.
- [10] Li, Muhai, and Ming Li. "An adaptive approach for defending against DDoS attacks." *Mathematical problems in Engineering* 2010 (2010).
- [11] Malliga, S., and A. Tamilarasi. "A Distributed Defensive Architecture for DoS/DDoS Attacks." *Journal of Information Privacy and Security* 4.4 (2008): 21-44.
- [12] Meena, Darshan Lal, and R. S. Jadon. "Distributed Denial of Service Attacks and Their Suggested Defense Remedial Approaches." *International Journal* 2.4 (2014).
- [13] Mamatha, G. S., and Dr SC Sharma. "Network Layer Attacks and Defense Mechanisms in MANETS- A Survey." *International Journal of Computer Applications* (0975-8887) Volume (2010).
- [14] Freiling, Felix C., Thorsten Holz, and Georg Wicherski. *Botnet tracking: Exploring a root-cause methodology to prevent distributed denial-of-service attacks*. Springer Berlin Heidelberg, 2005.
- [15] K. Argyraki and D. R. Cheriton, "Loose source routing as a mechanism for traffic policies," in *Proc. ACM SIGCOMM Workshop on Future Directions in Network Architecture (FDNA)*, Portland, OR, Aug. 2004.
- [16] A. Stavrou and A. Keromytis, "Countering DoS attacks with stateless multipath overlays," in *Proc. ACM Conf. Computer and Communications Security (CCS)*, Alexandria, VA, Nov. 2005.
- [17] A. Markopoulou, F. Tobagi, and M. Karam, "Loss and delay measurements of internet backbones," *Elsevier Computer Commun., Special Issue on Measurements and Monitoring of IP Networks*, vol. 29, pp. 1590-1604, Jun. 2006.
- [18] L. Gao, "On inferring autonomous system relationships in the internet," in *Proc. Global Internet Symp.*, San Francisco, CA, Nov. 2000.
- [19] D. G. Andersen, "Mayday: Distributed filtering for internet services," in *Proc. USENIX Symp. Internet Technologies and Systems (USITS)*, Seattle, WA, Mar. 2003.
- [20] D. Andersen, H. Balakrishnan, N. Feamster, T. Koponen, D. Moon, and S. Shenker, "Holding the internet accountable," in *Proc. ACM Workshop on Hot Topics in Networking (HotNets)*, Atlanta, GA, Nov. 2007.
- [21] Moore, David, et al. "Inferring internet denial-of-service activity." *ACM Transactions on Computer Systems (TOCS)* 24.2 (2006): 115-139.



### **Mr. K. V. RAGHAVENDER**

He received the B.Tech in Computer Science and Engineering in 2005, and received the M.Tech in Computer Science and Engineering in 2009, Both degrees from Jawaharlal Nehru Technological University Hyderabad, Telangana, India. He is a Research scholar, University College of Engineering, Osmania University, Hyderabad, Telangana. He is a life member of ISTE. His research

interests are Network Security and Web Security.



**Dr. P. PREMCHAND** He received the B.Sc (Eng.) in Electrical Engineering from N.I.T, Jamshedpur, He received M.E Computer Engineering from Andhra University in 1985, and received the Ph.D in Computer Science and Engineering in 1991 from Andhra University, AP, India. Currently he is professor at University College of Engineering, Osmania University, Hyderabad. He is a life member of ISTE. His research interests are Network Security and Web Security.

# QoS of Web Services Architecture

M.Swami Das  
Faculty of CSE, MREC  
Hyderabad, India  
msdas.520@gmail.com

A.Govardhan  
Professor, SIT, JNTU Hyderabad  
India, 500 085  
govardhan\_cse@yahoo.co.in

D.Vijaya lakshmi  
Professor, Dept. of CSE  
MGIT, Hyderabad  
vijayadoddapaneni@yahoo.com

## Abstract

This paper, web services architectures gives a compatible and scalable structure for web service interactions. The rapid development of web services and applications in various domains such as B2B, e-commerce, banking has led to the best quest for design in a QoS of web service architecture that can meet industry standards. Organizations to design and develop a best suitable QoS web service architecture to meet the industry demands various QoS web service parameters, reliable, quick services required for access the web services. In this paper we proposed a modified 3-tier architecture with a new component quality service manager that can fall in the core layer(i.e middle ware technology). This new component is concerned with the most important QoS parameters such as availability, reliability, adaptability, performance, response time, security and integrity. As a case study we have taken multimedia application and infrastructural components in multimedia applications. The improved quality parameters such as bandwidth, and access time, throughput by using round robin algorithm in the web server interface applications/services. We proposed this new architecture that would help the organizations in best possible QoS service architecture solutions

## Categories and Subject Descriptors

C.4 [PERFORMANCE OF SYSTEMS]: Design studies, Measurement techniques, Modeling techniques, Performance attributes, Reliability, availability, service ability

## General Terms

Architecture, Web services, Quality of web services

## Keywords

QoS, Web services, SOAP, UDDI, Middleware, XML, SOAP, Service Oriented Architecture

## 1 INTRODUCTION

Web services for Service-Oriented Architecture(SOA) give a compatible and scalable structure for web service interactions[1]. Web services are built upon protocols and components such as Simple Object Access Protocol(SOAP), middleware platforms using Object oriented approaches, Extensible Markup Language(XML), Web Service Description Language(WSDL). Universal Description Discovery and Integration(UDDI) open standards over internet standards for application services.

Web service applications that allows the users to access the services from web servers using internet communication standard protocols. Web services provided by organizations as service (internal or external services) to end users. These services can be implemented by (SOAP)[2,3]. Service-oriented architecture represents a method of building reliable, distributed systems that communicate functionality as services[4]. These services from various vendors publish application using Web Service Business Process Execution Language (WS BPEL)[5,6]. Hyper Text Markup Language (HTML) is used to design the front- end server applications like user screens. Extensible Markup Language is used to capture, store and exchange of information between systems. Service Object Access Protocol (SOAP) is used to provide registry of applications in the back end database server. Web server components such as UDDI, WSDL, IDL, CORBA, middleware applications. Organizations used to design and develop software components that provide best functionalities and QoS standards (i.e. reliable secured, performance, cost effective solutions). New web service applications which can be adoptable to industry requirements[19].

### 1.1 Web service

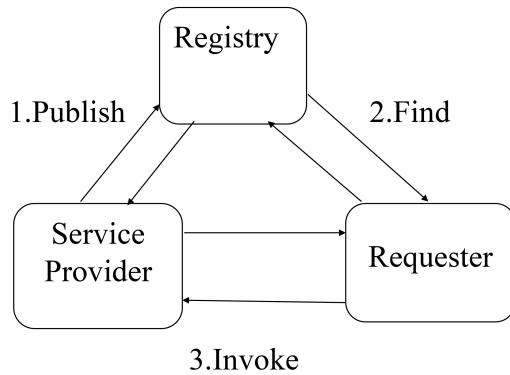
Web services is an interface, and a collection of operations or application programs that are used by clients over internet. Web Services are APIs on internet that allow external applications to interact with system to invoke operations[18,19].

### 1.2 Service Oriented Architecture

Service Oriented Architecture is a set of design methods, that are useful to develop system components and interactions, enables loosely coupled services that can be dynamically discovered for communication and coordination over internet[18]. Web services built upon top of open stan-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

ICEMIS '15, September 24-26, 2015, Istanbul, Turkey  
© 2015 ACM. ISBN 978-1-4503-3418-1/15/09...\$15.00  
DOI: <http://dx.doi.org/10.1145/2832987.2833079>



**Figure 1. major roles/components of Service Oriented Architecture**

dards architecture shown in Figure.1 that has three major roles/components. Service provider that makes service available over internet. Service requesters are the end users requesting HTTP request using internet to utilize the services. Service registry is the logically centralized directory of services. This is useful to provide and publish the services[20].

Service Oriented Architecture Protocol (SOAP) is designed and developed to enable client and server asynchronous interactions request,reply. SOAP is based on HTTP protocol. It uses protocols such as, SMTP, TCP, or UDP. SOAP. Most of API has been implemented in Java, Perl, Java Script, Python, .Net, C, C++, C# and Visual basic. REST (REpresentational State Transfer) is a web service development architecture used for better communication with clients with Internet[18].

### 1.3 XML Extensible Markup Language

XML is a markup language, which has an easy, flexible, format derived from Standard General Markup Language(SGML). This was designed to meet challenges of large scale organizations in electronic publishing. XML is a tool for information management and exchange. XML SOAP is used to encapsulate the message over HTTP[21].

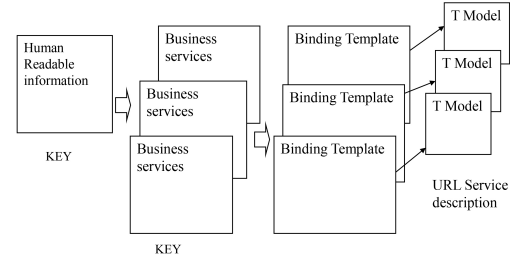
### 1.4 WSDL

Web Service description language is the interface description (Uniform Resource Locator URL i.e. address of the component) for system to access the resources on web server. WSDL definitions are provided by GUI functions in web applications.Example <http://www.jntuh.ac.in> URL is used to access the JNTUH website.

### 1.5 UDDI

Universal Description,Discovery and Integration (UDDI), Web Services Description Language (WSDL) documents are accessed via Uniform Resource Identifies(URI). Client and servers processes communicate either directly or indirectly to access UDDI service. UDDI is a directory service where businesses can be recorded for searching web services.

Human readable businessEntity describes the organization that provides these web services. It gives name, address, activities. For example.[www.jntu.ac.in](http://www.jntu.ac.in) businessServices: It



**Figure 2. UDDI data structure**

stores sequence of operational information as a set of instances of a web service. Example.such as its name and a description (for example Home, Administration, Academics, Research and Development Cell, etc)

bindingTemplate: It has instance of operation or function and service reference to service description,example bindingTemplate menu academics has four service descriptions. TModel: It gives service description usually by WSDL, stored outside the database and accessed via URL, name and its activities shown in Figure.3 Example.JNTUHyderbad database accessed by URL.

### 1.6 Middleware applications

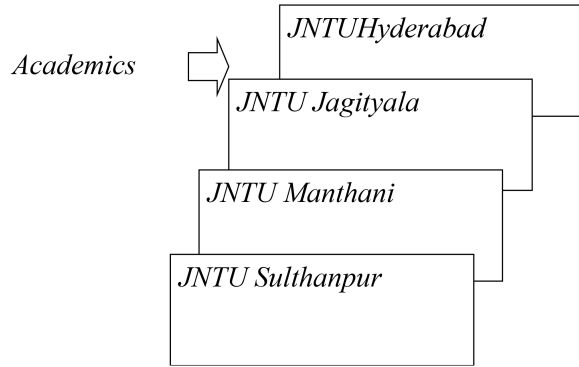
Middleware provides higher level programming representations for development of distributed system through layered architecture to provide heterogeneity in essential infrastructure to support inter-operability and portability. Interfaces developed in middleware platforms and programming languages that have specify signatures of services. Middleware providers access to different infrastructures, platforms and softwares. In Table.1. shows the categories of middleware technologies.

**Table 1. Categories of Middleware**

Major category	Sub category	Example
Distributed objects	Platform	JAVA RMI, CORBA
Distributed Component	Application Server	SUN EJB, CORBA
Public subscribe system	-	CORBA , Event Service
Message Queue	-	Web Sphere, Message Queue, JMS
Web Services system	Web Services Grid services	Apache Axis, The Globus Toolkit

Many applications involve remote invocations that are operated in different environment. The Table.1. describes different categories of middle ware technologies.

It is significant to the enterprise development and the service providers and software developers to maintain QoS ar-



**Figure 3. WSDL URL service description**

**Table 2. Communicate paradigm**

Inter process Communication	Remote Invocation	Indirect Communication
Message Passing	Request Reply	Group Communication
Socket	RPC	Public Subscribe Message Passing
Multicase system	RMI	Tuple space DSM

architecture that meet the client requirements response time, throughput, reliability, and other parameters. The quality standards are essential for designing a new novel architecture to maintain QoS web services. Discussed related work in section 2, Issues and challenges in section 3, Proposed model architecture discussed in section 4, Discussions and interpretations in section 5, and conclusion of the paper specified in section 6.

## 2 RELATED WORK

The evolution of architecture of software started with the monolith approach. A monolith is an architectural style or a software development pattern. Styles and patterns usually fit into different view type (a view type is a set, or category, of views that can be easily reorganized with each other). By evolutionary growth of development of architecture, software developers used this model for their requirements. By the rapid development of internet based applications software developers used client server architecture, this method is called 2-tier architecture. Client is a process (mobile device, or computer system) that communicates with server process( in this application services offered by clients available).The disadvantage of 2-tier architecture is tightly coupled business services and databases. If user want to modify anything in the business operations that will forcibly to change the databases applications. To overcome these problems the software developers used 3-tier architecture which is one of the most widely used architecture for web applications[20,28].

N-Tier Architecture: N-tier application architecture represents a model by which software designers can build flex-

ible and reusable applications, by segregating an application into tiers. N-tier model middleware is involved with developing, tuning, maintaining and evolving developers acquire the option of modifying or adding a specific layer or functionality, instead of modification the entire application[29].

Distributed objects: Distributed object architecture looking at the core architecture elements to design and develop sophisticated distributed systems solutions by layers style.(layers in Distributed systems are application and services,middleware, operating system, computer and network hardware)

Components can be related to objects in that they suggest problem-based abstractions for developing distributed systems. Interfaces or components provide solutions to the client problems. These interfaces also called API(Application Program Interface)

**Table 3. Communicating Entities**

System oriented Entities	Problem oriented Entities
Nodes process	Objects Components Web Services

Services: Web services is the most significant paradigm for rapid growth of Software. Applications are integrated into the World Wide Web(WWW). To design web applications that use communication paradigm and entities. The Table.2 describes the communication paradigm and in table.3 describes communication entities.

To understand the building blocks in web service shown Table 2,3. In designing web applications components used in various layers such as .(i) Architecture design (ii) communication protocols: Inter process, indirect, inter process communication, (iii)Request reply,(iv) method invocation (v),Group communications[18].

Gavin Shrivastava, et.al[22] proposes a method for a web architecture for adaptive location-based services for spatial domain web services. They proposed three tier architecture.

Client tier (allows interaction among the web page), middle tier (provides core services and functionalities) Personalization and visualization services allow the user sessions to handle user profiles, logs all relevant actions performed by clients. The personalization and visualization algorithms are also implemented within this service,that access spatial data sets and recorded interaction. Data sources also provides with spatial information, it has main component consists of a spatial DBMS, which stores spatial datasets.

C Aurrecoechea, et.al[23] proposed generalized frame approach for development of QoS architecture in distributed multimedia systems. They considered QoS assurance in distributed system platform. QoS assurance, complete data and media from remote server across the network points of delivery. Generalized frame work QoS architecture for research is to define a set of QoS configuration interfaces that formalize QoS in the end to end system and network. End to end QoS scenario continuous media flows in the applications area of telecommunication, computer communication and standard

communities using packet scheduling and flow control to improve data delivery and packet transmission.

D. Cotroneo M, et.al[23] proposed architecture that supports different classes of service by evaluating middle ware infrastructure, each with different quality attributes(QoS) concerning the network data delivery, availability, performance of network communications including web,video servers on demand services.

Marco Comuzzi, et.al[26] proposed negotiation broker to which both the consumer and the service provider approach by architecture notify the preferences on QoS attributes and negation strategies by specify value parameters. SOA architecture by service with price, availability and data quality as QoS attribute. Service invocation can be enriched by negotiation of the QoS attributes. SOA architecture by three different phases

Service publication provider who publishes as the service on architecture registry. Service selection consumer selects service by searching the service registry. Service invocation selected service has to be invoked by consumer by specified constraint on the QoS attribute a negotiation process between the consumer and the provider. They suggested architecture for QoS Negotiation (Web service Level Agreement Framework i.e WSLA). These approach selects QoS attributes during web services invocation.

M.Adel Serhani et.al[27] proposed a QoS broker- based architecture for web services architecture is to support the client in selecting web services based on user required QoS. The architecture is developed by four main participating roles the web service broker, the web service provider, the client and QoS- enabled UDDI registry services.

### 3 ISSUES AND CHALLENGES

The software developers design quality software which meets the requirements to overcome the problems of end users. Web services are to provide inter-operability across global internet, B2B applications, commerce etc. These software components must be reusable,efficient, cost effective, and must follow QoS parameters. Web applications interacting with server to access services using different methods, architectures, protocols. Heart of the web service is computing registry(service provider). Service provider will publish large number of function descriptions. All these provided by accessing service using HTTP protocols. Al-Masri et.al[10] suggested provided QoS of web services[8,9,12,13]

#### 3.1 Response time

The time taken to send and receive response from web server ( response time is measured in ms.)

$$RT = RCT - URT \quad (1)$$

where RT:Response time, RCT:Response Completion Time, URT:User Request Time

Software developers need to design a web service architecture that will minimize the response time.

#### 3.2 Availability

Number of successful invocations or total invocations are measured in (%).

$$AV = I/TI \quad (2)$$

where AV: Availability, I: Number of Invocations, TI: Total number of invocations

Availability of web services depends on middle ware, infrastructure, network performance and Service oriented architecture. We can improve the availability of service s by SOAP, UDDI[11]

#### 3.3 Throughput

Throughput is defined as average rate of successful messages received by communication channel per second (throughput is measured in invocationsseconds)

$$TP = TI/T \quad (3)$$

where TP:Throughput,TI:Total number of invocations, T:Period of time

Software kernel implementation is a high performance instruction set, which is required to improve throughput[30]

#### 3.4 Successability

Number of response over the number of requests (is measured in (%))

$$SUCC = Nresp/Nreq \quad (4)$$

where SUCC: Successability, Nresp: Number of response, Nreq:Number of requests

Software components are designed to improve successability

#### 3.5 Reliability

Reliability is measured as the number of failures over a period of time. It is derived from the unsuccessful invocations for a given period of measurement. ( reliability is measured in (%))

$$RE = Merror/Mtotal \quad (5)$$

where RE: Reliability, Merror: Error messages,Mtotal: Total Messages. Reliability to be improved by providing high availability of resources

#### 3.6 Compliance

Extends the WSDL document specification.(is measured in in (%)) Software developers provide insufficient information by WSDL file. WSDL is considered by input, output dependency, invocation sequence, hierarchical function description and concurrent specifications. WSDL are the standards,for the description of language written in XML and it is useful to locate the resource of a web service [14,15].

#### 3.7 Best Practices

The extension to which a Web service that follows, functionality, API, WS-I Basic Profile(measured in in (%)). Best practices like Enterprise Application Integration (EAI), Business to Business (B2B) Integration, SOAP etc.

#### 3.8 Latency

Time consumed for the server to process a given request is called latency (measured in ms)[6,10,16].

$$LAT = RT - Treq \quad (6)$$

where LAT: Latency, RT: Response time,Treq: Requested time. Latency performance depends on several implementations of software components like Middle ware technologies like Simple Object Access Protocol (SOAP), Java RMI, CORBA, UDDI, and Network Protocols.

### 3.9 Documentation

Measure of documentation (i.e. description of tags) in Web Service Description Language(WSDL) (measured in (%)).WSDL document is an XML document that defines a web service consuming of methods to invoke/access other web service applications. WSDL document is described among tags, XML specifications, and elements. WSDL gives the location of web services. This improves in the selection of service description[10,17].

### 3.10 WsRF

Web Service Relevancy Function: a rank for Web Service Quality (measured in (%)).This is used to measure quality of web services metrics (i to ix) to rank the web services described by Masri E[10] Service Classification 1. Platinum (High quality) 2. Gold 3. Silver 4. Bronze (Low quality) The categorization is based on the on the whole quality evaluation(rank) provided by Web Service Relevancy Function(WsRF). Using WsRF values attain for each Web services to a particular service group. The classification can be useful to distinguish between different services that offer the same functionality[10].

## 4 PROPOSED MODEL ARCHITECTURE

Web services are used to provide personalization, API customization, and support for multiple languages. They logically separate the layout Model ,View, and Controller in formats such as HTML, XML. Application centric services enable organizations to integrate applications, business process without constraints of proprietary infrastructure, platforms and operating system. These use open standards, including HTTP, XML, SOAP, WSD, UDDI. Application framework and web services. XML based web service architecture which allows programs written in different languages on different platforms to communicate with each other in a standard way as application frame[18,25].

**Functional Requirements:** In software development, functional requirement defines a function (services or operations) of the system and its components, described as set of inputs, performance and outputs.

**Non Functional Requirements:** Non functional requirements are the resources that specify to assess the operation of the system. For example network bandwidth. Adaptability: modify/update in system configurations and resource availability has been recognized as a further important aspect of service quality. Reliability and security issues are important in the design of most of the computer systems.

Performance is measured by quality of service, it was originally defined in terms of responsiveness and computational throughput, availability, but it has been redefined in terms of ability to meet timeliness, guarantees. Availability depends on communication resource that has guaranteed computing and communication resource.

We proposed a new QoS web service architecture by extending 3-tier Architecture adding new component Quality service manager depicted in Figure in 4.

### 4.1 Application Layer

End-users request applications by through this layer. Clients may use devices, web users and application users.

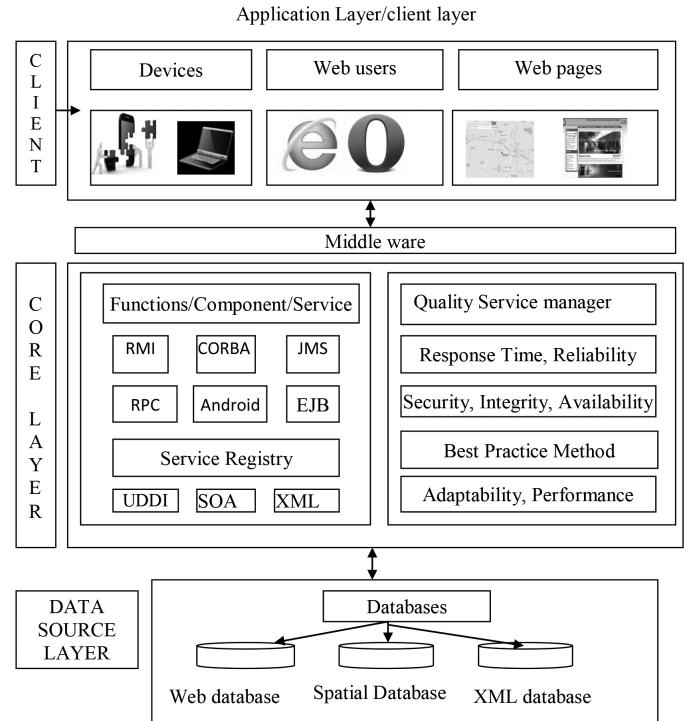


Figure 4. QoS of Web Service Architecture

### 4.2 Middleware

This layer is the core of the QoS architecture. The main concept of this layer to provide business functionality to the user. It has two parts such as Quality service manager, and Service manager.

#### 4.2.1 Quality service manager

Is a software component to select the QoS web service parameters that provide efficient, reliable, and secured services/components

#### 4.2.2 Service manager

Service Manger has various parameters, response time, reliability, availability, security, best practices and performance.

The performance of the services is measured by considering QoS parameters in functional and non functional requirements. To design an efficient Architecture of QoS services by extending three tier architecture[18].We proposed new Algorithm.1 in Quality Service Manager which will improve the QoS web service architecture with QoS web service parameters. We represented in Figure 5 as infrastructural components of multimedia application as a case study, using round robin algorithm in the Web Server interface applications will be improved the performance.

Multimedia applications run in networks of personal computers, they need resources from the work stations running applications(processor,buffer capacity,bus cycles) and network (physical transmission,routers,switches gateways). Workstations may support multimedia stream of data within individual/organization applications.

For multimedia application (non functional requirements, network communication, high bandwidth), and functional

### Algorithm 1 Quality Service Manager

- 1: **procedure** RANKING OF WEB SERVICES
- 2: Input: QWS data parameters related attributes  $D_i = \{A_1, A_2, A_3, A_k, A_i\}$
- 3: output: Component specify their QoS requirement from  $D_i$  to QoS manager to measure Rank of web service or (web service relevancy metrics quality metrics) i.e Service Classification Outcome =  $\{A, B, C, D\}$
- 4: *begin*
- 5: Read the QoS manager parameter  $D_i$  instances
- 6: Apply Classification Procedure to measure the Rank of Web services
- 7: Rank ClassificationMethod( $D_i, O$ )
- 8: If ( tuples will be equal required the rank after step 9 to step 11)
- 9: { Reserve these parameters as resource parameters;
- 10: Application will be proceeded with these parameters;
- 11: Software design notifies QoS manager will be used these parameters }
- 12: else
- 13: { Do not allow the application to be proceed
- 14: Negotiate or application refinement of QoS Manager parameters that will
- 15: Meet the required rank or service classification;
- 16: Give solution to improve the QoS parameters of different type of category;
- 17: To improve the efficiency by computer networks and operating system related parameters or Web services and software related parameters }
- 18: Repeat the steps step 7 to 11 until QWS data  $D_i$  instance (or tuples ) belonging to service classification.
- 19: *end*

Multimedia Service	Data rate (approximate)	Size bits	Size in frequency
Telephone speech	64 kbps	8 bits	8000/sec
CD Quality sound	1.4 Mbps	16 bits	44,000/sec
Standard TV video (uncompressed)	120 Mbps	Upto 640X480 Pixels 16 bits	24/sec
Standard TV video (MPEG-1) compression	1000-3000 mpbs	Variable	24/sec
HDTV video(Uncompressed)	1000-3000 mpbs	Upto 1920X1080 Pixels X 24 bits	24- 60 /sec
HDTV(MPEG-2/MPEG-4 Compression)	6-20 mpbs	Variable	24-60 /sec

Figure 6. Characteristics of Multimedia data stream

Component	In/out	Bandwidth	Latency	Loss rate	Resource required
Web camera	Out	20 frames / second raw video data	--	0	--
Service 1	In	20 frame second, raw video	Inactive	Low	20 ms CPU each 200 ms 20 mbytes RAM
Mixer	In	2X88 Kbps audio	Inactive	Very low	2 ms
	Out	1X88 kbps			2 mbytes RAM
Service 2	In	Various	In active	Low	10 ms CPU each 200 ms
	Out	100 frame s/ second			10 m bytes RAM
Service 3	Network connection	In/out	MPEG-1 stream , approximately 1.5 mbps	In active low	2.5 mbps , low loss stream protocol
Service 4	Network connection	Audio 88 Kbps	In active	Very low	88 kbps, very low stream protocol.

Figure 7. QoS specifications for components of the application shown in Figure.5.

requirements (business services, multitasking shared networks). The central process will perform tasks in the server process, in round robin algorithm scheduling for best effort to put among all the activities running in systems. Networks from different sources network topology, Ethernet, message protocols concepts manage the QoS web service parameters.

The key feature is the resource allocation scheme demanding availability of resources for more quickly completing the task. Round robin approach proposed is shown in Figure.5, best effort method for sharing the processor cycles and network bandwidth. This architecture given late delivery value is less. In order to achieve timely delivery, application need guarantees that the resource will be allocated and scheduled properly.

## 5 DISUSSIONS AND INTERPRETATIONS

From this study, QoS web services architectures, we assume the data as for Figure.6. It gives the characteristics of multimedia data stream the following observations are noted. Telephone speech is less data rate (i.e.bandwidth required or data transfer rate 64 kbps), and time taken to process information frequency size is 8000 seconds. HDTV(MPEG-2/MPEG-4 Compression) multimedia services requires the highest data transfer rate 6-20 Mbps, and time taken to process information frequency size is 24 to 60 seconds. The average case Standard TV video (uncompressed) requires data rate is 120 Mbps, and time taken to process information in frequency is 24 seconds. we proposed a Quality service man-

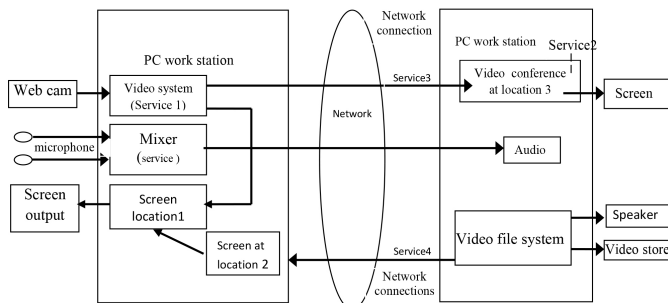


Figure 5. Infrastructural components in multimedia application

ager in the middleware of 3-tier architecture. This architecture suggests Algorithm.1, that gives the QoS parameters, such as response time, reliability and other parameters in the development of web applications.

## 6 CONCLUSION

In this paper we proposed a QoS web service architecture. This is a new model of QoS web service architecture for the rapid development of web services and applications in various domains such as B2B,e-commerce etc., organizations to design and develop the best suitable QoS Web service Architecture to meet the industry demands on various QoS web service parameters, which are reliable, quick services that required for access the web services. In this paper we presented web services fundamental concepts, HTTP,XML,UDDI,SOAP, issues in web services and discussed various architectures and implementation methods. To measure the performance of the services, we proposed a new QoS web service architecture that will meet the demands of organizations and individuals to design and develop QoS web service framework, ranking of web service algorithm which will be applicable for any platform and middleware technologies.

As a case study we have taken multimedia application and infrastructural components in multimedia application and improved quality parameters bandwidth, and access time, throughput by using round robin algorithm in the web server interface applications/services in Figure.7 The new proposed architecture will help the organizations to design and develop solutions using this new QoS service Architecture management and allocation of resource to give guarantees as quality of service management. In future the scope of this paper will extend the work to measure the various QoS web services [10]. Data set QWS from WSDL crawler and other web services data set to measure the performance of web services using various classification methods (i.e data mining techniques).

## 7 References

- [1] Tao Yu,Yue Zhang and Kwei-Jay Lin :Quality of service (qos) in web services: model, architecture and algorithms by University of California, Irvine, CA,ACM, Volume 1 Issue 1, May 2007,Article No. 6,1-26
- [2] [http://en.wikipedia.org/wiki/Serviceoriented\\_architecture](http://en.wikipedia.org/wiki/Serviceoriented_architecture)
- [3] <http://www.webopedia.com/TERM/W/Service.html>
- [4] <http://www.servicearchitecture.com>
- [5] Shrivastava, S.; Sharma, A. : An approach for QoS based fault reconfiguration in service oriented architecture, Information Systems and Computer Networks (ISCON), 2013 International Conference on Year: 2013,Pages: 180 - 184, DOI: 10.1109/ICISCON
- [6] <https://www.oasisopen.org/committees/download.php/10347/wsbpel-specification-draft-120204.htm>
- [7] Marco Crasso,Alejandro Zunino,Marcelo Campo :Easy web service discovery: A query-by-example approach., Elsevier, Science direct 144-163
- [8] <http://www.webperformancetoday.com/2011/06/30/revisiting-the-performance-equation/>
- [9] <http://www.lovelytool.com/>
- [10] QWS dataset <http://www.uoguelph.ca/qmahmoud/qws/>
- [11] <https://www.cs.kent.ac.uk/events/conf/2002/wads/Proceedings/cotroneo.pdf>
- [12] Zibin Zheng, Michael R. Lyu, QoS Management of Web Services, Springer Publications, 2013, 9-37
- [13] <https://www.oasis-open.org>
- [14] Tsai.W.T., Paul. R., Yamin Wang, Chun Fan and Dong Wang: Extending WSDL to facilitate Web services testing 7th IEEE International Symposium on High Assurance Systems Engineering, 2002, 71 - 172, DOI: 10.1109/HASE.2002.11
- [15] <http://www.w3schools.com/webservices>
- [16] Endo.T.,Kaneda.K.Taura.K., Yonezawa. A.High performance LU factorization for non-dedicated clusters IEEE International Symposium on Cluster Computing and the Grid, 2004. CCGrid 2004.678-685,DOI: 10.1109/CCGrid.2004.1336698
- [17] Ben Shil.A,Ben Ahmed.M.: Additional Functionalities to SOAP,WSDL and UDDI for a Better Web Services Administration 2nd. ICTTA '06. Volume: 1, 572-577, DOI: 10.1109/ICTTA.2006.1684434
- [18] George coloulouries, Jean Dollimore, Tim Kindberg and Gordon Blair : Distributed systems concept, 5th Edition,pearson publications
- [19] McArdle, Andrea Ballatore, Ali Tahir, Michela Bertolotto :An open-source web architecture for adaptive location-based services ,the International Archives of the Photogrammetry, RSSIS, Vol. 38, Part II
- [20] M.Swami Das, A. Govardhan, and D.Vijayalakshmi :A survey on web applications, QoS of web services and classification of web services by International conference on Computer and Communication Technologies (IC3T 2K14), CMR Technical Campus,112-116
- [21] <http://www.w3.org/standards/xml>
- [22] Gavin McArdle, Andrea Ballatore, Ali Tahir, Michela Bertolotto :An open-source web architecture for adaptive location-based services. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 38, Part II, IEEE, 296-301
- [23] Cristina Aurrecochea,Andrew T.Cambell, Linda Hauw : A survey of Qos architectures, Springer publications,138-151
- [24] <https://www.cs.kent.ac.uk/events/conf/2002/wads/.../cotroneo.pdf>
- [25] Gunjan Samtani, Dimple Sadhwan :Web Services and Application Frameworks (.NET and J2EE)
- [26] Marco Comuzzi and Barbara Pernici :An Architecture

for Flexible Web Service QoS Negotiation, EDOC Enterprise Computing Conference, 2005, Ninth IEEE international Conference.7079

- [27] M.Adel Serhani, Rachida Dssouli, Abdelhakim Hafid and Houari Sahraoui : A QoS broker based architecture for efficient web services selection, Ninth IEEE International,EDOC Enterprise Computing Conference, vol.1. 2005.113 - 120
- [28] <http://www.redbooks.ibm.com/redbooks/pdfs/sg246303.pdf>
- [29] [http://en.wikipedia.org/wiki/Multitier\\_architecture](http://en.wikipedia.org/wiki/Multitier_architecture)
- [30] Montgomery.D.Akoglu.A.: Methodology and Toolset for ASIP Design and Development Targeting Cryptography-Based Applications IEEE International Conference on ASAP.2007.,365 - 370, DOI: 10.1109/ASAP.2007.4459291



# Implementation of Attribute-Based Encryption for Secure data sharing in Cloud

O.L.Chandana<sup>1</sup>, Ajay. S<sup>2</sup>

<sup>1</sup>PG student, Department of CSE, Malla Reddy Engineering College, India.

<sup>2</sup>Associate professor, Department of CSE, Malla Reddy Engineering College, India.

*Abstract: Sharing of private medical records is associate rising patient central model of health data exchange, that's sometimes outsourced to store at the third party, like cloud suppliers. The confidentiality of the medical records is major downside once patients use industrial cloud servers to store their medical records as the results of it ought to be scan by everybody, to assure the patient's management over access to their own medical records; it's a promising methodology to cipher the files before outsourcing and access management got to be implemented although cryptography rather than role-based totally access management. There unit varied numerous problems like risks of privacy exposure, quality in key management, versatile access and economical user revocation, have remained the foremost vital challenges toward achieving fine-grained, cryptographically implemented information access management. To grasp fine-grained and scalable information access management for medical records detain semi-sure servers, we have a bent to leverage attribute-based totally secret writing (ABE) techniques to cipher every patient's story file. Throughout this paper, we have a bent to elucidate a replacement approach that allows secure storage and controlled sharing of patient's health information. We have a bent to explore key-policy attribute primarily based mostly secret writing and multi-authority attribute-based secret writing to enforce patient access management policy such everybody will transfer the information ,but solely authorize user will scan the medical records. This project along supports multiple*

*Owner eventualities and divides the users among the system into multiple security domains that greatly reduce the key management quality for home house owners and users. A high degree of patient privacy is secure by exploiting multi-authority ABE. Throughout this paper we have a bent to presents the detailed style of modules and implementation Packages of the planned framework.*

*Keywords: Multi-Authority Attribute-based Encryption, Key-Policy Attribute-based Encryption, Secure Sharing.*

## I. INTRODUCTION:

A PHR service permits a patient to make, manage and management her personal health info in one place through the web that has created the storage, retrieval, and sharing of the medical information extra economical. Especially, each patient is secure the whole management of her medical records and should share her health info with outsized degree vary of users, still as attention suppliers, relations or friends. Because of the high price of building and maintaining specialized info centers, many PHR services unit of measuring outsourced to or provided by third-party service suppliers, for example, Microsoft HealthVault<sup>1</sup>. Recently, architectures of storing PHRs in cloud computing unit of measuring projected in. While it's exciting to possess convenient PHR services for everybody; there a unit many



security and privacy risks which are able to impede its wide adoption. The foremost concern is regarding whether or not the patients might extraordinarily management the sharing of their sensitive personal health information (PHI), significantly once they are stick with it a third-party server that folk won't entirely trust. On the one hand, although there exist attention rules like HIPAA that's recently amended to incorporate business associates, cloud suppliers unit typically not coated entities. On the alternative hand, because of the high price of the sensitive personal health information, the third-party storage server's unit typically the targets of assorted malicious behaviors which may end in Expo- positive of the letter. Throughout this paper, we have a bent to endeavor to review the patient- central, secure sharing of PHRs stick with it semi-trusted servers, and ponder addressing the exhausting and exhausting key management issues. Therefore on defend the personal health info stick with it a semi-trusted server, we have a bent to adopt attribute-based committal to writing (ABE) as a results of the foremost committal to writing primitive. Using ABE, access policies unit expressed supported the attributes of users or info, that allows a patient to selectively share her PHR among a bunch of users by encrypting the file below a bunch of attributes, whereas not the requirement to grasp a whole list of users. The complexities per committal to writing, key generation and writing unit entirely linear with the number of attributes involved. However, to integrate ABE into a large-scale PHR system, necessary issues like key management quantifiability, dynamic policy updates, and economical on-demand revocation unit of measuring non-trivial to unravel, and keep for the

foremost half open up-to-date. To the current end, we have a bent to make succeeding main contributions: (1) we have a bent to propose a singular ABE-based framework for patient-centric secure sharing of PHRs in cloud computing environments, below the multi-owner settings. To address the key management challenges, we have a bent to conceptually divide the users among the system into a pair of sorts of domains, significantly public and personal domains. To boot, the framework enforces write access management, handles dynamic policy updates, and provides break-glass access to PHRs below emergence eventualities. (2) Among the possession, we have a bent to use multi-authority ABE (MA-ABE) to bolster the protection and avoid key instrument balk. Each attribute authority (AA) in it governs a disjoint set of user role attributes, whereas none of them alone is prepared to manage the protection of the whole system. We have a bent to propose mechanisms for key distribution and committal to writing thus as those PHR homeowners can specify bespoke fine-grained role-based access policies throughout file committal to writing. Among the non-public domain, homeowners directly assign access privileges for personal users and code a PHR file below its info attributes.

## II.METHODOLOGY

### *A.PHR Owner*

The main aim of our agenda is to provide economical key management and secure patient-centric PHR access at an equivalent time. in line with the many users' knowledge access wishes we have a tendency to tend to tend to separate the system into multiple



security domains (i.e., personal domains (PSDs) and public domains (PUDs)). The PUDs embody users World Health Organization build access supported their virtuoso roles, like nurses, doctors and medical researchers. The ultimate public domains unit generally mapped to associate separate sector among the society, as Associate in the Nursing example the government, health care or insurance sector. For every PSD, PSD users unit related to associate knowledge owner (such as shut friends or family members), that they build accesses to private health records supported the access rights appointed by the owner.

Every knowledge owner (example, patient) is that the sure authority of his/her own personal domains, World Health Organization uses a KPABE system to administer the access rights and secret keys of users in his/her PSD. But the PHR owner is aware of the users within the flesh, to acknowledge patient-centric access, on a individual basis the owner is at the most effective position to grant user access privileges. for private domains, knowledge attributes unit created public that unit determined the essential properties of the private health record knowledge, as Associate in Nursing example the type of a PHR file. For the operate of PSD access, every PHR file is termed with its knowledge attributes, as a results of the key size is believed with a amount of file classes a user will access. But the amount of users throughout a non-public domain is little, it reduces the load for the owner. Whereas encrypting the information for PSD, the owner ought to understand is that the essential knowledge properties.

#### ***B. Cloud Server***

In this paper, we consider the server to be the semi trusted, i.e., honest but curious. The server will attempt to detect as much secret information in the stored PersonaHealth Record files as possible, but they will directly follow the protocol in general. Alternatively, some users will try to access the files beyond their rights. For example, the pharmacy wants to obtain the prescriptions of patients for boosting and marketing its profits. To do so, they may join with the server or even either the other users. Additionally, we assume every user in our system is preloaded with a private public key pair, and entity authentication can be prepared by traditional challenge response protocols.

#### ***C. Characteristic-based Access Policy***

In our system, there are multiple owners, multiple SDs, multiple users and multiple AAs. Additionally, two ABE systems are involved. We name the users having written and read access as data readers and contributors respectively. We use Attribute-based encryption algorithm for it.

#### ***D. Data confidentiality***

Attribute-based Encryption- encrypted PHR files are upload to the server by the owners. Every owner's PHR file is ciphered both under a certain role-based and fine-grained access policy for users from the public domain to access and under a chosen group of data attributes that allows access from users in the personal domain. The PHR files can be decrypt by the authoritative users, excluding the server.

#### ***E. Sequence Diagram***



Attribute-based encryption. A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes function with one another and in order. It is a kind of a Message Sequence Chart. Sequence diagrams are also called as event scenarios, event diagrams and timing diagrams.

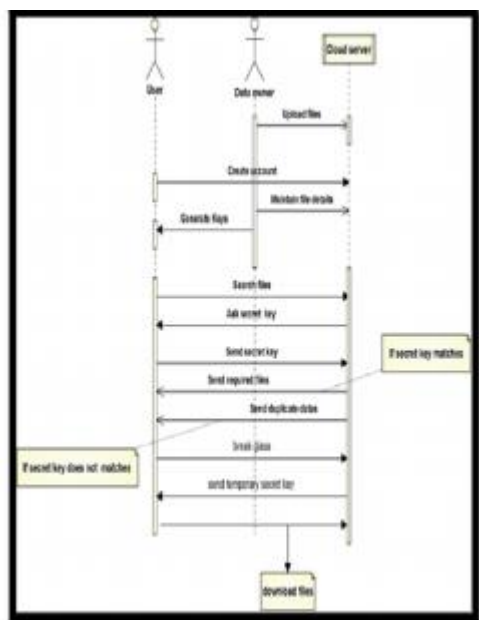


Fig 1 shows the sequence diagram for PHR using

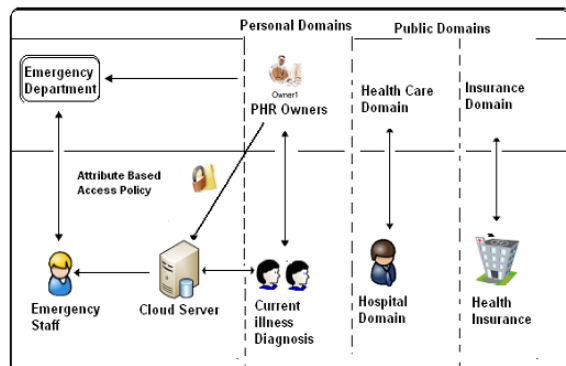
### III. PROPOSE FRAMEWORK

The main goal is to produce secure patient-centric PHR access and economical key management at the same time. In Fig. one the key found out is to divide the system into multiple security domains (namely, public domains (PUDs) and private domains (PSDs)) according to the absolutely fully totally different user's data access wishes. The PUDs contains users administrative unit turn out

access supported their skilled roles, like doctors, nurses and medical researchers. In observing, a course will be mapped to qualification freelance sector in the society, such as the health care, government or insurance sector. for every PSD, its users unit within the flesh associated with a data owner (such as relations or shut friends), that they turn out accesses to PHRs supported access rights assigned by the owner In each variety of security domains has to utilize ABE to know cryptographically enforced, patient-centric PHR access. Especially, in a course multi-authority ABE is used, during this there unit multiple—attribute authorities! (AAs), every governing a disjoint set of attributes. Role attributes unit written for PUDs, representing the skilled role or obligations of a course user. Users in PUDs acquire their attribute-based secret keys from the AAs, whereas not directly interacting with the homeowners. To management access from course users, homeowners unit absolve to specify role-based fine-grained access policies for her PHR files, whereas do not have to be compelled to understand the list of approved users once doing cryptography. Since the PUDs contain the bulk of users, it greatly reduces the key management overhead for each the homeowners and users. Every data owner (e.g., patient) may perhaps be a dependable authority of their own PSD, administrative unit uses qualification ABE system to manage the key keys and access rights of users in her PSD. Since the users unit within the flesh notable by the PHR owner, to know patient-centric access, the owner is at the sole position to grant user access privileges on an individual basis. For PSD, data attributes unit printed



that sit down with the intrinsic properties of the PHR data,



Such as the category of a PHR file. For the purpose of PSD access, each PHR file is labeled with its data attributes, while the key size is only linear with the number of file categories a user can access. Since the number of users in a PSD is often small, it reduces the burden for the owner. When encrypting the data for PSD, all that the owner needs to know is the intrinsic data properties. The Fig.2 represents the attribute hierarchy of files, leaf nodes are atomic file categories while internal nodes are compound categories.

#### IV. CONCLUSION & FUTURE WORK

A framework of secure sharing of non-public health records has been projected throughout this paper. Public and private access models unit of measuring designed with security and privacy enabled mechanism. The framework addresses the distinctive challenges brought by multiple PHR house owners and users; therein the complexness of key management is greatly reduced. The attribute-based committal to writing model is increased to support operations with MAABE. The system is improved to support dynamic policy management model. Thus,

Personal Health Records unit of measuring maintained with security and privacy. As future study, it's going to be fascinating to boost the HSN with a 3rd party auditor to verify the cloud server that stores and methodology the PHRs polymorphic Split key committal to writing will become extra improvement to verify the trustiness of the TPA.

#### REFERENCES:

- [1] Melissa Chase, "Multi-authority Attribute-based Encryption", TCC, volume 4392 of LNCS, pages 515–534, Springer, 2007.
- [2]X. Liang, R. Lu, X. Lin, and X. S. Shen, "Cipher text policy attribute-based encryption with efficient revocation," Technical Report, University of Waterloo, 2010.
- [3]S. Yu, C. Wang, K. Ren, and W. Lou, "Achieving secure, scalable, and fine-grained data access control in cloud computing," in IEEE INFOCOM'10, 2010.
- [4] V. Goyal, O. Pandey, A. Sahai, and B. Waters, "Attribute-based encryption for fine-grained access control of encrypted data," in CCS '06, 2006, pp. 89–98.
- [5] M. Li, W. Lou, and K. Ren, "Data security and privacy in wireless body area networks," IEEE Wireless Communications Magazine, Feb.2010.
- [6] A. Boldyreva, V. Goyal, and V. Kumar, "Identity-based encryption with efficient revocation," in ACM CCS, ser. CCS '08, 2008, pp.417–426.
- [7]. S. D. C. di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, and P. Samarati, "Over-encryption: management of access control evolution on outsourced data," in VLDB '07, 2007, pp. 123–134.
- [8]. A. Sahai and B. Waters. Fuzzy identity-based encryption. Advances in Cryptology {EUROCRYPT 2005, pages 457{473, 2005.
- [9]. Angelo De Caro and Vincenzo Iovino, "jPBC: Java Pairing-Based Cryptography" Computers and Communications (ISCC), 2011 IEEE Symposium on Digital Object Identifier: 10.1109/ISCC.2011.5983948 Publication Year: 2011 , Page(s): 850 – 855

# Big Data Technology Predictions and Trends

**Md. Rafeeq**

M.Tech (PhD), Assoc. Prof,  
CMRTC Kandlakoya-Hyderabad.

**Rajeshwar Rao Kodipaka**

Asst. Prof., Dept. of CSE,  
MREC Secunderabad.

**Dr. C. Sunil Kumar**

Professor, Dept. of ECM,  
SNIST Ghatkesar, Hyderabad

**Abstract** – In 2014 saw big data initiatives inside the enterprise increasingly move from test to production. In 2015, big data will push further into the enterprise with even more use cases specifically real-time use cases. We live in a world of data (both at a personal and professional level) and people express themselves through the work they do with it. Inside each organization, users want to be actively engaged with their data; however, they haven't yet had the technology to do so. In 2015, by having access to Business Intelligence (BI) solutions that allow true self-service, users will move from passively consuming data to actively using it to glean important information. Data is not only big but it is also fast. Data is rapidly being generated from mobile devices, factory sensors, wearables, retail transactions, on-line advertisements and many other sources. The insights from that fast, Big Data disappear as rapidly as they appear so taking action in "business time" is critical. Business time is the speed at which a business must service its customers, and it ranges from sub-seconds (real-time) to minutes and hours.

**Keywords** – Data Agility, Big Data Analytics, Hadoop-As-a-Service, The Ubiquity of YARN.

## I. INTRODUCTION

**Data Agility Emerges as a Top Focus**, Data agility has been one of the big drivers behind the development of big data technologies, as the processes around legacy databases and data warehouses have proven too slow and inflexible for many business needs. In 2015, Schroeder says data agility will become even more central as organization shift their focus from simply capturing and managing data to actively using it.[2] "The reality is that the tools are still emerging, and the promise of the [Hadoop] platform is not at the level it needs to be for business to rely on it," says Loconzolo. But the disciplines of big data and analytics are evolving so quickly that businesses need to wade in or risk being left behind. "In the past, emerging technologies might have taken years to mature," he says. "Now people iterate and drive solutions in a matter of months — or weeks." So what are the top emerging technologies and trends that should be on your watch list — or in your test lab? Computer world asked IT leaders, consultants and industry analysts to weigh in. Here's their list. With big data, analysts have not only more data to work with, but also the processing power to handle large numbers of records with many attributes, Hopkins says[1].

Traditional machine learning uses statistical analysis based on a sample of a total data set. "You now have the ability to do very large numbers of records and very large numbers of attributes per record" and that increases

predictability, he says. The combination of big data and compute power also lets analysts explore new behavioral data throughout the day, such as websites visited or location[3]. Hopkins calls that "sparse data," because to find something of interest you must wade through a lot of data that doesn't matter. "Trying to use traditional machine-learning algorithms against this type of data was computationally impossible. Now we can bring cheap computational power to the problem," he says. "You formulate problems completely differently when speed and memory cease being critical issues," Abbott says. "Now you can find which variables are best analytically by thrusting huge computing resources at the problem. It really is a game changer." [4].

"To enable real-time analysis and predictive modeling out of the same Hadoop core, that's where the interest is for us," says Loconzolo. The problem has been speed, with Hadoop taking up to 20 times longer to get questions answered than did more established technologies. So Intuit is testing Apache Spark, a large-scale data processing engine, and its associated SQL query tool, Spark SQL. "Spark has this fast interactive query as well as graph services and streaming capabilities.[4] .

## II. ORGANIZATIONS MOVE FROM DATA LAKES TO PROCESSING DATA PLATFORMS

In some ways, 2014 was the year of the data lake (or data hub), an object-based storage repository that stores raw data in its native format - whether structured, unstructured or semi-structured - until it's ready for use. Data lakes have a strong value proposition in that they represent a scalable infrastructure that's economically attractive (with a reduced per-terabyte cost) and extremely agile[6].

Schroeder says that the data lake will continue to evolve in 2015 with the capability to bring multiple compute and execution engines to the data lake to process the data in-place. That's not only more efficient, it creates a single point of governance and a single point of security[7].

**C-Level Executives Seeing Big Results From Big Data**, A study by Accenture finds that many organizations are putting big data initiatives into production and the vast majority of those that do are pleased with the results.

Between February and early April 2014, Accenture screened more than 4,300 CIOs, COOs, CMOs, CFOs, chief data officers (CDOs), chief analytics officers (CAOs) and other senior technology, data and analytics leaders from companies in 19 countries and across seven industries. Of that number, 36 percent said their company



Figure 1: Big Data Lake

had not completed and weren't currently pursuing a big data installation and four percent said their organizations were in the process of implementing their first big data project. Of the remaining 2,600 respondents that had applied big data to their businesses, Accenture focused on a sample of 1,007 executives.

"Businesses are at a transition point where instead of just talking about the potential results that can be achieved from big data, they are realizing actual benefits including increasing revenues, a growing base of loyal customers and more efficient operations," adds Narendra Mulani, senior managing director of Accenture Analytics. "They're recognizing that big data is one of the cornerstones of digital transformation." "Today, even the most basic items like water pipes can generate and provide data," adds Mulani. "While the Internet of Things is giving rise to massive sources and quantities of data, new big data technologies are emerging that help uncover crucial business insights from the data. Companies not implementing big data solutions are missing an opportunity to turn their data into an asset that drives business and a competitive advantage."

"In 2015, data lakes will evolve as organizations move from batch to real-time processing and integrate file-based, Hadoop and database engines into their large-scale processing platforms," he says. "In other words, it's not about large-scale storage in a data lake to support bigger queries and reports; the big trend in 2015 will be around the continuous access and processing of events and data in real time to gain constant awareness and take immediate action"[6].

### III. HADOOP VENDOR CONSOLIDATION: NEW BUSINESS MODELS EVOLVE.

In early 2013, Intel made a splash with the Introduction of its own Hadoop distribution, saying that it would differentiate itself by taking a ground-up approach in

which Hadoop was baked directly into its silicon. But just a year later, Intel ditched its distribution and threw its weight behind Hadoop distribution vendor Cloudera instead. At the time, Intel noted that customers were sitting on the sidelines to see how the Hadoop market would shake out. The number of Hadoop options were muddying the waters. Schroeder believes Hadoop vendor consolidation will continue in 2015 as the also-rans discontinue their distributions and focus elsewhere in the stack[5].



Fig.2. Hadoop vendor consolidation

"Hadoop is early in the technology maturity lifecycle with only 10 years passing since the seminal MapReduce white papers were published by Google," he adds. "Hadoop adoption globally and at scale is far beyond any other data platform just 10 years after initial concept. Hadoop is in the innovation phase, so vendors mistakenly adopting "Red Hat for Hadoop" strategies are already exiting the market, most notably Intel and soon EMC Pivotal"[6].

Schroeder believes 2015 will see the evolution of a new, more nuanced model of OSS that combines deep innovation with community development. "The open source community is paramount for establishing standards and consensus," he says. "Competition is the accelerant transforming Hadoop from what started as a batch analytics processor to a full-featured data platform." *The lines between IT and marketing are blurring in the age of digital marketing. For that reason, it's more important than ever that CIOs and CMO communicate and effectively.* Yet a strong partnership between IT and marketing is crucial if big data and analytics are to succeed. CMOs may increasingly take the lead on big data projects, but CIOs are the key to implementing, maintaining and scaling these solutions. "Marketing is the driver of the big data car," says Todd Merry, CMO of global hospitality and food service company Delaware North, "but it doesn't go anywhere without IT.[4]" Mainstream.

Previously, IT would be required to establish centralized data structure "Hadoop has made the enterprise comfortable with structure-on-read for some use cases. Advanced organizations will move to data bindings on execution and away from a central structure to fulfill

ongoing requirements. This self-service speeds organizations in their ability to leverage new data sources and respond to opportunities and threats."



Fig.3. Self-Service Big Data Goes

#### IV. ENTERPRISE ARCHITECTS SEPARATE THE BIG HYPE FROM BIG DATA

2015 will see enterprise architects take center stage as their improving understanding of the Hadoop technology stack leads to a better defined and more sophisticated statement of requirements for big data applications, including elements like high availability and business continuity. "As organizations move quickly beyond experimentation to serious adoption in the data center, enterprise architects move front and center into the big data adoption path," Schroeder says[8]. "IT leaders will be vital in determining the underlying architectures required to meet SLAs, deliver high availability, business continuity and meet mission-critical needs.

In 2014 the booming ecosystem around Hadoop was celebrated with a proliferation of applications, tools and components. In 2015 the market will concentrate on the differences across platforms and the architecture required to integrate Hadoop into the data center and deliver business results.

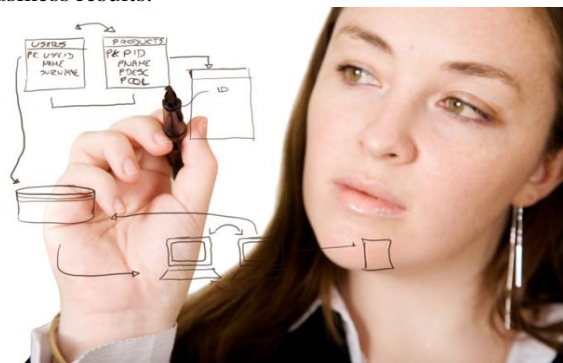


Fig. Enterprise Architecture

Big data technologies have evolved at a torrid pace that shows every sign of continuing in 2015. MapR CEO and co-founder John Schroeder predicts five major developments will dominate big data technology in the new Year[4].

#### More Distributions, More Competition:

The large enterprise vendors such as HP, Oracle, SAP, Software AG, and Tibco will create their own Hadoop distribution, as IBM and Pivotal already have. Specialized public and private cloud distributions also will emerged, generating plenty of competition and customer choice[7].

As the presence of the Internet of Things (IoT) — such as connected devices, sensors and smart machines — grows, the ability of things to generate new types of real-time information and to actively participate in an industry's value stream will also grow. Essentially, things become agents for themselves, for people and for businesses. Think of the car that alerts emergency services and an insurance company or the smart thermostat that schedules service. The added connectivity, communications and intelligence of things will make many of them agents for services that are currently requested and delivered via human intervention [5].

#### V. SECURITY AND MY DATA-DRIVEN PREDICTIONS

Security is a major concern with big data. To make more sense from the big data, organizations would need to start integrating parts of their sensitive data into the bigger data. To do this, companies would need to start establishing security policies which are self-configurable: these policies must leverage existing trust relationships, and promote data and resource sharing within the organizations, while ensuring that data analytics are optimized and not limited because of such policies.

What good is a prediction if its not based on data? When I look at our own internal data and the larger big data market in 2014, there are a few interesting trends that emerged that I believe are strong indicators for what will come in 2015. For an in-depth explanation of 2014 trends and my 2015 predictions. Big Data will be a business initiative. Historically, what we saw was that IT tried to service business, and built the infrastructure for them, but its clear now that business users themselves want to get their hands on big data. It's no longer that business sends IT off to find analytical results, the line of business wants to do that themselves.

There is a really strong demand for the democratization of data. Hadoop deployments will shift from being centralized in IT toward departmental deployments. We also saw that central, IT-driven deployments are shifting to department specific deployments[5].

#### VI. BIG DATA DEEP DIVE CONCLUSION: THE FUTURE OF ANALYTICS

As data volumes have grown, and as the complexity of data that is collected and analyzed has increased, new novel software architectures have emerged. What I'm most excited is the fact that software is open-source, and

---

3rd National Conference on Research Trends in Computer Science & Technology NCRTCST-2015

---

we're playing a key role in driving where that software is going. And what I'm most excited about on top of that is the commoditization of that software – I'm tired of talking about the container in which you put your data," Hammerbacher said. "No one should be required to invest tremendous amounts of money in their container anymore. They should be identifying novel data sources, algorithms to manipulate that data and the smartest people for using that data. Big Data technologies are changing all that: mathematical analysis makes it possible to weed out bad info in a highly efficient manner.



**Dr. C. Sunil Kumar**

Working as Professor in ECM ,SNIST- Hyd, having 15 years teaching and industry experience his domain is Data mining.

**ACKNOWLEDGEMENT**

I would like to extend thanks to the Director Dr. A. Raji Reddy, Chairman Mr. C. Gopal Reddy, CMR Technical Campus for encouragement of research work. and I thank to Dr. C. Sunil Kumar Professor, for his valuable suggestions and supervision.

**REFERENCES**

- [1] VinayakBorkar, Michael J. Carey, Chen Li, Inside "Big Data Management":Ogres, Onions, or Parfaits?, EDBT/ICDT 2012 Joint Conference Berlin, Germany,2012 ACM 2012, pp 3-14.
- [2] Jefry Dean and Sanjay Ghemwat, .MapReduce: Simplified data processing on large clusters, Communications of the ACM, Volume 51 pp. [2]Brad Brown, Michael Chui, and James Manyika, Are you ready for the era of „big data“?, Mc Kinsey Quaterly, Mckinsey Global
- [3] DunrenChe, MejdSafran, and ZhiyongPeng, From Big Data to Big Data Mining: Challenges, Issues, and Opportunities, DASFAA Workshops 2013, LNCS 7827, pp. 1-5, 2013.
- [4] [www.cio.com/.../big-data/5-big-data-technology-predictions-for-2015](http://www.cio.com/.../big-data/5-big-data-technology-predictions-for-2015)
- [5] Hadoop,“ Powered by Hadoop ”<http://wiki.apache.org/hadoop/PoweredBy>
- [6] PIG Tutorial, Yahoo Inc. <http://developer.yahoo.com/hadoop/tutorial/pigtutorial.html>
- [7] Apache: Apache Hadoop, <http://hadoop.apache.org> [8]amazon simple storage service (Amazon S3) <http://aws.amazon.com/s3/>

**AUTHOR'S PROFILE**



**Md. Rafeeq**

Doing external PhD in JNTUH- Hyd, Telangana, India. This work will be supported in doing research work as well as understanding the various predictions for big data. I would like thank to my guide Dr. Sunil Kumar, and co-guide Dr. Subhash Chandra for constant support for Understanding the subject.



**Rajeshwarrao. Kodipaka**

(ISTE, CSI Life Member), working as Asst.Prof in CSE, MREC since 1 year, Having 7 years teaching experience and interested domain is Data mining, cloud computing, Computer Networks.

# IMPLEMENTATION OF A ROBUST AUTHENTICATED APPROACH FOR CLOUD SERVICES

K.Rajeshwar Rao<sup>1</sup>, Mr.Amarendhar Reddy<sup>2</sup>

<sup>1</sup>Asst.Professor, CSE Dept, MREC(A) ,Secunderabad, Telangana, India

<sup>2</sup>Asst.Professor, CSE Dept, BVRIT(A), Narsapur, Telangana, India.

## ABSTRACT:

For the past few years, the technology of cloud computing has developed from a proficient business perception to one of the rapid increasing segments of infrastructure sector. The information on individuals as well as companies that are positioned within the cloud has been increased progressively however the issues will begin to rise concerning the safety of the environment. Access control is important in cloud system since it only approved users to have access towards valid service. As the single key distribution centre is not easy to continue due to several users that are maintained in cloud setting hence we focus that clouds must consider a decentralized approach while distribution of secret keys towards users. We make a study of an access control method which is decentralized and novel for protecting data in cloud storage that manages unspecified authentication process. It is rather common for clouds to contain multiple key distribution centres in various locations. Our proposed approach supports user revocation and prevents replay attacks and moreover manages in creation, alteration, and reading of the cloud data. It moreover will support authentication of privacy preserving, which is not managed by other techniques. Our authentication technique is decentralized as well as strong, when contrasting from several methods of access control that are intended for clouds which are centralized.

***Keywords: Cloud computing, Decentralized approach, Key distribution centre, Access control, Authentication, Privacy preserving, Centralized.***

## 1. INTRODUCTION:

Cloud computing which is an emerging technology assures to reduce the requirement for maintaining of high-priced computing devices. Hence cloud concept

ensures the owners regarding benefits of economy of scale and, simultaneously lessens the operating expenses for several services. In the cloud platform, data owner store up their data in the environment of cloud for benefitting from high quality applications [1]. Owner of the data will

send their data towards data centre and make use of the service that is provided by cloud service provider who will handle client data. One of major issues considered in the data storage of cloud system is verification of data integrity at entrusted servers. Privacy and security preserving in cloud system is an extremely important issue and this issue has been explored by numerous researchers. The cloud will make the user answerable for data which it has outsourced, and similarly cloud is responsible for the services it has offered. There is verification for user validity who store up the data. Privacy of user is an important requirement in order that the cloud users do not identify the user identity. For provision of effective and flexible data storage, the data requires encryption but the data is regularly modified and this property should be considered for designing the effective methods of secure storage [2][3]. Cloud accountability task concerns technical problems as well as law enforcement. We propose an access control method which is decentralized and novel for protecting data in cloud storage that manages unspecified authentication process. Our proposed method is decentralized as well as strong, contrasting from several methods of access control that are intended for clouds which are centralized. In the proposed decentralized approach, cloud proves the

series reliability devoid of recognizing of user identity prior to data storage. The proposed access control method has the additional trait of access control where proper users will mainly decrypt the stored up data.

## **2. METHODOLOGY:**

Consider an instance where a student, wants to reveal the information regarding some offense that was made by an authorities of a university to its entire professors, and higher authorities. She wants to stay on unidentified during the process of revealing the proof for offense and stores the necessary data within the cloud system. We consider access control as an important issue in this situation, in order that the users who are authorized can have permission to that data. It is in addition essential for verification of the information that whether it has come from a trustworthy source. The problems regarding the mechanism of access control, protection of confidentiality and authentication have to be worked out at the same time. Most of the data is being stored up in the cloud environment, and it is of sensitive information. Access control mechanism permits only approved users to hold permission towards valid service. The works that were made in earlier works regarding access control are for the most

part centralized in nature. Access control is usually applicable in health care services where clouds are used for storage of sensitive information regarding patients. Most of the research works has accepted a centralized approach in which only one key distribution centre will distribute the secret keys towards the users. But the single key distribution centre is not only a failure but it is not easy to continue due to numerous users that are maintained in cloud setting. Hence we focus that clouds must consider a decentralized approach while distribution of secret keys towards users and moreover it is quite normal for clouds to contain multiple key distribution centres in various locations. We propose an access control method which is decentralized and novel for protecting data in cloud storage that manages unspecified authentication process. Our proposed technique is tough for replay attacks, where a user can restore new data with the outdated one, although it no longer contains a convincing claim policy [4]. This is an essential asset since a user, who is revoked of attributes, might be not able to write to cloud. Our proposed method allows several times writing which was not allowed in the earlier efforts. The cloud proves the series reliability devoid of recognizing of user identity prior to data storage in the decentralized approach. Moreover it has the additional trait of

access control where proper users will mainly decrypt the stored up data. Our proposed authentication method is decentralized as well as strong, contrasting from several methods of access control that are intended for clouds which are centralized. Our proposed method will support authentication of privacy preserving, which is not managed by other techniques. Our method supports user revocation and prevents replay attacks and moreover manages in creation, alteration, and reading of the cloud data. In the proposed decentralized approach, there can be multiple key distribution centres for key management. The access control along with authentication is collusion resistant, where two users cannot collude and have access to data.

### **3. AN OVERVIEW OF PROPOSED SYSTEM:**

Protecting privacy and security in cloud system is an extremely important issue and this issue has been explored by numerous researchers. User privacy is an important requirement in order that the cloud users do not identify the user identity. It is important to control data access for the purpose of allowing of only authorized users to access the data. Access control permits only approved users to hold permission towards valid service and it is

of three different types such as attribute-based, user-based and role-based access control. In the user based method, list of access control contains the users who are permitted to access the data. Attribute-based method is more wide-ranging where the users are offered with specified attributes and only the users containing valid attributes, that assures the access policy, will have access to the data [5]. In our work we make several assumptions regarding our approach such as: Cloud is honest-but-curious, indication that administrators of the cloud can view the data of user but cannot change it. Honest-but-curious representations of adversary do not interfere with the data with the intention that they want to maintain the system functioning generally and stay on unnoticed. Users can contain permission to a read or else write operation or both of these operations accesses towards a file that are stored up within the cloud. We propose an access control method which is decentralized and novel for protecting data in cloud storage that manages unspecified authentication process. Our authentication method is decentralized as well as strong, different from several methods of access control that are intended for clouds which are centralized. Our authenticated method of access control will support authentication of privacy preserving, which is not managed by other techniques.

In our scheme a user have an ability to create a file and store up it safely in the cloud system. Our proposed system makes usage of the two protocols such as attribute-based signature as well as attribute-based encryption. Attribute-based encryption was introduced by Sahai and Waters in which a user contains set of attributes besides its unique identity. Attribute-based signature is a novel procedure that was proposed by Maji et al. in which users contain a claim predicate that is related to a message. The claim predicate will helps in identification of user as an approved one, devoid of disclosing its identity. The protocols of attribute-based signature as well as attribute-based encryption are combined to attain authentic access control devoid of exposing the user identity towards cloud. In our protocol as shown in fig1, three users such as User X, User Y, and User Z are present. User X will obtain a token from the honest trustee. On provision of the identity of User X, the trustee will present him a token. In the proposed decentralized approach, there can be multiple key distribution centres for key management. User X on provision of the token towards key distribution centres will obtain the keys for encryption or decryption process as well as signing. There is an encryption of message under access policy that decides the person

regarding the accession towards the data that is stored within the cloud system [6][7]. User X will make a decision on a claim policy, to confirm her accuracy and sign the message in this claim. The cipher-text through a signature is forward to cloud that makes sure signature and store up the cipher-text. When a reader desires to read, then the cloud system sends the cipher-text. When the user contains attributes that match up with the access policy, it will decrypt and retrieves its original message [8].

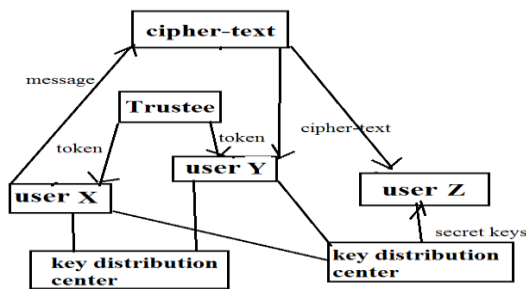


Fig1: Our proposed protocol.

#### 4. CONCLUSION:

The knowledge of cloud computing is a popular choice for renting of storage services. This expertise will allow for efficient computing by means of centralizing memory, processing and storage. The earlier works made in literature regarding access control are for the most part centralized in nature in which only one key distribution centre will distribute the secret keys towards the users. As it is not simple to continue due

to numerous users that are maintained in cloud setting, we mainly focus that clouds must consider a decentralized approach while distribution of secret keys towards users. We have introduced an access control method which is decentralized and novel for protecting data in cloud storage that manages unspecified authentication process. Our projected method supports authentication of privacy preserving, which is not managed by other techniques. Attribute-based signature in addition to attribute-based encryption is combined to attain authentic access control devoid of exposing the user identity towards cloud. Our proposed authentication technique is decentralized in which there can be multiple key distribution centres for key management as well as strong, when contrasting from several methods of access control that are intended for clouds which are centralized. The proposed method has the additional trait of access control where proper users will mainly decrypt the stored up data and supports user revocation and prevents replay attacks and manages in creation, alteration, and reading of the cloud data.

#### REFERENCES

- [1] S. Jahid, P. Mittal, and N. Borisov, "EASiER: Encryption-Based Access Control in Social Networks with Efficient Revocation," Proc. ACM Symp. Information, Computer and Comm. Security (ASIACCS), 2011.

[2] R.L. Rivest, A. Shamir, and Y. Tauman, "How to Leak a Secret," Proc. Seventh Int'l Conf. Theory and Application of Cryptology and Information Security (ASIACRYPT), pp. 552-565, 2001.

[3] X. Boyen, "Mesh Signatures," Proc. 26th Ann. Int'l Conf. Advances in Cryptology (EUROCRYPT), pp. 210-227, 2007.

[4] H.K. Maji, M. Prabhakaran, and M. Rosulek, "Attribute-Based Signatures: Achieving Attribute-Privacy and Collusion-Resistance," IACR Cryptology ePrint Archive, 2008.

[5] H.K. Maji, M. Prabhakaran, and M. Rosulek, "Attribute-Based Signatures," Topics in Cryptology - CT-RSA, vol. 6558, pp. 376-392, 2011.

[6] A. Beimel, "Secure Schemes for Secret Sharing and Key Distribution," PhD thesis, Technion, Haifa, 1996.

[7] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-Policy Attribute-Based Encryption," Proc. IEEE Symp. Security and Privacy, pp. 321-334, 2007.

[8] X. Liang, Z. Cao, H. Lin, and D. Xing, "Provably Secure and Efficient Bounded Ciphertext Policy Attribute Based Encryption," Proc. ACM Symp. Information, Computer and Comm. Security (ASIACCS), pp 343-352, 2009.

# **A Refuge Business model for Cloud Computing using novel AES Approach**

**Ch.Ramesh Babu**

Assoc Prof, MREC(A),

[chramesh522@gmail.com](mailto:chramesh522@gmail.com)

## **Abstract:**

Enterprises usually store data in internal storage and install firewalls to protect against intruders to access the data. They also standardize data access procedures to prevent insiders to disclose the information without permission. In cloud computing, the data will be stored in storage provided by service providers. Service providers must have a viable way to protect their clients' data, especially to prevent the data from disclosure by unauthorized insiders. Storing the data in encrypted form is a common method of information privacy protection. If a cloud system is responsible for both tasks on storage and encryption/decryption of data, the system administrators may simultaneously obtain encrypted data and decryption keys. This allows them to access information without authorization and thus poses a risk to information privacy. This study proposes a business model for cloud computing based on the concept of separating the encryption and decryption service from the storage service. Furthermore, the party responsible for the data storage system must not store data in plaintext, and the party responsible for data encryption and decryption must

delete all data upon the computation on encryption or decryption is complete.

A CRM (Customer Relationship Management) service is described in this paper as an example to illustrate the proposed business model. The exemplary service utilizes three cloud systems, including an encryption and decryption system, a storage system, and a CRM application system.

One service provider operates the encryption and decryption system while other providers operate the storage and application systems, according to the core concept of the proposed business model.

This paper further includes suggestions for a multi-party Service-Level Agreement (SLA) suitable for use in the proposed business model.

## **1. Introduction**

Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network (typically the Internet). The name comes from the use of a cloud-shaped symbol as an abstraction for the complex infrastructure it contains in system diagrams. Cloud computing entrusts remote services with a user's data, software and computation. The word Cloud is used as a metaphor for the internet. End users access

cloud-based applications through a web browser or light-weight desktop or mobile app.

With cloud computing, companies can scale up to massive capacities in an instant without having to invest in new infrastructure. Cloud computing is benefit to small and medium-sized businesses. Basically consumers use what they need on the Internet and pay only for what they use. So we say it as pay-as-you go model of business.

Cloud computing eliminates the costs and complexity of buying, configuring, and managing the hardware and software needed to build and deploy applications, these applications are delivered as a service over the Internet (the cloud).

Cloud computing is an where an application doesn't access resources it requires directly, rather it access them through something like a service. So instead of talking to a specific hard drive for storage, and a specific cpu for computation etc., it talks to some service that provides these resources.

The service then maps any requests for resources to its physical resources, inorder to provide for the application.

Cloud computing incorporates infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS) as well as Web 2.0

## **2. Types of services:**

### **Software as a service:**

Software as a service(saas) deliver software as a service over the internet, eliminating the

need to install and run the application on the customer's own computers and simplifying maintenance and support.

This is what almost everyone has already used in the form of gmail, yahoo, wordpress.com(hosting this blog), the rest of google apps, the various search engines, Wikipedia, etc., some company hosts an application in the internet that many users sign-up for and use with out any concern about where, how, by whom the compute cycles and storage bits are provided. So in our application any user can sign-up from anywhere and use the application. Here the cloud provides software as a service.

Advantages of saas are:

1.Activities that are managed from central locations rather than at each customer's site, enabling customers to access applications remotely via the web.

2.centralized feature updating, which obviates the need for downloadable patches and upgrades.

### **Platforms as a service:**

Platform as a service(paas) provides operating systems, development frameworks, applications. This is the newest entry where an application platform is offered to developers in the cloud. Developers write their application to a more or less open specification and then upload their code into the cloud where the app is run magically somewhere, typically being able to scale up automatically as usage for the app grows

Examples are google app engine, force.com

### **Infrastructure as a service:**

The whole cloud infrastructure viz. servers, routers, hardware based load-balancing, firewalls storage & other network equipments are provided by the Infrastructure as a service(IaaS) provider. This is the base layer of the cloud stack. It serves as a foundation for the other two layers, for their execution. The keyword behind this stack is virtualization. The application will be executed on a virtual computer(instance). You have the choice of virtual computer, where you can select a configuration of cpu, memory & storage that is optimal for your application.

#### **4. User data privacy concerns in a cloud computing environment**

In a cloud computing environment, the equipment used for business operations can be leased from a single service provider along with the application, and the related business data can be stored on equipment provided by the same service provider. This type of arrangement can help a company save on hardware and software infrastructure costs, but storing the company's data on the service provider's equipment raises the possibility that important business information may be improperly disclosed to others.

Some researchers have suggested that userdata stored on aserviceprovider's equipment must be encrypted .Encrypting data prior to storage is a common method of data protection, and service providers may be able to build firewalls to ensure that the decryption keys associated with encrypted user data are not disclosed to outsiders. However, if the decryption key and the encrypted data are held by the same service provider, it raises the possibility that high-level administrators within the service provider would have access to both the decryption key and the encrypted data, thus presenting a risk for the unauthorized disclosure of the user data.

#### **4.1 Existing methods for protecting data stored in a cloud environment**

Common methods for protecting user data include encryption prior to storage, user authentication procedures prior to storage or retrieval, and building secure channels for data transmission. These protection methods normally require cryptography algorithms and digital signature techniques, as explained below:

#### **5. AES Encryption algorithm**

The Advanced Encryption Standard (AES) is a specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001. Originally called Rijndael, the cipher was developed by two Belgian cryptographers, Joan Daemen and Vincent Rijmen, who submitted to the AES selection process.

AES has been adopted by the U.S. government and is now used worldwide. It supersedes the Data Encryption Standard (DES), which was published in 1977. The algorithm described by AES is a symmetric-key algorithm, meaning the same key is used for both encrypting and decrypting the data.

AES is based on a design principle known as a substitution-permutation network, and is fast in both software and hardware. Unlike its predecessor DES, AES does not use a Feistel network. AES is a variant of Rijndael which has a fixed block size of 128 bits, and a key size of 128, 192, or 256 bits. By contrast, the Rijndael

specification *per se* is specified with block and key sizes that may be any multiple of 32

## 5.1 High-level description of the algorithm

**KeyExpansion**—round keys are derived from the cipher key using Rijndael's key schedule

### 1. Initial Round

1. AddRoundKey—each byte of the state is combined with the round key using bitwise xor

### 2. Rounds

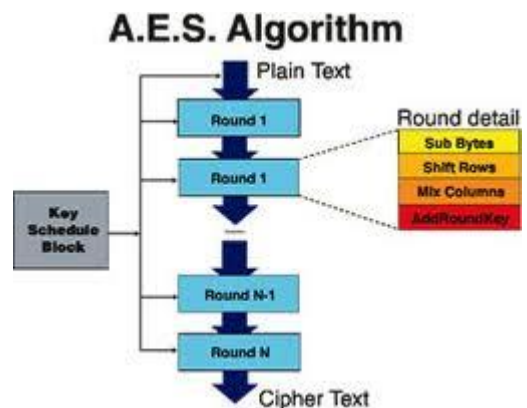
1. SubBytes—a non-linear substitution step where each byte is replaced with another according to a lookup table.
2. ShiftRows—a transposition step where each row of the state is shifted cyclically a certain number of steps.
3. MixColumns—a mixing operation which operates on the columns of the state, combining the four bytes in each column.
4. AddRoundKey

### 3. Final Round (no MixColumns)

1. SubBytes
2. ShiftRows
3. AddRoundKey

In the AddRoundKey step, the subkey is combined with the state. For each round,

a subkey is derived from the main key using Rijndael's key schedule each subkey is the same size as the state. The subkey is added by combining each byte of the state with the corresponding byte of the subkey using bitwise XOR.



The use of passwords as an authentication process is more familiar to general users, but messages sent by the user are vulnerable to surreptitious recording by hackers who can then use the data in the message to log into the service as the user.

In more advanced authentication systems, the system side will generate a random number to send the user a challenge message, requesting the user to transmit an encrypted response message in reply to the challenge message, thus authenticating that the user has the correct encryption key. Without this key, the user will not be allowed access. In the process of challenge and response, the client's encrypted key uses the client's password to convert a derived value and. In this program, each communication between the client and server is unique, and a hacker using an old message would fail to access the system. In addition, the One-Time Password (OTP) authentication system differs from most peoples' conception of a password. Most people understand a password to be a password chosen by the user to be

meaningful, and can be used again and again. The emphasis of OTP, however is the single-use nature of the password.

After receiving authentication from the user, the system side must create a secure transmission channel to exchange information with the user. The Secure Sockets Layer (SSL) is a common method of building secure channels, primarily using RSA encryption to transmit the secret keys needed for the both sides to encrypt and decrypt data transmitted between them.

## **6. A REFUGE BUSINESS MODEL FOR CLOUD COMPUTING BASED ON NOVEL AES APPROACH.**

### *A. Core Concepts*

This study proposes a Business Model for Cloud Computing Based on a Separate Encryption and Decryption Service using AES. The concept is based on separating the storage and encryption/decryption of user data, as shown in Fig. In this business model, Encryption/Decryption as a Service and Storage as a Service (SaaS) are not provided by a single operator. In addition, the SaaS provider may not store unencrypted user data and, once the provider of Encryption/Decryption as a Service has finished encrypting the user data and handed it off to an application (e.g. a CRM system), the encryption/decryption system must delete all encrypted and decrypted user data.

Figure Encryption/Decryption as an independent service

The concept of dividing authority is often applied in business management. For example, responsibility for a company's finances is divided between the accountant and cashier. In business operations, the accountant is responsible for keeping accounts, while the cashier is responsible for making payments. By keeping these two functions separate, the company can prevent the accountant from falsifying accounts

and embezzling corporate funds.

In a cloud computing environment, the user normally uses cloud services with specific functions, e.g., Salesforce.com's

CRM service, SAP's ERP services, etc. Data generated while using these services is then stored on storage facilities on the cloud service. This study emphasizes the addition of an independent encryption/decryption cloud service to this type of business model, with the result that two service providers split responsibility for data storage and data encryption/decryption.

To illustrate the concept of our proposed business model,

Fig. presents an example in which the user uses separate cloud services for CRM, storage and encryption/decryption.

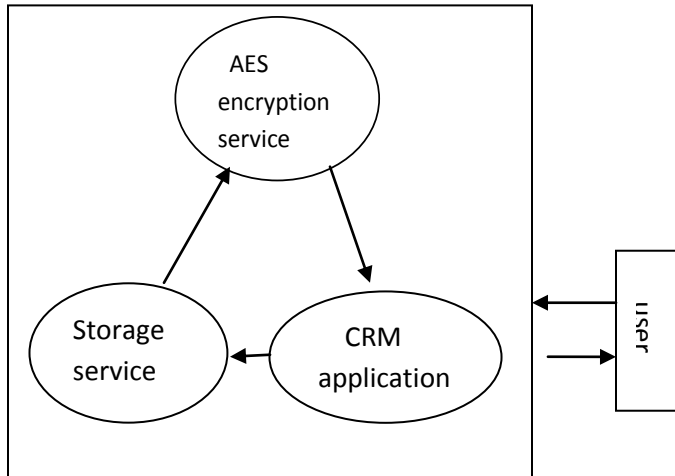
Business model concept integrating separate cloud services for data encryption/decryption, CRM and storage.

Prior to the emergence of an emphasis on the independence of encryption/decryption services, CRM, ERP and other cloud services would simultaneously provide their users with storage services. This study emphasizes that Encryption/Decryption Cloud Services must be provided independently by a separate provider.

## **7. Operating examples of the AES Encryption/Decryption as a Separate Cloud Service Business Model**

This section presents a CRM application service as an example of the new business model.

After the user logs into the CRM system, if the CRM Service System requires any client information, it will execute a Data Retrieval Program. When this data needs to be saved, it will execute a Data Storage Program. The Data Retrieval Program is illustrated in and is explained below.



Data retrieval diagram

When a user wants to access the CRM Cloud Service, he must first execute the Login Program as shown in Step 1. This step can use current e-commerce or other services which have already securely verified the user's registration, such as symmetric key-based challenge and reply login verification, or through a One-Time Password.

After the user's login has been successfully verified, if the CRM Service System requires client information from the user, it sends a request for information to the Storage Service System, as shown in Step 2. In this step, the CRM Service System transmits the user ID to the Storage Service System where it searches for the user's data. This data is encrypted so, once found, a request must be sent to the Encryption/Decryption Service System along with the user ID. Step 3 shows the Storage Service System executing the transmission of encrypted client data and the user ID to the Encryption/Decryption Service System.

Since the Encryption/Decryption Service System can serve multiple users and the encryption/decryption for each user's

data requires a different key, therefore each user's unique ID and keys are stored together. Therefore, in Step 4, the Encryption/Decryption Service System uses the received user ID to index the user's data decryption key, which is then used to decrypt the received data. Using the correct decryption key to decrypt the data is critical to restoring the data to its original state.

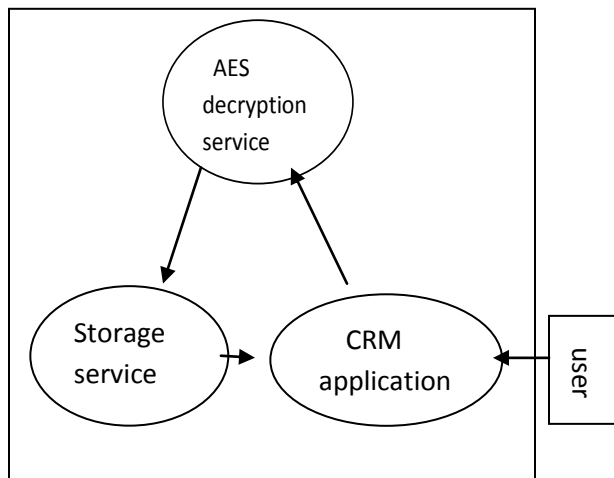
After the Encryption/Decryption Service System has decrypted the client's data, in Step 5 the decrypted client data is provided to the CRM Service System which then displays the client data to the user in Step 6, completing the Data Retrieval Program. Prior to sending the decrypted client data, the Encryption/Decryption Service System and the CRM Service System can establish a secure data transmission channel (e.g., a Secure Sockets Layer connection) to securely transmit the decrypted client data. After the decrypted client data is sent, the Encryption/Decryption Service System is not allowed to retain the decrypted data and any unencrypted data must be deleted to prevent the encrypted data and the decryption key from being stored in the same system. This is a critical factor in ensuring the privacy of user data.

The above-mentioned Data Retrieval Program requires the collaboration of three different cloud service systems.

we describe the Data Storage Program, as shown in Fig.

This program also involves the collaboration of three cloud service systems: CRM Service System, AES Encryption/Decryption Service System, and Storage Service System.

Data storage diagram



Step 1 of Fig. 6 shows the client sending a Data Storage Request to the CRM Service System which then initiates the Data Storage Program, requesting data encryption from the Encryption/Decryption Service System as shown in Step 2. In Step 2, the CRM Service System and the Encryption/Decryption Service System establish a secure data transfer channel to transmit the user ID and the data requiring storage from the CRM Service System to the Encryption/Decryption Service System. As the encryption of data from different users requires different keys, in Step 3 the Encryption/Decryption Service System initiates data encryption, which involves using the received user ID to index the user's encryption key which is then used to encrypt the received data. Following this study's emphasis on the principle of divided authority, once the client data is encrypted by the Encryption/Decryption Service System it must be transferred to the Storage Service System where the user ID and encrypted data are stored together.

Therefore, when the Encryption/Decryption Service System executes Step 4, it must transfer the user ID and encrypted client data to the Storage Service System. Step 5 shows the Storage Service System receiving the user ID paired with the data for storage. In this business model, the following the completion of Step 4 at the Encryption/Decryption Service System, all unencrypted and decrypted user data must be deleted.

Step 6, the final step of the Data Storage Program, transmits a Data Storage Complete message from the Storage Service System to the CRM Service System, at which point the CRM Service System may confirm that the client data has been stored. If it doesn't receive a Data Storage Complete message.

## REFERENCES

- [1] A. Weiss, "Computing in the clouds", netWorker, vol. 11, no. 4, pp. 16-25, December 2007.
- [2] C. S. Yeo, S. Venugopal, X. Chu, and R. Buyya, "Autonomic metered pricing for a utility computing service", Future Generation Computer Systems, vol. 26, issue 8, pp. 1368-1380, October 2010.
- [3] M. Baker, R. Buyya, and D. Laforenza, "Grids and grid technologies for wide-area distributed computing," International Journal of Software: Practice and Experience, vol.32, pp. 1437-1466, 2002.
- [4] B. R. Kandukuri, V, R. Paturi and A. Rakshit, "Cloud security issues," in Proceedings of the 2009 IEEE International Conference on Services Computing, pp. 517-520, September 2009.
- [5] R. Sterritt, "Autonomic computing," Innovations in Systems and Software Engineering, vol. 1, no. 1, Springer, pp. 79-88. 2005.
- [6] R. Buyya, CS. Yeo, S. Venugopal, J. Broberg, and I.

Brandic, “Cloud computing and emerging IT platforms: vision.

# Textured Based Combined Approach for Plant Leaf Classification

<sup>1</sup>Anil Kumar.Muthevi, <sup>2</sup>Dr U Ravi Babu

<sup>1</sup>Research Scholar, Acharya Nagarjuna University, Assoc. Professor-SSAIST, Surampalem, A.P, India

<sup>2</sup>Professor, Dept. of CSE, Malla Reddy Engg College(Autonomous),T.S, India

[lettertoanil@gmail.com](mailto:lettertoanil@gmail.com)

[uppu.ravibabu@gmail.com](mailto:uppu.ravibabu@gmail.com)

**Abstract**—Although there is no strict definition of the image texture, texture classification has broad applications in the field of agriculture and medicine. As plants are vitally important for environmental protection, it is more important to identify from the images to classify them. Even this approach is applied on plant leaves it could use in different applications and classify them correctly. So far, there have been so many methods proposed to solve this problem. Plant leaf classification is a technique where leaf is classified based on its different morphological features. The purpose of this article is to introduce a novel method which explains the process on the images of textures based on Local Binary pattern and Gray Level Co-occurrence Matrix and finally, extracting the statistical features from the images. Finally apply some classified techniques and compare results with other methods.

**Index Terms**— Texture classification, Local Binary pattern, Gray Level Co-occurrence Matrix, statistical features

## I. INTRODUCTION

Plants play a vital role in protecting earth ecology and environment by maintaining a healthy atmosphere and providing substance and shelter to many insect and animal species [1][5][6]. In addition, plants have plenty of use in foodstuff, botanical and many herbal industries [2][3][4]. Texture classification is the process to classify different textures from the given images. Since 90s with the progress of the computer vision and advancements in image processing, texture classification has a close relationship with the sciences such as machine learning, it functions in areas like leaf classification, pattern recognition, tracking of the objects, to find out skin diseases defect detection, face recognition etc. The main two phases in texture classification process are first, gathering and extracting the best features for texture description. Second is selecting corresponding types of classifiers. The extraction of relevant features from the given image is called feature extraction methods. The Local Binary Pattern (LBP), Gray Level Co-Occurrence Matrices (GLCM) are the two popular extraction techniques.

The remaining part of this article is organized as follows. Section 2 is related to the description of LBP algorithm and the way of process. Section 3 is explained with procedure of GLCM algorithm. Section 4 deals with the proposed method and finally results, conclusions & future work are given in section 5 and 6 respectively.

## II. LOCAL BINARY PATTERN (LBP)

The original Local Binary Pattern invented by Ojala and Pietikainen back in 90's. It plays an important role in texture analysis process [8]. The LBP calculates a value that reflects the relationship within a 3x3 neighborhood that is multiplied with the corresponding binomial values. The 8 neighbor values are placed using binary numbers either 0 or 1 by comparing with the central pixel. If the observed gray value is lesser than the gray value of the central pixel, then it is taken as 0 else it is 1.

$$d_i = 0 \text{ if } I(x_i, y_i) < I(x_0, y_0)$$

$$= 1 \text{ otherwise}$$

Where,  $d_i$  represents binary value,  $d_i$  represents the original pixel value at position  $i$  and  $d_0$  is the central pixel value. As each of LBP has one of two possible values (0 or 1), with the combination of all the eight elements results in  $256(2^8)$  possible Local binary patterns ranging from 0 to 255 (These are also called as texture units) to the corresponding pixels. The obtained value is then multiplied with weights.

We can obtain the LBP by totaling the obtained values  $L_{LBP} = \sum_{i=1}^8 d_i 2^{i-1}$

The Figure 1 shows an example on how to compute LBP. The original 3x3 neighborhood is given in Fig 1(a). The central pixel value is considered as a threshold value in order to assign a binary value to its neighbors. Fig 1(b) shows the result of thresholding the 3x3 neighborhood

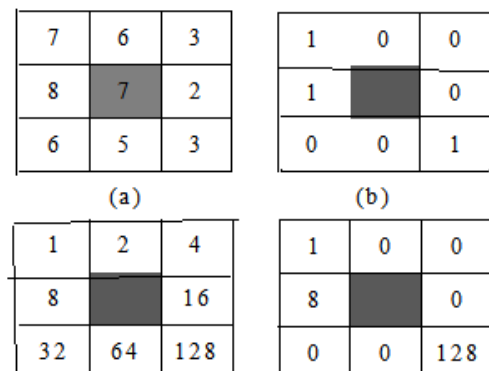


Fig: 1 LBP approach

$$LBP = 1 + 8 + 128 = 137$$



The obtained values given in the Fig1(c) the result is given in Fig1 (d). The sum of the resulting values gives the LBP measure (137). In this case the central pixel (7) replaced by the obtained value (137). Now, a new LBP image is constructed by processing each pixel and its 3x3 neighbors in the original image. The LBP approach is used to represent a significant breakthrough in texture analysis, outperforming earlier methods in many applications. LBP has tolerance against changing of intensity values. And it is computationally very simple approach, which makes possible to analyze images in challenging real-time environment. The concept of LBP also used in applications such as face recognition and age classification.

### III. GRAY LEVEL CO-OCCURRENCE MATRICES (GLCM)

Grey Level Co-occurrence Matrices (GLCM) is one of the statistical methods which is used to determining parameters for texture analysis. It is an efficient and widely used feature extraction method for texture classification which is proposed in [9] by Haralick and Shanmugam. GLCM is one of the most important techniques to extract the texture information. The extracted texture features are called Haralick Features [9]. The GLCM of an  $N \times M$  image is a two dimensional matrix. The elements of the matrix are joint occurrence of intensity levels  $X$  and  $Y$  at a certain distance  $d$  and an angle  $\Theta$ . Hence many occurrences of matrices are possible. For Ex Fig 2 describes how to compute the GLCM. The template (1, 1) is present four times in the image. Hence the entry for (1,1) is 4. There no combination for (2,2) in the image and hence the entry for (2,2) is 0. The max gray scale value in the given image is 3. Therefore the co-occurrence matrix is 3x3. Each entry of the co-occurrence matrix indicates the presence of the template in the given image. For ex in Fig 2, there is no combination of (3,3) together. Therefore its value is zero. On the other hand (1,1) is present four times in the image. So the entry for (1,1) in the co-occurrence matrix is 4. In this manner the values of co-occurrence matrix can be obtained. The features are then extracted from the co-occurrence matrices. Some of the important features are Angular second moment, contrast, correlation, Entropy, Inverse difference moment.

1	3	2
1	1	1
1	1	1

Fig: 2 (a)

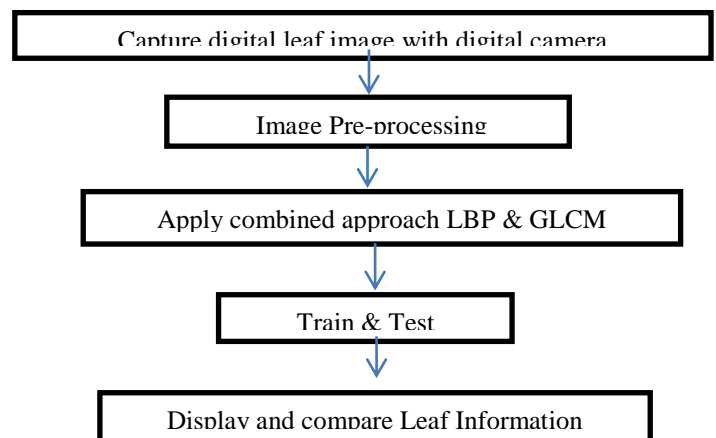
	1	2	3
1	4	0	1
2	0	0	0
3	0	1	0

Fig: 2(b)

Coming to the practical approach in the most of the applications combining the GLCM with other approaches

### IV. PROPOSED METHOD

The approach which is proposed is



The preliminary step for any plant leaf classification is image acquisition which consists of collecting a suitable leaf and then, taking the digital image of leaf by using a digital camera, called as an input image [11, 12].

In the next step, this image is pre-processed to enhance the required features. This step may sometimes include conversion in to suitable formats like gray scale or binary format, and applied by the process of image smoothing. The purpose of image pre-processing is to improve image data so that it can reduce unnecessary distortions and enhances the image features that are relevant for further processing [11,12]. Color image of leaf is converted to gray scale image because there may be changes in environmental changes and season causes the color feature having lower reliability. Thus, it is better to work with gray scale image.

### V. RESULTS

In this paper we introduced a new method for leaf classification. Thus to get results and observe the efficiency of proposed approach 3 kinds of leaves, 60 images were taken by digital camera. That is 20 images of every model of leaves. The images after pre-processing submitted to the proposed combined method. To be simplified process and dataset's dimensions only GLCM



was estimated in the 0'. Then there has been a dataset made consisting 60 instances, and 7 attributes.

Finally, by using some of the classifiers such as KNN, Naïve Bayes, and LAD Tree, and by using N. Fold method, the accuracy of the dataset has been made for leaf textures is computed, which is shown in the table 1. In the 2nd and 3rd rows of the table, the leaf textures were classified just based on LBP and GLCM images to compare the combined method with some of the earlier schemes in terms of accuracy. As it is shown in the, accuracy of the proposed combined approach is higher than earlier methods. The main advantages of the proposed approach specified in this paper can be mentioned in two points first one is, corresponding with near of all classifiers, and the second one is introducing a series of new features which are able to be computed for different applications.

<b>Classifier(→) Approach(↓)</b>	<b>3NN</b>	<b>5NN</b>	<b>Naive Bayes</b>	<b>LAD Tree</b>
<b>Combined method</b>	<b>90.3</b>	<b>87.7</b>	<b>89.6</b>	<b>85.8</b>
<b>LBP</b>	<b>88.4</b>	<b>82.6</b>	<b>87.5</b>	<b>81.4</b>
<b>GLCM</b>	<b>86.3</b>	<b>85.2</b>	<b>80.2</b>	<b>79.8</b>

## VI. CONCLUSION & FUTURE WORK

The main purpose of this article was to offer an approach for leaf classification. In this respect, in the previous sections have algorithms of LBP, GLCM for the images and combined approach has been offered and described thoroughly. As it was described in combined method first input images were pre-processed then combined method applied as shown in the figure, then the it submitted to Train & Testing phases finally by using some classifiers we have classified all the items. so we see that all features which have computed by the combined approach can use for classifying all kinds of leafs by accuracy.

In future, we are going to apply statistical pattern recognition methods which take noise into consideration. Our system has an advantage of its ability of classifying and recognizing the plant from a small part of the leaf without depending neither on the shape of the leaf or on its color features, since the system essentially depends on the textural features. Hence, the system is useful for the botany researchers when he wants to recognize a damaged plant, since this can be carried out depending only on a small part of the damaged plant. When the botany researcher has a damaged plant and wants to put it into its position in the classification, our system is the solution because it depends on the textural features not on the color features which is naturally changing during the seasonal succession.

The Authors would like to say thanks to the management, principal and the staff for their support and encouragement.

## REFERENCES

- [1] Jyotisma Chaki and Ranjan Parekh, "Designing an Automated System for Plant Leaf Recognition", International Journal of Advances in Engineering & Technology, Vol 2, Issue 1, Jan 2012, pp. 149-158.
- [2] N. Valliammal and Dr. S.N. Geethalakshmi, "Analysis of the Classification Techniques for Plant Identification through Leaf Recognition" Ciit International Journal of IJCSN International Journal of Computer Science and Network, Vol 2, Issue 2, 2013 ISSN (Online): 2277-5420
- [3] M. Z. Rashad, B. S. el-Desouky, and Manal S. Khawasik, "Plants Images Classification Based on Textural Features using Combined Classifier", International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No. 4, August 2011, pp. 93-100.
- [4] Abdul Kadir, Lukito Edi Nugroho, and Paulus Insap Santosa, "Leaf classification using shape, color, and texture", International Journal of Computer Trends & Technology (IJCTT), July-August 2011, pp. 225-230.
- [5] C. S. Sumathi and A. V. Senthil Kumar, "Edge and Texture Fusion for Plant Leaf Classification", International Journal of Computer Science and Telecommunications, Vol 3, Issue 6, June 2012, pp. 6-9
- [6] T. Beghin, J. S. Cope, P. Remagnino, & S. Barman, "Shape and texture based plant leaf classification", Advanced Concepts for Intelligent Vision Systems (ACVIS), Vol 6475, 2010, pp. 45-353.
- [7] C. H. Arun, W. R. Sam Emmanuel, and D. Christopher Durairaj, "Texture Feature Extraction for Identification of Medicinal Plants and Comparison of Different Classifiers", International Journal of Computer Applications (0975-8887), Vol 62, No. 12, January 2013, pp. 1-9.
- [8] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", Pattern Recognition 29(1), 1996, pp 51-59.
- [9] R.M. Haralik, K. Shanmugam and I. Dinstein, "Textural Features for Image Classification", IEEE Trans On Systems, Man and Cybernetics, 1973, pp 610-621.
- [10] Moulay, A. Akhlouki, Xavier Madague and Wael Ben Larbi "A New Color-Texture Approach for Industrial Products Inspection", Journal of multimedia, Vol 3, No 3, JULY 2008.
- [11] T. Beghin, J. S. Cope, P. Remagnino, & S. Barman, "Shape and texture based plant leaf classification", Advanced Concepts for Intelligent Vision Systems (ACVIS), Vol 6475, 2010, pp. 45-353.
- [12] Jyotisma Chaki and Ranjan Parekh, "Designing an Automated System for Plant Leaf Recognition", International Journal of Advances in Engineering & Technology, Vol 2, Issue 1, Jan 2012, pp. 149-158.

**Anil Kumar. Muthevi** obtained his M.Sc Computer Science (CS) from D.N.R College, Andhra University in the year 2002 and M.Tech Degree from J.N.T.U.K Kakinada in the year 2010. He pursuing Ph.D from Acharya Nagarjuna University, Guntur in Computer Science & Engineering under guidance of Dr. U. Ravi Babu. He has published research papers in various National, International conferences, proceedings. He is working as Associate professor in Sri Sai Aditya Institute Of Science & Technology from Oct 2011 to till date.

**Ravi Babu. Uppu** received his M.Tech Degree from RV Deemed University in the year 2003. He received his Ph.D degree from Acharya Nagarjuna University, Guntur in the year 2014 in Computer Science & Engineering. He is IBM Certified Database Associate. He has published more than 20 research papers in various National, International conferences, proceedings. His Research interests includes Image processing, Digital Water marking, Pattern recognition, Stenography and Image retrieval. He is currently working as Professor in Department of Computer Science and Engineering, Malla Reddy Engineering College.

## ACKNOWLEDGMENT

# Prediction of the post-operative life expectancy using Data Mining technique

Janaki Sharan Pahariya  
Department of Information  
Technology,  
Rustamji Institute of  
Technology(RJIT),  
BSF Academy, Tekanpur,  
Gwalior, India  
talktopahariya@gmail.com

Anil Kumar Dudyala<sup>\*</sup>  
Department of Computer Science  
and Engineering,  
National Institute of Technology  
Patna (NIT Patna),  
Patna, India.  
anilkumard001@gmail.com

Sabhapathy Myakal  
Department of Information  
Technology,  
Malla Reddy Engineering College,  
Maisammaguda, Dhulapally,  
Secunderabad, India  
sabha.delight@gmail.com

**Abstract**— Prediction of the post-operative life expectancy is one of the crucial aspects which every healthcare organization has to do with. Effective prediction of life expectancy may save life of patient. In the urge to provide an efficient technique for prediction of life expectancy, we have used Thoracic Surgery Data. The data set used is a classification data of the post-operative life expectancy of lung cancer patients. Different Data mining techniques using hold-out and ten-fold cross validation methods have been applied exhaustively on the Thoracic Surgery Data to find out the best intelligent technique for post-operative life expectancy prediction. It has been found from the conducted experiments that the results obtained by the Random Forest have given the superior results compared to other models.

**Keywords**— *Decision Tree (DT), Support Vector Machine (SVM), Logistic Regression(LR), Random Forest(RF), Data Balancing.*

## I. INTRODUCTION

Life is a precious gift given by god through which all the living beings enjoy in this world. Human-being being intelligent enough to identify various sources and reasons for the abrupt ending of life, strives to protect and prolong the life span of that person who has been predicted as the non-survival due to various reasons. Proceeding in the similar lines, we tried to provide a prediction system which could tell accurately well in advance about the survival of post-operative life of a patient. The data used for training the prediction algorithm cannot be used directly as it is unbalanced. Normally, raw data gathered from any sources cannot be used directly for the training process due to various reasons like, some attributes having missing values[1], sequential nature of delivering data [2], or disproportions in class distribution [3]. In the literature, the third issue is known as imbalanced data problem. In general, every dataset with unequal class distribution can be considered as imbalanced dataset[4]. The disproportion between classes effects decision making capability of a classifier while training with typical methods by biasing toward majority class. When a classifier is trained using traditional methods and is classifying all objects only to the majority class, it is deemed as extreme imbalance data problem. Applying techniques which deal with the problem of

unbalanced data is essential for learning decision models with high predictive accuracy [5]. The problem of imbalanced data is widely observed in medical decision making, particularly in post-operative risk evaluation domain. The misclassification related to treating deceases as survivals is much more troublesome than the decision mistake made in opposite direction. The second important issue connected to the problem of post-operative life expectancy prediction is the interpretability of the decision model. Using so-called “black box” models in the considered application is not recommended. It is mainly caused by the patient's fear about being treated by machines with hidden and difficult to understand inference process and distrust among doctors of being supported by vaguely-working models. To dissipate the doubts it is necessary to propose a method to extract knowledge in the form of decision rules or trees from “black box” models.

The objective of this paper is to evaluate the capability of different intelligence techniques in predicting of life expectancy in the context of Thoracic Surgery Data. The compared models are one statistical classifier technique: Logistic Regression (LR), two tree structured classifier technique: Decision Trees (DT), Random Forest (RF), one Neural Network: Multilayer Perceptron (MLP) and Support Vector Machine (SVM).

The rest of this paper is organized as follows, In section 2 Literature survey done in the field of prediction is discussed. Section 3 overviews the data description and data preparation, the techniques applied in this paper is discussed in section 4, section 5 presents the experiment methodology, section 6 describes about results and discussions and finally, section 7 gives the conclusion of the paper.

## II. LITERATURE SURVEY

Lot of work has been done on imbalanced data. Few of the related work are mentioned in this literature survey. During the experiments it has been observed that the classifiers tend towards the majority class if they are trained with imbalanced data. To deal with this various techniques like undersampling,

oversampling[6, 7] and SMOTE (Synthetic Minority Over-sampling Technique) are used. SMOTE is one of the popular balancing technique[8], that generates artificial samples lying on the path connecting two neighbors from minority class. Borderline SMOTE is an extension of SMOTE, which incorporates in the sampling procedure only the minority data points located in the neighborhood of the separating hyper plane [9]. On the other hand, non-informative majority examples can be eliminated to obtain more balanced data in the under sampling procedure. The group of methods solves the problem of uneven data directly during training phase, and it includes ensemble-based approaches which make use of under and over sampling methods to construct diverse and balanced base classifiers. The most common approach in this group is SMOTEBoost [10] that combines the benefits of boosting with multiple sampling procedure using SMOTE. Alternatively, synthetic data sampling can be included in the process of constructing bagging based ensemble [11]. Some of ensemble solutions make use of under sampling techniques to construct class unbiased base learners [12, 13].

Khoshgoftaar et al. investigated the use of the neural network as a model for predicting software quality. They used large telecommunication system to classify modules as fault prone or not fault prone [14]. Disadvantage of neural network is learning in form of weights and human cannot interpret any knowledge from the weights. Khoshgoftaar et al. applied regression trees with classification rule to classify fault-prone software modules using a very large telecommunications system as a case study [15]. The advantage of Decision Tree is giving the, if then rules that is human readable and understandable format. By using these rules we can design early warning prediction system for the post-operative life expectancy of a patient[16].

In recent years, a number of alternative modeling techniques have been proposed for prediction. Alternative models include artificial neural networks, analogy-based reasoning, fuzzy systems and ensemble techniques. Unfortunately the accuracy of these models is not satisfactory so, there is always a scope for new prediction technique.

### III. DATA DESCRIPTION AND DATA PREPARATION

We have used Thoracic Surgery Data which is a classification data related to the post-operative life expectancy in the lung cancer patients. This dataset is publicly accessible from UCI Machine Learning Repository. The dataset contains total 470 numbers of instances, each instances contains 16 Attributes (independent variables) and the associated dependent Boolean variable which has the following two values: Risk (death within one year after surgery) or survival. Table 1 summarizes some main characteristics of Attributes.

TABLE I. DATA SET DESCRIPTION

S.No	Attribute Name	Description
1	DGN	Diagnosis
2	PRE4	Forced vital capacity

3	PRE5	Volume that has been exhaled at the end of the first second of forced expiration
4	PRE6	Performance status
5	PRE7	Pain before surgery
6	PRE8	Haemoptysis before surgery
7	PRE9	Dyspnoea before surgery
8	PRE10	Cough before surgery
9	PRE11	Weakness before surgery
10	PRE14	original tumour size
11	PRE17	Type 2 diabetes mellitus
12	PRE19	MI up to 6 months
13	PRE25	peripheral arterial diseases
14	PRE30	Smoking
15	PRE32	Asthma
16	AGE	Age at surgery
17	Risk	1 year survival period

The dataset contains 85% survival and 15% risk instances. It is evident that the dataset is highly unbalanced in terms of proportion of survival vs. risk. Consequently, before supplying this data to the intelligence techniques, balancing has to be done so that the classifier while predicting will not be influenced by the majority class, that is, survival in this case. The balancing techniques such as under-sampling, and over-sampling have been applied for data preparation.

Under-sampling is a technique in which some of the samples belonging to the majority class are removed and combined with the minority class samples. For example, under-sampling (25%) means that the majority class is reduced to 25% of its original size[16]. Over-sampling is a technique in which the samples belonging to the minority class are replicated a few times and combined with the majority class samples. For example, over-sampling (100%) means that the minority class is replicated once. After the dataset is balanced using one of the above-mentioned two methods, intelligence techniques are applied. All the experiments have been carried out using Hold-out and ten-fold cross validation methods. In Hold-out method the complete data is partitioned into 70: 30 ratio. The 70 percent data is used for training an algorithm and 30 percent data is used for testing.

The prediction capability and the quality of the various classifiers are measured using sensitivity, specificity, accuracy and Area Under ROC Curve (AUC) which are defined as follows[17]:

Sensitivity is the measure of proportion of the true positives, which are correctly identified.

$$\text{Sensitivity} = \text{True positive} / (\text{True positive} + \text{False negative})$$

Specificity is the measure of proportion of the true negatives, which are correctly identified.

$$\text{Specificity} = \text{True negative} / (\text{True negative} + \text{False positive})$$

Accuracy is the measure of proportion of true positives and true negatives, which are correctly identified.

Accuracy = (True positive + True negative) / (True positive + True negative + False positive + False negative).

#### IV. OVERVIEW OF TECHNIQUES EMPLOYED

The following section discusses about the overview of the techniques applied in this paper.

##### 4.1 Random Forest(RF)

Random Forests works by growing many classification trees. To classify a new sample from an input vector, the random forest puts the new sample down each of the trees in the forest. Each tree gives a classification with the help of "votes" for that class. The random forest classifies based on the maximum votes over all the trees in the forest.

Each tree is grown as follows:

1. If the number of cases in the training set is N, random sampling of these N are done *with replacement*, from the original data.
2. If there are K input variables, a number  $k \ll K$  is specified such that at each node, k variables are selected at random and the best split on these k is used to split the node. The value of k is held constant during the forest growing.
3. Each tree is grown to the largest extent possible. No pruning is done[22].

##### 4.2 Decision Tree (DT)

A Decision Tree is a tree based knowledge representation methodology used to represent classification rules. J48 Algorithm (A WEKA implementation of c4.5 algorithm) is a widely used one to construct Decision Trees. The procedure of forming the Decision Tree and exploiting the same for feature selection is explained as follows.

1. The set of features available at hand forms the input to the algorithm; the output is the Decision Tree.
2. The Decision Tree has leaf nodes, which represent class labels and other nodes associated with the classes being classified.
3. The branches of the tree represent each possible value of the feature node from which they originate.
4. The Decision Tree can be used to classify feature vectors by starting at the root of the tree and moving through it until a leaf node, which provides a classification of the instance, is identified.
5. At each decision node in the Decision Tree, one can select the most useful feature for classification using appropriate estimation criteria. The criterion used to identify the best feature invokes the concepts of entropy reduction and information gain[18].

##### 4.3 Support Vector Machine (SVM)

The SVM is a powerful learning algorithm based on recent advances in statistical learning theory proposed by Vapnik [19]. SVM is a learning system that uses a hyper plane

constructed from the points lying on the border of the two class samples. The support vectors are conventionally determined by solving a quadratic programming (QP) problem.

The new SVM learning algorithm is called Sequential Minimal Optimization (or SMO). SMO is a simple algorithm that can quickly solve the SVM QP problem without any extra matrix storage and without using numerical QP optimization steps at all. SMO decomposes the overall QP problem into QP sub-problems, using Osuna's theorem to ensure convergence. The advantage of SMO is that the amount of memory required for SMO is linear in the training set size, which allows SMO to handle very large training sets[20].

##### 4.4 Logistic Regression(LR)

Logistic regression is used to predict a categorical (usually dichotomous) variable from a set of predictor variables. With a categorical dependent variable, discriminant function analysis is usually employed if all of the predictors are continuous and nicely distributed; logit analysis is usually employed if all of the predictors are categorical; and logistic regression is often chosen if the predictor variables are a mix of continuous and categorical variables and/or if they are not nicely distributed (logistic regression makes no assumptions about the distributions of the predictor variables). For a logistic regression, the predicted dependent variable is a function of the probability that a particular subject will be in one of the categories [21].

#### V. EXPERIMENT METHODOLOGY

The experiments have been carried out exhaustively by using sixteen different datasets and five different classifiers. The data set is balanced using two different data balancing techniques as discussed in the Data Preparation section. Thus we ended up with eight different datasets. The original dataset is also referred as actual data in this paper.

The experiments have been carried out as follows; a single dataset is taken at an instance and is supplied to one of the classifier. The parameters of the classifier are set to some initial values and the classifier is run in one of the two methods (i.e., holdout or ten-fold cross validation).

In holdout method we have adapted stratified random sampling for dividing the data into 70:30 ratio. The first part that is 70 percent of data is used for training and second part that is 30 percent of data is used for testing.

As there is a possibility of missing a good feature learning in the holdout method, we adopted ten-fold cross validation also. In this technique, the whole data set is divided into ten parts and in first iteration the first nine parts are supplied as training and the tenth part is supplied as testing. In the next iteration, one of the ninth part is taken out for testing and the tenth part is included in the nine parts as training. So, this substitution goes on until all the parts of the dataset have

been trained and tested. WEKA (<http://www.softpedia.com/get/Internet/Servers/Database-Utils/Weka.shtml>) tool has been used to carry out the experiments. The single algorithm is run for 'N' times for each dataset by changing the parameters in each run. The results have been aggregated in a single table and the best run is selected based upon the sensitivity, specificity, accuracy and AUC. Thus obtained best runs are tabulated as the results based upon the classifier used, method applied and dataset used.

Each table of the results represents the experiments carried out on one of the datasets that is either the actual data or one of the balanced data. Therefore there are totally eight tables which represent the results of the experiments been carried out. Labelling had been done accordingly.

## VI. RESULTS AND DISCUSSION

Exhaustive experiments have been carried out using the actual and balanced data and two of the methods mentioned in experimental methodology. Out of which the best results are presented below, the rows which are represented in italic font in each table indicate the top performing algorithm for that data.

TABLE II. ACTUAL DATA

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	0.581	0.992	0.889	90.72
	LR	0.194	0.967	0.808	80.79
	MLP	0.871	0.975	0.973	95.36
	RF	<i>1</i>	<i>1</i>	<i>1</i>	<i>100</i>
	SVM(SMO)	0.032	1	0.516	80.13
Ten-fold cross validation	DT	0.086	0.925	0.567	80
	LR	0.043	0.963	0.651	82.55
	MLP	0.186	0.913	0.568	80.42
	RF	0.029	0.985	0.61	84.25
	SVM(SMO)	0	1	0.5	85.1

TABLE III. 25 PERCENT UNDERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	0.694	0.943	0.909	84.03
	LR	0.7	0.959	0.953	91.52
	MLP	1	0.98	0.998	98.3
	RF	<i>1</i>	<i>1</i>	<i>1</i>	<i>100</i>
	SVM(SMO)	0.7	0.939	0.819	89.83
Ten-fold cross validation	DT	0.429	0.83	0.641	66.47
	LR	0.486	0.79	0.681	66.47
	MLP	0.507	0.7	0.65	61.17
	RF	0.4	0.77	0.674	61.76
	SVM(SMO)	0.371	0.8	0.586	62.35

TABLE IV. 50 PERCENT UNDERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	0.688	0.969	0.915	91.35
	LR	0.438	0.938	0.828	83.95
	MLP	0.75	1	0.993	95.06
	RF	<i>1</i>	<i>0.985</i>	<i>1</i>	<i>98.76</i>
	SVM(SMO)	0.313	0.985	0.649	85.18
Ten-fold cross validation	DT	0.329	0.845	0.556	71.11
	LR	0.243	0.9	0.654	72.96
	MLP	0.357	0.83	0.621	70.74
	RF	0.257	0.885	0.63	72.22
	SVM(SMO)	0.171	0.94	0.556	74.07

TABLE V. 75 PERCENT UNDERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	0.542	0.959	0.792	87.6
	LR	0.375	0.969	0.842	85.12
	MLP	0.792	0.897	0.941	87.6
	RF	<i>0.917</i>	<i>1</i>	<i>0.999</i>	<i>98.34</i>
	SVM(SMO)	0.208	0.99	0.599	83.47
Ten-fold cross validation	DT	0.329	0.845	0.556	71.11
	LR	0.071	0.927	0.651	76.48
	MLP	0.243	0.887	0.554	76.48
	RF	0.257	0.885	0.63	72.22
	SVM(SMO)	0.014	0.967	0.49	78.64

TABLE VI. 100 PERCENT OVERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	0.906	1	0.971	98.14
	LR	0.375	0.931	0.813	82.09
	MLP	0.844	0.985	0.984	95.67
	RF	<i>1</i>	<i>1</i>	<i>1</i>	<i>100</i>
	SVM(SMO)	0.344	0.931	0.637	81.48
Ten-fold cross validation	DT	0.521	0.873	0.827	78.14
	LR	0.214	0.913	0.718	73.14
	MLP	0.671	0.875	0.841	82.22
	RF	0.936	0.95	0.967	94.62
	SVM(SMO)	0.207	0.953	0.58	75.92

TABLE VII. 200 PERCENT OVERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	0.582	1	66.66	80.5
	LR	0.493	0.892	57.37	77.4
	MLP	0.596	1	67.75	89.2
	RF	0.719	1	77.59	99.3
	SVM(SMO)	0.349	0.973	47.54	66.1
Ten-fold cross validation	DT	0.871	0.82	83.77	90.4
	LR	0.4	0.863	70.32	74.3
	MLP	0.805	0.858	83.93	90.0
	RF	0.99	0.953	96.55	99.6
	SVM(SMO)	0.324	0.87	68.19	59.7

TABLE VIII. 300 PERCENT OVERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	1	0.951	0.997	96.07
	LR	0.585	0.853	0.802	79.9
	MLP	0.951	0.908	0.985	91.66
	RF	1	1	1	100
	SVM(SMO)	0.537	0.847	0.692	78.43
Ten-fold cross validation	DT	0.852	0.815	0.894	82.78
	LR	0.521	0.813	0.754	69.26
	MLP	0.889	0.825	0.923	84.14
	RF	1	0.948	1	96.91
	SVM(SMO)	0.489	0.823	0.656	68.52

TABLE IX. 400 PERCENT OVERSAMPLE

Method	Algorithm NAME	SENSITIVITY	SPECIFICITY	AUC	ACCURACY
Hold-out	DT	1	0.961	0.998	96.88
	LR	0.622	0.783	0.792	75.11
	MLP	0.956	0.844	0.977	86.66
	RF	1	1	1	100
	SVM(SMO)	0.622	0.806	0.714	78.88
Ten-fold cross validation	DT	0.991	0.825	0.923	90.26
	LR	0.6	0.748	0.752	67.86
	MLP	0.926	0.775	0.909	84.53
	RF	1	0.923	1	95.86
	SVM(SMO)	0.603	0.785	0.694	70

The above results show that the Random Forest has outperformed all other techniques by giving 100 percent sensitivity and specificity using holdout method for Actual data, 100 percent oversampling, 300 percent oversampling, 400 percent oversampling, 25 percent undersampling and 50 percent undersampling datasets. This gives a scope for overfitting of an algorithm used.

To ensure that the algorithm is not overfitted, a tenfold cross validation is also done, in which again random forest ranked top among all classifiers only for oversampled data. Hence, random forest can be considered as a best

classification algorithm for classifying the post-operative life expectancy of a patient with high accuracy.

When we look at the results of the ten-fold cross validation method, here also Random Forest is dominating all other intelligent techniques only for oversampled data sets like 100 percent oversample, 200 percent oversample, 300 percent oversample and 400 percent oversample. So, from the above results, we can firmly say that random forest is the best algorithm for predicting the post-operative life expectancy of a patient with high accuracy.

The second perform algorithm among all the algorithms for all the datasets is MLP which again maintained consistency across all the datasets.

## VII. CONCLUSION AND FUTURE WORK

It is observed that Random Forest has outperformed all the other intelligent algorithms when experimented using holdout method for all types of data sets. When we look at the results of the ten-fold cross validation method, here also Random Forest is dominating all other intelligent techniques only for oversampled data sets. So, we can firmly say that random forest is the best algorithm for predicting the post-operative life expectancy of a patient with high accuracy.

As a future work an ensemble model for predicting software defect prediction can be developed.

## ACKNOWLEDGMENT

The authors would like to thank the UCI Machine Learning Repository for providing the Thoracic Surgery Data set.

## REFERENCES

- [1] P.J. Garcia-Laencina, J.L. Sancho-Gomez, A.R. Figueiras-Vidal, "Pattern classification with missing data: a review", Neural Computing and Applications 19 (2009)263–282.
- [2] T. Dietterich, "Machine learning for sequential data: a review", Springer, Berlin Heidelberg, 2002, pp. 227–246.
- [3] H. He, E.A. Garcia, "Learning from imbalanced data", IEEE Transactions on Knowledge and Data Engineering 21 (2009) 1263–1284.
- [4] Maciej Zieba, Jakub M. Tomczak, Marek Lubicz and Jerzy Swiatek, "Boosted SVM for extracting rules from imbalanced data in application to prediction of the post operative life expectancy in the lung cancer patients", Applied Soft Computing, Volume 14, Part A, January 2014, Pages 99–108
- [5] Bartosz Krawczyk, Michał Woźniak and Gerald Schaefer, "Cost sensitive decision tree ensembles for effective imbalanced classification", applied soft computing, Volume 14, Part C, January 2014, Pages 554–562.
- [6] H. He, E.A. Garcia, "Learning from imbalanced data", IEEE Transactions on Knowledge and Data Engineering 21 (2009) 1263–1284.
- [7] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, F. Herrera, "A review on ensembles for the class imbalance problem: bagging, boosting, and hybrid-based approaches", IEEE Transactions on Systems,

- [8] N.V. Chawla, K.W. Bowyer, L.O. Hall, "SMOTE: Synthetic Minority Over-sampling Technique", *Journal of Artificial Intelligence Research* 16 (2002) 321–357.
- [9] H. Han, W.-Y. Wang, B.-H. Mao, "Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning", in: D.S. Huang, X.P. Zhang, G.B. Huang (Eds.), *Advances in Intelligent Computing*, 2005, pp. 878–887.
- [10] N. Chawla, A. Lazarevic, L. Hall, K. Bowyer, "Smote boost: improving prediction of the minority class in boosting", *Knowledge Discovery in Databases: PKDD 2003*, Springer, Berlin Heidelberg, 2003, pp. 107–119.
- [11] S. Wang, X. Yao, "Diversity analysis on imbalanced data sets by using ensemble models", *IEEE Symposium on Computational Intelligence and Data Mining Proceedings*, 2009, pp. 324–331.
- [12] E. Chang, B. Li, G. Wu, K. Goh, "Statistical learning for effective visual information retrieval", *IEEE Proceedings of the 2003 International Conference on Image Processing*, vol. 3, 2003, pp. 609–613.
- [13] D. Tao, X. Tang, X. Li, X. Wu, "Asymmetric bagging and random subspace for support vector machines based relevance feedback in image retrieval", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28 (2006) 1088–1099.
- [14] T. Khoshgoftaar, E. Allen, J. Hudepohl and S. Aud, "Application of neural networks to software quality modeling of a very large telecommunications system", *IEEE Transactions on Neural Networks* 8(4), 1997, pp. 902–909.
- [15] T. Khoshgoftaar, E. Allen and J. Deng, "Using regression trees to classify fault-prone software modules" *IEEE Transactions on Reliability* 51 (4), 2002, pp. 455–462.
- [16] D. Anil Kumar and V. Ravi, "Predicting credit card customer churn in banking using data mining", *international journal of data analysis techniques and strategies*, vol. 1(1), 2008, pp. 4–28.
- [17] T. Fawcett, "An introduction to ROC analysis", *Pattern Recognition Letters*, Vol. 27, 2006, pp. 861–874.
- [18] N. Saravanan, K.I. Ramachandran, "Fault diagnosis of spur bevel gear box using discrete wavelet features and Decision Tree classification", *Expert Systems with Applications* vol. 36 (2009), pp. 9564–9573.
- [19] V.N. Vapnik, "Statistical Learning Theory", John Wiley, New York, 1998.
- [20] J.S. Pahariya, V. Ravi, M. Carr and M. Vasu, "Computational Intelligence Hybrids Applied to Software Cost Estimation", *International Journal of Computer Information Systems and Industrial Management Applications*, ISSN: 2150-7988 Vol. 2, 2010, pp. 104–112.
- [21] <http://core.ecu.edu/psyc/wuenschk/MV/Multreg/Logistic-SPSS.PDF>  
date accessed on 20/05/2014.
- [22] [http://www.stat.berkeley.edu/~breiman/RandomForests/cc\\_home.htm#overview](http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm#overview)  
date accessed on 13/06/2014.

**Department of  
Humanities & Sciences**



# NUMERICAL INTEGRATION (QUADRATURE) METHOD FOR STEADY –STATE CONVECTION- DIFFUSION PROBLEMS

K. Sharath Babu<sup>1</sup>, N. Srinivasacharyulu<sup>2</sup>, J.V. Brammam<sup>3</sup>

<sup>1</sup>Professor of Mathematics, Malla Reddy Engineering College (Autonomous), Maisammaguda ,  
Secunderabad (India)

[sharathkandayrao@mrec.ac.in](mailto:sharathkandayrao@mrec.ac.in)

<sup>2</sup>Associate Professor of Mathematics, National Institute of Technology, Warangal

<sup>3</sup>Research Scholar, Krishna University, Machilipatnam, Krishna District, Andhra Pradesh (India)

## ABSTRACT

*In this paper, a numerical integration method is presented for solving a general steady-state convection problem or singularly perturbed two-point boundary value problem. The governing second-order differential equation is replaced by an approximate first-order differential equation with a small deviating argument. Then the Simpson one-third formula is used to obtain the three term recurrence relationship. The proposed method is iterative on the deviating argument. To test and validity of this method we have solved several model linear problems with left-end boundary layer or right-end boundary layer or an internal layer and offered the computational results.*

**Keywords:** *Singular Perturbation; Boundary Layer; Peclet Number; Two-Point Boundary Value Problem..*

## I. INTRODUCTION

Convection-diffusion problems occur very frequently in the fields of science and engineering such as fluid dynamics, specifically the fluid flow problems involving large Reynolds number and other problems in the great world of fluid motion. The numerical treatment of singular perturbation problems is far from trivial because of the boundary layer behavior of the solution. However, the area of convection-diffusion problems is a field of increasing interest to applied mathematicians.

The survey paper by Kadalbajoo and Reddy [1], gives an intellectual outline of the singular perturbation problems and their treatment starting from Prandtl's paper [2] on fluid dynamical boundary layers. This survey paper will remain as one of the most readable source on convection-diffusion or singular perturbation problems. For a detailed theory and analytical discussion on singular perturbation problems one may refer to the books and high level monographs: O'Malley [3], Nayfeh [4], Bender and Orszag [5], Kevorkian and Cole.

In this paper , a numerical integration method is presented for solving general singularly perturbed two-point boundary value problems .The main advantage of this method is that it does not require very fine mesh size. The inventive second-order differential equation is replaced by an approximate first-order differential equation with a small differing argument. Then , the Simpson one-third formula is used to obtain the three term recurrence relationship. Thomas Algorithm is applied to solve the resulting tridiagonal algebraic system of equations. The proposed method is iterative on the deviating argument. The method is to be repeated for different choices of the deviating argument until the solution profile stabilizes. To examine the applicability of the proposed method, we have solved several model linear problems with left-end boundary layer or right –end boundary layer or an internal layer and presented the numerical results. It is observed that the numerical integration method approximates the exact solution extremely well.

## II. NUMERICAL INTEGRATION METHOD

For the sake of convenience we call our method the ‘ Numerical Integration Method’. To set the stage for the numerical integration method , we consider the following Governing linear Convection-diffusion ( singularly perturbed two-point boundary value problem:

$$\varepsilon y''(x) + a(x) y'(x) + b(x)y(x) = f(x) ; 0 \leq x \leq 1 \quad (1)$$

$$\text{With } y(0) = \alpha \text{ and } y(1) = \beta \quad (2)$$

Where  $\varepsilon$  is a small positive parameter called diffusion parameter which lies in the interval  $0 < \varepsilon \leq 1$  ;  $\alpha$  and  $\beta$  are given constants;  $a(x)$  ,  $b(x)$  and  $f(x)$  considered to be sufficiently continuously differentiable functions in  $[0,1]$ . Furthermore , we assume that  $a(x) \geq M > 0$  throughout the interval  $[0,1]$ , where  $M$  is some positive constant. This assumption merely implies that the boundary layer will be in the neighborhood of  $x=0$ .

Let  $\delta$  be a small positive deviating argument ( $0 < \delta \leq 1$ ). By applying Taylor series expansions in the neighborhood of the point  $x$ , we have

$$y(x - \delta) \approx y(x) - \delta y'(x) + \frac{\delta^2}{2} y''(x) \quad (3)$$

And consequently, Eq. (1) is replaced by the following first-order differential equation with a small deviating argument.

$$\frac{\delta^2}{2} y''(x) = y(x - \delta) - y(x) + \delta y'(x) \Rightarrow y''(x) = \frac{2}{\delta^2} [y(x - \delta) - y(x) + \delta y'(x)] \text{ So that}$$

$$(1) \Rightarrow \frac{2\varepsilon}{\delta^2} [y(x - \delta) - y(x) + \delta y'(x)] + a(x) y'(x) + b(x) y(x) = f(x) ; 0 \leq x \leq 1$$

$$\Rightarrow 2\varepsilon y(x - \delta) - 2\varepsilon y(x) + 2\varepsilon \delta y'(x) + a(x) y'(x) \delta^2 + b(x) y(x) \delta^2 = \delta^2 f(x)$$

$$\Rightarrow [2\varepsilon \delta + a(x) \delta^2] (y'(x)) + [b(x) \delta^2 - 2\varepsilon] y(x) = \delta^2 f(x) - 2\varepsilon y(x - \delta)$$

$$\Rightarrow y'(x) = \frac{\delta^2 f(x) - 2\varepsilon y(x - \delta)}{2\varepsilon \delta + a(x) \delta^2} y(x - \delta) + \frac{(2\varepsilon - b(x) \delta^2)}{2\varepsilon \delta + a(x) \delta^2} y(x)$$



$$\Rightarrow y'(x) = \frac{-2\varepsilon}{2\varepsilon\delta + a(x)\delta^2} y(x - \delta) + \frac{2\varepsilon - b(x)\delta^2}{2\varepsilon\delta + a(x)\delta^2} y(x) + \frac{\delta^2 f(x)}{2\varepsilon\delta + a(x)\delta^2} \quad (4)$$

(4) Can be re-written as

$$y'(x) = p(x) y(x - \delta) + q(x) y(x) + r(x) \text{ for } \delta \leq x \leq 1 \quad (5)$$

Where

$$p(x) = \frac{-2\varepsilon}{2\varepsilon\delta + \delta^2 a(x)} \quad (6)$$

$$q(x) = \frac{2\varepsilon - \delta^2 b(x)}{2\varepsilon\delta + \delta^2 a(x)} \quad (7)$$

$$r(x) = \frac{\delta^2 f(x)}{2\varepsilon\delta + \delta^2 a(x)} \quad (8)$$

We now divide the interval [0,1] in to N equal parts with mesh size h, i.e.,  $h=1/N$  and  $x_i = ih$  for  $i=1,2,3,\dots,N$ . Integrating equation (5) in  $[x_{i-1}, x_{i+1}]$  we get

$$y(x_{i+1}) - y(x_{i-1}) = \int_{x_{i-1}}^{x_{i+1}} [p(x) y(x - \delta) + q(x)y(x) + r(x)] dx \quad (9)$$

By making use of the Newton-Cotes formula when  $n=2$  i.e. applying Simpson's 1/3 rule approximately, we obtain

$$\begin{aligned} y(x_{i+1}) - y(x_{i-1}) &= \frac{h}{3} [p(x_{i+1})y(x_{i+1} - \delta) + 4p(x_i)y(x_i - \delta) + p(x_{i-1} - \delta) \\ &+ (p_{i+1} + p_{i-1}) [y(x_{i+1} - \delta) + y(x_{i-1} - \delta)] + q(x_{i+1})y(x_{i+1}) + q(x_{i-1})y(x_{i-1}) + q(x_{i+1})y(x_{i+1}) \\ &+ 4q(x_i)y(x_i) + q(x_{i-1})y(x_{i-1}) + r(x_{i+1}) + 4r(x_i) + r(x_{i-1}) + r(x_{i+1}) + r(x_{i-1})] \end{aligned} \quad (10)$$

Again by means of Taylor's series expansion, we have

$$y(x - \delta) \cong y(x) - \delta y'(x)$$

and then by approximating  $y'(x)$  by Linear Interpolation method we get

$$\begin{aligned} y(x_i - \delta) &\cong y(x_i) - \frac{\delta [y(x_{i+1}) - y(x_{i-1})]}{2h} \\ &= y(x_i) + \frac{\delta}{2h} y(x_{i-1}) - \frac{\delta}{2h} y(x_{i+1}) \end{aligned} \quad (11)$$

similarly

$$y(x_{i-1} - \delta) \cong (1 + \frac{\delta}{h}) y(x_{i-1}) - \frac{\delta}{h} y(x_i) \quad (12)$$

$$y(x_{i+1} - \delta) = (1 - \frac{\delta}{h}) y(x_{i+1}) + \frac{\delta}{h} y(x_i) \quad (13)$$

Hence making use of (11),(12),(13) (10) can be written as

$$\begin{aligned}
 y_{i+1} - y_{i-1} &= \frac{h}{3} [p_{i+1}[(1 - \frac{\delta}{h})y_{i+1} + \frac{\delta}{h}y_i] + 4p_i[y_i - \frac{\delta}{2h}y_{i+1} + \frac{\delta}{2h}y_{i-1}] + p_{i-1}[(1 + \frac{\delta}{h})y_{i-1} - \frac{\delta}{h}y_i] \\
 &+ (p_{i+1} + p_{i-1})[(1 - \frac{\delta}{h})y_{i+1} + \frac{\delta}{h}y_i + (1 + \frac{\delta}{h})y_{i-1} - \frac{\delta}{h}y_i + 2q_{i+1}y_{i+1} + 2q_{i-1}y_{i-1} + 4q_iy_i + 2r_{i+1} + 4r_i + 2r_{i-1} \\
 &[-1 - \frac{2p_i\delta}{3} - \frac{h}{3}p_{i-1}(1 + \frac{\delta}{2h}) - \frac{h}{3}(p_{i+1} + p_{i-1})(1 + \frac{\delta}{h}) - \frac{2h}{3}q_{i-1}]y_{i-1} + [\frac{\delta p_{i-1}}{3} - \frac{\delta}{3}p_{i+1} - \frac{4hp_i}{3} \\
 &- \frac{4hq_i}{3}]y_i + [1 - \frac{h}{3}p_{i+1}(1 - \frac{\delta}{h}) + \frac{2p_i\delta}{3} - \frac{h}{3}(p_{i+1} + p_{i-1})(1 - \frac{\delta}{h}) - \frac{2h}{3}q_{i+1}]y_{i+1} \\
 &= \frac{2h}{3}[r_{i+1} + 2r_i + r_{i-1}]
 \end{aligned} \tag{14}$$

can be written in the standard form as

$$A_i y_{i-1} + B_i y_i + C_i y_{i+1} = D_i \tag{15}$$

where

$$A_i = -1 - \frac{2p_i\delta}{3} - \frac{h}{3}p_{i-1}(1 + \frac{\delta}{2h}) - \frac{h}{3}(p_{i+1} + p_{i-1})(1 + \frac{\delta}{h}) - \frac{2h}{3}q_{i-1} \tag{16}$$

$$B_i = \frac{\delta p_{i-1}}{3} - \frac{\delta}{3}p_{i+1} - \frac{4hp_i}{3} - \frac{4hq_i}{3} \tag{17}$$

$$C_i = 1 - \frac{h}{3}p_{i+1}(1 - \frac{\delta}{h}) + \frac{2p_i\delta}{3} - \frac{h}{3}(p_{i+1} + p_{i-1})(1 - \frac{\delta}{h}) - \frac{2h}{3}q_{i+1} \tag{18}$$

$$D_i = \frac{2h}{3}[r_{i+1} + 2r_i + r_{i-1}] \tag{19}$$

Here  $y_i = y(x_i)$ ,  $p_i = p(x_i)$ ,  $q_i = q(x_i)$  and  $r_i = r(x_i)$ . Equation (16) gives a system of (N-1) equations with (N+1) unknowns  $y_0$  to  $y_N$ . The two given boundary conditions ( ) together with these (N-1) equations are then sufficient to solve for the unknowns  $y_0, y_N$ . The solution of the Tri-diagonal system (15) can be obtained by using an efficient algorithm called 'Thomas Algorithm'. In this algorithm we set a difference relation of the form

$$y_i = W_i y_{i+1} + T_i \tag{20}$$

Where  $W_i$  and  $T_i$  corresponding to  $W(x_i)$  and  $T(x_i)$  are to be determined from (20) we have

$$y_{i-1} = W_{i-1} y_i + T_{i-1} \tag{21}$$

Substituting (21) in (15) we get

$$y_i = \frac{C_i}{B_i - A_i W_{i-1}} y_{i+1} + \frac{A_i T_{i-1} - D_i}{B_i - A_i W_{i-1}} \quad (22)$$

By comparing (20) and (22), we can get

$$W_i = \frac{C_i}{B_i - A_i W_{i-1}} \quad (23)$$

$$T_i = \frac{A_i T_{i-1} - D_i}{B_i - A_i W_{i-1}} \quad (24)$$

To solve these recurrence relations for  $i=1,2,3,\dots,N-1$ ; we need to know the initial conditions for  $W_0$  and  $T_0$ . This can be done by considering (2)

$$y_0 = \alpha = W_0 y_1 + T_0 \quad (25)$$

If we choose  $W_0=0$ , then  $T_0 = \alpha$ . With these initial values, we compute sequentially  $W_i$  and  $T_i$  for  $i=1,2,3,\dots,N-1$ ; from (24) and (25) in the forward process and then obtain  $y_i$  in the backward process from (20) using (2).

Repeat the numerical scheme for different choices of  $\delta$  (deviating argument, satisfying the conditions  $(0 < \delta \leq 1)$ , until the solution profiles do not differ significantly from iteration to iteration. For computational point of view, we use an absolute error criterion, namely

$$|y(x)^{m+1} - y(x)^m| \leq \rho, 0 \leq x \leq 1 \quad (26)$$

Where  $y(x)^m$  is the solution for the  $m^{\text{th}}$  iterate of  $\delta$ , and  $\rho$  is the prescribed tolerance bound.

### III. LINEAR PROBLEMS

Here we are considered the applicability of the numerical integration method, we have applied it to linear singular perturbation problems with left-end boundary layer. These examples have been chosen because they have been widely discussed in the literature and because approximate solution is available for comparison.

#### Example 1.

Consider the following homogeneous Singular value perturbation problem from Kevorkian and Cole [6, p.33, Eqs.(2.3.26) and (2.3.27)] with  $\alpha=0$ :

$$\varepsilon y''(x) + y'(x) = 0, 0 \leq x \leq 1 \text{ with } y(0) = 0 \text{ and } y(1) = 1$$

The exact solution is given by

$$y(x) = \frac{(1 - \exp(-x/\varepsilon))}{(1 - \exp(-1/\varepsilon))}$$

The computational results are presented in Table 1(a) and (b) for  $\varepsilon = 10^{-3}, 10^{-4}$  respectively.

**Table 1 Computational Result for Example 1**

X	y(x)			Exact
---	------	--	--	-------

				solution
$\varepsilon=0.001, h=0.01$	$\delta=0.008$	$\delta=0.009$	$\delta=0.007$	
0.00	0.00000000	0.00000000	0.00000000	0.00000000
0.02	0.9876486	0.9899944	0.9917358	1.0000000
0.04	0.9998419	0.9998944	0.9999319	1.0000000
0.06	0.9999925	0.9999934	0.9999995	1.0000000
0.08	0.9999945	0.9999945	1.0000000	1.0000000
0.10	0.9999946	0.9999948	1.0000000	1.0000000
0.20	0.9999954	0.9999952	1.0000000	1.0000000
0.40	0.9999964	0.9999964	1.0000000	1.0000000
0.60	0.9999976	0.9999976	1.0000000	1.0000000
0.80	0.9999988	0.9999988	1.0000000	1.0000000
1.00	1.00000000	1.00000000	1.00000000	1.0000000

(b)  $\varepsilon = 10^{-4}$  and  $h = 0.01$

0.00	0.00000000	0.00000000	0.00000000	0.00000000
0.02	0.9998016	0.9998477	0.9998792	1.0000000
0.04	0.9999999	1.0000000	1.0000000	1.0000000
0.06	1.0000000	1.0000000	1.0000000	1.0000000
0.08	1.0000000	1.0000000	1.0000000	1.0000000
0.10	1.0000000	1.0000000	1.0000000	1.0000000
0.20	1.0000000	1.0000000	1.0000000	1.0000000
0.40	1.0000000	1.0000000	1.0000000	1.0000000
0.60	1.0000000	1.0000000	1.0000000	1.0000000
0.80	1.0000000	1.0000000	1.0000000	1.0000000
1.00	1.0000000	1.0000000	1.0000000	1.0000000

## Example 2

Consider the following homogeneous Spp from Bender and Orsag[2,p.480. problem 9.17]

with  $\alpha = 0$  :

$$\varepsilon y''(x) + y'(x) - y(x) = 0, \quad 0 \leq x \leq 1 \text{ with } y(0) = 0 \text{ and } y(1) = 1$$

The exact solution is given by

$$y(x) = \frac{(e^{m_2} - 1) e^{m_1 x} + (1 - e^{m_1}) e^{m_2 x}}{(e^{m_2} - e^{m_1})} \quad \text{where}$$

$$m_1 = \frac{-1 + \sqrt{1 + 4\varepsilon}}{2\varepsilon} ;$$

$$m_2 = \frac{-1 - \sqrt{1 + 4\varepsilon}}{2\varepsilon}$$

**Table 2 Computational Results for Example 2 .**

X	y(x)			Exact solution
$\varepsilon=0.001, h=0.01$	$\delta=0.008$	$\delta=0.009$	$\delta=0.007$	
0.00	<b>1.0000000</b>	<b>1.0000000</b>	<b>1.0000000</b>	<b>1.0000000</b>
0.02	<b>0.3834784</b>	<b>0.3819605</b>	<b>0.3808348</b>	<b>0.3756784</b>
0.04	<b>0.3834410</b>	<b>0.3833556</b>	<b>0.3832939</b>	<b>0.3832599</b>
0.06	<b>0.3910826</b>	<b>0.3910290</b>	<b>0.3909866</b>	<b>0.3909945</b>
0.08	<b>0.3989720</b>	<b>0.3989188</b>	<b>0.3988770</b>	<b>0.3988851</b>
0.10	<b>0.4070216</b>	<b>0.4069688</b>	<b>0.4069269</b>	<b>0.4069350</b>
0.20	<b>0.4497731</b>	<b>0.4497210</b>	<b>0.4496799</b>	<b>0.4496879</b>
0.40	<b>0.5492185</b>	<b>0.5491707</b>	<b>0.5491330</b>	<b>0.5491404</b>
0.60	<b>0.6706514</b>	<b>0.6706123</b>	<b>0.6705816</b>	<b>0.6705877</b>
0.80	<b>0.8189330</b>	<b>0.8189092</b>	<b>0.8188905</b>	<b>0.8188942</b>
1.00	<b>1.0000000</b>	<b>1.0000000</b>	<b>1.0000000</b>	<b>1.0000000</b>

## REFERENCES

- [1] Carl M. Bender ,Steven A. orszag Advanced Mathematical methods for Scientists and Engineers, Asymptotic methods and perturbation theory, Springer.
- [2] L.E. El'sgol'ts, S.B.Norkin, Introduction to the Theory and Application of Differential Equations with Deviating Arguments, Academic Press, New York, 1973.
- [3] K.Eriksson, D. Estep, P.Hansbo and C. Johnson (1996) , Computational Differential Equations, Cambridge University Press, Cambridge.
- [4] A. M. Il'in (1969), 'A difference scheme for a differential equation with a small Parameter multiplying the highest derivative', *Mat. Zametki* **6**, 237–248.  
equations', *Comput. Methods Appl. Mech. Engrg.* **190**, 757–781.

- [5] Introduction to singular Perturbation problems by Robert E.O's Malley,Jr, Academic press.
- [6] Martin Stynes (2005) ‘ Steady-state convection-diffusion problems ‘, Acta Numerica(2005) , pp. 445-508, Cambridge University Press.
- [7] K. W. Morton(1996) , Numerical solution of Convection-Diffusion Problems, **Vol 12** of Applied Mathematics and Mathematical Computation, Chapman & Hall, London.
- [8] Mikhail Shashkov (2005) ‘Conservative finite difference methods on General grids ‘ , CRS Press(Tokyo).
- [9] A.H. Nayfeh, Perturbation Methods, Wiley, New York, 1979.
- [10] G.D. Smith, ‘Partial Differential equations ‘, Oxford Press.
- [11] N. Srinivasacharyulu, K. Sharath babu (2008) , ‘Computational method to solve steady-state convection-diffusion problem’ , International Journal of Mathematics, Computer Sciences and Information Technology, Vol. **1** No. 1-2, January-December 2008,pp.245-254.
- [12] M. Stynes and L. Tobiska (1998), ‘A finite difference analysis of a streamline diffusion method on a Shishkin meshes’, Numer. Algorithms **18**, 337-360.

**Department of  
Master of Business  
Administration**



# Corporate Restructuring Dynamics: A Case Study Analysis

**Dr.K.VEERAAIAH**

MBA, M.COM, M.Phil, UGCNET, Ph.D.

**Professor & Head of the Department**

**Malla Reddy Engineering College (Autonomous)**

**Misammaguda, Dhulapally (Post Via Kompally), Secunderabad-500100.**

**Mob No: 9885650478, Email ID:veeru.k88@gmail.com**

## **ABSTRACT**

*Corporate restructuring has recently faced widespread criticism due to its low-turnaround success rate and the sharp increase in insolvency filings. Our analysis extends beyond this discussion by investigating the inherent complexity of restructurings as a critical success factor with a particular focus on portfolio restructuring. Based on theoretical findings, we develop a System Dynamics model for a case study in portfolio restructuring. Thereafter, we evaluate this research approach in an outlined context and devise a roadmap for subsequent research. Results of the analysis indicate that the restructuring process can be reasonably simplified with a System Dynamics model. Further, this model enables a comprehensive sensitivity analysis that allows management to develop an understanding of the underlying dynamics. The key implication is that managers should consider a System Dynamics model as a complement to conventional modeling, as the conceptual and numerical benefits can outweigh the related costs.*

## **Keywords**

Corporate Restructuring, Portfolio Restructuring, System Dynamics, Modeling, Simulation.

## **INTRODUCTION**

### **1. MOTIVATION AND BACKGROUND**

Corporate restructuring is a key area in strategic management, finance, and organizational theory. Although various fields have contributed to the literature, numerous restructurings have failed in practice, which has resulted in vast criticism of the process.<sup>1</sup> Results from empirical performance investigations of restructurings reveal a diverse spectrum of conclusions. While some companies have been very successful in their restructuring efforts, others have destroyed shareholder value.<sup>2</sup> This article examines the hypothesis that the tension between the need for a rigorous business analysis and the urge to deliver vital performance enhancements is crucial for this ambiguous performance. The reasoning coincides with Bowman and Singh's findings that define the process's complexity as a key characteristic of corporate restructurings.<sup>3</sup> Consequently, prudent corporations have approached the reconfiguration with a suitable approach to complexity management. Recently, information technology and computational resources have become more powerful and widely used by corporations.<sup>4</sup> A variety of modern modeling and simulation (M&S) techniques provide management with new forms of information, but research indicates that the mere use of computational power is only one component of a valuable decision-support tool.<sup>5</sup> In addition, it should be based on sound concepts that enable efficient

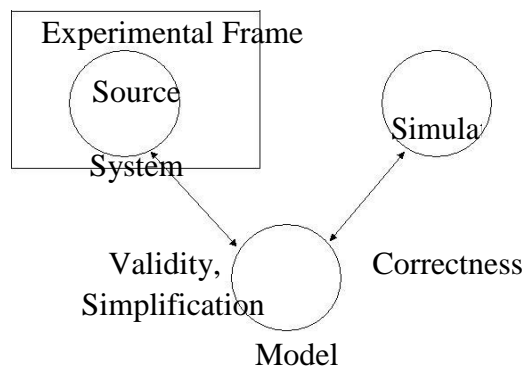
communication.<sup>6</sup> The System Dynamics approach is one of the most prominent and effective modeling approaches that meets this requirement.<sup>7</sup> By combining both research areas, the goal of this article is to close the gap in the literature by analyzing a System Dynamics approach to portfolio restructuring. Two elements focus this study. First, the analysis is restricted to portfolio restructuring although corporate restructuring includes a broad spectrum of activities.<sup>8</sup> However, because several examples of successful applications of System Dynamics for accounting-based models exist, portfolio restructuring is the most promising branch in corporate restructuring due to its accounting-related nature.<sup>9</sup> The second focus is a case study by which we investigate a spin-off. This article contains four sections following this introduction. Section 2 contains an overview on modeling, while Section 3 provides background information and theory on restructuring. The System Dynamics approach is applied to an illustrative case study in Section 4. Finally, in Section 5, we discuss the implications of our findings and provide suggestions for further research.

## 2. MODELING AND SIMULATION

As humans are constrained by cognitive limitations, effective decision making in an uncertain environment requires simplifications.<sup>10</sup> This can be achieved by reducing the complexity of models and assessing different scenarios via simulations.<sup>11</sup> The effectiveness of these approaches depends on several critical factors.

### 2.1 A Relational Framework for Modeling and Simulation

A generic framework for integrating M&S includes the source system, the experimental frame, the model, and the simulator as depicted in Figure 1.<sup>12</sup>



**Figure 1: Entities and Relationships in Modeling and Simulation. Based on Zeigler, Praehofer, and Kim (2000).**

While all entities are interrelated, the most relevant relationships are the simplification of the source system in the model, the validity of the model with respect to the source system, and the simulator's adequate behavior with respect to the model. Hence, the key to modeling is to solve the dilemma of valid simplifications. Both aspects — simplification and validity — are prerequisites for deriving

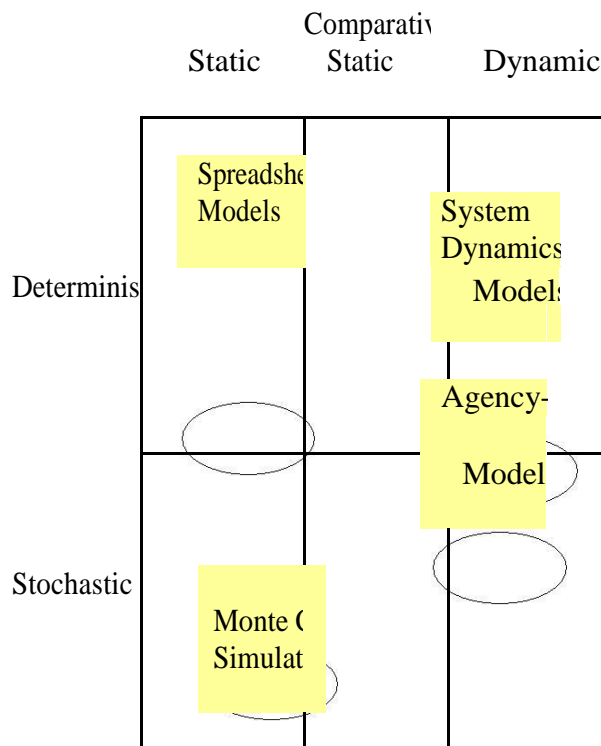
reliable policy implications. Based on this theoretical framework of the relationships between modeling and simulation, the respective definitions can be derived. Modeling is generically defined as the process of designing an image of reality with required characteristics.<sup>13</sup> Hannon and Ruth (1994) characterize modeling as an iterative process that is initiated by real -world events or virtual events. The underlying interactions are combined to enable a higher level of observation by identifying key elements and processes. Next, the model is further constructed by defining the underlying variables and their relationships. At this stage, the model enables the simulation of different paths that allow us to derive predictions and appropriate policy implications. The comparison of these predictions with real-world observed events leads to a revision of the model, which initiates a new iteration of the modeling process.<sup>14</sup> Another aspect of modeling is emphasized by Van den Bosch and Van der Klauw (1994) who claim that modeling is “the art of creating mathematical descriptions of ... phenomena which appear in reality.”<sup>15</sup> All these aspects have shaped the view on modeling in this article which defines modeling as the process of creating a sound and valid simplification of a source system in a relevant experimental frame. Simulation is defined as a “tool for obtaining responses of ... models to understand their dynamic behavior,” thus emphasizing the dynamics of models.<sup>16</sup>

## **2.2 The Motivation for Modeling and Simulation Techniques**

After outlining the genesis of M&S, the next logical step is to analyze its potential for enhancing decision making. Van den Bosch and Van der Klauw (1994) identify two predominant motives of M&S: Designing control systems and enhancing decision quality.<sup>17</sup> Models can be used as controls to monitor actual systems. As it is easy to update these models with new data, a subsequent risk-management strategy can readily be developed.<sup>18</sup> The calculation of the Value at Risk (VaR) is an example of a successful risk management simulation.<sup>19</sup> In the corporate restructuring context, the quality enhancement of decisions is more important than the controlling function. Thinking ahead of the current situation or crisis has become a necessary step and a competitive advantage in the dynamic environment of corporations.<sup>20</sup> M&S allows management to develop hypotheses, to make decisions, and to evaluate the respective consequences without actually having to bear any negative effects.<sup>21</sup> M&S can reveal even unintuitive relationships and assist organizations in making optimal decisions for a given information set.<sup>22</sup>

## **2.3 Classification of Modeling and Simulation Techniques**

The various M&S methods can be clustered by different parameters.<sup>23</sup> As the model's underlying time horizon is of particular interest in the corporate restructuring context, it is useful to distinguish static, comparatively static, and dynamic models.<sup>24</sup> Depending on the research perspective, models are furthermore separated into deterministic and stochastic models.<sup>25</sup> The combination of both characteristics allows us to develop a classification of M&S techniques; some representative techniques are outlined in Figure 2.



**Figure 2: Types of Modeling and Simulation Techniques. Based On Ruth and Hannon (1997) and Hartmann (1996).**

While complex stochastic simulations account for uncertainty in the decision-making model, they require a vast amount of computation power, whereas deterministic simulations quickly deliver intuitive results;<sup>26</sup> the required computational power is a disadvantage of stochastic methods, even though recent developments in the IT industry have resulted in an enormous processing capacity even on home computers.<sup>27</sup> With this classification in hand, managers can choose the appropriate M&S technique depending on their concerns, underlying assumptions and the idiosyncratic decision context.

## 2.4 The Ideal Modeling and Simulation Process

Steinhausen (1994) recommends a seven-step procedure for model design. According to this procedure, identifying a problem is the basis for all further steps, which implies that the source system and the examination frame are defined.<sup>28</sup> In the next phase, relevant characteristics of the real world are abstracted by tools such as causal diagrams, flowcharts and mathematical descriptions.<sup>29</sup> Relevant information and their interaction patterns, which include input and output variables, are identified.<sup>30</sup> In the next step, data is collected to calibrate the parameters.<sup>31</sup> Then the abstract model is converted into an executable program with the help of an M&S technique that is chosen according to the classification scheme.<sup>32</sup> Step five is the validity check of the model, which is achieved by a reconstruction of historic developments.<sup>33</sup> Once the validity of the model is confirmed, scenarios are simulated and validated in a sensitivity analysis.<sup>34</sup> The last step is an analysis of the results and their interpretation.

## **2.5 System Dynamics Modeling**

After having outlined the theoretical process of modeling, we now focus on System Dynamics. System dynamic models provide a wide range of opportunities to model variables and their relationships. Accordingly, we can set state variables, control variables, and transforming variables.<sup>35</sup> The relationships between these variables are mainly dominated by feedback loops in which reinforcing positive processes are distinguished from rebalancing negative ones.<sup>36</sup> The variables and feedback processes are integrated within a model to simulate a system's behavior over time. Basic model types are stimulus-response, self-referencing, goal-seeking, and goal-setting models.<sup>37</sup> Due to the fact that many relevant elements and relationships can be integrated, System Dynamics models offer a scalable complexity, which is valuable for modeling corporate restructuring decisions and their implications.

## **3. MOTIVATION, THEORY, AND CONCEPTS OF PORTFOLIO RESTRUCTURING**

This chapter provides background information on motives, theories, and concepts of the experimental frame of corporate restructuring in general, and portfolio restructuring in particular.<sup>38</sup>

### **3.1 Corporate Restructuring**

Corporate restructuring is one of the most complex and fundamental phenomena that management confronts.<sup>39</sup> Each company has two opposing strategies from which to choose: to diversify or to refocus on its core business. While diversification represents the expansion of corporate activities, refocusing characterizes a concentration on its core business. From this perspective, corporate restructuring is a reduction in diversification.<sup>40</sup> Corporate restructuring entails a range of activities including portfolio restructuring, financial restructuring, and organizational restructuring.<sup>41</sup> Accordingly, portfolio or asset restructuring involves the redeployment of corporate assets through divestitures of business lines that are considered peripheral to the core business strategy. Significant changes in a corporation's capital structure are termed financial restructuring. In organizational restructuring, the focus of change is on management and internal corporate governance structures.

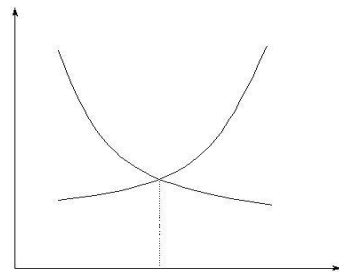
This study primarily focuses on portfolio restructuring which is construed as the elimination of elements from a business portfolio.<sup>42</sup> It is important to note, however, that the effects of restructuring activities are not restricted to one class of restructuring. In contrast, although portfolio restructuring primarily affects the asset side of the balance sheet, it cannot be accomplished without adjusting the liability side. Therefore, portfolio restructuring should not be seen as an isolated process, but rather as a multidimensional and long-term process in a series of corporate restructuring activities.<sup>43</sup>

### **3.2 Motives for Restructurings**

The motivations for restructuring are manifold and depend on the particular set of problems and circumstances facing firms. Nevertheless, we can identify the shareholder value principles as a common rationale for all restructuring processes. Therefore, we analyze the particular motives in greater detail after outlining the underlying shareholder value paradigm.

The genuine function of a corporation is the subject of ongoing scholarly debate.<sup>44</sup> While some researchers support the stakeholder value rationale, others favor the shareholder value rationale. Three factors support the shareholder view within this analysis. First, shareholders exert a high degree of influence on corporations via corporate governance mechanisms.<sup>45</sup> Therefore, we can reasonably assume that management decisions are focused on shareholder interests. A second factor supporting the shareholder-value perspective is that shareholder dividends are the residual after all other stakeholders' needs are satisfied. Therefore, optimal shareholder -value management implies the maximum value

creation for all stakeholders. Finally, empirical evidence supports the hypothesis that shareholder-oriented firms outperform others.<sup>46</sup>



**Figure 3: The Optimal Point for Diversification. Based on Markides (1995).**

Following the shareholder-value perspective, portfolios should be restructured and some divisions must be suspended, if they contribute no value.<sup>47</sup> Alternatively, a change in the corporate strategy might require an optimization of the corporate structure, since empirical evidence suggests that even profitable business units should be sold if they do not directly support a corporation's general strategy.<sup>48</sup> In both cases, the focus on core competencies creates value if the marginal benefits of the diversification are below marginal costs.<sup>49</sup> Beneficial synergies result either from an enhancement of revenues or from a reduction of costs.<sup>50</sup> Both can be accomplished by economies of scale, economies of scope, learning effects, and other effects that improve a firm's profitability.<sup>51</sup> Negative synergies, in contrast, destroy shareholder value due to inflexible structures or high overhead costs.<sup>52</sup> The most important drivers of such disadvantages are coordination, compromise, transaction, and inflexibility costs.<sup>53</sup> Dismantling a value-negating conglomerate in which the benefits fail to outweigh the costs is a central motive for corporate restructuring.<sup>54</sup>

### ***3.3 Concepts of Corporate Restructuring***

To achieve the outlined shareholder value goals in a corporate restructuring transaction, firms have three major options:<sup>55</sup> arranging a sell-off to a strategic buyer, externalizing an independent entity, and liquidating. Divestitures, or sell-offs as they are called, describe the sale of an affiliate company or business unit to a strategic buyer. In these cases the buyer normally purchases the whole business unit in order to exert full control.<sup>56</sup> The second group of restructuring concepts involve the separation of an independent entity, and include spin-offs, split-offs, split-ups, subsidiary initial public offerings (IPOs) and equity carve-outs. Spin-offs describe the change in ownership of the affiliate company from the parent corporation directly to its shareholders. Each owner receives a stake in the affiliate, which depends on the equity participation in the holding corporation. After the transaction, the shareholders own two different corporations and the subsidiary is legally independent from the holding corporation. Split-offs differ from spin-offs in that shareholders must choose between participating in the original corporation or in the newly independent subsidiary. Split-ups describe the break-up of a firm into two or more independent companies. While the holding corporation is liquidated, its former shareholders receive a participation in the new companies. Subsidiary IPOs describe the partial sale of an affiliated company via sales of shares in the stock market in which the holding corporation may lose its control over the affiliate.<sup>57</sup> Equity carve-outs describe the sale of a minority participation of an affiliate on the stock market. After this transaction, the holding firm remains in control of the affiliate.<sup>58</sup>

Finally, the third class of restructuring options is liquidation, which is different from the first two options because the individual assets of the unit are sold and the proceedings are distributed among its shareholders.<sup>59</sup>

### **3.4 Empirical Evidence of Restructuring Potential**

The two most important theories to prove the value potential of corporate restructurings are the management efficiency hypothesis and the information hypothesis.<sup>60</sup> The management efficiency hypothesis posits that the management of large corporations is generally unable to address the unique peculiarities of each segment in diversified corporations. Its performance is therefore inferior to smaller specialized firms. Consequently, management should restructure the corporation by bundling corporate resources in its core expertise. The information hypothesis posits that the information that investors get about the separate businesses of conglomerates is low. Consequently, financial market participants often penalize them with a conglomerate discount in the valuation.<sup>61</sup> After a restructuring, more information about the individual business unit is processed, which lowers the conglomerate discount and increases shareholder value.<sup>62</sup> These two hypotheses have been tested extensively by a variety of researchers in the financial fields, indicating that corporate restructuring offers an opportunity for companies to create shareholder value.<sup>63</sup> First, Achleitner (2000) and Charifzadeh (2002) summarize a cluster of studies that confirm both hypotheses. Similarly, Bowman, Singh, Useem, and Badhuri (1999) summarize a set of studies and find positive effects of portfolio restructuring based on both hypotheses. However, following them, spin-offs create more value than sell-offs. Diverging from this school of thought, Gaughan (1999) summarizes studies that indicate abnormal returns result from sell-offs, spin-offs and voluntary liquidation. Finally, Weston, Chung and Siu (1998) arrive at similar conclusions in their review of studies on divestitures, spin-offs, split-ups, and equity carve-outs. Müller- Stewens, Schäfer, and Szeless (2001) confine the outlined findings by emphasizing that the effects from restructuring vary substantially depending on the given context. This insight is the trigger for analyzing the opportunities and limitations of idiosyncratic restructuring options with System Dynamics modeling.

## **4. SYSTEM DYNAMICS MODELING IN CORPORATE RESTRUCTURING**

The previous sections have outlined the theoretical framework. Based on these insights we develop a System Dynamics model for a spin-off. The analysis is therefore limited to the case of a spin-off, i.e., the case of running a business unit separately. It is furthermore abstracted from other factors that exist in reality, such as debt financing. Once the general idea is clear, we can extend the model and include more details.<sup>64</sup> Thereby, we elucidate the general application of System Dynamics in the outlined research context and devise a roadmap required for subsequent research projects.

### **4.1 Case Study Description**

In the following analysis we apply data from industrial corporation A, which consists of two business units X and Y. Although the corporation is profitable, it operates in a difficult environment. Therefore, its management is considering a spin-off in order to concentrate on its core competencies. As a consequence, the CEO of corporation A initiates a consulting project based on a System Dynamics approach to evaluate the portfolio restructuring strategy. Management uses a simplified discounted earnings approach to determine the alternative corporate values.<sup>65</sup> Following the ideal modeling process, we have formulated the underlying problem and continue our analysis by identifying the vital variables, relationships, and dynamics to design the model.<sup>66</sup> Once the model is validated, we must run

the sensitivity analysis to derive the respective implications for the outlined decision problem. In corporation A, the two business units exhibit three vital relationships. X and Y share the same general administration. Hence, both benefit from savings in administrative overhead costs. Similarly, they work with the same marketing department, which has ambiguous implications. On the one hand, this cooperation reduces marketing expenses, but it negatively affects the market share due to imperfectly adapted marketing strategies. Finally, both divisions purchase the raw materials from the same supplier, which lowers their costs through combined purchasing power. In this basic setting of the restructuring model, we assume that all other departments such as production are the responsibility of each unit.

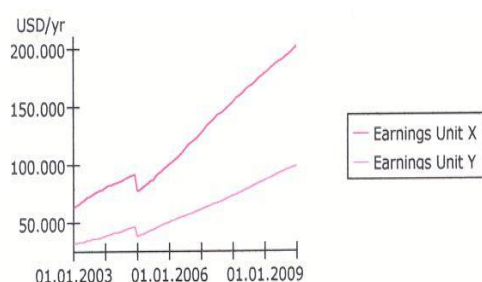
## 4.2 Limiting Variables

To select the variables for the model's design, the complexity of the decision conditions must be reduced. Since shareholder value is defined as the ultimate corporate goal, the modeling must identify all variables that greatly impact company value.<sup>67</sup> In this case, it is important to compare the earnings and the related value of both alternatives. These are assumed to be affected by raw materials, production costs, market size, market share, and the discount and constant-growth rates. In the given experimental frame, the most important relationships influencing the value drivers are those that result from the potential externalization of a business unit: the costs of raw materials and the business unit market shares. In addition, the changes in administration costs and marketing expenses must be included in the model since they are also affected by a restructuring decision. Other costs and expenses are assumed to indirectly impact the final decision. Since both units are equity financed, no interest is paid. All earnings are distributed to shareholders as dividends and no taxes are paid.<sup>68</sup>

## 4.3 The System Dynamics Model for Portfolio Restructuring

In this section, the System Dynamics model is presented to derive reasonable estimations on alternative earnings and the respective company values. Therefore, the earnings and the equity values of the two business units and the entire corporation are calculated for a seven-year horizon. The core of the model is the distinction between a spin-off restructuring and a conglomerate performance mode.<sup>69</sup> In the designed model it is possible to use the restructuring button at the top of the model to switch between the two situations before and after a restructuring. This will automatically show changes in the variables affected by a restructuring decision. The model allows us to change some of the underlying variables that are not directly affected by a restructuring in order to analyze the decision under different scenarios. To perform a sensitivity analysis, the annual total demand of market X can be varied between \$0 and \$4 million, while the annual demand of market Y can be changed between \$0 and \$2 million. Production costs can range between 25% to 50% of sales revenues.

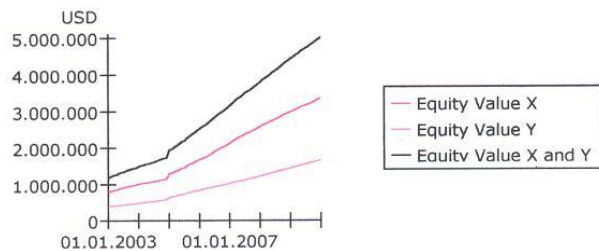
Based on empirical market surveys, the most realistic setting for A is to assume that the size of X's market is \$2 million, the size of Y's market is \$1 million and all production costs are 35% of sales revenues. If we analyze this base case scenario with a discounted earnings approach, the model derives the following results.



**Figure 4: Earnings Forecast for the Business Units. Based on own analysis.**

In the earnings forecast, the upper line represents the annual earnings of unit X, which in the first two years is a part of the corporation and the following five years, an individual entity. In contrast, the lower line reveals the development of unit Y. Under our assumptions, we can identify a linear increase in earnings until the restructuring date. The spin-off leads to a decrease in earnings and to an increase in earnings growth for both units after the restructuring. The decrease results from an immediate increase in administration and marketing expenses, whereas the increase in earnings growth derives from higher growth rates. line closest to the bottom of the graph illustrates the equity value of unit Y. In the first two years the cumulative value represents the value of the conglomerate corporation, whereas after the second year it represents the value of two independent entities.

In the representation of the residual equity value, the upper line represents the sum of the values of unit X and unit Y at each point in time in the future. The line in the middle of the chart indicates the equity value of unit X, whereas the



**Figure 5: Equity Value Forecast for the Business Units**

We can see that the equity value of the corporation increases slowly before the restructuring. Then, in the second year, the restructuring leads to an instantaneous increase in the corporate value because of a higher growth rate and greater market share.

#### **4.4 Policy Recommendations from the Model**

This basic model illustrates the general mechanics that support the decision-making process in corporate restructuring. In the specified configuration, the restructuring decision is the superior alternative based on an analysis of the created value. Consequently, based on this systems dynamics analysis the dominant strategy is to pursue the spin-off restructuring. In a next step, we can extend the basic model to a more realistic version by introducing further deterministic or new probabilistic and stochastic elements. Although these specifications might increase design costs, they increase the robustness and the explanatory potential of the model.

#### **4.5 Evaluation of System Dynamics Modeling in Corporate Restructuring**

After an analysis of the theory and the case study, we can obtain the most relevant opportunities and limitations of the application of System Dynamics modeling to portfolio restructuring.

##### **4.5.1 Opportunities**

A general advantage of modeling results from the comparison of the outcomes of different scenarios. Hence, it is possible to run a bulk of scenarios and to determine which decision is most

advantageous under a given set of data. This valuable experience improves the quality of the decision-making process.<sup>70</sup> System Dynamics are particular with respect to this aspect, because they foster an understanding of the underlying feedback loops affecting the situation. Thereby, the System Dynamics model mimics the complex interactions that take place in the real world far better than spreadsheet models.

A further vital advantage of modeling techniques is closely related to the first aspect of understanding the dynamics of the situation: the process of setting up a reliable model forces management to cooperate and to communicate. Only if all functional experts contribute to the design process will the resulting models represent the underlying mechanisms. This prompts management to develop a common representation of reality.<sup>71</sup> Communication is a particular advantage of the System Dynamics approach, as the technique works with a refined set of tools that support the conflict-oriented development of models such as causal loop diagrams.<sup>72</sup> These facilitate efficient communication between members of the management team and other experts, which is vital to corporate restructuring. A third advantage of the System Dynamics approach is its scalability, which is of particular interest in corporate restructuring. Making decisions requires a comprehensive analysis of the situation within a short period of time. System Dynamics modeling closes this gap and provides a trained analyst within a short period of time with reliable information, if the model design and the input data are accurate.

#### ***4.5.2 Limitations***

Despite these advantages, modeling approaches in general, and System Dynamics in particular, have specific limitations.<sup>73</sup> In the development of any model, the designer should be reminded that a perfect replication of the complexity of a process or situation, in reality, is inefficient and useless. A perfect model that contains all relevant information of the original system is not easier to analyze than reality. It ignores the cognitive limitations of humans and does not help simplify the system's complexity.<sup>74</sup> This observation concerning the optimal level of complexity in designing a model is at the core of modeling. The process of restructuring, in particular should not be seen as an isolated event, but as a multidimensional and long-term process leading to a series of transactions.<sup>75</sup> The simplification process must be addressed with the required respect as it represents a form of art, which should balance the minimum required level of complexity and the maximum level of simplification. A second disadvantage of modeling is the problem of ex-ante assumptions about the underlying system before a model is designed. These assumptions often bias the model and influence the final outcome of the analysis towards the desired outcome.<sup>76</sup> Therefore, corporate restructuring decisions should not be based solely on one model. Furthermore, the System Dynamics approach should be viewed as a complementary tool that enhances the quality and effectiveness of conventional approaches.

Finally, modeling processes require considerable time and monetary resources.<sup>77</sup> Therefore; a cost-benefit analysis should be the first step in the design of a model so as to clarify the goals, upside potential, and costs. This evaluation is a key step to convince the top management to support the project. Corporate management's participation, in turn, is vital for the overall success of the restructuring and determines the related level of trust. This analysis demonstrates the high potential for System Dynamics modeling for corporate restructuring if the limitations outlined here are respected.

## 5. CONCLUSIONS AND OUTLOOK

This final section summarizes the main findings of the research based on the theoretical analysis and concludes with suggestions for subsequent research questions.

### 5.1 Conclusions

This article outlines the potential role of System Dynamics modeling in the context of portfolio restructuring. The goal of this research has been to determine whether and how the System Dynamics approach can be applied in corporate restructuring processes in order to improve the quality of decision making. After establishing the theoretical background, the System Dynamics model was applied to a portfolio restructuring strategy.

The analysis demonstrates that sophisticated decision-support tools can be required in portfolio restructuring and that the System Dynamics approach fulfills this requirement for a variety of reasons.<sup>78</sup> It allows the simple implementation of a comprehensive sensitivity analysis for different sets of assumptions and alternative decisions.<sup>79</sup> Thereby, it assists management in developing an understanding of nonlinear dynamics in restructuring.<sup>80</sup> Finally, the analysis of the opportunities and limitations identifies the scalability and the communication potential of the System Dynamics approach. Consequently, a key part of the complexity of the restructuring process can be reduced to a minimum required level with a System Dynamics model. On the other hand, the system analyst must consider the increasing marginal costs for the information that must be compared to the decreasing insights resulting from more complex models. This trade-off is outlined in the case study and determines the optimal level of complexity for the model design.

Overall, System Dynamics models should be regarded as complements to conventional instruments in corporate restructuring as they enhance the quality of the decision-making process.

### 5.2 Outlook

The scope of this paper is limited to specific research aspects to enable a detailed analysis of System Dynamics modeling in portfolio restructuring. This focus and the findings of the analysis should inspire similar research avenues. Although several advantages of the System Dynamics approach have been outlined in this analysis, direct benchmark analyses with other tools would be of great value for managers. Such an analysis should particularly investigate the compatibility of different modeling approaches as it is of great interest for the practice, whether it is possible to import and export data from conventional applications.

Within the research on corporate restructuring, this article has emphasized the portfolio perspective. Subsequent investigations should assess whether the insights can be transferred to financial and organizational restructuring.<sup>81</sup> Despite of the strength of the outlined discounted-earnings approach, it can be difficult to implement it in a highly uncertain environment.<sup>82</sup> In such a setting, the capital budgeting literature suggests the application of the real options theory.<sup>83</sup> Thereby, it is possible to quantify flexibilities with a marked-based approach, if the required data is available.<sup>84</sup> In consequence, further research should investigate the possibility of combining the option pricing valuation methodology with systems dynamics modeling.

Finally, an empirical study could contribute invaluablely to the research in this area. A survey could examine whether the managers in corporate restructurings have recognized the potential of modeling in

general and of System Dynamics in particular. Moreover, such an empirical analysis might be able to identify some hurdles that prevent managers from applying the System Dynamics approach. A common problem of all outlined research projects is data availability. While it should be relatively easy to conduct a survey with a representative sample size, other information, such as details on the restructuring strategy and motivation, are difficult to obtain. Particularly, as managers in restructuring projects are extremely short of time for obvious reasons, they might not be willing to participate in a mail survey. On the other hand, possibly, the advantages and research projects outlined here could induce a broader audience to apply System Dynamics models to corporate restructuring. Consequently, positive feedback dynamics might foster the further distribution of the underlying concepts in management practice and thereby facilitate access to important, required data.

## 6. REFERENCES

1. Achleitner, A.-K. 2000. Handbuch Investment Banking. Wiesbaden: Gabler. 2000.
2. Basty, G. 1988. Die Interessen der Gläubiger in einem künftigen Sanierungs-/Reorganisationsverfahren. Munich: VVF. 1988.
3. Berztiss, A. 1996. Software Methods for Business Reengineering. New York: Springer. 1996.
4. Bianchi, C. 2002. Introducing SD modeling into planning and control systems to manage SME's growth: a learning-oriented perspective. System Dynamics Review. 18. 3. Fall. 2002.
5. Bowman, E.H., H. Singh, M. Useem, and R. Badhury 1999. When does Restructuring Improve Economic Performance?. California Management Review 2. 33-54. 1999.
6. Bowman, E. H. and H. Singh 1993. Corporate Restructuring: Reconfiguring the Firm. Strategic Management Journal 14. 5-14. 1993.
7. Brealey, R.A. and S.C. Myers 2000. Principles of Corporate Finance. New York: McGraw-Hill. 2000.
8. Brüchner, T. 1999. Reorganisationsstrategien für insolvenzbedrohte Unternehmen. Frankfurt am Main: Lang. 1999.
9. Bruppacher, P.R. 1990. Unternehmensverkauf als strategischer Erfolgsfaktor. In: H. Siegart, J.I. Mahari, I.G. Caytas and B.-M. Rumpf. Meilensteine im Management: Mergers and Acquisitions. Stuttgart: Schäffer. 1990.
10. Charifzadeh, M. 2002. Corporate Restructuring: Ein wertorientiertes Entscheidungsmodell. Lohmar: Josef Eul Verlag. 2002.
11. Chongfu, H. 2001. Fuzzy Risk Analysis Vs. Probability Risk Analysis. In: D. Ruan, J. Kacprzyk and M. Fedrizzi. Soft Computing for Risk Evaluation and Management: Applications in Technology, Environment and Finance. Heidelberg: Physica. 2001.
12. Clark, J.J., J.T. Gerlach, and G. Olson 1996. Restructuring Corporate America. Fort Worth: Dryden Press. 1996.
13. Copeland, T., T. Koller, and J. Murrin 2000. Valuation: Measuring and Managing the Value of Companies. New York: John Wiley & Sons. 2000.
14. Dillerup, R. 1998. Strategische Optionen für vertikale Wertschöpfungssysteme. Frankfurt am Main: Lang. 1998.

15. Drucker, P.F. 1999. Management im 21. Jahrhundert. Munich: Econ Verlag. 1999.
16. Fogel, D.B., K. Challapilla, and P. J. Angeline 2002. Evolutionary Computation and Economic Models: Sensitivity and Unintended Consequences. In: Shu-Heng Chen. Evolutionary Computation in Economics and Finance. Heidelberg: Physica. 2002.
17. Forrester, J.W. 1956. Dynamic models of economic systems and industrial organizations. In: System Dynamics Review. 19. 4. 2003.
18. Forrester, J.W. and others 2000. Road Maps: A Guide to Learning System Dynamics. Cambridge: MIT. 2000.
19. Gaughan, P.A. 1999. Mergers, Acquisitions, and Corporate Restructurings. New York: Wiley & Sons. 1999.
20. Gouillart, F.J. and J.N. Kelly 1999. Business Transformation : Die besten Konzepte für Ihr Unternehmen. Vienna: Ueberreuter. 1999.
21. Hammer, M. and J. Champy 1993. Reengineering the Corporation: A Manifesto for Business Revolution. New York: HarperCollins. 1993.
22. Hannon, B. and M. Ruth 1994. Dynamic Modeling. New York: Springer. 1994.
23. Hartmann, C. 1996. Mathematische Modelle der Wirklichkeit: Von der Theorie zum Computerexperiment. Thun: Deutsch. 1996.
24. Jansen, S.A. 1999. Mergers and Acquisitions. Wiesbaden: Gabler. 1999.
25. Kahraman, C. 2001. Capital Budgeting Techniques Using Discounted Fuzzy Cash Flows. In: D. Ruan, J. Kacprzyk and M. Fedrizzi. Soft Computing for Risk Evaluation and Management: Applications in Technology, Environment and Finance. Heidelberg: Physica. 2001.
26. Kaufmann, A.H. 1994. Software Reengineering: Analyse, Restrukturierung und Reverse-Engineering von Anwendungssystemen. Munich: Oldenbourg. 1994.
27. Kautt, G. and F. Wieland 2001. Modeling the Future: The Full Monte, the Latin Hypercube and Other Curiosities. Journal of Financial Planning 12. 78-88. 2001.
28. Kautt, G. and L. Hopewell 2000. Modeling the Future. Journal of Financial Planning 10. 90-100. 2000.
29. Leinenbach, S. 2000. Interaktive Geschäftsprozessmodellierung: Dokumentation von Prozesswissen in einer Virtual Reality-gestützten Unternehmensvisualisierung. Wiesbaden: Gabler. 2000.
30. Linsmeier, T.J. and N.D. Pearson 2000. Value at Risk. Financial Analysts Journal 2. 47-67. 2000.
31. Markides, C.C. 1995. Diversification, Refocusing and Economic Performance. Cambridge: MIT Press. 1995.
32. Press. 1995.
33. Müller-Stewens, G., M. Schäfer and G. Szeless 2001. Wertschaffung durch strategische Desinvestitionen. Mergers and Acquisitions 1. 13-18. 2001.
34. Odum, H.T. and E.C. Odum 2000. Modeling for all Scales: An Introduction to System Simulation. San Diego: Academic Press. 2000.
35. Oehrich, M. 1999. Strategische Analyse von Unternehmensakquisitionen: Das Beispiel der pharmazeutischen Industrie. Wiesbaden: Gabler. 1999.

36. Papageorgiou, A. and S. Paskov 1999. Deterministic Simulation for Risk Management. *Journal of Portfolio Management* 5. 122-127. 1999.
37. Perridon, L. and M. Steiner 1999. *Finanzwirtschaft der Unternehmung*. Munich: Vahlen. 1999.
38. Petersen, M.C. 1995, Erfolgreiche Unternehmensakquisition durch strategisches Wert-Management: Analyse, Planung und Steuerung von Unternehmenskäufen. Zürich: Füssli. 1995.
39. Porter, M.E. 1987. Diversifikation – Konzerne ohne Konzept. *Harvard Manager* 4. 30-49. 1987.
40. Räss, H.E. 1993. Die Restrukturierung von Unternehmen aus der Sicht der kreditgebenden Bank. Bern: Haupt. 1993.
41. Ruth, M. and B. Hannon 1997. *Modeling Dynamic Economic Systems*. New York: Springer. 1997.
42. Safavi, A. 2000. Choosing the Right Forecasting Software and System. *The Journal of Business Forecasting Methods & Systems* 3. 6-11+14. 2000.
43. Scott, R.W. 2000. Betting on Monte Carlo. *Accounting Technology* 9. 67-71. 2000.
44. Senge, P. M. 1990. *The Fifth Discipline: The Art & Practice of The Learning Organization*. New York: Currency Doubleday. 1990.
45. Shleifer, A. and Vishny, R. 1997. A survey of corporate governance. *Journal of Finance*. 52. 737-783. 1997.
46. Singh, H. 1993. Challenges in Corporate Restructuring. In: *Journal of Management Studies*. 30. 1. 147-172. 1993. Sommer, E. 1996. *Theory Restructuring: A Perspective on Design & Maintenance of Knowledge-Based Systems* Dortmund: Universität Dortmund. 1996.
47. Steinhausen, D. 1994, *Simulationstechniken*. München: Oldenbourg. 1994.
48. Sterman, J. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. New York: Irwin/McGraw-Hill. 2000.
49. Sterman, J. 2002. All models are wrong: reflections on becoming a systems scientist. *System Dynamics Research*. 18. 4. Winter. 2000.
50. Trigeorgis, L. 1996. *Real Options: managerial flexibility and strategy in resource allocation*. Cambridge MA: MIT Press. 1996.
51. Troitzsch, K.G. 1990. *Modellbildung und Simulation in den Sozialwissenschaften*. Opladen: Westdeutscher Verlag. 1990.
52. Van den Bosch, P.P.J. and A.C. Van der Klauw 1994. *Modelling, Identification and Simulation of Dynamical Systems*. Boca Raton: CRC Press. 1994.
53. Voit, F. 1999. *Methodik zur ereignisdiskreten Modellierung und effizienten Simulation komplexer Systeme*. Düsseldorf: VDI. 1999.
54. Weston, J.F., K.S. Chung, and J.A. Siu 1998. *Takeovers, Restructuring, and Corporate Governance*. Upper Saddle River: Prentice-Hall. 1998.
55. Zeigler, B.P., H. Praehofer and T.G. Kim 2000. *Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems*. San Diego: Academic Press. 2000.
56. Zhu, Y. and T. Backx 1993. *Identification of Multivariable Industrial Processes for Simulation, Diagnosis and Control*. London: Springer. 1993.

# A study on Technological impact on rural markets in India

Ch.Bala Raju<sup>1</sup> Dr.K.Veeraiah<sup>2</sup>

<sup>1</sup> Mr.CH.Bala Raju Professor & H.O.D, School of Management, CMR Technical Campus, Affiliated to JNTU, Hyderabad. Ph No.9246848418

<sup>2</sup>Dr.K.Veeraiah, Professor & HOD, Malla Reddy Engineering College (Autonomous) Affiliated to JNTU, Hyderabad.

[<sup>1</sup>bcraju@gmail.com](mailto:bcraju@gmail.com)

[<sup>2</sup>veeru.k88@gmail.com](mailto:veeru.k88@gmail.com)

## Abstract—

The vast country side in India is where the future markets will be. Rural Markets offer lucrative opportunities to invest in, as it is fast expanding in its size and scope. Rising affluence, good monsoons, use of modern technologies in farming and rural areas have shifted the focus of marketing gurus from the urban saturated centers to the rural potential markets. The advent of green marketing concept and organic farming techniques have added to the increase in the scope for investments in the rural markets. The high demand base coupled with elevated market potential is culminating into developing fruitful technologies that will help to enhance the market of the new millennium. The regional diversity and the growing techno-savvy consumer base throughout India, makes it the best battleground for the multinationals to showcase their innovative offerings to the rural consumers. India is growingly becoming a better choice for global marketers to pilot test their innovative products and services, before they can cater to other parts of the world. Big brands like Nokia, Godrej, Tata, HUL, ITC, P&G and the like have successfully tested and crafted their products to the rural consumers.

The present article focuses on analyzing the impact of information technology in improving rural marketing, providing an insight into opportunities available and enabling strategies for further technological advancement of rural markets. Finally the exploratory study suggests a proposed model for e-Rural Marketing in India.

**Keywords—** *Rural Marketing, e-rural marketing, rural marketing chain, Rural CRM.*

## I. INTRODUCTION

India is predominantly an agrarian economy. Rural markets offer immense potential for market growth and expansion. The market demand for agricultural input such as Farm products (Seeds, fertilizers, pesticides, etc.), equipment (including tractors) has risen by leaps and bounds. Rural markets contribute to seventy percent of the population spanning across 6, 27, 000 villages in the country. The

growth potential of the rural markets is depicted in the table below:

Products	Amount in Billion (INR)
FMCG	6500
Consumer Durables	500
Agriculture Inputs	4500
Automobiles(2 & 4 Wheelers)	500
<b>Total</b>	<b>12300</b>

Source: Mc Kinsey Global Institute

An understanding of the term “Rural Marketing” is essential to gain insights into the present article in its perspective. Rural Marketing is any marketing activity which identifies and serves the needs of the consumers living in villages. Technological advancement will prove successful if and when marketers dwell in rural areas, live rural lives and understand the nuances of Rural Consumer Behaviour. In recent times, the focus of marketing has been channelised on rural markets, as they offer huge opportunities that Multi Nationals can ill afford to ignore. Green Revolution has ushered in the intelligent use of fertilizers, herbicides, high yielding agricultural crops, modern implements and improved use of technology which has resulted in increase in disposable income of the farmers, substantial use in purchasing power, decent lifestyle and a better standard of living of the large rural population.

The growing rural markets envisages tremendous marketing opportunities for growth and expansion as they are vast, attractive, scattered, homogeneous and is nearly three times the urban markets.

## THE INDIAN RURAL SCENARIO: KAL, AAJ AUR KAL

Rural India as depicted in the 1960 to 1990's was characterized by a number of features which have been analyzed below from a dual perspective:

Consumer's Side	Marketer's Side
<ul style="list-style-type: none"> <li>• Low Connectivity (Physical Communication, infrastructure, Roads, Electricity Transport and Logistics)</li> <li>• Underdevelopment, low levels of literacy and low standard of living</li> <li>• Scattered and uneven spread of villages</li> <li>• Inaccessibility</li> <li>• Low Per Capita income and spending power</li> <li>• Seasonal demand</li> </ul>	<ul style="list-style-type: none"> <li>• Inapproachability</li> <li>• Focus on the attractive urban markets</li> <li>• Myopia towards urban markets</li> </ul>

The rural markets may be beset with problems but are alluring and are virgin markets waiting to be explored. They pose challenges of availability, affordability, acceptability and awareness (4 A's).

Tapping and capturing these markets can prove a mckinsey's gold for the future marketing gurus. Here the focus of the marketer should be initially change the mindset of the rural folk to think urban. Innovation is the key. Companies need to innovate profusely with their products and services. They need to design them specific to the rural community. Pricing and credit facilities also poses a great threat. If companies want to sell to rural consumers, they need to extend credit till harvest time and accept payment in grains and also have oodles of patience when it is receivable time.

Technology and communications have penetrated into the hinterland of the villages. Education and employment have resulted in a shift from ignorance to awareness, poverty to plenty and greater access to products and services like never before.

**The Present Scenario:** The trends in the rural communities have shown a marked change.

Technologies and strategies being propagated in the rural areas are urban based and ill suited to the needs of the rural folks. The products and services provided by the technology developers only cater to the needs of the elite farmers as a result of which the poorer section of the rural masses and are left behind and neglected, adding to their woes. Moreover, the mindset of the rural fraternity needs to be made in tune with that of the technology developers.

The hall mark of rural markets should be innovation. Technologies developed should be backed by good after sales service, convenience, affordable prices and they have to be tailor made to suit the requirements of the rural areas. An industry-funding agency interface can benefit the lower strata of the population and can also act as a catalyst to employment generation.

## Future of E-Rural Marketing in India: An Overview

Rural India accounts for 833 million living in India, (census 2011). They have been exposed to new communication technologies and a gamut of services. The rural segment is growing at a pace of 9 – 10% per annum (Technopak report) and is anticipated to add new consumption of US \$90

billion – 100 billion over 2012-2017. Moreover according to ASER (2012) report, private school enrollment in rural India has risen by 5.5% over the past six years and the literacy rate has gone up by 68.91% of rural India.

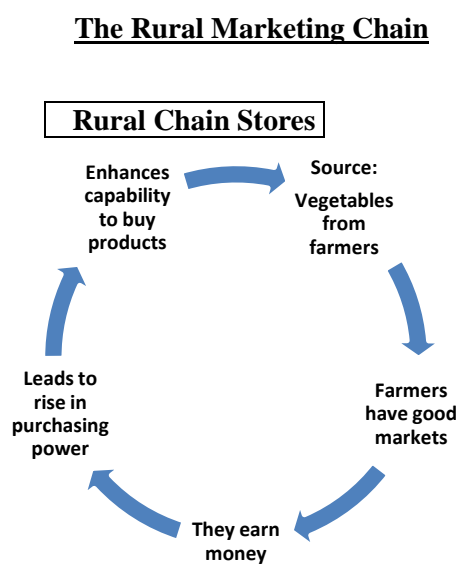
The future of rural marketing is in retail, e.g.: DCM Shiram's rural retail arm "*Hariyali Kisan Bazaar (HKB)*" is planning to open 10 outlets in 2012, covering 230 stores across 8 states.

FMCG: big time players as Dabur, ITC, Colgate, and Godrej have penetrated into the heartland of rural areas.

Automobiles: Rural India with its vast size is attracting the automobile makers in a huge way. They account for almost 35% of the automobile industry sales. Players like Ford, Toyota, Hyundai, and General Motors are in the fray to offer financing services, retail banking services and rural reach to attract its rural fraternity.

Rural marketing efforts have also led to corporate social marketing initiatives like providing employment opportunities to rural people and thereby increasing their purchasing power.

Chain stores have stormed the rural bastion, which will ultimately lead to a cycle or chain as evolved below:



The rural chain stores will not only prove to be a boon to the farmer in selling his agricultural produce but also will generate employment opportunities, thereby resulting in improving the standard of living of the rural folks.

### **A Proposed Model for E-Rural Markets:**

Information technology is the magic wand for the future of E-Rural Marketing in India. Technology can provide timely information for providing effectiveness to the 4P's of marketing in the rural areas. Farmers can be facilitated to sell their products effectively at remuneration prices, gain access to credit facilities at the right place and time and gain access to newer and greener avenues of selling agricultural products.

### **A Workable Model of E-Rural Marketing:**

The proposed model provides linkages in six areas

**Rural Re-Engineering Initiatives:** Innovative practices to develop the rural areas are envisaged in this initiative. The back to the roots ideology is one such move which encourages the rural youngsters to educate themselves outside their village (earn foreign degrees) but they should endeavor to come back and serve their rural community. This will lead to the wellbeing and development of the place. This practice will ensure that there is a check on brain drain in India. Central Government is encouraging Entrepreneurship in the rural areas by floating many a employment schemes like, PMRY(Prime Minister's Rozgar Yojana), SGSY(Swarnajayanti Gram Swarozgar Yojana), REGP (Rural Employment Generation Plan), etc, to name a few.

Initiatives by the state governments are also well framed to benefit weaker sections of the society and the downtrodden. The recent example of Andhra Pradesh government promoting the Pavala Vaddi Scheme (Interest Subsidy) during the year 2004-05 with an objective to provide interest subsidy on the

bank loans taken by the Self Help Groups in Andhra Pradesh to reduce the financial burden on them is a good example of promotional schemes in the different states.

**Rural Marketing Information system:** In order to provide timely and needed pertinent information an effective information system is the need of the hour. Rural India needs to be connected via Broadband (Co Options technologies and Walden international are already in the fray to connect 2740 rural co-operative banks on its technology solution). Connectivity via broadband using World Wide Web, mobile phones can be used for information gathering and dissemination via world space satellite radio. Remote areas can be connected deploying the services of a mobile communication node (MCN) which consists of a computer/laptop, solar panel, adaptor, portable printer and the world space satellite radio receiver. Providing marketing information services can be of great help to the small rural producers and the big and small businesses.

The integration of all the rural markets across India through a Management Information system will provide and benefit the Multinational companies to fasten their product testing and generation process, thus benefitting the global society at large.

A thorough training and inculcation of Information Systems in all functional areas and the extensive use of Enterprise Resource Planning Systems (ERP) along with Business Intelligence should be provided to the rural youth to ensure efficient and effective marketing solutions to the rural markets.

**Rural-Corporate Social Responsibility** Rural social initiatives undertaken by the corporates help the poor people to have spendable income on product categories Hindustan Unilever Limited has created an income generating opportunity called Project Shakti and Shaktimaan which provides employment opportunities, distribution network to rural entrepreneurs. The ITC'S e-choupal initiative is the sole largest information

technology based initiative which provides enriching and customized information to improve the farmers decision making abilities, securing better quality and productivity price. It also facilitates a direct

marketing channel, is logistics efficient and generates considerable income for farmers spanning one sixth of rural India. There are other corporate initiatives of TATA( Tata Kissan Sansar), Indian Oil, Reliance Rural Hub and Future Groups Aadhar which enable the farmer to sell his produce , get credit and insurance facilities and non fuel products.

The Aditya Birla Rural Technology Park at Renukoot, Uttar Pradesh fosters people empowerment through capacity building. The company has built a huge campus which houses many training centers mostly dedicated to the agriculture and veterinary research activities, like, animal husbandry, cattle rearing and the latest farming techniques. Apart from that the company also has a vocation training centre and a hospital which ensures the health of the rural masses thus guaranteeing a better nation.

Corporate social responsibility (CSR) can be turned into a business opportunity too. Infosys Technologies is experimenting with the BPO centres in rural areas to exploit the workforce which is available in abundance in rural India. Infosys is targeting FMCG Companies (fast moving consumer goods companies), which are increasingly developing and growing by venturing into the rural areas.

**Green Marketing:** Synonymously known as Ecological Marketing or Environmental marketing or sustainable marketing. Green marketing is fast picking up pace in the rural areas and more so in the back bone sector of our economy, the agricultural sector. Farmers are coming up with innovative ideas with the help of many companies dedicated to the cause of the environment as a part of their Corporate Social Responsibility initiatives. The lift irrigation programme, which ensures the availability of water

throughout the year, irrespective of the weather conditions is a rural integrated development system taken up by the corporate major, TATA in conjunction with GVK under the Gram Vikas Kendra project, in Jamshedpur. This has ensured that the farmers have substantial incomes throughout to invest in farming for the whole year and also has improved the fertility of the soil and thus made it more arable.

Mass planting of trees and avoiding soil erosion, the use of drip irrigation, the advent of herbal pesticides (Herbicides) and organic fertilizers, innovative water harvesting methods and conservation of natural resources are few of the successfully implemented initiatives in the rural development scenario. Green warehousing, vermicomposting, electricity generation from biological waste or bio-fuel manufacturing, use of tidal waves and wind energy in the production of electricity apart from solar energy utilization for farming purposes and production activities and the like are the few of the commendable initiatives taken up in the rural areas for higher productivity and a greater cause of the preservation of environment. For example, Jolly TV, a local TV Manufacturing company in Uttar Pradesh is producing televisions which run on rechargeable battery systems, to ensure uninterrupted entertainment to the consumers, despite many power cuts in rural India.

**E-Rural Social Marketing:** Using the electronic media to communicate the advantages of products or brands that have the specific objective of creating, executing, achieving and further controlling an influence of social change or overall societal well being are included in this arena. The different commercials aired by the government and the NGO's (non government organizations)/NPO's (non-profit organizations), about the responsible consumption patterns in the rural households, the educative advertisements on the soil testing and fertilizer usage patterns to protect the soil from erosion; establishment of 24 hour, toll free call centers for counseling of the farmers on various issues relating to farming practices; consumer forums and consumer protection cells for the various

issues relating to the retailer and consumer transactions, the legal provisions available for the same, and so on are included in the electronic rural social marketing.

Most of the organizations are increasingly using this medium to attract and bring about a change in the mindset of the society.

**E-Rural Marketing Chain:** The distribution network is one of the vital factors in rural marketing and it is observed that most of the companies are shifting their focus onto building a reliable network of supply chain which is custom made to suit to the requirements of the rural producers/ rural marketers. Most of the rural distribution chains are uncomplicated and mostly include only the key success factors.

HUL, the FMCG giant and the maker of soaps such as Lux and Lifebuoy, launched a project named Shakti in the year 2001, employing mostly women for last-mile distribution in the rural areas. **Short Supply Chains:** typically includes all forms of marketing involving direct links between food producers and consumers. Some of the famous examples of the same include: Sales at the farm, farm shops, Haats, Mandis, Rythu Bazaars (Farmer's Market), etc.

#### **Farmer/Producer → Consumers**

V.G. Ram Kumar, Founder, Arunteccame up with the idea of building a two way business channel between rural entrepreneurs and buyers in urban markets using Information technology; which included online deal of products developed by rural artisans like, weavers, potters, handicraft makers as well as agricultural produce.

Amul is a classic example of a firm that has built on its strength and integrated technology for the purpose of reaching a larger customer base and building a strong supplier network. The introduction and successful implementation of Cooperative model to reach its consumers in the

rural areas and urban areas, including various nations globally, is the forte of Amul. Activities at the village level comprised of developing and servicing the Village cooperative societies, increasing milk collection, procuring milk, and transporting it to the chilling and processing units twice a day. The villages with no chilling processing units were mainly catering to the local demands. The successful integration of the main stakeholders, the cooperative society, producers and suppliers ensured a great supply chain management model for Amul.

**E-Rural Services Marketing and C.R.M:** Services and relationship management are the key to building bridges and opening doors in the rural bastion. Some of the prominent initiatives planted by corporates are:

**Digital Gangetic Plan (DGP) - a Wi-fi project** connecting some villages in Uttar Pradesh has been successful in providing internet access to the rural population. The people are benefited by the various modules like Bimari Jankari (information relating to health and diseases) and Digital Mandi (information pertaining to agricultural commodities, knowledge of relative prices, etc.) of the project.

Launch of e-governance services through E-sevas in rural Andhra Pradesh in 2003, enabling the people to pay their utility bills, passport services and receive certificates and the like at their doorstep. Vihaan Networks Ltd. (VNL) has come up with GSM towers called the WorldGSM which work entirely on solar energy, whose technology can help many networking organizations to foster their business model of GSM, and broadband services in the rural markets. The trial version of the project has been set up in Bhiwandi tehsil in Rajasthan.

Samsung Electronics Company Limited has launched a trial version of SOLAR GURU, a made in India mobile that operates on solar power. The shortage of power supply does not hinder the recharge of mobiles for the consumers.

Kisan Credit Cards have been issued to over 41 million people in rural India. Maintaining cordial and continual relationship with them will be beneficial to companies. Moreover, thanks to technology, net based initiatives are penetrating the rural soil offering galore opportunities to develop create demand for products, services and generate incomes.

## Conclusion

Change is the order of the day and business organizations need to improvise and adapt in order to survive in the market. Since the urban markets have reached the saturation point, it is imperative that rural marketing will offer the much needed push in the form of wider consumer base, gain volumes and profitability. Simple technological solutions can create networks in rural areas which can be highly useful to farmers in farming purposes, selling their produce effectively. This will ultimately improve the economic condition of the rural folks. Rural markets are the rich investment hub for future. Marketing Gurus need to innovate, make marketing plans and programmes to fit the rural market.

As quoted by the marketing guru, Philip Kotler, there are only two basic businesses of any organization, to market the produce and to innovate. Organisations which innovate shall flourish, while the others are likely to perish.

**“The rural markets are the indicators of the health of a nation at its grass root level; the stronger the roots, the healthier the nation in terms of economy, prosperity and affluence.”**

## References:

1. Dr R L Varshney, Dr.S L Gupta, “Marketing Management- An Indian Perspective”, Sultan Chand and Sons, 2000, Pg- 674 -684.
2. Biplab .S.Bhose, “Marketing Management”, Himalaya Publishing House, 2009, Pg. 626-659

3. <http://www.123eng.com/forum/f12/rural-retail-marketing-innovation-creativity-97928/>
4. <http://www.livemint.com/2010/03/21210409/Marketeers-flock-to-rural-Indi.html>
5. <http://www.pagalguy.com/2011/11/rural-marketing-offers-limitless-opportunities-for-innovation-pradeep-kashyap-ceo-of-mart/>
6. [http://en.wikipedia.org/wiki/Rural\\_marketing](http://en.wikipedia.org/wiki/Rural_marketing)
7. <http://www.seminarprojects.com/thread-emerging-trends-in-rural-marketing>
8. C N Sontakki, "Marketing Management", Kalyani Publishers, 2007, Pg 620-626
9. Anil .K Rajvanshi, "Development of Technologies for Rural Areas- Need for new Thinking", CAPART, Moving Technology, Vol. 7, No.1, March 1992, Pg.2-5
10. <http://www.moksha-yug.in/pdf/livemintMYA.pdf>
11. <http://csridentity.com/philanthropy/rural.asp>
12. [http://www.brandchannel.com/images/papers/522\\_2011-03\\_India\\_Rural\\_Marketing.pdf](http://www.brandchannel.com/images/papers/522_2011-03_India_Rural_Marketing.pdf)
13. T.P. Gopalaswamy, Rural Marketing: Environment, Problem, and Strategies. New Delhi:Wheeler Publishing ,2000, Pg 1-130
14. Philip Kotler, K. Keller, A.Koshy and M.Jha 2007. Marketing Management: A South Asian Prespective. New Delhi: Pearson Prentice Hall, 13<sup>th</sup> Edition, Pg no.12
15. <http://www.rural-impact.net/bestanden/81512198102.html>
16. <http://www.ibef.org/industry/consumermarket/ruralmarket.aspx>
17. Menda.A, ICT for improved Crop Marketing in Rural Tanzania, [www.icconnect.online.org](http://www.icconnect.online.org), 2005
18. <http://www.ifad.org/events/lectures/tradenet/index.htm>

**"A STUDY ON COMPARATIVE ANALYSIS OF SELECT ETFs' AND INDEX BASED MUTUAL FUNDS WITH NIFTY"**

Mr. P.S.Viswanadh, Assistant Professor, SMS, SNIST and  
Ms. V.V.L.Sameera, Assistant Professor, SMS, MREC

**ABSTRACT**

*The present study entitled "A Study on Comparative Analysis of Select Exchange Traded Funds (ETFs') and Index based Mutual Funds with NIFTY" is objected to measure and compare the risks and returns of the Select Equity Exchange Traded funds and Index Based Mutual Funds and to analyze the correlation of Select Equity Exchange Traded funds and Index Based Mutual Funds with NIFTY. The data associated to Net Asset Values' (NAV) of Select ETFs' and Index based mutual funds is collected for the period of three years from April 2012 to March 2015. The statistical tools that are used to attain the desired objectives are Mean return, Standard deviation, Correlation, Wilcoxon signed rank test and F test. The study concludes that the performances of select index based mutual funds are better when compared to ETFs' since the Index based mutual fund's returns are high and risk is less as compared to ETFs'.*

*Key words: Exchange Traded funds (ETFs'), Index based mutual funds, Net Asset Value.*

**Introduction:**

Globalization has created tremendous growth in the industrial sector in India. It has provided many more opportunities to boost up the growth of Indian economy as well as thrown many more challenges to industrial sector in general and Stock Markets in particular. In these volatile conditions, the investors showed interest to invest in Index based mutual funds and Exchange Trade Funds (ETFs') to minimize the risk of their investments. The launch of new products and the entry of new categories of investors will boost the Indian exchange-traded fund (ETF) market<sup>1</sup>. It is also expected that Indian ETF market will grow 20 per cent every year<sup>2</sup>. Growth in new products and allowing new investors such as pension funds will drive this growth. On the other hand Index based mutual funds were also performing well in our country. The statistics reveals that the investors preferred to invest in ETFs' to that of Index based mutual funds in the recent days due to various reasons like flexible investment vehicle, low expense ratio, size and time horizons. In this regard the researcher thought to make a comparative analysis between ETFs' and Index based mutual funds.

**Review of Literature:**

- P.Krishna Prasanna (2012) examined the performance of ETFs' in India
- Reena Agarwal (2012) conducted a study on growth of global ETFs and regulatory challenges.
- S.Kevin (2013) in review of integrative business and economics research analysis empirically the price behavior of stock market index, index futures and index ETFs' in the Indian securities market.
- Yanyan Li in (2013) examined the performance evaluation of exchange traded funds in emerging markets.
- HortenseBioy (2013) conducted a study on comparative analyses of the costs of investing in ETFs' and index funds across different asset classes.
- Kapil Arora (2014) in his article "An analysis of exchange traded funds performance in India stock market"

<sup>1</sup>.said by Deborah Fuhr, managing partner of ETFGI on 27<sup>th</sup> October, 2015.

<sup>2</sup>Interview with the editor of Smart investor business – Standard by By Samie Modak .

- Dr.P.vudhyapriya and Dr.M.Mohansundari (2014) made a study on the performance of gold ETF in India published by Indian Journal Applied of Research.
- Mukesh Kumar Mukul, Virat Kumar (2014) studied comparative performance analysis of Gold ETFs' on monthly basis.
- Stoyi.I.Ivanov (2014) has studied the influence of ETFs' on the price discovery of gold, silver and oil.
- Itzhak Ben-David, RabihMoussawi (2015) has analyzed the impact of ETFs' on volatility of equity stocks.

#### **Research Gap:**

The earlier studies were carried out on conceptual frame work, performance and cost analysis of Exchange Traded funds and Index Based mutual funds independently and they have not focused on comparative analysis both categories of funds. Therefore the researcher made an attempt to compare the relative performance with risk and return characteristics of ETFs' and Index Based Mutual funds.

#### **Objective of Study:**

1. To measure and compare the risks and returns of the Select Equity Exchange Traded funds and Index Based Mutual Funds.
2. To analyze the correlation of Select Equity Exchange Traded funds and Index Based Mutual Funds with NIFTY.

#### **Hypothesis:**

H<sub>01</sub>: There is no significant relation between NIFTY and selected Index funds.

H<sub>02</sub>: There is no significant relation between NIFTY and Selected ETFs'

#### **Scope of Study**

The study has considered six ETFs' and Index Based Mutual funds such as Quantum ETF, Goldman Sachs NIFTY BeES ETF(NBES), Motilal Oswal Most Shares Midcap 100 ETF(M100 NIFTY), Motilal Oswal Most Shares Midcap 50 ETF(M50 NIFTY), Kotak PSU Bank ETF, Goldman Sachs CNX NIFTY Shariah BeES ETF (SBET), HDFC Index Fund, IDFC NIFTY Fund, Reliance Index Fund, Franklin India Index Fund, Principal Index Fund, SBI NIFTY Index Fund.

#### **Source of Data:**

The present study is based on secondary data, collected from various websites like Association of Mutual Fund Industry (AMFI), NSE India, Yahoo Finance etc.

#### **Period of the study:**

The study has considered 3 years data from the period of April 2012 to March 2015.

#### **Tools and Techniques used for the study:**

The statistical tools used for the study are Mean for Return calculation, Standard Deviation for risk measurement, correlation test to find relationship between fund's NAV's and NIFTY Index, Wilcoxon signed rank test for comparing the return and risks of ETFs' and Index based mutual funds and F-test for hypothesis test. The researcher used Statistical package for Social Sciences (17 version) for statistical calculations.

#### **Limitation of the Study**

- The major limitation of this study is that the sample size was reduced considerably due to the non-availability of data.
- The results obtained cannot be generalized.

### Data Analysis

The present study entitled "A Study on Comparative analysis of the Select ETFs' and Index funds with NIFTY" is an attempt made by the researcher to analyze and compare the return of ETFs' and index funds, for that he has considered the select ETFs' and index funds.

In order to make effective presentations the researcher used the techniques like Mean, standard deviation, correlation, T-test. The brief discussion about the techniques is as follows:

#### Mean Return:

Mean return is the technique used to calculate the aggregate return of certain variable for the select sample. Following the formula to find out the mean return:

$$\text{Mean Returns in Percent} = ((P_1 - P_0)/P_0) \times 100$$

Where  $P_1$  is Last date's Net Asset value (NAV) of certain fund,  $P_0$  is the beginning day's NAV of the same fund.

#### Standard Deviation:

Standard deviation is also known as historical volatility and is used by investors as a gauge for the amount of expected volatility. In simple words to calculate the deviations and the variations (Risk) of the stocks prices in different indices to get the Maximum returns. The symbol for Standard deviation is S (The Greek letter sigma).

The Formula of Standard deviation:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}, \text{ where } \mu = \frac{1}{N} \sum_{i=1}^N x_i.$$

$\sigma$  is standard deviation,  $x_i$  is each value of dataset,  $\bar{x}$  (with a bar over it) is the arithmetic mean of the data (This symbol will be indicated as mean from now),  $N$  is the total number of data points,  $\sum (x_i - \text{mean})^2$  = The sum of  $(x_i - \text{mean})^2$  for all data points.

#### Correlation:

It is a statistical measure of how two securities move in relation to each other. Correlations are used in advance portfolio Management. Correlation is computes into what is known as correlation coefficient, which ranges between -1 to +1.

Perfect positive correlation implies that as one security moves, either up or down, the other security will move in lockstep, in the same direction. Alternatively, perfect negative correlation means that securities are moved in opposite directions. If the correlation is 0, the movements of the securities are said to have no correlation, they are completely random. The symbol for correlation is 'r'.

The mathematical formula for computing  $r$  is:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

#### Wilcoxon signed rank test:

This test is a non parametric test used to rank variables. This technique has been used in the present to compare the returns and risks of Index based mutual funds and ETFs' and to find the better performing option among them. The results of test can be interpreted as the variable, whose mean rank is high, then, mentioned statement could be true and the variable's whose mean rank is less, that statement could be considered as false statement.

**F Statistic:**

It is used to prove the hypothesis. The test statistic significant value or probability value is less than 5 percent or 0.05 the researcher can reject the null hypothesis where as it is more than 5 percent, one will not have enough evidence to reject null hypothesis.

**Return analysis of Selected ETFs'**

Table no: 1

S.No	Name of the fund	Returns in percent
1	Quantum ETF	40.90881
2	Goldman Sachs NIFTY BeES ETF (NBES)	37.97361
3	Motilal Oswal Most Shares Midcap 100 ETF(M100 NIFTY)	41.23541
4	Motilal Oswal Most Shares Midcap 50 ETF(M50 NIFTY)	10.2284
5	Kotak PSU Bank ETF	1.6124866
6	Goldman Sachs CNX NIFTY Shariah BeES ETF (SBET)	33.90386

Source: Compiled data

The above table reveals that the chosen ETFs' are having positive returns in the selected period. Motilal Oswal Most Shares Midcap 100 ETF (M100 NIFTY) has offered highest positive return (41.23541) among the selected funds. The second highest return is by Quantum ETF and Goldman Sachs NIFTY BeES ETF (NBES) with the returns of (40.90881) and (37.97361) respectively but the least performance is by Kotak PSU Bank ETF since it is having 1.6124866 returns.

**Risk Analysis of Selected ETFs'**

Table no: 2

sl. no	Name of the fund	Standard Deviation
1	Quantum ETF	266.0878
2	Goldman Sachs NIFTY BeES ETF (NBES)	235.9176
3	Motilal Oswal Most Shares Midcap 100 ETF(M100 NIFTY)	3.943393
4	Motilal Oswal Most Shares Midcap 50 ETF(M50 NIFTY)	6.061885
5	Kotak PSU Bank ETF	4.3253722
6	Goldman Sachs CNX NIFTY Shariah BeES ETF (SBET)	44.56767

Source: Compiled data

From the above table it is clear that Motilal Oswal Most Shares Midcap 100 ETF (M100) relatively possess less risk (3.943) and then the second lowest risk is by Kotak PSU Bank ETF (4.325) when compared to other funds. Among the remaining funds Quantum ETF and Goldman Sachs NIFTY BeES ETF (NBES) posses high risk (266.08) and (235.9176) respectively.

Return Analysis of Select Index Funds

Table no: 3

S.No	Name of the fund	Returns in Percent
1	HDFC Index Fund	39.99391
2	IDFC NIFTY Fund	41.00276
3	Reliance Index Fund	39.79001
4	Franklin India Index Fund	39.06787
5	Principal Index Fund	39.27236
6.	SBI NIFTY Index Fund	38.4494

Source: Compiled data

The above table reveals that the chosen Index Funds are having positive returns in the selected period. IDFC NIFTY Fund has offered highest positive return (41.00276) among the selected funds. The second highest return is by HDFC Index Fund with the returns of (39.99391) but the least performance is by SBI NIFTY Index Fund since it is having (38.4494) returns.

Risk Analysis of Selected Index Funds

Table no: 4

Sl.No	Name of the fund	Standard Deviation
1	HDFC Index Fund	21.64913
2	IDFC NIFTY Fund	5.055248
3	Reliance Index Fund	3.982143
4	Franklin India Index Fund	19.03489
5	Principal Index Fund	16.71056
6	SBI NIFTY Index Fund	20.09265

Source: Compiled data

From the above table it is clear that the Reliance Index funds possess less risk (3.9821) when compared to other funds. The second lowest risk is by IDFC NIFTY Fund with risk (5.0552). Among the remaining funds, HDFC and SBI NIFTY possess high risk (21.649) and (20.092) respectively.

**Comparative Analysis of Return and Risk between selected ETFs' and Index based Mutual Funds**

**Table no: 5**

		N	Mean Rank	Sum of Ranks
Return of Index base mutual funds - Return of ETFs'	Negative Ranks	2 <sup>a</sup>	1.50	3.00
	Positive Ranks	4 <sup>b</sup>	4.50	18.00
	Ties	0 <sup>c</sup>		
	Total	6		
Risk of Index Based mutual funds - Risk of ETFs'	Negative Ranks	3 <sup>d</sup>	5.00	15.00
	Positive Ranks	3 <sup>e</sup>	2.00	6.00
	Ties	0 <sup>f</sup>		
	Total	6		

Return of Index base mutual funds < Return of ETFs', b. Return of Index base mutual funds > Return of ETFs', c. Return of Index base mutual funds > Return of ETFs', d. Risk of Index Based mutual funds < Risk of ETFs', e. Risk of Index Based mutual funds > Risk of ETFs', f. Risk of Index Based mutual funds = Risk of ETFs'.

Source: SPSS output

the above table (Wilcoxon Test statistics) can be interpreted that the variable, whose mean rank is high, on that statement could be true and the variable's whose mean rank is less, that statement could be considered as false statement.

the above table reveals that, the returns of Index based mutual funds are better than that of ETFs' and the risk of ETFs' are high when compared to Index Based Mutual funds since the mean rank of 4<sup>b</sup> (4.5) is greater than that of mean rank of 2<sup>a</sup> (1.50) and mean rank of 3<sup>d</sup> (5.0) is greater than that of mean rank of 3<sup>e</sup> (2.0).

**Analysis of co-relation with select ETFs' and Index Funds**

**Table no: 6**

S.no	Description	Correlation	Significant Value
1	NIFTY & Quantum	0.978	0.000
2	NIFTY & NBES	0.977	0.000
3	NIFTY & M100	0.947	0.000
4	NIFTY & M50	0.634	0.000
5	NIFTY & Kotak PSU	0.466	0.004
6	NIFTY & SBET	0.967	0.000
7	NIFTY & HDFC	0.977	0.000
8	NIFTY & IDFC	0.978	0.000
9	NIFTY & Reliance	0.979	0.000
10	NIFTY & Franklin	0.977	0.000
11	NIFTY & Principle	0.978	0.000
12	NIFTY & SBI	0.976	0.000

the above table reveals that there exists highest positive co-relation among various selected funds with NIFTY except Motilal Oswal Most Shares Midcap 50 ETF and Kotak PSU (0.634) and (0.466) respectively which presents the selected funds returns where fluctuating when there exists fluctuation in NIFTY.

# A Study on Women Entrepreneurs in Ranga Reddy District, A.P., India.

MS A.LATHA, MS.GV VENELA, MS.P.RAJITHA

ASSOCIATE PROFESSOR, MBA, MREC

ASSOCIATE PROFESSOR, MBA, MREC

ASSISTANT PROFESSOR, MBA, MREC

LATHA\_VENNI@GMAIL.COM

PRAJITHA\_KJKC@GMAIL.COM

## ABSTRACT

Women entrepreneur development is an essential part of human resource development. Entrepreneurship amongst women has been a recent concern. Women have become aware of their existence their rights and their work situation. However, women of middle class are progressing through single business in urban areas.

This paper focuses on women entrepreneur. The paper talks about the status of women entrepreneurs and the problems faced by them when they ventured out to carve their own niche in the competitive world of business environment. Our study focused on sample of 50 women entrepreneurs in ranga reddy district, where the information is gathered through questionnaire and analyzed on socio economic changes.

## INTRODUCTION

“If women didn't exist, all the money in the world would have no meaning” a quoted by **Aristotle Onassis**, it is true in the current scenario, the women became the strongest entrepreneur all over the world. Women entrepreneurs are fast becoming a force to reckon with in the business world and are not only involved in business for survival but to satisfy their inner urge of creativity and to prove their capabilities.

The development of women entrepreneurship is very low in India, especially in the rural areas. Gender equality is a prerequisite for poverty reduction because of the contribution women make and the role they play in society and in the economic well-being of the family and communities. Be it in rural or urban areas, be it in micro or medium and large enterprises, women must be an integral part of development, not only as beneficiaries, but also as decision-makers and agents of change.

Both Women and girls consist three-fifths of the world's poor. Their poverty level is worse than that of men as clear gender disparities in education,

employment opportunities and decision making power exist. A large number of women are mainly engaged in subsistence agriculture as well as in micro and small-scale enterprises (MSE).

In addition to their economic and income-generating activities, women assume multi-faceted roles in society, i.e. as breadwinner of family, unpaid family workers, service providers in the communities and mother/care-taker of the family. To respond to the needs of women to materialize their economic potential and thereby to improve their standard of living, it is necessary to design programmes by applying a mainstreaming strategy. This requires devising measures to integrate women as decision-makers, participants and beneficiaries in all relevant development activities, irrespective of the sector or type of activity. It is also necessary to address the totality of problems women face as entrepreneurs, due to the wide spectrum of elements affecting the equitable participation of women in development. A plan or strategy must be designed and implemented in close collaboration with various development partners in different specialized areas, notably: education, health, human rights as well as environment and energy.

## “Women” as an entrepreneur in India:

Entrepreneurship has gained currency across the sphere and female- entrepreneurship has become an important module. India is one of the fastest emerging economies and the importance of entrepreneurship is realized across the gamut. “Women Entrepreneur” is a person who denies the role of their personal needs to participate and be accepted economically independent. Strong desire to do something positive is a high-quality women entrepreneur who contributes to the position values of family and social life. “An enterprise owned and controlled by a women having a minimum financial interest of 51% of capital and giving at least 51% of the employment generated by the enterprise to women.”- **Government of India**

According to **Kamala Singh**," A women entrepreneur is a confident, innovative and creative woman capable of achieving economic independence individually or in collaboration generates employment opportunities for others through initiating establishing and running an enterprise by keeping pace with her personal, family and social life.

According to **Medha Dubhanshi Vinze**," a women entrepreneur is a person who is an enterprising individual with an eye for opportunities and an uncanny vision, Commercial acumen, with tremendous perseverance and above all a person who is willing to take risk with the unknown because of the adventures spirit she possesses."

In the words of Former President **APJ Abdul Kalam** "empowering women is a prerequisite for creating a good nation, when women are empowered, society with stability is assured. Empowerment of women is essential as their thoughts and their value systems lead to the development of a good family, good society and ultimately a good nation."

**Pandit Jawaharlal Lal Nehru** has remarked "When women move forward, the family moves, the village moves and the Nation moves." Thus a women entrepreneur is one who starts business and manages it independently and tactfully takes all the risks, accepts challenging role to meet her personal needs and become economically independent. A strong desire to do something positive is an inbuilt quality of entrepreneurial women, who is capable contributing values in both family and social life and is one who faces the challenges boldly with an iron, will to succeed.

## REVIEW LITERATURE

**Snehalatha and Reddy (1998)** revealed that the income generating activities taken up by the women were dairy, forest nursery, sericulture, vegetable cultivation and petty business, which were carried out as a group activity.

**Savitha (2004)** in her study revealed that high participation was seen in regular activities in all the three enterprises, viz., dairy (88.70%), poultry (100%) and goat rearing (83.60%).

**Dasaratharamaiah et al. (2006)** studied the income generation activities through the Development of Women and Children in Rural Areas (DWCRA) and found that ram lamb rearing was under taken by 8 per cent beneficiaries. Vegetable vending by 26 per cent, basket making 20.67 per cent, milk vending by 29.33 and 4.67 per cent petty trade, 1.33 per cent beneficiaries under took leaves making and 60 per

cent of beneficiaries under took the fruit vending activity.

**Mrs. Geetha Reddy** stated that the Government's industrial investment promotion policy 2005-10 provided special incentives for women in areas like infrastructure at their doorstep, VAT and stamp duty reimbursement and power subsidy

**Sarah (2004)** conducted study on critical analysis of pro-poor initiatives for empowerment of rural women through South Asia Poverty Alleviation Programme (SAPAP) revealed that, majority (40.00%) of the SAPAP beneficiaries belonged to middle age category, followed by young aged (37.80%) and old aged (22.20%) categories.

## OBJECTIVES

- To analyze the status of Women Entrepreneurs in Ranga Reddy District.
- To Study the Socio-economic changes in Women Entrepreneur

## SCOPE:

The study is limited to only five areas of Ranga Reddy District, Andhra Pradesh and focused upon the middle class women entrepreneurs.

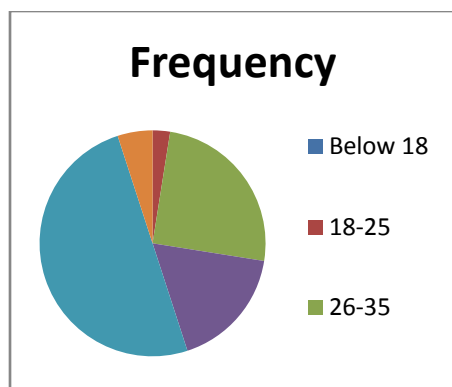
## RESEARCH DESIGN

The study is based on the primary data has been obtained directly from the sample of 50 Women Entrepreneurs with the help of questionnaire. The secondary data is gathered from Journals, Magazines, Books, and Web-sites. The research methodology used to collect information is done through Random Sampling Technique. Analysis is done with the help of Percentages and Mean. The time frame of study is short from January 2012 to June 2012

## Data Analysis

### (A) Age Category:

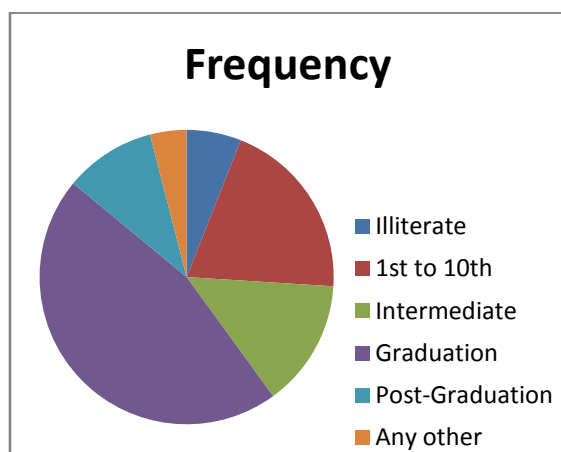
Age Category	Frequency	Percent	Cumulative Percent
<b>Below 18</b>	<b>0</b>	<b>0</b>	<b>0</b>
18-25	2	4	4
26-35	20	40	24
36-45	14	20	64
46-55	10	28	92
55 & above	4	8	100



**Interpretation:** The research study examined that at what age that most of the women involve in the entrepreneurial activities and from this we got to know that the most of the women start their own business at the age category of 36-45 i.e., 40%.

**(B) Education :**

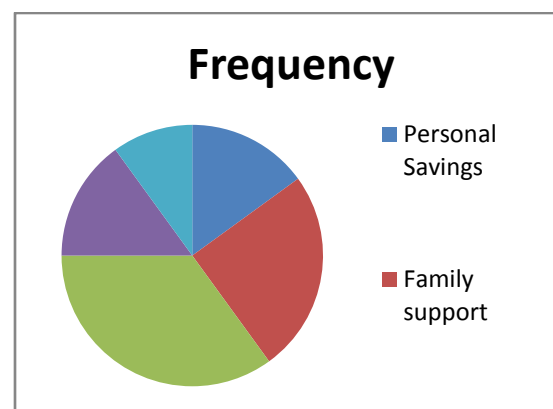
Range	Frequency	Percent	Cumulative Percent
Illiterate	3	6	6
1st to 10th	10	20	26
Intermediate	7	14	40
Graduation	23	46	86
Post-Graduation	5	10	96
Any other	2	4	100



**Interpretation:** The research study examined, the average category of women are graduates i.e., 46% and least category are under any others (i.e., professional in fashion designing etc) the percentage is 4%.

**(C) Financial Support**

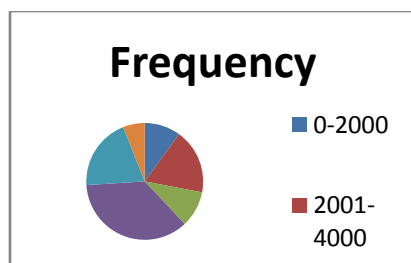
Range	Frequency	Percent	Cumulative Percent
Personal Savings	15	15	15
Family support	25	25	40
Local Private Financiers (money lenders)	35	35	55
SGHs	15	15	90
Banks	10	10	100



**Interpretation:** The above analysis states that, the sources of funds invested are especially from personal savings, local private financiers, Banks and SGHS. The average funds are from local private financier i.e., 35%, the next are from family support i.e., 25% and the least is from Banks i.e., 10%

**(D) Income:**

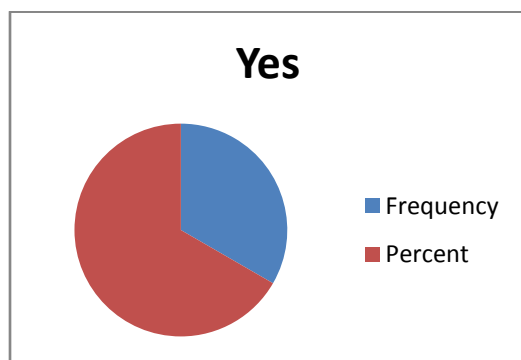
Income level	Frequency	Percent	Cumulative Percent
0-2000	5	10	10
2001-4000	9	18	28
4001-6000	5	10	38
6001-8000	18	36	74
8001-10000	10	20	94
above 10000	3	6	100



**Interpretation:** The above analysis states that, the most of the women entrepreneurs fall under the income level of 6001 to 8000 and the 2<sup>nd</sup> average fall under the income of level of 8000-10000. The least are at income level of above 10000.

**(F) Government Awareness & Development Programmers**

Range	Frequency	Percent
Yes	10	20
NO	40	80

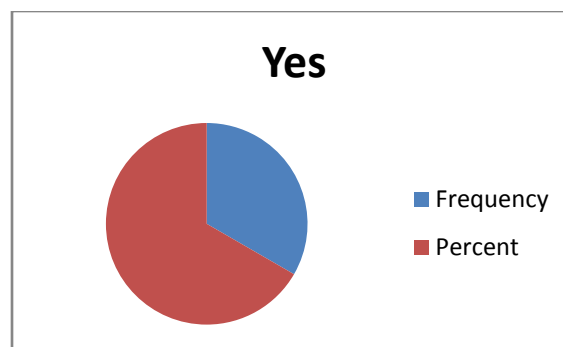


**Interpretation:** The study states that, 80% of the women are not aware of Government initiated funds and Development Programmes.

**Socio Economic Performance of the Women Entrepreneurs:**

**(G) Decision Making Ability**

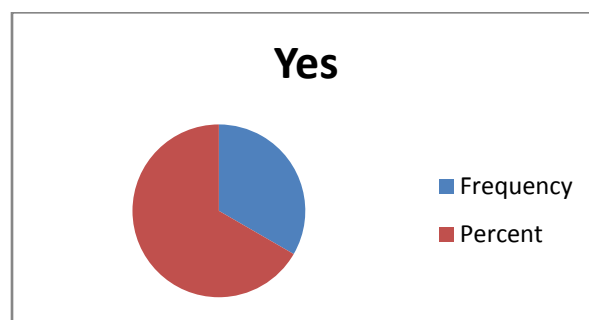
Range	Frequency	Percent
Yes	35	70
NO	15	30



**Interpretation:** The study states that most of the women entrepreneurs have increased in taking decisions at house hold as well in business.

**(H) Savings**

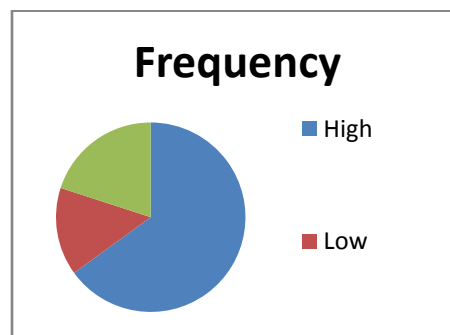
Range	Frequency	Percent
Yes	40	80
NO	10	20



**Interpretation:** The research study examined that most of the women average saving have increased to 80%

**(I) Recognition in family and Society**

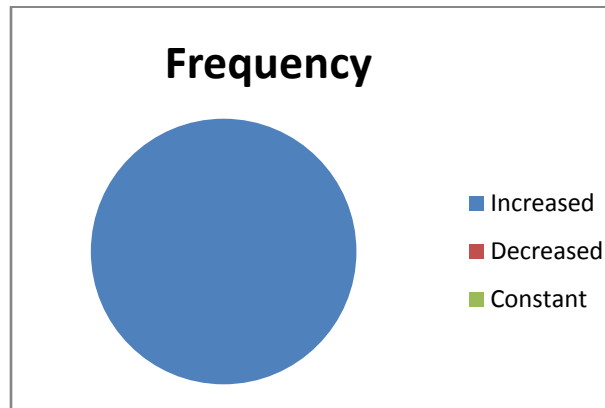
Range	Frequency	Percent
High	65	65
Low	15	15
No Recognition	20	20



**Interpretation:** They are actively participating in voting, training sessions; welfare activities in local areas which have lead them for high recognition in their family and society which is 65% (32members)

#### (J) Children Education

Range	Frequency	Percent
Increased	100	100
Decreased	0	0
Constant	0	0



**Interpretation:** Women entrepreneurs especially started focusing on children education, which has increased to 100%. The number of children of women entrepreneurs is not more than three and the average age group is 15.

#### FINDINGS

1. Most of women entrepreneurs are sole traders where the type of business undertaken are tailoring, flower boutiques, beauty parlour, curry point, home made foods, general stores, Private chit funds, textiles, vegetable shop, baby care centers, tutorials.
2. The main reason for starting the business is to financially support to their family.
3. They are emerging as an employer and employing the people in local areas in turn leading to the development of the society.
4. Savings of the women entrepreneurs has increased rapidly. Achieving economic independence has increased
5. No growth in business due to lack of capital, training, awareness of govt. funds, family support and health problem.
6. Gender inequalities have reduced to certain extent.
7. Due to multi-faceted roles, many are suffering from health issues.

#### SUGGESTIONS

1. The Government should help them in enhancing their education qualification by professional courses and technical courses by imparting training and awareness programmes.
2. A local women entrepreneur society has to be formed and the inner creativity has to be identified by the interaction with the related business areas.
3. Personal counseling has to be conducted to their family members for the support.
4. Access to formal banking sector has to be increased.
5. In order to avoid risk their business has to be insured.
6. Women entrepreneurs can start up with new projects such as collective marketing, bulk purchasing, sharing of machinery and equipment, group lending., group owned enterprises;

#### CONCLUSION

Empowering women entrepreneurs is must for achieving goals of sustainable development of society on large. The society and the government should play an effective role to eliminate gender discrimination which allows 'women' to be an entrepreneur equivalent to men. Training sessions, mentoring, trade fairs and exhibitions has to be provided by forming local women entrepreneurial society which acts as a source of development programme.

#### REFERENCES

1. Dasaratharamaiah, K., Naidu, M. C. and Jayaraju, M., 2006, Women's empowerment through DWCRA – An empirical study. *Social Welfare*, **52**(12): 33-38.
2. Savitha, S.S., 2004, Role of rural woman in Animal husbandry. *M.Sc. (Agri.) Thesis*, Univ.Agric. Sci., Dharwad.
3. Snehalatha, M. and Reddy, M.N., 1998, Impact of thrift and credit groups in income generation of rural women. *J. Extn. Edu.*, **9**(2): 2031-2032.
4. Sarah Kamala, T., 2004, A critical analysis of pro-poor initiatives for empowerment of rural women through south Asia poverty alleviation programme (SAPAP). *Ph.D. Thesis*, Andhra Pradesh Agric. Univ., Hyderabad.
5. The Hindu Saturday, Sep 01, 2007, ePaper
6. International Refereed Research Journal [www.researchersworld.com](http://www.researchersworld.com) Vol.– II, Issue – ,January 2011

**Author: A. LATHA**

Title of Work: Women Empowerment through Microfinance: A Study on SHGs in Ranga Reddy District

Affiliation: Associate Professor, MBA Department, Malla Reddy Engineering College, Maisammaguda, Dhullapally, Secunderabad, Andhra Pradesh. Pin 500014.

Email: [latha.venni@gmail.com](mailto:latha.venni@gmail.com)

Telephone: 9848557305

Postal Address: H.no- 1-5-631/124, Ambedkar Nagar, Old Alwal, Secunderabad.

Pin-500010

## **Abstract**

Microfinance has been present in India in one form or another since 1970s and is now widely accepted as an effective poverty alleviation strategy. Over the last five years, the microfinance industry has achieved significant growth in part due to the participation of commercial banks. The primary objective of this study is to understand the depth of financial inclusion in Ranga Reddy district with the help of SHGs promoted by Government Agencies. The study selected five blocks of SHGs consisting of 533 members. The study employed primary data and the required data for the study is collected thorough field observation and structured questionnaire, survey method with the help of discussions and interviews with the local NGOs and workers. To support and substantiate information and results drawn from the primary data, the study also uses secondary data like the data from records of SHGs , NGOs and Government Agencies etc., Finally the study suggests that there is a dearth of research and empirical evidence about the borrowings habits of the world's poor. As when government suggests any policy framework towards financial inclusion it should necessarily address three parties ie. Microfinance Institutions, MFI regulators and Investors, Supported with proper amount of awareness in the industry.

## **INTRODUCTION**

Microfinance is emerging as a powerful instrument for poverty alleviation in the new economy. In India, Microfinance scene is dominated by Self Help Group (SHGs)-Bank Linkage Programme as a cost effective mechanism for providing financial services to the "Unreached Poor" which has been successful not only in meeting financial needs of the rural poor women but also strengthen collective self help capacities of the poor ,leading to their empowerment. Rapid progress in SHG formation has now turned into an empowerment movement among women across the country.

A well understood but poorly articulated reality of development is the role of women. The UN commission on status of women observed ' women who contribute half of the world's population by virtue of an accident of birth, perform two-thirds of the world's work, receive one-tenth of its income and owns less than one-hundredth of its property'. In India, women produce 30 percent of all food commodities consumed but get only 10 percent of the property or wealth of the country. There is a growing realization that rural women have been underestimated and discriminated against all walks of life, despite their substantial contribution to the household economy and in turn, the national economy as such. As it is clear

that the rights and protection of women from social inequalities in statute books are not good enough, some practical solutions are to be acknowledged and most importantly, implemented. Women have been deprived of economic independence. The empowerment of women and improvement of their status and economic role needs to be integrated into economic development programs, as the development of any country is inseparably linked with the status and development of women. Given the gender division of labour that prevails in India, Nutrition, Child health, and related matters typically depend mostly on women's actions and decisions. Experience has shown that promotion of enterprise creation and income generating activities among women would transform them from 'being alive' to 'living with dignity'. One of the powerful approaches to women empowerment and rural entrepreneurship is the formation of Self Help Groups (SHGs) especially among women. This strategy had fetched noticeable results not only in India and Bangladesh but world over. Women being central to the entire development process and at the precursor of social transformation can be demonstrated with many examples that could include Grameen Bank's success, SHGs of ICICI Bank, Shakthi Ammas at HLL, Cemex, Amul, the success of Avon, Mary Kay, and Tupperware in US and

other parts of the world. Micro Finance as a tool of poverty alleviation and women empowerment has gained acceptance in development dialogue the world over. There is an acute need among the poor for credit, both for consumption and production, which often forms the declining line between survival and succumbing to poverty. It has been found that besides food, credit is also needed for health, housing education. These needs are also critical for survival (Zeller, 2000). The success of SHGs as a development tool depends on the availability of Micro Finance.

This paper tries to link all the pertinent points discussed above namely Micro Finance, SHGs and Women Empowerment United Nations (2001) defines empowerment as the processes by which women take control and ownership of their lives through expansion of their choices. Kabeer's (1998, 1999) view of empowerment refers to the processes by which those who have been denied the ability to make choices acquire such ability. The fundamentals of empowerment have been defined as agency (the 1 Key among them are: the ratification of the Convention on Elimination of All Forms of Discrimination against Women (CEDAW) in 1993, the Mexico Plan of Action (1975), the Nairobi Forward looking Strategies (1985) and the Beijing

Declaration as well as the Platform for Action (1995)

### **Objective of this study :**

- The Objective of this study is to analyze the role of microfinance through Self Help Groups has an impact on the empowerment of women members. More specifically, in this work, we seek to explore if the SHG approach has been successful in the empowerment of rural women living in the highly patriarchal and traditional societies of the Ranga Reddy District of Telangana.
- The prime objective of this study is to know the depth of financial inclusion in Ranga Reddy district with the help of SHGs promoted by Government Agencies.
- To evaluate the different socio-economic factors this has changed the living environment of members of Self Help Groups.

### **Research Methodology:**

**Data collection:** The study employed primary data and the required data for the study is collected thorough field observation and structured questionnaire, survey method with the help of discussions and interviews with the local NGOs and workers. To support and substantiate

information and results drawn from the primary data, the study also uses secondary data like the data from records of SHGs , NGOs and Government Agencies etc.,

**Sample Size :** The study selected 50 SHGs groups consisting of 533SHG members spread over Bolaram, Kowkooor, Surya nagar, Risalabazar and Old Alwal .

**Sampling Technique:** The study applied Random sampling technique from Ranga Reddy District in Telangana.

**Statistical Tools:** This study employed Mean, correlation, standard deviation wherever required for analyzing data

Selection of study area and Sample Units

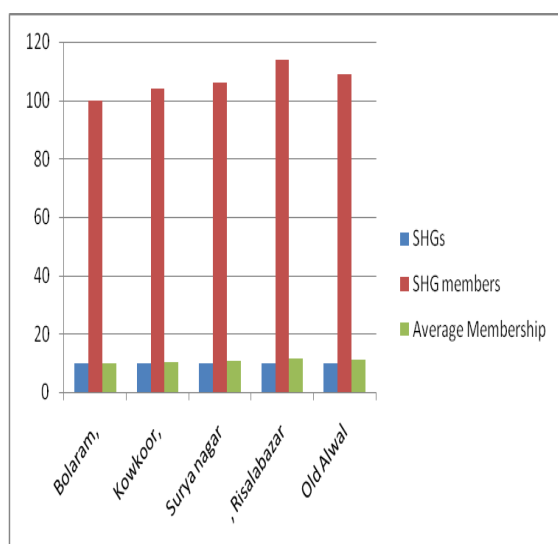
The study was carried out in selective clusters spread over five blocks of Ranga Reddy District in Telangana

It is noteworthy to mention here that the Self-Help Groups in Ranga Reddy District are promoted by NGOs as well as government agencies. Due to time constraint the researcher has selected the SHGs promoted by Government Agencies.

**Table 1**  
**Sampling Frame of the study**

Block	SHGs	SHGmembers	Average Membership
Bolaram	10	100	10
Kowkooor	10	104	10.4
Surya nagar	10	106	10.6
Risalabazar	10	114	11.4
Old Alwal	10	109	10.9
Total	50	533	10.66

*sources: Compiled by the author.*



## Data Analysis

### Structure of SHGs

The structure and characteristics of Self-Help Groups in the study area is presented in Table 1. It is noted that the average membership per SHG was 10.66. The study reveals that in

the Risalabazar block, the average membership of the Self-Help Group is the Highest (11.4) and Bolaram block had lowest membership (10). Most of the members agreed that their motives in joining SHGs were to save money; some said that they joined the SHG to get credit to meet the unexpected cash demand for consumption and other purposes. Few opined that it led to social empowerment.

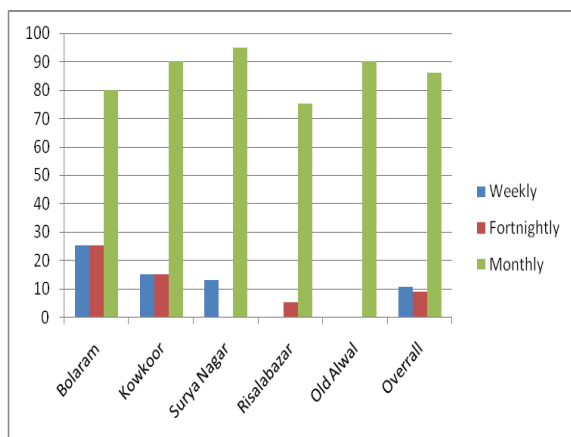
From Table 2, The Frequency of group meeting by SHG indicated that monthly meetings were most common followed by weekly and fortnightly. Meetings are arranged regularly by NGO and Groups. It is held in the middle of the street or in any members home. The absentee member has to pay a fine of Rs.10. The Overall meetings attended by the Group are more in weekly is Bolaram and lowest in Old Alwal. The percentage of members who attend the meeting monthly is more in Surya nagar followed by Old Alwal and Kowkooor. The Overall Percentage of attending meetings mothly is 86%.

**Table-2: Frequency of Group**

### Meeting (percentage)

Item	Bola ram	Kow kooor	Surya nagar	Risala bazar	OldAl wal	Overall
Weekly	25	15	13	0	0	11.6
Fortnightly	25	15	0	5	0	10
Monthly	80	90	95	75	90	86

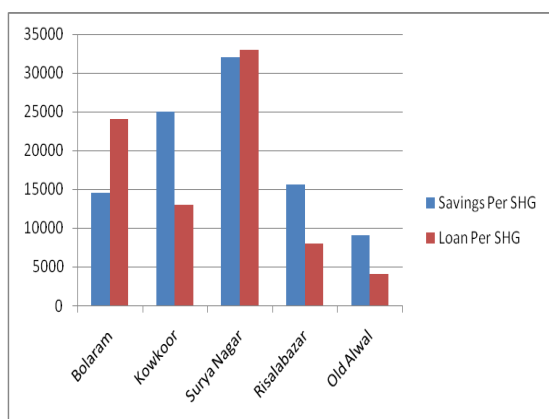
## frequency of meeting



**Table-3**  
**Structure of SHG in the Study Area**

Item	Bola ram	Kow koor	Surya Nagar	Risala bazar	Old Alwal	Over all
Savings Per SHG	14500	25000	32000	15569	9041	19222
Loan Per SHG	24000	13000	33000	8000	4000	16400

## Structure of SHG in the study area



The average savings per SHG was Rs. 19222, The average loan amount per SHG was Rs. 16400, The highest saving being Rs.32000 in Surya Nagar block and the lowest of Rs. 9041 in Old Alwal block. The highest loan per SHG is in Surya Nagar is Rs.33000 and the lowest is in Old Alwal is 4000.

## Profile of SHG Members

## Profile of SHG Members

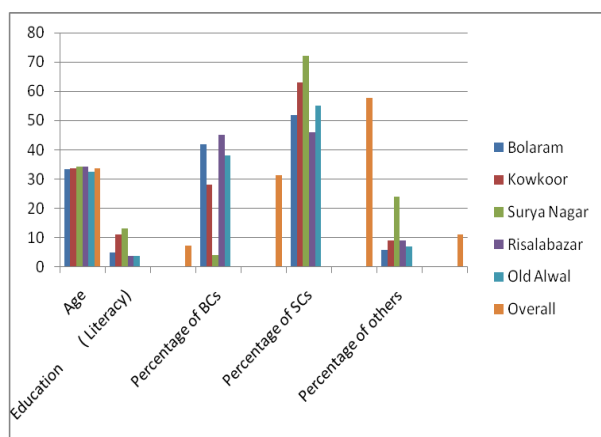
The average age of SHG members in the study area is presented in Table 3 it is observed that the average age of SHG members was 33.62years, lowest being 32.43 years in Old Alwal block. Regarding the Caste profile of SHG members, the table shows that majority of members belong to Schedule Caste Community. In Surya nagar block it is 72% followed by 63 % in Kowkoor. Educational background of the SHG members shows that most of them are Illiterate. Only 7% of them have studied up to primary level. So far as the occupation of the members are concerned, majority of them are engaged in tailoring activities. As regards to average income per SHG, it was around

Rs. 7918. It was highest in Surya nagar block ( Rs. 10057) followed by Kowkooor block ( Rs. 9223) and lowest in Old Alwal block (Rs. 5861).

**Table 3**  
**Profile of SHG Members**

Item	Bola ram	Kow kooor	Surya Nagar	Risal a baza r	Old Alwal	Ove rall
Age	33.37	33.77	34.18	34.40	32.43	33.62
Education	5.11	11.23	13.26	3.69	3.88	7.43
% of BCs	42	28	04	45	38	31.4
% of SCs	52	63	72	46	55	57.6
% of others	06	9	24	09	07	11.0
Income	7867	9223	10057	6582	5861	7918

*Sources: Complied by the author.*



## Purpose of credit Demanded and Utilised

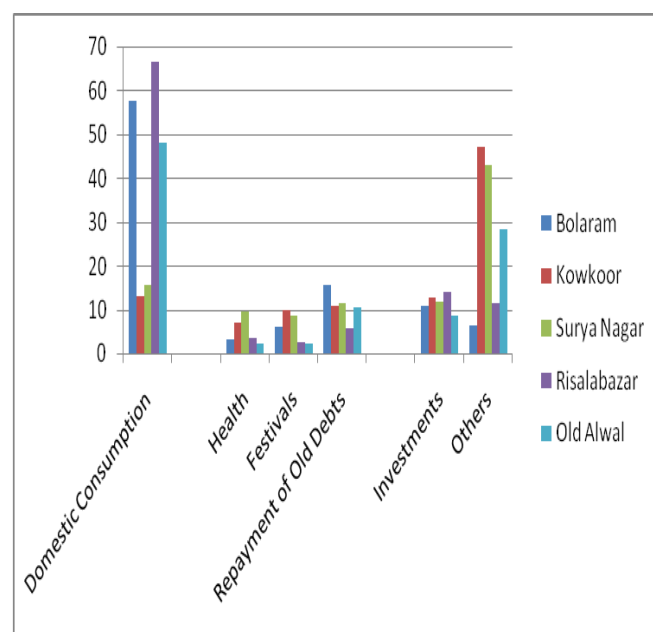
Generally, after six months operation of saving account, the saving is

pooled and used for internal lending among the members. The amount of loan and number of loans are decided by the members themselves depending on their need and urgency.

**Table 4**  
**Purpose of credit Demanded by SHG Members**

*Sources: Complied by the author.*

Item	Bola ram	Kow kooor	Surya Nagar	Risala bazar	Old Alwal	Over all
Domestic Consumption	57.66	12.91	15.57	66.54	47.89	30.53
Health	3.22	6.87	9.50	3.56	2.27	4.37
Festivals	6.11	9.78	8.71	2.39	2.18	5.83
Repayment of Old Debts	15.67	10.66	11.52	5.68	10.59	10.82
Investments	10.89	12.57	11.66	14.05	8.67	11.56
Others	6.45	47.21	43.04	11.34	28.4	27.28



The purpose –wise credit demanded by the SHG members from the SHGs is given in Table 4. It is observed that most of the members have demanded credit for domestic consumption purposes. It is highest in Risalabazar block (67%) followed by Bolaram ( 57.66%) . Reasonable proportion of SHG members has demanded credit for other purposes. This percentage is highest in Surya nagar (43.04%) and lowest in Bolaram block (6.45%). About 11.56 percentage of credit was demanded for investment purposes. A proportion of credit was demanded by the SHG members for payment of old Debts. From the Table it is clear that a larger share of credit demanded by SHG members is being utilized for domestic Consumption purposes followed by repayment of debts and others.

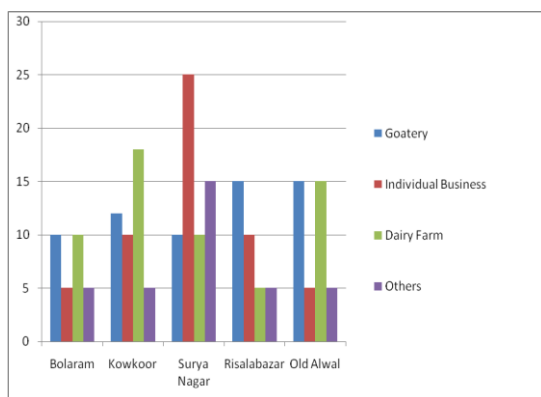
### **Economic Activities Covered by SHG members**

Table 5 reveals that most of the SHG members are engaged in the collection and processing of minor forest products. These products include broom

making, cashew, Turmeric, Tamarind, Khalli ( leaf plates), amla,etc. Some of the members are engaged in individual business like preparing pickle, papad , Haldi powder , Wax , making bags, poultry, vegetable business, tailoring, pan shop, etc. They are also engaged in poultry, dairy and goatery business. Some are engaged in other activities.As there is a good demand for milk products, they are preparing sweets with milk, ghee, etc, and are getting good price. They earn about Rs. 600 to Rs. 1000 per month through these activities.

**Table -5**  
**Economic Activities covered by SHG members**

Item	Bola ram	Kow koor	Surya Nagar	Risala bazar	Old Alwal	Overall
Goatery	10	12	10	15	15	12.4
Individual Business	5	10	25	10	5	11.00
Dairy Farm	10	18	10	5	15	12.6
Others	5	5	15	5	5	7.00



## Loan Support to SHGs by Banks

There is a bank linkage programme established to SHGs. The SHG members opened their accounts in various nationalized banks such as State Bank of Hyderabad, Indian Overseas Bank, Central Bank of India, Syndicate Bank, Andhra Bank, etc.

SHG members are getting both internal loans and external loans under the security of NGOs. They are paying Rs.0.25 as interest per Rs. 100. They are also maintaining cashbook, membership register, loan register, Individual passbook register, etc. They are taking loans for both production and consumption purposes, State Bank of Hyderabad has advanced loans of 1,30,000 to Surya Nagar Block,

followed by Rs. 80,000 to Old Alwal block.

## Benefits Derived by SHG Members

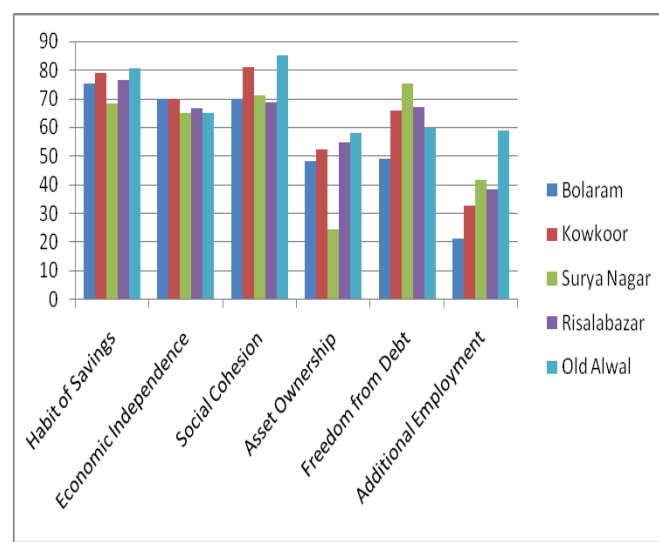
Varieties of benefits are derived by the members of SHGs as presented in Table 6. Members of SHG increased their Habit of savings by 76% followed by Economic Independence by 67.3%, Self – Confidence by 81.54%, Social Cohesion by 75.17%, Asset Ownership by 47.65% , Freedom from debt is 63.36% and additional employment by 38.6 %

**Table-6**  
**Benefits Derived by SHG members**  
**(multiple Reasons)**

Item	Bola ram	Kow koor	Surya Nagar	Risalab azar	Old Alwal	Over all
Habit of Savin gs	75.4 4	79.1 1	68.22	76.66	80.66	76.0 0
Econo mic Indep enden ce	70.0 0	70.0 0	65.00	66.80	65.00	67.3 6
Social Cohes ion	70.0 0	80.8 8	71.11	68.88	85.00	75.1 7
Asset Owne rship	48.3 3	52.3 4	24.40	54.99	58.22	47.6 5
Freed om from Debt	49.0 0	66.0 0	75.08	67.00	59.76	63.3 6
Additi onal Empl oyment	21.1 1	32.8 0	41.80	38.40	58.90	38.6 0

*Figures in the table indicate percentages.*

*(sources: Compiled by the author)*



## **FINDINGS**

1. The SHGs have made a lasting impact on the lives of the women particularly in the areas of Ranga Reddy District which drastically influenced the quality of living
2. Increased self-reliance and self confidence have improved the ability of women to mobilize various public services for their benefit.
3. They could develop their skills and abilities in various productive activities which increased their income, savings and consumption expenditure.
4. Meetings are regularly conducted where in the members are getting a chance to exchange their views, solve

their problems and be able to develop their group strength by interactions.

5. The members started focussing on new business for financial support to their family

## **SUGGESTIONS**

1. Women who have limited skills, do not have autonomous access to resources and are dependent on husbands are those that benefit the least. The policy lesson here is that to ensure that loans fulfill their potential in terms of empowering women, institutions ought to first provide skills and get information on the existing lives of these women.
2. Training in legal literacy, rights and gender awareness are important complements to micro-credit for the empowerment of women. The members should be given necessary training and guidance for the successful operation of the group.
3. Personal counseling has to be conducted to their family members for the support.
4. The members of the SHG should be more active, enthusiastic and dynamic

to mobilize their savings by group activities. In this process NGO act as facilitator.

5. Marketing facilities for the sale of products of SHG may be created.
6. Uniformity should be maintained in formation and extension of financial assistance to them by banks in all blocks.
7. Periodical exhibition at block-level may be organized where the product of SHG can be displayed.

## REFERENCES

1. Micro Finance, Self Help Groups and Women Empowerment – current issues and concerns(V.P.Sriraman, Faculty, Bharathidasan Institute of Management, Tiruchirapalli).
2. Kabeer, N. 1998. Can't buy me love? Re-evaluating gender, credit and empowerment in rural Bangladesh. IDS Discussion Paper No. 363. Institute of Development Studies, Sussex.  
  
Kabeer, Naila. 1999. Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. Development and Change Vol. 30, Institute of Social Studies, Oxford, UK.
3. United Nations Development Programme. 2001. Participatory governance, people's empowerment and poverty reduction. UNDP Conference Paper Series.
4. Government of India: Economic survey (various years). Government of India New Delhi.
5. Financial Inclusion an abiding virtue, the Hindu Monday, Jan 16, 2006.
6. [www.iibf.org.in](http://www.iibf.org.in), [www.rbi.org.in](http://www.rbi.org.in)

# **A study on Risk – Return Analysis of banking sector with respect to public and private sector**

**A.LATHA<sup>1</sup> P.BUELA PRASANNA KUMARI<sup>2</sup> V.VENU KUMAR REDDY<sup>3</sup>**

*<sup>1</sup>Associate Professor  
MBA Department, MREC*

*<sup>2</sup>Associate Professor  
MBA Department, MREC*

*<sup>3</sup>Assistant Professor  
MBA Department, MREC*

## **ABSTRACT:**

This study empirically investigates the risk-return dynamics of the Four Firms of banking industry for the period of 12 months. The objective of study is to ascertain risk and returns of four banking firms. This study employed standard deviation and average return of each of the firm to analyze where to invest. This study revealed that the sizes of risks are different in the selected firms. They varied positively with the sizes of returns. In addition, correlation of public sector banks is negative and the correlation of private sector banks is also negative. But when we correlate public and private sector banks, they are positive at one point and negative at another. Most of the firms' risks (standard deviation) are greater than Unity, which imply higher risk as compared to Market Portfolio. Higher the risk, higher the return. The outcome of this study conformed to similar studies in the emerging stock markets.

**Keywords:** Risk-Return Characteristics, the four firms of banking sector, standard deviation.

## **Introduction**

Stock investment is basically a long –term investment. Every investment carries one risk or the other. This existential reality is more pronounced in the quest for wealth through investment in stock markets. The stock market offers investors the opportunity to invest in securities of quoted firms such investment could be in fixed income security e.g. debentures, etc. or they could be in equities. Each of these securities offer returns to investors, depending on firms' risk and the nature of the stock invested on. Generally, higher risk, higher returns, all things being constant.

The risk and return of security in the stock market may differ because of different factors affecting securities, such as differences in structure and managerial capacity of different firms, different sectors in which they operate, the state of the

economy, government policies as well as internal corporate policies, themselves<sup>1</sup> (Oludoyi , 2003)

The risk/return relationship is a basic concept in not only financial analysis, but in every part of life. If decisions are to lead to gain maximization, it is necessary that individuals/institutions consider the combined influence on expected return as well as on risk. Return is the amount which an investor actually earned on an investment during a certain period. Return comprises of interest, dividend and capital gains; while risk represents the uncertainty related with a particular task. In financial terms, risk is the chance or probability that a certain investment may or may not deliver the actual/expected returns.

The risk and return trade off says that the potential return rises with an increase in risk. It is important for an investor to decide on a balance between the desire for the lowest possible risk and highest possible return. In this study, the risk return analyses of five banking firms are to be analyzed.

The focus of the study is to protect the interest of the investor and show him the best investment alternative where he has to invest among the five banks. This is done with the help of statistical tools of standard deviation, variance and average return.

## **OBJECTIVES OF THE STUDY:**

1. The primary objective of the study is to analyze the risk-return of four banks in the banking sector.
2. To analyze the best bank for the investment purpose in these banks.

## REVIEW OF LITERATURE:

L.C.Gupta (1981)<sup>2</sup> in his report it is concluded that the rate of return provided by equities are unsatisfactory because 20% of the returns for various holding periods are negative. The returns provide only a partial hedge against inflation and the fluctuations in returns even within a year are large enough to conclude that time has an important bearing on realized return. The risk is considerable even when investment is made in a portfolio of securities and for long periods of time.

Srivastava (1984)<sup>3</sup>, he analyzed a cross-sectional study of 327 firms of Bombay Stock Exchange for the year 1982-83, and interpreted higher market prices of securities are the result of dividend rates associated with the organization. The study rejected the MM approach of negative importance of dividend in explaining the rates of return in Indian context.

**S. Natarajan , T.Muthukumaran And M. Rengasamy**, In their paper, an attempt has been made to measure the impact and recovery of Global Financial Crises on Indian economy in general and on the Indian Stock Market in particular. Additionally, this study intends to discuss the role of RBI in the present context. This paper employs the data from April 2001 to March 2013 to investigate the impact on Indian stock market and the integration of Indian stock market and US stock market. The result suggests that the Indian Stock market was affected by the Global financial crunches and also shows that the recovery from the slowdown during the global financial crisis is well underway.

## SCOPE OF THE STUDY:

The research is primarily concentrated on four stocks of banking sectors i.e., HDFC, ICICI, SBI and UBI. This paper covers risk and return analysis and the calculation of covariance between the public banking securities and private banking securities. The beta values of the each stock helps the investor to measure the risk and return helps in the investment for the particular security. The scope of the study is limited to one year.

## LIMITATIONS OF THE STUDY

Due to certain factors there are certain limitations to the analysis. Those factors are limited to secondary data which is historical in nature and of four banking firms two of public sector and private sector and the time period of study is only of one year.

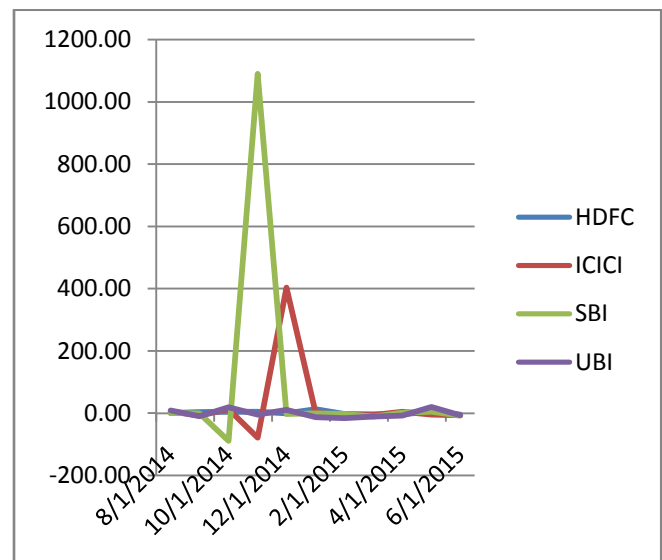
## RESEARCH METHODOLOGY

The information which we used for the study consists of secondary data of NSE, various publications, magazine, news papers and reports prepared by research scholars. The period of study for analysis is one year. Statistical tools used are average returns, risk is determined with the help of standard deviation, variance, covariance and co-relation of different securities of public and private banks in banking sector.

## DATA ANALYSIS:

Returns of the selected four banks in banking sector :

DATES	HDFC	ICICI	SBI	UBI
6/1/2015	-4.28	-6.78	-7.46	-6.32
5/1/2015	6.25	-4.20	4.42	20.58
4/1/2015	-3.32	4.96	1.01	-8.11
3/2/2015	-2.97	-5.89	-10.95	-10.75
2/2/2015	-2.17	-7.06	-2.93	-16.16
1/1/2015	13.21	2.15	-0.93	-12.64
12/1/2014	-0.58	403.16	-2.97	11.22
11/3/2014	4.97	-78.43	1089.13	-4.59
10/1/2014	4.49	13.47	-88.95	19.49
9/1/2014	3.45	-7.92	-0.61	-9.62
8/1/2014	1.15	5.81	0.88	9.23



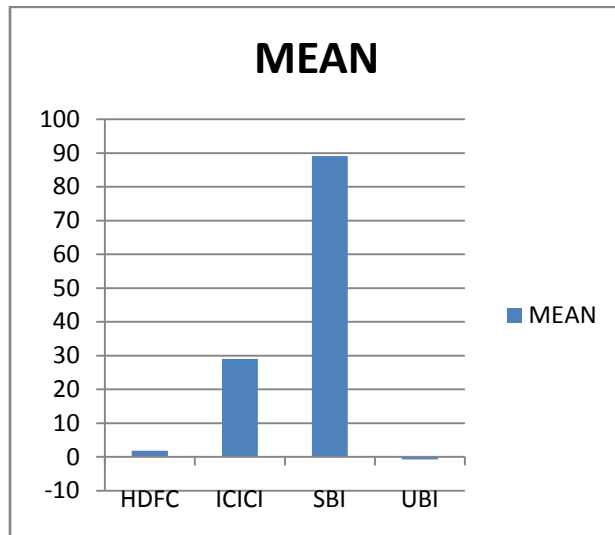
The graph reveals that the total returns of SBI is higher than of other banks succeeded by ICICI bank and the least returns from UBI.

#### MEAN:

Calculation of mean helps us to analyze the returns of four different banks which help to calculate the variance and standard deviation

The following table depicts the returns of four banks.

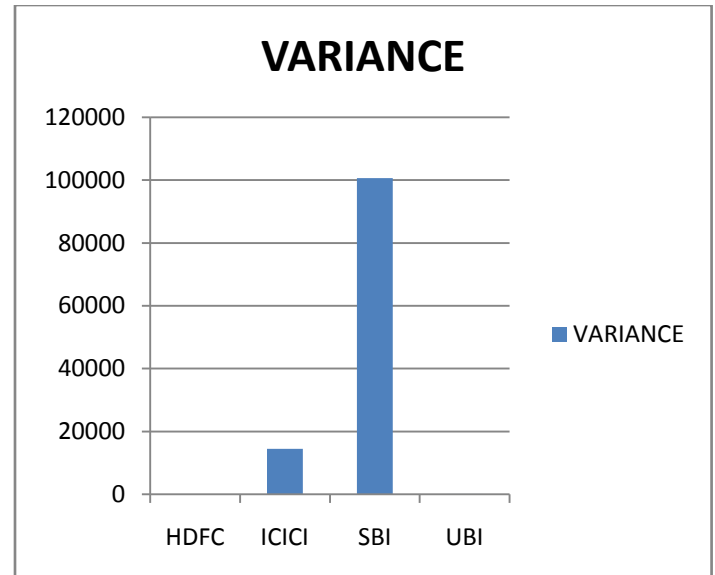
INDEX	HDFC	ICICI	SBI	UBI
MEAN	1.84	29.03	89.15	-0.69



The individual mean of four different banks states that the return of SBI is high, since it is One of the oldest and largest financial institutions in India according to rediff.com, SBI's assets in March 2012<sup>3</sup>. The return of 2<sup>nd</sup> position in the selected banks is ICICI since ICICI Bank is India's largest private sector bank with total assets of Rs. 6,461.29 billion (US\$ 103 billion) at March 31, 2015 and profit after tax Rs. 111.75 billion (US\$ 1,788 million) for the year ended March 31, 2015. ICICI Bank currently has a network of 4,050 Branches and 12,919 ATM's across India<sup>4</sup> followed by HDFC bank. Among them the least or negative returns of UBI is -0.69. Operating expenses of the financial year 2014-2015 is high due to which the current rate of return has decline.

#### VARIANCE:

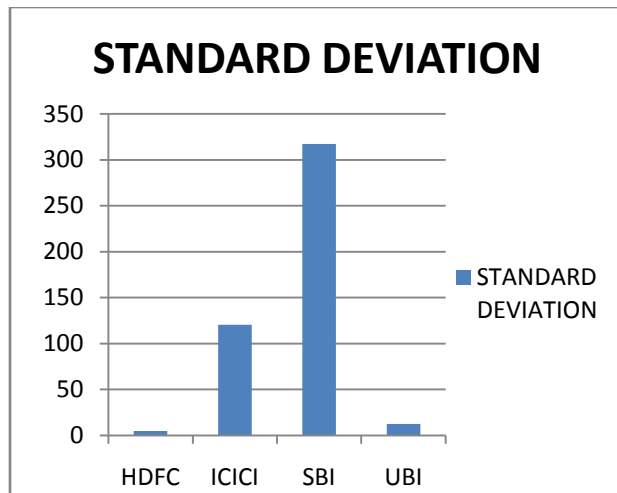
INDEX	HDFC	ICICI	SBI	UBI
VARIANCE	25.25	14534.23	100629	160.44



Variance of means is higher in case of SBI, followed by ICICI, UBI and the least is HDFC. Variance is non-negative, high variance indicates that the data points are very spread out around the mean and from each other. a small variance indicates that the data points tend to be very close to the mean (expected value) and hence to each other. In this analysis SBI Bank has high variance and HDFC Bank has lowest variance.

#### STANDARD DEVIATION:

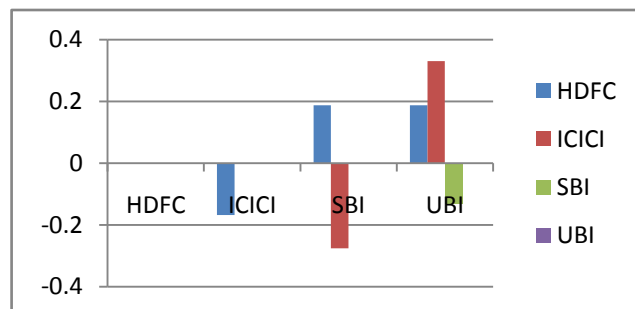
INDEX	HDFC	ICICI	SBI	UBI
STANDARD DEVIATION	5.025	120.56	317.22	12.67



For every investment we have risk and return in this research the measurement of risk is determined by standard deviation. Standard deviation is very high in SBI than that of other four banks, but the return of SBI is also high. Followed by ICICI where the risk is high but the returns are comparatively low. Though the risk is seen in UBI, the returns are in negatives. In HDFC, though the risk is less the returns are also less. Therefore more the risk yields more returns.

#### CORRELATION:

INDEX	HDFC	ICICI	SBI	UBI
HDFC		-0.168	0.188	0.188
ICICI			-0.276	0.331
SBI				-0.1303
UBI				



The above graph reveals that HDFC is positively correlated with SBI and UBI i.e. 0.188 and negatively correlated with ICICI.

ICICI is negatively correlated with SBI i.e. - 0.276 and positively correlated with UBI i.e. 0.331.

The result of correlation is negative for public banks, SBI and UBI i.e. - 0.1303.

#### FINDINGS:

1. As per the study the total returns of SBI is higher since it is the oldest and eminent bank in India than of other banks succeeded by ICICI, HDFC and the negative returns from UBI.
2. Mean helps to know the average return of each bank. In this analysis when we compare between four different banks the return of SBI is high, 2<sup>nd</sup> position is ICICI since ICICI Bank is India's largest private sector bank followed by HDFC bank. Among them the least or negative returns of UBI is -0.69.
3. Variance of means is higher in case of SBI, followed by ICICI, UBI and HDFC.
4. Risk is very high in SBI than that of other four banks, but the return of SBI is also high. Followed by ICICI where the risk is high but the returns are comparatively low. Though the risk is seen in UBI, the returns are in negatives. In HDFC, though the risk is less the returns are also less. Therefore more the risk yields more returns.
5. When we see the correlation both the private and public sector banks are negatively correlated. Whereas when we correlate one private bank i.e. HDFC and SBI, UBI are positively correlated i.e. 0.188 ICICI is negatively correlated with SBI i.e. - 0.276 and positively correlated with UBI i.e. 0.331.

#### SUGGESTIONS:

We would like to suggest that UBI bank has to reduce the operating expenses so that it would earn higher returns in future.

As per analysis ICICI bank returns are higher than HDFC bank since it is the first private sector banking in India.

#### CONCLUSIONS:

With the above analysis in nutshell I would conclude that all the returns are positive except UBI. It would be the best option to invest in SBI bank to fetch more returns for the investment.

Reserve Bank of India has identified State Bank of India and ICICI Bank as systemically important banks in India, requiring these lenders to keep aside additional capital to cover risk. Their capital requirement rises by 0.6% and 0.2% respectively, but this may not have much significance because banks in India in general maintain more than two to three percentage points more capital than regulatory stipulation<sup>5</sup>.

#### REFERENCES:

1. Oludoyi, S.B. (2003), An Empirical Analysis of Risk Profile of Quoted Firms in the Nigerian Stock Market. Ilorin Journal of Business and Social Sciences, 8 (1&2): 9-19
2. L.C.Gupta (1981)<sup>2</sup> Rates of Return on Equities: The Indian Experience, Oxford Publishing, New Delhi
3. R. M.Srivastava(1984), "Testing MM's Dividend Valuation Model in Indian Context- A case study of 327 Joint Stock Companies", Management Accountant, Vol.19(11), Nov.
4. S. Natarajan , T.Muthukumaran And M. Rengasam **AN ANALYTICAL STUDY OF RISK AND RETURN ON INDIAN STOCK MARKET**, Indian Streams Research Journal. Volume: III, Issue : VI, July - 2013
5. [Rahul Gusain](#): **Working in Bank is the aim of many Job seeker. This article provide you list of top 10 Bank to work in best place. Scroll down to know more about these top 10 bank**". June 1, 2015
6. Atmadip Ray, RBI has identified SBI and ICICI Bank as systemically important banks in India, ET Bureau
7. [www.ICICBank.com](http://www.ICICBank.com).

# Customer response Credit card usage

S.Eswarareddy  
Asst.Prof, MREC

**Abstract:** *The survey conducted purpose is how frequently Employees they are using their credit cards .How they will pay the Bills purpose. Now a day's Employees and professional highly they are using their credit cards.*

*Credit cards helped shape modern society and have paved the way for faster, easier, convenient, and more secure financial transactions. Our credit articles are not only about the benefits of credit cards. They also give basic information like definition of your credit rating, credit scores and what risk they fall into, and how to choose the perfect credit card. Build your credit score through creative ways and find means to receive an unsecured loan even when you have bad credit. All of these and more, you can find here in our credit category.*

**Key words:** *what range people they are using credit card in India. I have done my survey in Hyderabad.*

## Introduction

In recent years, there has been a dramatic growth in credit card usage among employers and business people. This increase in the number of employers holding credit cards and incurring credit card debt has generated concern that these employers are overextended and unaware of the long-term consequences associated with severe indebtedness. When other debt is added to this, such as entertainment shopping, the concern becomes even greater.

If used responsibly, credit cards can provide a number of advantages to employers. Credit cards can be a convenient means of payment, a useful tool for learning financial responsibility, a resource in case of emergencies, a means to establishing a good credit history and a way to gain greater access to credit in the future. If credit cards are mismanaged or misused, however, the disadvantages can result in severe financial Consequences. The convenience of credit may tempt employees to live beyond their means. Excessive credit card debt and late payments can damage employees 'credit ratings and make it more difficult for them to obtain credit down the road. In addition, Inexperience with credit and a lack of personal financial knowledge are likely to place some students at greater financial risk for having large and

perhaps unmanageable, debt burdens when they graduate.

## Research Objectives

In the fall of 2014, Employees and professional were surveyed to determine their usage of credit cards. The purpose of this publication is to provide a detailed description of credit card usage and financial practices of employees and professional. Most of the employee's usage of a credit card shorts the financial problems.

.Research indicates that formal financial services play an important role in reducing the financial management problems of employees. The next section describes the survey methodology and characterizes the employee sample. The remaining sections provide detailed analysis of employee credit card behavior and financial practices.

## Methodology

The main aim of the study is how frequently employees and professional they are using the credit card and also which purpose they are using the credit card purpose i have done my analysis .

For this purpose I have done my survey in Hyderabad I have taken 50 Samples. Data has been analyses using MS-Excel and Spss. To analysis purpose. The survey contained 15 questions and also sample size it should be 50.online survey purpose I am using the www.docs.com I used.

**Table I: Age of the respondents**

	Frequency	Percent	Valid Percent	Cumulati ve Percent
Valid	25-30	27	54.0	54.0
	30-40	18	36.0	90.0
	40-50	5	10.0	100.0
	Total	50	100.0	100.0

The above demographic table Explain 54 % respondents are 25-30 age below groups highly using the credit cards.30-40 in between Age group people using credit cards 36% and also the

The above table represents least usage of a credit card.

**Table II : Level of education of respondents**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	High school	6	12.0	12.0	12.0
	Graduate	6	12.0	12.0	24.0
	Post Graduate	38	76.0	76.0	100.0
	Total	50	100.0	100.0	

Most of the employees are post graduate level (76 percent) reported that they were.

12 percent reported that they were high school and graduate level of Employees.

**Table III : Occupation of the respondents**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Salaried	46	92.0	92.0	92.0
	Professional	4	8.0	8.0	100.0
	Total	50	100.0	100.0	

This table represented that majority (92 %) of the employees they are depending the job Remaining it should be a professional (8 %).

**Table IV: Monthly income of the respondents**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Up to 10000	4	8.0	8.0	8.0
	10000-20000	18	36.0	36.0	44.0
	20000-30000	8	16.0	16.0	60.0
	30000-50000	20	40.0	40.0	100.0
	Total	50	100.0	100.0	

Table 4 indicates monthly income more than 30000 thousand people two third of the people is there. After 30000 thousand to 10000 in between people 60 % is there.

**Table V: Number of credit cards held by respondents**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	26	52.0	52.0	52.0
	2	6	12.0	12.0	64.0
	3	9	18.0	18.0	82.0
	4	4	8.0	8.0	90.0
	5 and above	5	10.0	10.0	100.0
	Total	50	100.0	100.0	

Table 5 shows that slightly more than half (52 percent) of the employees they are using only one credit card they reported that .18 percent of the peoples they are using 3 credit cards and also above 5 credit usage is 10 percent of the peoples is there.

**Table VI: Number of visits for shopping and entertainment**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5	24	48.0	48.0	48.0
	10	7	14.0	14.0	62.0
	15	14	28.0	28.0	90.0
	25	5	10.0	10.0	100.0
	Total	50	100.0	100.0	

In this table represent that survey is monthly 5 times they are using credit card for shopping and entertainment purpose is 48 percent. Then 10 % of the peoples regularly they are using the credit card for the purpose of Entertainment and shopping.

**Table VII Card usage per month**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5	25	50.0	50.0	50.0
	10	16	32.0	32.0	82.0
	15	5	10.0	10.0	92.0
	20	4	8.0	8.0	100.0
	Total	50	100.0	100.0	

In this table shows that monthly usage of credit card per month 5 times is a50 % respondents are reported. 30 percent had to use credit cards for month 10 times they are using. A few people they are using credit card regularly

**Table VIII: Avg monthly spending through credit card**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1000-2000	15	30.0	30.0	30.0
2000-3000	3	6.0	6.0	36.0
3000-5000	22	44.0	44.0	80.0
5000 and above	10	20.0	20.0	100.0
Total	50	100.0	100.0	

This table shows the information that is spending the amount through credit card 3000 -5000 respondents are 44 percent.1000 to 2000 in between spending the people is 30 percent.

**Table IX: Minimum purchase amount**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Up to 1000	5	10.0	10.0	10.0
1000-2000	17	34.0	34.0	44.0
2000-5000	14	28.0	28.0	72.0
5000 and above	9	18.0	18.0	90.0
5	5	10.0	10.0	100.0
Total	50	100.0	100.0	

Minimum purchase amount of credit card through 1000 to 2000 peoples are 34 %.2000 To 5000 respondents are 28 %.up to 1000 purchase amount people is 10%.

**Table X: Recently purchased Product using credit card**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Electronic Good	29	58.0	58.0	58.0
Garmments	21	42.0	42.0	100.0
Total	50	100.0	100.0	

Majority of the peoples recently using the credit card is electronic goods purchase purpose 58 % percent of respondents and also garments purpose using the credit card is 42 percent.

**Table XI: Credit card used to make payment of bills**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Electricity Bills	7	14.0	14.0	14.0
Mobile bills	16	32.0	32.0	46.0
Travel tickets	19	38.0	38.0	84.0
Movie tickets	8	16.0	16.0	100.0
Total	50	100.0	100.0	

Credit card through pay the travel tickets 38 percent of the peoples pay their amount. Mobile bill make a payment is 32 percent. Remaining few people make a payment for electricity bills and movie tickets purpose.

**Table XII: Credit Card usage to buy products online**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Clothes	20	40.0	40.0	40.0
Books	16	32.0	32.0	72.0
Cosmetics	14	28.0	28.0	100.0
Total	50	100.0	100.0	

Table 12 shows that 40 percent of respondents reported that online clothes purchase purpose they are using their credit cards. Online books purchase purpose 32 percent and cosmetics purpose 28 percent they are using.

**Table XIII: Amount payer monthly on credit card**

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Minimum due	11	22.0	22.0	22.0
Full payment	36	72.0	72.0	94.0
Part payment	3	6.0	6.0	100.0
Total	50	100.0	100.0	

more than half (72 percent) of Employees full payment with at least one credit card ( Part of pay)pay only a portion on their bills with 6 percent paying only the minimum due Respondents are 22 percent respondents reported paying their bills pay in each month.

**Table IV Credit card bill payment made by others**

		Frequen cy	Percent	Valid Perce nt	Cumula tive Percent
Valid	Never	30	60.0	60.0	60.0
	Once in a while	8	16.0	16.0	76.0
	Frequently	12	24.0	24.0	100.0
	Total	50	100.0	100.0	

More than half (60 percent) of the employees reported that their parents or others never paid their credit card bills. About frequently (16 percent) reported that their parents paid the bills once in a while, and a Third (24 percent) reported that their parents paid the bills on a more regular basis.

**Table XV: Where did you apply your credit card**

		Frequenc y	Perce nt	Valid Perce nt	Cumula tive Percent
Valid	Applied through salary account	20	40.0	40.0	40.0
	Applied in bank branch	8	16.0	16.0	56.0
	Applied Online	12	24.0	24.0	80.0
	Applied for it through an application received in the mail	10	20.0	20.0	100.0
	Total	50	100.0	100.0	

The most frequent ways they obtained their first credit card were from through their salary account or through a applied through online application. Few Employees Acquired their first card at bank branch few members Applied for it through an application received in the mail.

**Table XVI :Age of the respondents \* Level of education of respondents Cross tabulation**

Count		Level of education of respondents			Total
		High school	Graduate	Post Graduate	
Age of the respondents	25 - 30	6	0	21	27
	30 - 40	0	1	17	18
	40 - 50	0	5	0	5
Total		6	6	38	50

Above table represents the demographic profiles and employee Education survey tell that most of the young stars (25-30) in Between age group people are Post Graduates.40 to 50 age group people are 6 respondents they are Graduate and also 30-40 Age group people are all are they completed their Post graduate .

**Table XVII: Occupation of the respondents \* Monthly income of the respondents Cross tabulation**

Count		Monthly income of respondents				Total
		Up to 10000	10000 to 20000	20000 to 30000	30000 to 50000	
Occupation of the respondents	Salaried	0	13	5	14	32
	Professional	3	0	0	0	3
Total		3	13	5	14	35

In this table tell represent the most of the respondents are employers salaries based they are working majority of the employers they are 10000 to 20000 respondents are 13 members 30000 to 50000 respondents are 14 members a few members are Professional.

**Table XVIII:**

**Number of visits for shopping and entertainment \*  
Card usage per month Cross tabulation**

Count

		Card usage per month				Total
		5	10	15	20	
Number of visits for shopping and entertainment	5	13	3	0	0	16
	10	0	5	0	0	5
	15	0	4	4	3	11
	20	3	0	0	0	3
	25					
Total		16	12	4	3	35

In this survey monthly usage of credit card for shopping and entertainment 13 respondents are monthly 5 times they are using their credit card. A few respondents are regularly they are using their credit cards for shopping and entertainment purpose.

**Table XIX: Monthly income of the respondents \* Avg monthly spending through credit card Cross tabulation**

Count

		Avg monthly spending through credit card				Total
		1000-2000	2000-3000	3000-5000	5000 and above	
Monthly income of the respondents	Up to 10000	0	0	3	0	3
	10000-20000	8	0	5	0	13
	20000-30000	0	0	2	3	5
	30000-50000	3	2	6	3	14
Total		11	2	16	6	35

Who are getting Salaries 10000 to 20000 respondents are there usage of credit card 1000 to 2000 respondents are 8 members their usage of a credit card is high after 30000 to 50000 who are getting monthly salaries those respondents are members 6 members they do the usage of credit card 3000 to 5000.

**Table XX: Avg monthly spending through credit card \* Amount paid monthly on credit card Cross tabulation**

Count

		Amount paid monthly on credit card			Total
		Minimum due	Full payment	Partial payment	
Avg monthly spending through credit card	1000-2000	4	4	3	11
	2000-3000	1	1	0	2
	3000-5000	4	12	0	16
	5000 and above	0	6	0	6
Total		9	23	3	35

In this survey who are spending bills through credit card 3000 to 5000 again they will pay the bills in full payment 12 respondents .An and average majority of the peoples they will make a full payment every Month. A few members they partially they are paying the amount.

#### Suggestions and Recommendation:

- In this survey when you have job or not in recession period at least you survive 3 months to usage of credit card.
- Credit cards can be helpful, but should only be used for emergencies.
- Credit cards are really useful and convenient.
- Most of the debt on my card belongs to my parents, who are in a debt consolidation program that doesn't allow them to have credit. They needed help so they used my cards.
- I don't think a credit card is necessary, but it is useful in sticky situations. I plan on

applying for one soon just in case of emergencies. Just haven't found the right one yet.

- I am a very responsible credit card holder and so are many others.
- I usually pay off my credit card with my income, but sometimes, very seldom, do I use my loan to pay for.
- I use it often to accrue cash back money awards.
- Nice to have a way to pay large sums of money slowly through credit card.
- I use my credit card for internet purchases mostly because it is safer than using my debit card.
- I think credit cards should only be used for emergency purchase.
- Use a credit card instead of cash, because it is easier to keep track of my spending.
- The credit card is useful when traveling because you don't have to keep cash on you in what may be a dangerous situation. However, it can be tempting to use it when you see something that you really want, but don't have the money for.

#### **Conclusion:**

when you have credit card and also you pay the your credit card bills in time you get certain benefit .when you delay pay the credit card bill you must pay high interest it's a disadvantage. When you have credit card you don't have cash itself certain time period you can survive. Debt is created by poor spending habits and irresponsibility. Debt consolidation is a great opportunity for employers to try, but payments must always be on time!

#### **Bibliography References:**

1. [www.google.com](http://www.google.com)
2. [www.axisbank.com](http://www.axisbank.com)
3. [www.ibm.com/spss](http://www.ibm.com/spss)
4. [www.wikihow.com](http://www.wikihow.com)
5. [www.hdfcbank.com](http://www.hdfcbank.com)
6. [www.google.com](http://www.google.com)
7. [www.amazon.com](http://www.amazon.com)
8. [www.kotlermarketing.com](http://www.kotlermarketing.com)
9. [www.goodreads.com](http://www.goodreads.com)

# A Study on Impact of Select Exchange Rate Fluctuations on Indian Stock Market

Mr. P.S.Viswanadh, Mrs. V.V.L. Sameera

*Asst.Professor, SMS, SNIST*  
*Asst.Professor, MBA Dept, MREC*

VISWANATHP@REDIFFMAIL.COM

LAKSHMIFEB24@GMAIL.COM

## Abstract:

This study is aimed to examine the long and short run impact of select currency exchange rate fluctuations on Indian Stock Market. In order to examine in impact, the researcher has considered US Dollar (USD), Euro and Yuan currency fluctuations against Indian rupee as independent variables and National Stock Exchange Index for Fifty (NIFTY) as a dependent variable. The study has been conducted for the period of five years from September 2010 to August 2015. The researcher has used the Unit root test, Johansen Co-integration test, Granger Causality Test, Vector Error Correction Model (VECM) and Regression analysis in order to meet the desired objectives. The study concludes that the chosen variables such as NIFTY, USD, Euro and Yuan were co-integrated, but the Euro and Yuan were not having long run association with NIFTY. The study also concludes that 54.59 percent variations can be examined in NIFTY with the select currency fluctuations.

## Key Words:

Globalized Economy, Currency Fluctuations, Volatility, Exchange rate, Devaluation of currency, Stationery

## Introduction:

In today's globalized economy and the international financial system, the effect of macroeconomic factors on stock markets is increasing day to day. These macroeconomic factors were dominating the fundamentals and technicalities prevailed in the domestic markets. In the recent past, Indian Stock Markets have witnessed biggest crash of around 7 lakh crores worth of investments have evaporated in a single day on 24<sup>th</sup> August, 2015, the SENSEX fell as much as 1,741 points to its one-year low of 25,625, while the 50-share Nifty tumbled to 7,769<sup>1</sup> for the reason that of China's decision towards devaluation of their currency<sup>2</sup>, we have also witnessed the volatility in global markets for Euro crisis<sup>3</sup>. In this regard the

researcher felt that there is a need of analyzing the short run and long run relation and the impact of US dollar, Euro and Yuan fluctuations on Indian Stock Market in order to examine the current market situations and to reveals the facts to the Investors.

## Review of Literature:

**a. Vivek Bhargava, Daniel Konku** examined the impact of exchange rate fluctuations on US stock market returns, with reference to S&P 500. Using the returns on S&P 500 index as the dependent variable and changes in the US dollar-Euro exchange rates as the independent variable, regression technique has used to determine whether stock market returns depend on exchange rate. The paper also investigates the correlation between exchange rate and stock market returns to determine whether there are any interdependencies between the two. Using different GARCH models their results show the dollar appreciation impacts the S&P 500 negatively. Exchange rate volatility also increases the volatility of the S&P and reduces returns.

**b. Amare and Mohsin** (2000) inspected the long-run association among stock prices and exchange rates for nine Asian countries such as Japan, Hong Kong, Taiwan, Singapore, Thailand, Malaysia, Korea, Indonesia, and Philippines. They use monthly data from January 1980 to June 1998 and employed co-integration technique. The long-run relationship between stock prices and exchange rates was found only for Singapore and Philippines. They attributed this lack of co-integration between the said variables to the bias created by the "omission of important variables". When interest rate variable was included in their co-integrating equation they found co-integration between

---

the collapse of financial institutions, high government debt and rapidly rising bond yield spreads in government securities. The European sovereign debt crisis started in 2008, with the collapse of Iceland's banking system, and spread to primarily to Greece, Ireland and Portugal during 2009. The debt crisis led to a crisis of confidence for European businesses and economies. In June 2015 Greece, with divided political and fiscal leadership and a continued recession, was facing a sovereign default.

---

<sup>1</sup>. NDTV Profit news about SENSEX and NIFTY

<sup>2</sup> **"Devaluation"** means official lowering of the value of a country's **currency** within a fixed exchange rate system, by which the monetary authority formally sets a new fixed rate with respect to a foreign reference **currency**.

<sup>3</sup> The European sovereign debt crisis occurred during a period of time in which several European countries faced

stock prices, exchange rates and interest rate for six of the nine countries.

**Golaka C. Nath\* and G. P. Samanta** examined the dynamic linkages between the foreign exchange and stock markets for India. Empirical results show that generally returns in these two markets are not inter-related, though in recent years, the return in stock market had causal influence on return in exchange rate with possibility of mild influence in reverse direction. These results have opened up some interesting issues regarding the exchange rate and stock price causal relationship.

**Reza Ahmadi, 1Mehdi Rezayi and 2Mehrzad Zakeri** investigated the effect of exchange rate volatility on stocks return in different industries of listed companies in Tehran Stock Exchange using GJR-GARCH model. The data were collected monthly including seven fields of industries: Automotive, metals, machinery, cement, pharmacy, food and chemistry over 1384-1389. They find strong evidence of exchange rate exposure in all three aspects. This implies that the entire currency risk actually faced by firms is not fully captured by the traditional “exchange rate exposure coefficient” alone.

**Peleg Litali Ambunya** This study set to establish the relationship between exchange rate movement and stock market returns volatility at the Nairobi Securities Exchange. The study used secondary data collected from the Nairobi Securities Exchange and the Central bank of Kenya for the period 2007-2011. The study regressed stock market returns volatility against exchange rate movement. From the regression output, the study established that exchange rate movements greatly affected the stock market return volatility owing to its information content to the investors. With high fluctuations in the exchange rates, the exchange rates movement became bigger accompanied by a huge stock market return volatility. Study concludes that there is a strong relationship between exchange rate movement and stock market returns volatility. This is especially carried through the information content of exchange rate movement on the security's business. The study concludes that exchange rate movement also affects the stock market performance greatly through its spiral effects.

#### **Gap Analysis:**

The above mentioned researchers have conducted their studies related to past and they have not focused on the current issues prevailed in the global markets such as Euro Crisis, Chinese' devaluation of currencies. The retail investors were afraid of their investments due to the current trends prevailed in the Indian stock markets. In this regard the researcher thought to focus on the current issues and made an attempt to find out the long and short run relationship between the chosen currencies and NIFTY.

#### **Research Methodology:**

##### **I. Objectives of the Study:**

1. To study the long run and short run association of USD, Euro and Yuan with NIFTY.
2. To analyze the impact of fluctuations on USD, Euro and Yuan on volatility of NIFTY.

##### **II. Hypothesis:**

H01: The select variables such as NIFTY, USD, Euro and Yuan are not stationery.

H02: The variables are not co-integrated with each other.

H03: There is no long run association among the chosen variables with NIFTY.

H04: There is no short run association between select variables with NIFTY.

H05: There is no impact of USD, Euro and Yuan on NIFTY.

##### **III. Scope of the study:**

The study is conducted only for three country currencies such as US Dollar, Euro and Yuan in order to concentrate on the current market conditions. The study is also limited to the period of five years daily data from September, 2010 to August 2015.

##### **IV. Source of data:**

The present study is depending upon the secondary data only. The daily data of NIFTY and various currency values against Indian rupee has been taken from the website of National Stock Exchange, Yahoo Finance and Investing.

##### **V. Sample Size:**

The study has considered 1252 trading days' data from September 2010 to August 2015.

##### **VI. Tools and Techniques:**

In order to meet the desired objectives of this study the Researcher has used the E-Views software as a tool for calculations, and he has considered the following statistical techniques for analysis.

**a. Unit Root Test:** The data used for time series analysis must be stationery, which means the data throughout the period must be in proper order, and then only it will be useful for time series analysis. One can use the Unit Root Test for understanding the stationery of the data. When the available data is non stationery at 'level' after making the unit root test, then we should convert in to the first difference order and conduct the unit root test again, even then the data is non stationery then we should go for second difference order in order to get the desired data.

One can analyze the results of Unit root test with the help of probability value. If the probability value is less 5 percent then we can consider that the data is stationery and if the probability value is greater than 5 percent then we can consider that the data is not stationery.

**b. Johansen Co-Integration Test:** This test helps us to know whether the variables are co-integrated or not. If the variables are co-integrated then we can use Vector Error Correction mode (VECM) for analyzing the long run and short run causality where as if the variables are not co-integrated then one have to use Vector Auto Regressive

Sl. No	Name of the Variable	P Value at Level data	P Value at First Difference
1	NIFTY	0.8442	0.00
2	USD	0.8998	0.00
3	Euro	0.4110	0.00
4	Yuan	0.3627	0.00

(VAR) model. This test must be conducted with the 'level' data only. The test will generate various Null Hypothesis such as none is co-integrated, at most one is co-integrated etc., based on the number of variables we have chosen. It will also give us the probability value for various null hypotheses. If the probability value is less than 5 percent then we can reject the null hypothesis.

**c. Granger Casualty Test:** It is used for testing the long run association between variables. When the variables are not granger caused then we can concluded that there is no long run association between them. The results of the test will generate different null hypothesis between variables. With the help of probability value in the results, one can conclude that, there exists a long run association between the variables or not.

**d. Vector Error Correction Model (VECM):** This model helps us to find out long run and short run association between the dependent and independent variables. Initially it generates an equation between dependent and independent variables. In that equation we can observe the long run and short run associated variables. One can observe the long run association with the help of co-efficient term and its probability value, where as when we wanted to find out the short run association, we should make another test named as 'Wald test'. This will generate an F and Chi-square Statistic. If the probability value of F and Chi-square statistic is less than 5 percent then we can conclude that the independent variables will have short run impact on dependent variable, where as if the probability value is greater than 5 percent then independent variables doesn't have short run impact on dependent variable.

**e. Regression Analysis:** The regression analysis helps us to measure impact of Independent variables on Dependent variables. The  $R^2$  value in the regression analysis calculation measure the total impact of Independent variables on dependent variable. The analysis also

generates the co-efficient and corresponding probability values. If the probability value is less than 5 percent of a particular variable then we can consider that the independent variable has a significant impact on the dependent variable. If the probability value is greater than 5 percent then one can conclude that the corresponding independent variable is not significant to explain the dependent variable.

## VI. Limitations of the Study:

Though there are other stable currencies and macroeconomic variables are there in economy the researcher has restricted the study to three currency fluctuations in order present the recent crashes in the Indian Stock markets.

### Analysis of First Hypothesis ( $H_{01}$ )

#### Results of Unit Root Test:

Source: E-views out put

The above table reveals that the chosen variables such as NIFTY, USD, Euro and Yuan at level data are not stationery since the Unit Root Test 'P' Value is greater than 0.05 or 5 percent. Therefore we have no evidence to reject the null hypothesis rather we accept the Null hypothesis ( $H_{01}$ ) i.e., the variables are not stationery.

Since the non stationery data is not suitable for time series analysis, one must convert the data into stationery with the help of first difference technique. Therefore the data has been modified into first difference and tried to measure the 'P' value for Unit Root test. These values indicate that the chosen variables are stationery at first difference since 'P' value is less than 5 percent.

### Analysis of Second Hypothesis ( $H_{02}$ )

#### Results of Johansen Co-Integration Test:

Unrestricted Co integration Rank Test (Trace)				
Hypothesized		Trace	0.05	
No. of CE(s)	Eigen value	Statistic	Critical Value	Prob .**
None *	0.126613	184.1403	47.85613	0.00
At most 1	0.006224	14.64834	29.79707	0.8023
At most 2	0.005441	6.831436	15.49471	0.5973
At most 3	8.61E-07	0.001078	3.841466	0.9736

Source: E-views out put

The above table reveals that chosen variables are co-integrated with each other since the Probability value of Johansen Co-Integration test at 'None' is less than 0.05 or 5 percent and the Probability value 'at most 1', 'at most 2', and 'at most 3' variables' are greater than 5 percent, therefore we have enough evidence to reject the null hypothesis i.e., variables are co-integrated.

#### Analysis of Third Hypothesis ( $H_{03}$ )

##### Results of Granger Casualty Test:

Source: E-views out put

Pair wise Granger Causality Test			
Sample: 9/01/2010 8/28/2015			
Lags: 4			
Null Hypothesis:	Observations	F-Statistic	Prob.
DUSD does not Granger Cause DNIFTY	1252	2.49002	0.0417
DNIFTY does not Granger Cause DUSD		0.96085	0.428
DEURO does not Granger Cause DNIFTY	1252	1.44561	0.2167
DNIFTY does not Granger Cause DEURO		10.2416	4.00E-08
DYUAN does not Granger Cause DNIFTY	1252	1.5143	0.1956
DNIFTY does not Granger Cause DYUAN		2.75393	0.0268
DEURO does not Granger Cause DUSD	1252	1.73725	0.1394
DUSD does not Granger Cause DEURO		74.3941	1.00E-56
DYUAN does not Granger Cause DUSD	1252	0.02749	0.9985
DUSD does not Granger Cause DYUAN		0.45125	0.7715
DYUAN does not Granger Cause DEURO	1252	0.83022	0.5059
DEURO does not Granger Cause DYUAN		0.60619	0.6582

The above table reveals that the US Dollar has the long run association with NIFTY since the probability value of Granger causality test is less than 0.05 or 5 percent. The other variables such Euro and Yuan were not having long run association with NIFTY. It is also observed that NIFTY has long run impact on Yuan since the probability value (0.0268) is less than 5 percent at DNIFTY Vs DYuan. Therefore we have enough evidence to reject the Null Hypothesis about long run association between USD

and NIFTY although we have no evidence to reject the Null Hypothesis between Euro, Yuan and NIFTY.

#### Analysis of Fourth Hypothesis ( $H_{04}$ )

##### Results of Vector Error Correction Model (VECM):

Wald Test:			
Test Statistic	Value	df	Probability
F-statistic	1.375664	(12, 1234)	0.1709
Chi-square	16.50797	12	0.1691
Null Hypothesis: C(6)=C(7)=C(8)=C(9)=C(10)=C(11)=C(12) =C(13)=C(14)=C(15)=C(16)=C(17)=0			

Source: E-views out put

The above table concludes that there is no proper evidence to reject the null hypothesis about the short causality between NIFTY and other chosen variables, since the chi-square probability value (0.1691) of Wald test is greater than 0.05 or 5 percent.

#### Analysis of Fifth Hypothesis ( $H_{05}$ )

### Results of Regression Analysis:

Dependent Variable: NIFTY				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2145.092	227.102	9.4455	0
			-	
EURO	-93.05208	5.084052	18.30274	0
USD	189.9055	7.799506	24.3484	0
			1.21047	
YUAN	38.92833	32.15953	6	0.2263
R-squared	0.545953	Mean dependent var		6297.335
Adjusted R-squared	0.544866	S.D. dependent var		1190.581
S.E. of regression	803.2094	Akaike info criterion		16.21829
Sum squared resid	8.08E+08	Schwarz criterion		16.23463
Log likelihood	-10189.19	Hannan-Quinn criter.		16.22443
		Durbin-Watson stat		0.021702
F-statistic	502.2087			
Prob (F-statistic)	0			

Source: E-views out put

The R-square value in the above table reveals that the independent variables such as USD, Euro and Yuan together can predict the NIFTY at the rate of 54.59 percent, and the rest of 45.41 percent of variations were dependent on other factors that are not included in this study.

USD, Euro and Yuan together are significant variables to explain the changes in NIFTY in other words we have enough evidence to reject the null hypothesis ( $H_{05}$ ) since the probability value of F statistic (0.00) is less than 5 percent. Independently the USD and Euro were significant variables and Yuan cannot be considered as significant variable to explain the changes in NIFTY since its probability value (0.2263) is greater than 5 percent.

The above table also reveals that for every change in one unit of Euro, USD and Yuan leads to -93.05, 189.9055 and 38.928 points of change respectively on NIFTY.

### Findings:

1. The select variables such as NIFTY, USD, Euro and Yuan are not stationery data at level and these became stationery at first difference.

2. The chosen variables are co-integrated with each other.
3. Except the USD, there are no other variables have long run association with NIFTY.
4. The study also found that NIFTY will show an impact on Yuan in long run.
5. There is no short run causality or relationship between chosen variables and NIFTY.
6. The USD, Euro and Yuan together can predict the NIFTY at the rate of 54.59 percent and the rest of 45.01 percent of variations in NIFTY are due to other reasons that are not included in the study.
7. The study also reveals that the NIFTY will witness about -93.05, 189.9 and 38.928 points changes for every one unit of change in Euro, USD and Yuan respectively.

### Conclusion:

The study concludes that the USD has long run association and it is having greater impact on NIFTY where as Euro and Yuan were not having long and short run association with NIFTY. The chosen variables will have only 54.59 percent impact on NIFTY, and the further volatility is due to other reasons. Therefore the investors need not be panic to invest in the market just by observing the recent fluctuations in market because the study found that the variations are only a temporary phenomenon but not a long run quandary.

### References: I. Text Books:

- a. "Financial Econometrics with E views" written by Roman Kozhan, Ventus Publishing Aps, ISBN No. 978-87-7681-427-4.
- b. "Eviews Illustrated for version 8" written by Richard Startz, University of California, Santa Barbara.
- c. Punithavathi, Pandian, (2001) "Security Analysis and Portfolio Management", Vikas Publishing House, PP 109-127.
- d. S.Kevin "Security Analysis and Portfolio Management", 2<sup>nd</sup> edition, Printice Hall Publishing House. Prasanna Chandra "Investment Analysis and Portfolio Management", 3<sup>rd</sup> edition, Tata Mc Graw Hill Publishers.
- e. Deepak Chawla and Neena Sondhi "Research Methodology, Concepts and Cases", 1<sup>st</sup> edition, Vikas Publishing House, PP 460-479.

### II. Websites:

- a. [www.sayedhossain.com](http://www.sayedhossain.com)
- b. [www.nseindia.com](http://www.nseindia.com)
- c. [www.investing.com](http://www.investing.com)
- d. [www.yahoofinance.com](http://www.yahoofinance.com)

## About MREC

Intellectuals are not born but are made. They live a trail with unmatched perseverance, rocky commitment in their principled endeavors. Visionaries' logical and reasonable encounters in their walks of life, become values, these conscientious values are irradiating millions of student fraternity and kindled their careers.

This millennium poses numerous challenges to educational institutions, who chant quality mantra. An institution of par excellence, has reached unflinching success, in its strides of imparting quality education recognized. This feat was achieved by the Chairman of MRGI, in all his humble, mesmerizing and charismatic ways.

Owing to all this, MRGI has become a house-hold name; it has become the destination point for scintillating success and unparallel careers. As the largest Hi-tech, State-Of-art facilitated private campus in Southern-India. It is catering to the needs of infotainment or technical explosion.



**Sri CH. MALLA REDDY**  
Founder Chairman-MRGI  
Hon'ble M.P, Malkajgiri Constituency

## About PRINCIPAL

Dr. S. Sudhakara Reddy obtained his B.Tech Degree in Mechanical Engineering (Production) from Kakatiya University, Warangal, M.Tech Degree in Design and Production Engineering (Machine Tools) from NIT, Warangal and M.B.A (Human Resource Management) from Dr. B.R. Ambedkar Open University, Hyderabad. He obtained his Ph.D. in Mechanical Engineering from S.V. University Tirupathi.

The research work related to his Ph.D was carried out at Defence Metallurgical Research Laboratory (DMRL) and Advanced Research Center for Powder Metallurgy and New Materials (ARCI) at Hyderabad. His area of research includes Tribology, CAD/CAM & Advanced Manufacturing Technology.

He is having 2 years industrial and 24 years of teaching as well as administrative experience at Under-Graduate and Post-Graduate level at various institutions. He has authored 4 textbooks, the most recent book is titled "Effect of Laser treatment on Tribological Alsi alloys" is being published by LAMBERT Academic Publisher.

In addition to teaching he also worked as Associate NCC Officer, NSS Office, Training and Placement Officer and Head of the Mechanical Engineering Department.



**Dr. S. Sudhakara Reddy**  
Principal - MREC(A)



**Malla Reddy Engineering College** (Autonomous)

Maisammaguda, Dhulapally (Post via kompally), Secunderabad - 500100

Ph: 040-64634234, 040-65864982, 040-32914085, Mob: 9348161125, 9052945735

Email: principal@mrec.ac.in

[www.mrec.ac.in](http://www.mrec.ac.in)

EAMCET/ICET code: **MREC**

