

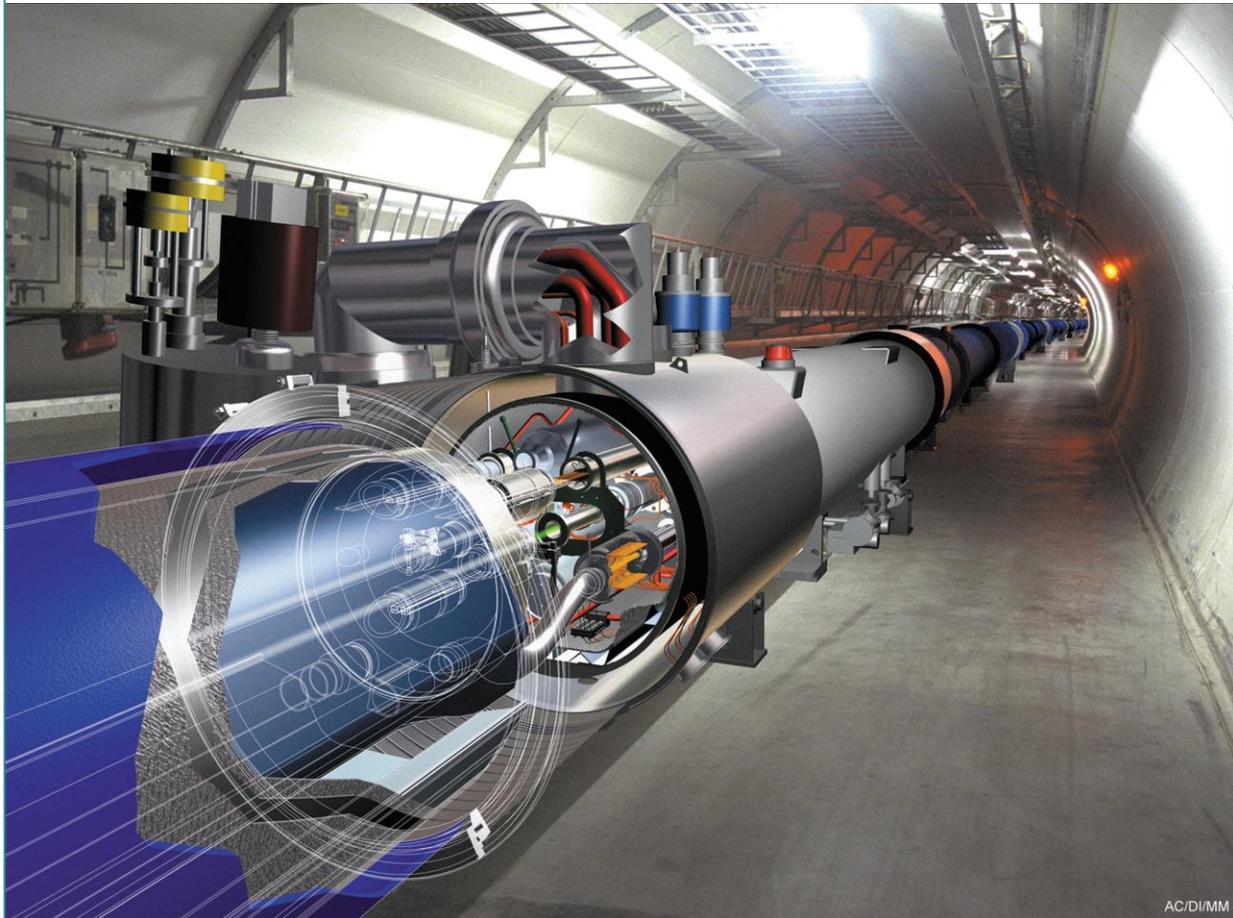
International Journal of computational Engineering Research (IJCER)

ISSN: 2250-3005

VOLUME 2

May-June 2012

ISSUE 3



Email: ijceronline@gmail.com

Url : www.ijceronline.com

International Journal of computational Engineering Research (IJCER)

Editorial Board

Editor-In-Chief

Prof. Chetan Sharma

Specialization: Electronics Engineering, India
Qualification: Ph.d, Nanotechnology, IIT Delhi, India

Editorial Committees

DR.Qais Faryadi

Qualification: PhD Computer Science
Affiliation: USIM(Islamic Science University of Malaysia)

Dr. Lingyan Cao

Qualification: Ph.D. Applied Mathematics in Finance
Affiliation: University of Maryland College Park,MD, US

Dr. A.V.L.N.S.H. HARIHARAN

Qualification: Phd Chemistry
Affiliation: GITAM UNIVERSITY, VISAKHAPATNAM, India

DR. MD. MUSTAFIZUR RAHMAN

Qualification: Phd Mechanical and Materials Engineering
Affiliation: University Kebangsaan Malaysia (UKM)

Dr. S. Morteza Bayareh

Qualificatio: Phd Mechanical Engineering, IUT
Affiliation: Islamic Azad University, Lamerd Branch
Daneshjoo Square, Lamerd, Fars, Iran

Dr. Zahéra Mekkioui

Qualification: Phd Electronics
Affiliation: University of Tlemcen, Algeria

Dr. Yilun Shang

Qualification: Postdoctoral Fellow Computer Science
Affiliation: University of Texas at San Antonio, TX 78249

Lugen M.Zake Sheet

Qualification: Phd, Department of Mathematics
Affiliation: University of Mosul, Iraq

Mohamed Abdellatif

Qualification: PhD Intelligence Technology
Affiliation: Graduate School of Natural Science and Technology

Meisam Mahdavi

Qualification: Phd Electrical and Computer Engineering
Affiliation: University of Tehran, North Kargar st. (across the ninth lane), Tehran, Iran

Dr. Ahmed Nabih Zaki Rashed

Qualification: Ph. D Electronic Engineering

Affiliation: Menoufia University, Egypt

Dr. José M. Merigó Lindahl

Qualification: Phd Business Administration

Affiliation: Department of Business Administration, University of Barcelona, Spain

Dr. Mohamed Shokry Nayle

Qualification: Phd, Engineering

Affiliation: faculty of engineering Tanta University Egypt

CONTENTS :

S.No.	Title Name	Page No.
1	Image Analysis Techniques for Fingerprint Recognition Vidyadevi G Biradar, H Sarojadevi	606-615
2	Computational studies of swirl effects on instabilities and pollutions due to non-premixed turbulent combustion Luthenda Gamany, Taha Janan Mourad, Agouzoul Mohamed	616-627
3	Analysis of Skew Bridges Using Computational Methods Vikash Khatri, P. R. Maiti, P. K. Singh, Ansuman Kar	628-636
4	A New Omni-directional Monopole Antenna for Interference Reduction T .S. Ghouse Basha, K.Tulasi Krishna, C.Chandrakala, V.Kishore, D.Aruna	637-641
5	A Novel Two Stage Binary Image Security System Using (2,2) Visual Cryptography Scheme Mr. Rohith S, Mr. Vinay G	642-646
6	Numerical Investigation of Secondary Flow In An Axial Flow Compressor Cascade T. Suthakar, Akash Dhurandhar	647-656
7	A Modified Marker Controlled Watershed Algorithm With Linear Convolution For Medical Image Segmentation Brijesh Shah, Jigar Modh, Satish Shah	657-661
8	Smart Lighting And Control Using Msp430 & Power Line Communication Sanjay Belgaonkar, E. Elavarasi, Gurjeet Singh	662-666
9	The e-Health scenario with latest trends in EMR applications: A Review of EMR techniques with healthcare framework Onkar S Kemkar, Dr P B Dahikar	667-670
10	E-Governance Initiative In Technical Institute – A Case Study Of Dr.B.R.Ambedkar Institute Of Technology, Andaman & Nicobar Islands (India) Utpal Sharma, Sunil Kumar Chakraborty	671-680
11	Optimal Power Flow Using Differential Evolution Algorithm With Conventional Weighted Sum Method Rohit Kumar Verma, Himmat Singh, Laxmi Srivastava	681-685
12	Minimization of Reactive Power Using Particle Swarm Optimization Vivek Kumar Jain, Himmant Singh, Laxmi Srivastava	686-691
13	Speaker Recognition System Using Combined Vector Quantization and Discrete Hidden Markov model Ameen Khan A, N V Uma Reddy, Madhusudana Rao	692-696
14	Design, Implementation and Performance Analysis of an Integrated Vedic Multiplier Architecture Ramachandran.S, Kirti.S.Pande	697-703
15	Performance Comparison Study Of AODV, OLSR And TORA Routing Protocols For MANETS Manjeet Gupta, Sonam Kaushik	704-711

16	EIA for Ramapada Sagar (Polavaram) Irrigation Project using the Model of RS and GIS Sreeramulu. Y, Murali Krishna.I.V	712-719
17	Speaker Features And Recognition Techniques: A Review Mrs. Sharada Vikram Chougule, Dr. M S. Chavan	720-728
18	Computing Over a Multi Cloud for MTC Applications Challa Vanitha Reddy, Battula Sudheer Kumar	729-731
19	An Overview Of Smart Antennas And Its Techniqes Beamforming And Diversity Suraya Mubeen, Dr.A.M.Prasad, Dr.A.Jhansi Rani	732-736
20	A Comparative Analysis of Fuzzy C-Means Clustering and K Means Clustering Algorithms Mrs. Bharati R.Jipkate, Dr. Mrs.V.V.Gohokar	737-739
21	Performance Analysis of Timing Attack on Elliptic Curve Cryptosystem Mr. Praful V. Barekar, Prof. K. N. Hande	740-743
22	Hand Held Emergency Wireless Telemedicine System Suganthi.J. N.V.Umareddy, Sridharan.B	744-747
23	Secure and Reliable Data Transmission in Wireless Sensor Network: A Survey Rudranath Mitra, Rudranath Mitra, Tauseef Khan	748-754
24	Detection and Classification of Epileptic Seizures using Wavelet feature extraction and Adaptive Neuro-Fuzzy Inference System Dr. D. Najummissa, Dr. T. R. Rangaswamy	755-761
25	Customer and User Requirements Modeling Enhanced Software Development Tawfik Saeed Zeki	762-765
26	Enhanced Clusterhead Selection Algorithm Using LEACH Protocol for Wireless Sensor Networks Rudranath Mitra, Anurupa Biswas	766-770
27	Implementing VGA Application on FPGA using an Innovative Algorithm with the help of NIOS-II Ashish B. Pasaya, Kiritkumar R. Bhatt	771-775
28	Vibration Analysis Of Glass Fiber Reinforced Composites R. K. Mishra	776-789
29	Comparitive Study of Advanced Database Replication Strategies A. Pramod Kumar, B.Sateesh	790-794
30	Vehicular Number Plate Recognition Using Edge Detection and Characteristic Analysis of National Number Plates Bharat Raju Dandu, Abhinav Chopra	795-799
31	Comparative Analysis of Image Registration using SIFT and RANSAC method Riddhi J Ramani, N.G. Chitaliya	800-805
32	Assessment of Radiation Emission from Waste Dumpsites in Lagos State of Nigeria Olubosedo. O, Akinnagbe .O.B., Adekoya O	806-811

33	Performance Evaluation of different α value for OFDM System Dr. K.Elangovan	812-818
34	Image Segmentation Using Active Contour Model Abhinav Chopra, Bharat Raju Dandu	819-822
35	Sbpgp Security Model Using Iodmrp Meenakshi Mehla, Himani Mann	823-828
36	Design of Neural Architecture in 0.35 μ mTechnology Using Analog VLSI Mr.Maulik B.Rami, Prof.H.G.Bhatt, Prof.Y.B.shukla	829-837
37	Combining Multimedia Building Blocks In Image Analysis P. Shanmugam, Dr. C. Loganathan	838-842
38	An Intrinsic Dislocation Density – Finite Element Formulation Of Metal Plasticity Njoroge K. D., Mutuli S. M., Kihui J. M	843-850
39	Dynamic Analysis Of Dislocation Cores In the a- Fe Lattice Using The Embedded Atom Model K. D. Njoroge, G. O. Rading, J. M. Kihui, M. J. Witcomb, L. A. Cornish	851-859
40	Performance Enhancement in Mobile Computing Using Replicated Cache Agent Meenakshi Mehla, Reena Dahiya	860-867
41	Oscillation Test Methodology for Built-In Analog Circuits Ms. Sankari.M.S and Mr.P.SathishKumar	868-877
42	A new approach Data hiding in 2D data matrix and tilt correction algorithm Kimmy Ghanaiya, Gagandeep Kaur	878-881
43	Estimation of the Periods of Occurrence of Spread– F Over Ouagadougou Tomwa, A. C, Adimula, I. A	882-890
44	Application Specific Quality of Service (QoS) Centric Parameters Simulation in Wireless Mobile Ad hoc Networks Padmashree S, Manoj P B	891-897
45	The Implementation OF Prosthetic Index Finger Based On EMG Signals Amanpreet Kaur, Gagandeep Kaur	898-900
46	Design of an Intelligent SMS Based Remote Metering System for AC Power Distribution to HT and EHT Consumers Mrs. Mahalakshmi N, Mr.KrishnaiahParamesh, Ms. Elavarasi E	901-911
47	A study on strength properties of roller compacted concrete with 30% replacement of OPC 53 grade cement by flyash Ganapati Naidu. P, B. Jagannadh Kumar, S.Adishes, P.V.V.Satayanarayana	912-918
48	Motion Detection Method to Compensate Camera Flicker Using an Algorithm Alam Inder Singh, Gagandeep Kaur	919-922
49	An Efficient Approach for Mining Frequent Itemsets with Large Windows K Jothimani, S. Antony Selvadoss Thanmani	923-926

50	Recent Developments In Preparation Of Non Conventional Activated Carbons Prof Shantini Bokil, Prof.Dr.R.K.Rai, Prof.Dr.S.N.Kaul	927-934
51	Credit Card Fraud: Bang in E-Commerce Khyati Chaudhary, Bhawna Mallick	935-941
52	Efficient Moving Object Detection Based On Statistical Background Modeling Kusuma.U, S.T Bibin Shalini	942-945
53	Designing Continuous-Time Observers for Linear Hybrid Systems with Application to Three Tank Model Mohammad Amin Zahedi Tajrishi, Behrouz Rezaie, Reza Ghaderi	946-954
54	Single Precision Floating Point Divider Design Serene Jose, Sonali Agrawal	955-958
55	Improvement in Enhanced Privacy Protection in LBS Rashmi Samele, Vivek Singh Senger	959-963
56	Reliability Analysis: The Mathematical Expression Pooja Parnami, Ruchi Dave, Neha Singha, Ankur Dutt Sharma	964-967
57	Simulation of Beamforming Solution of Intereference Reduction for High Altitude Systems Bharati.L.Rathod, Mr.Hemanthkumar.P, Mr.Aaquib.Nawaz.S	968-971
58	Floating Point Unit Implementation on FPGA Deepa Saini, Bijender M'dia	972-976

Image Analysis Techniques for Fingerprint Recognition

Vidyadevi G Biradar¹, H Sarojadevi¹

1. Nitte Meenakshi Insitute of Technology, Bangalore , Karnataka, India – 560 064

Abstract

Fingerprint recognition is a method of biometric authentication that uses pattern recognition techniques based on fingerprints image of the individual. Fingerprint patterns are full of ridges and valleys and these structures provide essential information for matching and classification. The steps for fingerprint recognition include image acquisition, preprocessing, feature extraction and matching. A number of pattern recognition methods have been used to perform fingerprint matching. In this paper a survey of fingerprints matching methods are presented, they have been classified into approaches based on minutiae, image transform and hybrid approaches, among them minutiae based methods are widely used, and Hybrid methods are used for more reliable matching with an additional computational cost. Comparison and contrasting of all these methods reveals that a lot of emphasis is put into the design of accurate fingerprint features extractor to improve the classification accuracy.

Keywords- Fingerprint analysis, Biometrics, Minutiae, Image Transforms, Gabor filter

1. Introduction

Biometrics recognition refers to the use of distinctive anatomical and behavioral characteristics or identifiers such as Fingerprints, Face, Iris, Voice, Hand geometry etc. for automatically recognizing a person. The human biometrics recognition technologies are not only rapidly developed but also widely applied to a variety of works such as safeguarding work, legal affairs, personnel authentication, personnel identification, etc.

The advantages of biometric recognition are :1)users do not need to memorize any password,2) users do not need to carry any identification card, 3) users cannot deny their biometric identifications, and 4) reduction in a large amount of expenses on making personal ID cards or relative documents. Among all the biometrics, fingerprint is the most mature and proven technique. A fingerprint is a pattern of ridges and valleys on the surface of the fingertip.

The formation of fingerprints depends on the initial conditions of embryonic development, and their ridge pattern is unchanged throughout the entire life. Both the immutability and the uniqueness properties have determined the use of fingerprint matching as one of the most reliable techniques of people identification [1]. A biometric system can be operated in two modes:1) verification mode and 2) identification mode. A biometric system operating in verification mode either accepts or rejects user claimed identity while a biometric system operating in the identification mode establishes the identity of the user without claimed identity information.

The steps in fingerprint recognition include image acquisition, preprocessing, feature extraction and matching. A fingerprint is the pattern of ridges and valleys on the surface of fingertip. Ridges and valleys of the fingerprint often run in parallel and sometimes they bifurcate and sometimes they terminate. The fingerprint pattern, when analyzed at different scales, exhibits different types of features.

The paper is organized logically into seven parts: Section II and III discuss fingerprint features and gives brief overview of minutiae based matching methods respectively. Section IV discusses the methods based on image transforms. Section V deals with Hybrid methods. Section VI reports comparisons and VII indicates conclusions.

2. Fingerprint Features

The ridge details of the fingerprint are generally described in a hierarchical order at three different levels i.e., level 1, level 2

and level 3 as explained below.

2.1 Level 1 features

At the global level, the ridge line flow delineates a pattern similar to one of those shown in Figure 1. Singular points, called loop and delta (denoted as squares and triangles, respectively in Figure 1), act as control points. Singular points and coarse ridge line shape are useful for fingerprint classification and indexing, but their distinctiveness is not sufficient for accurate matching.

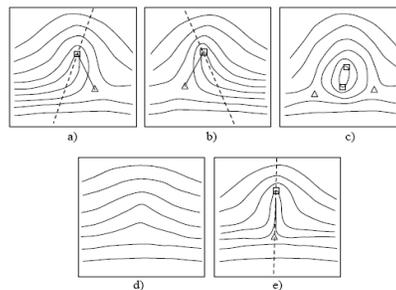


Figure1. Loop and delta (singular points)

2.2 Level 2 features

At the local level, a total of 150 different local ridge characteristics, called minute details, have been identified so far [1]. These local ridge characteristics are not evenly distributed. Most of them depend heavily on the impression conditions and quality of fingerprints and are rarely observed in fingerprints. The two most prominent ridge characteristics, called minutiae (see Figure 2) are: ridge endings and ridge bifurcations. A ridge ending is defined as the ridge point where a ridge ends abruptly. A ridge bifurcation is defined as the ridge point where a ridge forks or diverges into branch ridges. Minutiae in fingerprints are generally stable and robust to fingerprint impression conditions.

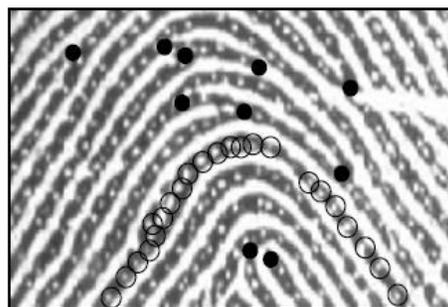


Figure2. Minutiae and sweat pores

2.3 Level 3 features

At the very-fine level, intra-ridge details can be detected. These include width, shape, curvature, edge contours of ridges as well as other permanent details such as dots and incipient ridges. One of the most important fine-level details is the finger sweat pores (see Figure 2), whose positions and shapes are considered highly distinctive [2]. However, extracting very-fine details including pores is feasible only in high-resolution (e.g., 1,000 dpi) fingerprint images of good quality and therefore this kind of representation is not practical for non-forensic applications.

The uniqueness of a fingerprint can be determined by the overall pattern of ridges and valleys and local ridge anomalies. Although fingerprints possess the discriminatory information, designing a reliable automatic fingerprint matching algorithm is a very challenging. As fingerprint sensors are becoming smaller and cheaper, automatic identification based on fingerprints is becoming more attractive alternative, complement to traditional methods of identification. The critical factor in the

widespread use of fingerprints is in satisfying the performance (e.g., matching speed and accuracy) requirements of emerging civilian identification applications.

Various approaches of automatic fingerprint matching have been proposed in the literature. They include correlation based, minutiae based approaches, and image-based approaches as the most prominent classes of fingerprint matching methods.

Fingerprint matching depends on the similarity measure between representative features of fingerprints. Generally, fingerprint recognition is based on a set of relevant local characteristics, such as ridge ending and bifurcation (minutiae). Fingerprint classification is based on fingerprint global features, such as core and delta singularity points. Minutiae-based matching algorithms are the most well-known and widely used for fingerprint matching. For each minutia three features are usually extracted: type, coordinates and orientation. Generally, fingerprint matching is affected by some common problems: too small regions of interest, loss of genuine minutiae, and false minutiae detection. Many approaches have been developed to deal with these issues. Following sections will briefly summarize the most effective solutions.

Many approaches to fingerprint recognition have been presented in the literature. Yet, it is still an actively researched field. Existing approaches can be summarized into four different categories: Minutiae based methods, approaches based on filter, image transform based approaches, hybrid approaches and others. However, many new approaches sprouting out are mainly beyond categorization or a combination of the categories. The steps for fingerprint recognition include image acquisition, preprocessing, feature extraction and matching.

2.4 Performance metrics

Commonly used performance metrics are percentage of recognition accuracy, speed, percentage of False Acceptance Ratio(FAR) and percentage of False Rejection Ratio(FRR) and Receiver Operating Characteristics(ROC) curves. FAR is the probability that the system incorrectly matches the input fingerprint to a non-matching template in the database. It measures the percent of invalid inputs which are incorrectly accepted. FRR is the probability that the system fails to detect a match between the input fingerprint and a matching template in the database. It measures the percent of valid inputs which are incorrectly rejected. ROC plot is a visual characterization of the trade-off between the FAR and the FRR.

3. Brief Overview of Minutiae Based Matching Methods

Minutiae based techniques are most widely used automatic fingerprint recognition techniques. These first locate the minutiae points and then match their relative placement in a given finger and the stored template. Minutiae are the local ridge anomalies – ridge bifurcation or ridge endings etc., [5].The most common minutiae are shown in the Figure 3.

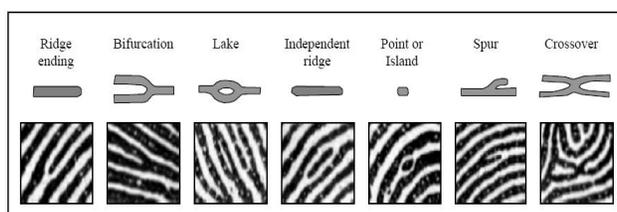


Figure3. Seven most common minutiae

A majority of the existing fingerprint recognition algorithms are based on matching minutia features. Therefore, minutiae extraction is one of the critical steps in fingerprint verification algorithms. Due to different properties of fingerprint sensors and different conditions under which a fingerprint is scanned, the quality of a fingerprint image can vary greatly. Poor quality

fingerprint images lead to missing and spurious minutiae that degrade the performance of the matching system.

Minutiae extraction algorithms [5] produce a large number of spurious minutiae such as break, spur, bridge, merge, triangle, ladder, lake, island, and wrinkle, as shown in Figure 4. Therefore, reliably differentiating spurious minutiae from genuine minutiae in the post-processing stage is crucial for accurate fingerprint recognition. The more spurious minutiae are eliminated, the better the matching performance will be. In addition, matching time will be significantly reduced because of the reduced minutiae number. This is very important since the execution time is a critical parameter in an automatic fingerprint recognition system. Lot of emphasis is given to minutiae post processing techniques.

			
Break	Spur	Merge	Triangle
			
Multiple breaks	Bridge	Break & merge	Ladder
			
Lake	Island	Wrinkle	Dot

Figure4. Examples of spurious minutiae

Researchers have proposed ideas to reduce spurious minutiae. Paper [3] presents a set of algorithms for the extraction of minutiae from skeletonized binary images which perform well on dirty areas and spurious minutiae are reduced by bridge cleaning based on ridge positions. Several post-processing techniques to efficiently remove spurious minutiae including bridge, triangle, ladder, and wrinkle all together are proposed in [4], the experimental results showed that false minutiae rate drops 56.5% after post-processing. Approaches based on neighborhood information are dealt in [5] where matching of two fingerprints is done based on minutiae neighborhood information using directionally selective steerable wedge filters, and Gabor expansion. The area around the core point is treated as the area of interest for extracting the minutiae features as there are substantial variations around the core point as compared to the areas away from the core point. Paper [6] considered the probability distribution of minutiae and achieved classification accuracy of 86.57%.

Authors of paper [7] have combined texture information and neighboring minutiae to obtain a descriptor which is used for minutiae matching. Paper [8] presents a concept of Eigen-codebook for the minutiae verification mechanism of fingerprint image. Principal component analysis (PCA) is applied to find the optimal projection bases for true minutiae regions and false minutiae regions.

Minutiae are difficult to be extracted robustly in low quality images and easily lead to false recognition, to effectively strengthen the performance of fingerprints matching other discriminatory features can be combined with minutiae feature. In [9] minutiae features and their invariant moments are used as a feature set to match the template and input fingerprint. Paper [11] introduces a fingerprint matching algorithm using both ridge features (ridge count, ridge length, ridge curvature direction, and ridge type) and the minutiae feature to increase the recognition performance against nonlinear deformation in fingerprints. Paper [10] proposes an algorithm which divides fingerprint images into two concentric circular regions – inner and outer – based on the degree of distortion. The algorithm assigns weightages for a minutiae–pair match based on the region in which the pair exists. Minutia extractor is built in [12], minutia marking is done by the average inter-ridge width, minutia unification by decomposing a branch into three terminations and matching in the unified x-y coordinate system and in combination alignment based elastic matching algorithm used which is capable of finding the correspondences between minutiae without resorting to

exhaustive research.

Many other approaches exist in the literature - paper[13] introduces a method of robust minutiae matching based on comprehensive minutiae and binary relation between minutiae; paper [14] proposes a method of minutiae extraction considering additional attributes like orientation field and quality map. Minutiae scoring technique is introduced in [15] for fingerprint; paper [16] presented a method of feature extraction using minutiae spectral features; paper [17] has presented a method of fingerprint matching based on five nearest neighbors of center minutiae in the first stage and second stage matching is based on fuzzy logic. Gabor basis functions are used to extract discontinuous points in the fingerprint image and classify the points into core, deltas and minutiae in [18].

4. Methods Based on Image Transforms

Minutia based approaches are the most popular ones being included in almost all contemporary fingerprint identification and verification systems. Although rather different from one other the minutiae-based approaches require extensive preprocessing operations in order to reliably extract the minutia features. The preprocessing operations include image enhancement, orientation flow estimation, ridge segmentation, ridge thinning, and minutiae detection. In addition, a minutiae purification stage is also required in order to reduce the number of false minutiae erroneously detected in noisy fingerprint images. Image-based approaches do not use the minutiae features for fingerprint matching. They are usually applied directly onto the grayscale fingerprint image without pre-processing, and hence they may achieve higher computational efficiency than minutiae- based methods. In addition, the image-based approaches may be the only choice to match fingerprints which have too low image quality to allow reliable minutiae extraction. The main disadvantage of image-based approaches consists in their limited ability to track with variations in position, scale and orientation angle. Usually the variation in position between the two fingerprints is cancelled by choosing a reference point in each fingerprint. Such reference point may be the core point which can be detected using for example methods like those proposed in [23]. Image based approaches include methods based on optical correlation and transform based features. The following papers discuss mainly on image transform based methods.

Fingerprint recognition based on features extracted from the wavelet transform of the discrete image is achieved in [19] [21]. A rectangular region around a core point is cropped and wavelet transformation is applied, the energy of different sub bands give information regarding the ridge spatial frequency as well as the ridge orientation and this information is represented based on standard deviation of each wavelet sub image. [20] proposed an algorithm for fingerprint identification using wavelet packet analysis as best basis selection. Each fingerprint is decomposed using two directional wavelet packet family corresponding to different scales. The energy distribution or the fingerprint in each sub band is extracted as a feature for identification. Critical wavelet coefficients are selected to form a feature vector of the fingerprint. In [22] algorithm based on the wavelet transform, and the dominant local orientation which is derived from the coherence and the gradient of Gaussian. Multiresolution descriptors for fingerprint recognition are presented in [23]. Firstly it computes the Discrete Fourier Transform (DFT) for the given fingerprint image then transform it to polar coordinate (r, θ) using the centre of mass of the pattern as origin, then apply the Fourier transform along the axis of polar angle θ and the wavelet transform along the axis of radius r . The features thus obtained are invariant to translation, rotation, and scaling. Haar wavelet decomposition [24] is taken on the fingerprint images and then, after the decomposition, the matching is carried out over the Gabor features extracted from the detailed sub-images. [27] Developed fingerprint recognition method using Haar wavelet transform and achieved verification rate of 82.08 overall verification rate (Genuine Acceptance Rate) even by rotating each fingerprint image from 0° to 360° in steps of 10° each with FAR of 0.5. Authors have noted that features extraction time per fingerprint image is 0.58 sec. Fingerprint matching is also experimented using combined approaches where minutiae features are combined with the wavelet

based features such a method is applied in [25]. An approach proposed in [26] towards fingerprint recognition is based on wavelet domain features. The 64-subband structure of the FBI (Federal Bureau of investigation) fingerprint compression standard is used to directly extract the wavelet features of the fingerprint image. Paper [28] describes the fingerprint verification based on wavelet transform and the local dominant orientation. Daubechies wavelet is used to decompose the fingerprint image. The local dominant orientation is computed using the coherence, results show 85 percent genuine acceptance rate at 6 percent FAR.

Three different types of transforms, Discrete Cosine Transform (DCT), Fast Fourier Transform (FFT) and Discrete Wavelet Transform (DWT) are used in [29] to create feature vector for fingerprints and are used in matching. First a core point is identified and then around it a image of size 64X64 is cropped in the fingerprint image The transform is applied on the cropped image without any pre-processing. The transform coefficients are arranged in specific manner and are used to obtain the feature vector in terms of standard deviation. The fingerprint matching is based on the minimum Euclidean distance between two feature vectors. Comparison of all the three transforms is presented and it is observed that DCT and DFT gives better result as compared DWT.

5. Hybrid Approaches

The advent of solid-state fingerprint sensors presents a fresh challenge to traditional fingerprint matching algorithms. These sensors provide a small contact area (= 0.6" x 0.6") for the fingertip and, therefore, sense only a limited portion of the fingerprint. Thus multiple impressions of the same fingerprint may have only a small region of overlap. Minutiae based matching algorithms, which consider ridge activity only in the vicinity of minutiae points, are not likely to perform well on these images due to the insufficient number of corresponding points in the input and template images.

A hybrid matching algorithm is presented in [30] to reduce the effect of nonlinear deformation in the fingerprint that uses both minutiae information and texture information for matching the fingerprints. A bank of Gabor filters is used to extract texture features from the template and input images. Results shows that at 1% FAR, the hybrid matcher gives a Genuine Accept Rate of **92%** while the minutiae-based matcher gives a Genuine Accept Rate of **72%**. Gabor filter bank are in [31] to extract fingerprints features. It is found that performance of feature extraction is dependent on mask size and Gaussian deviation value. For Gaussian deviation value 4 and filter mask size 32X32 the percentage of FAR is 6.4%.

In paper [32] a localized texture based representation scheme is presented that relies on visual content for identification and the advantage of this method is it does not require absolute alignment. [33] Proposes a hybrid fingerprint verification system based on local texture pattern obtained using Gabor filtering and wavelet global features obtained by Multiresolution analysis of a fingerprint. The Experimental results show that the system is efficient and suitable for real-time authentication applications with a small size database. Statistical texture analysis of a fingerprint using Spatial Grey Level Dependence method (SGLDM) for discrimination and personal verification is proposed in [34]. Other statistical approach is presented in [35], it reduces multi-spectral noise by enhancing a fingerprint image to accurately and reliably determine a reference point and then extract a 129 X 129 block, making the reference point its center. From the 4 co-occurrence matrices four statistical descriptors are computed. Experimental results show that the proposed method is more accurate than other methods the average FAR is 0.62%, the average FRR is 0.08%, and the Equal Error Rate (EER) is 0.35%.

6. Comparisons

Major categories of fingerprint methods are compared based on the characteristics given in the literature and are tabulated as below (Table 1).

Methods Metrics	Minutiae based	Filter based	Image transform	Hybrid
Reliability	Difficult to extract minutiae in low quality image	Reliable	Reliable	More reliable
Limiting factors	Image quality	Accurate reference point is required	Accurate reference point is required,	depends
Accuracy	High	Less accurate	Less	High
Time delay	High	Less	Less	High
Algorithmic complexity	High	Less	Less	High
Overhead of preprocessing and post processing steps	High	Not applicable	Not applicable	High
Storage space	Efficient	Less	Less	High
Data size	Not suitable for large databases	Suitable for large databases	Suitable for large databases	Not suitable for large databases
Matching techniques	Sophisticated	Simple	Simple	Sophisticated
Applications	Suitable for forensic applications	Suitable for civilian applications	Suitable for civilian applications	Suitable for forensic applications

Table 1. Comparison of Fingerprint Recognition Approaches

7. Conclusion

Fingerprint is an important and challenging research area of Digital Image processing so this field still attracts many researchers even after many years of research. The demand of research in this area is growing as the world is now much more concerned about safety and security than ever. A fingerprint matching and recognition system not only forbids the unauthorized access to some facility but in case of security breach, allows tracking the criminals. The requirement of fingerprint recognition is now in every field of our daily life. This paper presents a study of various fingerprint recognition algorithms from a different viewpoint - quality, performance and application. By classifying fingerprint recognition approaches into - Minutiae based methods, filter based approaches, Image transform based approaches, and Hybrid approaches, we can clearly see the trade-off between various approaches with respect to their suitability towards finger print applications. Many literature have stated their own sets of encouraging results. However, most of them cannot be compared against each other; either the system or the database is different. It would be interesting for future research to evaluate each of the approaches under a standardized test bed.

Acknowledgement

The authors would like to thank Dr. H.C.Nagaraj, Principal, Nitte Meenakshi Institute of Technology, Bangalore, for kind support to the publish the paper.

References

- [1] D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer, New York. 2003
- [2] Mayank Vatsa, Richa Singh, Afzel Noore, Sanjay K. Singh, *Quality Induced Fingerprint Identification using Extended Feature Set*, IEEE transactions.
- [3] Alessandro Farina, Zsolt M. Kova cs-Vajna, Alberto Leone, *Fingerprint minutiae extraction from skeletonized binary images*, pp. 877-889, Pattern Recognition, 1999.
- [4] Feng Zhao, Xiaou Tang, *Duality-based Post-processing for Fingerprint Minutiae Extraction*, pp.36-39, InfoSecu'02, Shanghai, China, July 2002.
- [5] Sharat Chikkerur, Venu Govindaraju, Sharath Pankanti, Ruud Bolle, Nalini Ratha, *Novel Approaches for Minutiae Verification in Fingerprint Images*, Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision (WACV/MOTION'05).
- [6] Jayant V Kulkarni, Jayadevan R, Suresh N Mali, Hemant K Abhyankar, and Raghunath S Holambe, *A New Approach for Fingerprint Classification based on Minutiae Distribution*, pp. 253-258, Vol.1, 2006.
- [7] Jianjiang Feng, *Combining minutiae descriptors for fingerprint matching*, pp.342-352, Pattern Recognition, 2006.
- [8] Ching-Tang Hsieh, Shys-Rong Shyu, *Principal Component Analysis for Minutiae Verification on Fingerprint Image*, 2007.
- [9] JuCheng Yang, JinWook Shin, ByoungJun, Sook Yoon, *Fingerprint Matching using Global Minutiae and invariant Moments*, pp.599-602, Congress on Image and Signal Processing, 2008.
- [10] Neeta Nain, Deepak B M, Dinesh Kumar, Manisha Baswal, Biju Gautham , *Optimized Minutiae-Based Fingerprint Matching*, Proceedings of world congress on engineering, vol-I, WCE-2008.
- [11] Heeseung Choi, Kyoungtaek Choi, Jaihie Kim, *Fingerprint Matching Incorporating Ridge Features with Minutiae*, December 28, 2010.
- [12] Ashwini R. Patil, Mukesh A. Zaveri, *A Novel Approach for Fingerprint Matching using Minutiae*, pp.317-322, Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, 2010.
- [13] Xiaguang He, Jie Tian, Liang Li, Yuliang He, Xiayang, *Modeling and Analysis of Local Comprehensive Minutia Relation for Fingerprint Matching*, IEEE transactions on System, Man and Cybernetics, vol 37, no.4, August 2007.
- [14] Gwo-Cheng Chao, Shyh-Kang Jeng, Shung-Shing Lee, *A fingerprint matching algorithm based on alignment using LPD and GCD minutia descriptors*, 978-1-4244-1597-7/07, 2007 IEEE.
- [15] Ravi J, K B Raja, Venugopal K R, *Fingerprint Recognition using Minutiae Score Matching*, International Journal of Engineering Science and Technology, vol. 1(2), 2009,35-42.
- [16] Haiyun Xu, Raymond N. J. Veldhuis, Tom A. M. Kevenaer, and Ton A. H. M. Akkermans, *A Fast Minutiae-Based Fingerprint Recognition System*, IEEE SYSTEMS JOURNAL, VOL. 3, NO. 4, DECEMBER 2009
- [17] Khurram Yasin Qureshi, Shoab A. Khan, *Effectiveness of a Novel Feature and Confidence Levels Assignment to Classifiers in Fingerprint Matching*, 978-1-4244-4520-2/09, 2009 IEEE
- [18] Chih-Jen Lee, Tai-Ning Yang, Chun-Jung Chen, Chun-Lin Lo, and Li-Wei Chiang, *Apply Principal Gabor Basis Functions to Extract Discontinuous Points from Low-Resolution Binarized Fingerprint Images*, International Symposium on Computer, Communication, Control and Automation, 2010
- [19] Marius Tico, Eero Immonen, Pauli Ramo, Pauli Kuosmanen, and Jukka Saarinen, *Fingerprint Recognition Using Wavelet Features*, pp.21-24, Vol-II,0-7803-6685-9/01, 2001 IEEE.
- [20] Ke Hunng clid Selbz Aviyerire, *Choosing Best Basis in Wavelet Packets for Fingerprint Matching*, International Conference on Image Processing (ICIP), 2004.
- [21] Yik-Hing Fung and Yuk-Hee Chan, *Fingerprint Recognition with Improved Wavelet*, Proceedings of International Symposium on Intelligent Multimedia. Video and Speech Processing, October 2004 .

- [22] Woo Kyu Lee, *Fingerprint Recognition Algorithm Development Using Directional Information In Wavelet Transform Domain*, pp.1201-1204, IEEE International Symposium on Circuits and Systems, June 1997.
- [23] I. A. Ismail, M. A. Ramadan, T. El danf, A. H. Samak, *Multiresolution Fourier- Wavelet Descriptor for Fingerprint Recognition*, pp.951-955, International Conference on Computer Science and Information Technology , 2008.
- [24] XU Cheng, Cheng Xin-Ming, *An Algorithm for Fingerprint Identification Based on Wavelet Transform and Gabor Feature*, pp.827-830, Third International Conference on Genetic Algorithms and Evolutionary Computing, 2009.
- [25] Umair Mateen Khan, Shoab Ahmed Khan, Naveed Ejaz, , Riaz ur Rehman, *A Fingerprint Verification System using Minutiae and Wavelet based Features*, pp. 291-296, International Conference on Emerging Technologies, 2009.
- [26] Linlin Shen and Alex Kot, *A New Wavelet Domain Feature for Fingerprint Recognition*, Vol.14,pp.55-60, Biomedical soft Computing and Human Sciences, 2009.
- [27] Priti S. Sanjekar, Prof. Priyadarshan S. Dhabe, *Fingerprint Verification Using Haar Wavelet*, Vol.3, pp.361-365, IEEE International Conference on Computer Engineering and Technology, 2010.
- [28] Bhushan D. Patil, Jayant V. Kulkarni, Raghunath S. Holambe, *Fingerprint verification using wavelet and local dominant Orientation*.
- [29] M. P. Dale, M. A. Joshi, *Fingerprint Matching Using Transform Features*.
- [30] Anil Jain, Arun Ross, Salil Prabhakar, *Fingerprint Matching Using Minutiae And Texture Features*, pp.282-285, IEEE, 2001.
- [31] Yang Xu , Xuedong Zhang, *Gabor Filterbank and Its Application in the Fingerprint Texture Analysis*, Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'05).
- [32] Sharat Chikkerur, Sharath Pankanti, Alan Je, Nalini Ratha and Ruud Bolle, *Fingerprint Representation Using Localized Texture Features*, The 18th International Conference on Pattern Recognition (ICPR'06).
- [33] Shankar Bhausaheb Nikam, Pulkit Goel, Rudrajit Tapadar, Suneeta Agarwal, *Combining Gabor Local Texture Pattern and Wavelet Global Features for Fingerprint Matching*, International Conference on Computational Intelligence and Multimedia Applications, pp.412-417, 2007.
- [34] Zahoor Ahmad Jhat Ajaz Hussain Mir Simeen Rubab, *Fingerprint Texture Feature for Discrimination and Personal Verification*, 2009 Third International Conference on Emerging Security Information Systems and Technologies. pp.230-236, 2009.
- [35] Mohammed S. Khalil, Dzulkifli Muhammad, Qais AL-Nuzaili, *Fingerprint Verification using the Texture of Fingerprint Image*, pp.230-235, Second International Conference on Machine Vision, 2009.

Computational studies of swirl effects on instabilities and pollutions due to non-premixed turbulent combustion

Luthenda Gamany¹, Taha Janan Mourad² and Agouzoul Mohamed¹

¹ Equipe de Recherche et Développement : Modélisation et Multimédia Mécanique (ERD3M)

UFR : Modélisation et Calcul Informatique en Conception Mécanique (MCICM)

Ecole Mohammedia d'Ingénieurs – EMI, University Mohammed V Agdal (UM5A)

Postal address: Avenue Ibn Sina, B.P 765, Agdal, Rabat 10000, Morocco

² Laboratoire de Mécanique Procédés et Processus Industriels (LM2PI)

Ecole Normale Supérieure de l'Enseignement Technique – ENSET,

Postal address: Avenue de l'Armée Royale, Madinat Al Irfane 10100, B.P. 6207 Rabat-Instituts, Rabat 10000, Morocco

Abstract

Considerable effort is currently being extended by means of open CFD analysis to examine and fight against mechanisms responsible of combustion instabilities and environment pollution due to CO₂ and NO productions. To achieve that, the present paper suggests a system based on injection of a secondary air swirling flow in a non-premixed turbulent combustion chamber fed by fuel oil n°2. Computational studies are based on analysis of swirl intensity impact using OpenFOAM's solver named reactingFoam to compare the recommended system to a basic combustor of drying furnace. Data allowing discussions are temperatures and concentrations of unburned species and products of gas combustion calculated at transversal and longitudinal sections of the combustion chamber. The fact that results obtained reveal no risk of flashback or blowing phenomena, fast diminution of unburned products, significant thermal losses near walls, reduction of CO₂ production combined to a rise of NO formation pushes us to investigate more about the proposed apparatus.

Keywords: Coflow Non-premixed turbulent combustion OpenFOAM reactingFoam Swirling flow Swirl number

1. Introduction

The need for increased fuel consumption efficiency and environmental protection regulations are imposing stricter requirements on high intensity industrial combustion systems. In particular, the need to limit energy consumption costs, the restrictions on NO and soot emissions and the competitive placement in the market have encouraged the introduction of innovative solutions for better combustion control and the extrapolation of new concepts to practical designs. In non-premixed turbulent combustion area several efforts are made to accord experimental and numerical studies but still are insufficient compared to requirements becoming more and stricter. Gupta A. K. et al. [9] will be among the first to reveal (by means of experimental results) the important effects of swirl on promoting flame stability, increasing combustion efficiency and controlling emission of pollutants from combustion. Even if CFD (Computational Fluids Dynamics) analysis will later confirm this thesis through advanced modeling and simulation, the prediction of flames stabilized by swirling jet continues to be a research area under exploration. Due to the importance of this type of fluid flow, considerable investigations have been performed in this regard. Many projects were conducted to increase interaction between experimentalists and numerical analysts in order to reach good agreements between theory and realistic phenomena. Among these projects two references can be typically cited: CORIA (COmplexe de Recherche Interprofessionnel en Aérothermochimie) [5, 25] and TNF (the International workshop on measurement and computation of Turbulent Non-premixed Flames) [4]. To boost TNF researches, experimental simulations devices like PRECCINSTA [1] and TECFLAM [1, 11] swirl burners were developed but two limitations can be identified.

The first one concerns the nature of fuels used by PRECCINSTA and TECFLAM. These are more suitable for gas fuels (methane, natural gas, etc.) with direct applications in gas turbine combustors or in aircraft engines [18]. In our case, the fuel is a preheated liquid (fuel oil n°2) used in industrial drying furnaces [14]. According to CFD literature, it's less complicated to predict flames based on gaseous hydrocarbons than heavy oils [5, 16]. This difficulty comes a part from strong influence of high viscosities on turbulence and combustion phenomena. And this is clearly seen during modeling and simulation of instabilities and mixtures.

Another limitation is about PRECCINSTA and TECFLAM basic configurations. In most cases swirl jet is involved in primary air flows [1, 13] while this one is located in secondary air flows (Fig.4). In these kinds of situations the size of confinement plays an important role in the aerodynamics of flame [5].

In addition, the problem here is complicated by 3D assumption according to the complexity of systems and phenomena.

Despite all this, giving predictions that can satisfy industrial requirements (performance and security of systems), standards of environment protection (fight against pollution) and computational efficiency (need of precision, speed and robustness) still be a topical challenge in CFD.

2. Description of the problem

2.1. Geometry

In this study, the combustion chamber geometry is a typical case of those used in industrial drying furnaces (Fig. 1) [14]. To obtain a non-premixed turbulent flame, the fuel oil n°2 is injected through the internal tube of the burner after being preheated and sprayed. At the same time, primary air at ambient temperature is introduced through the annular space. It is assumed that the two considered flows are coaxial. Finally, secondary air is introduced through 24 circular orifices on the edges of the front of our firebox (Fig. 1). To obtain a swirling flow (Fig. 3), system of axial and tangential injection of air [5] is used by mean of exhaust fans [14].

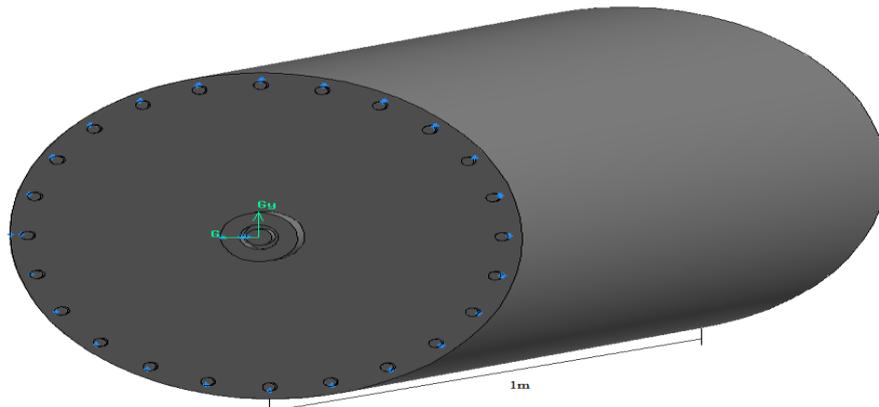


Fig. 1: Combustion chamber

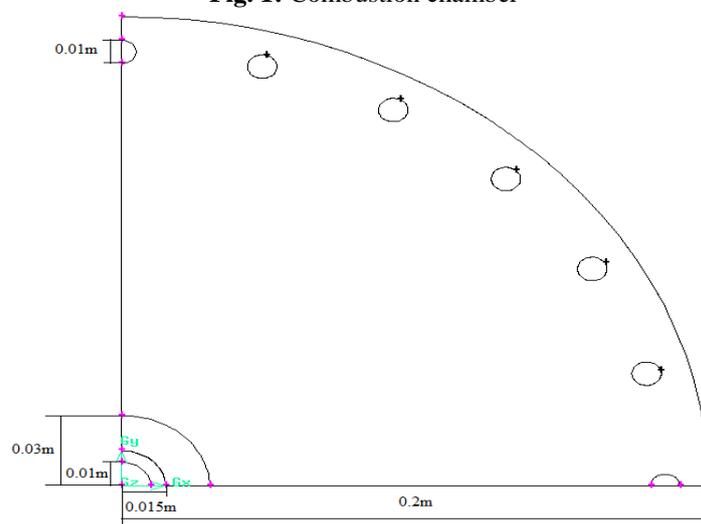


Fig. 2: Face measures

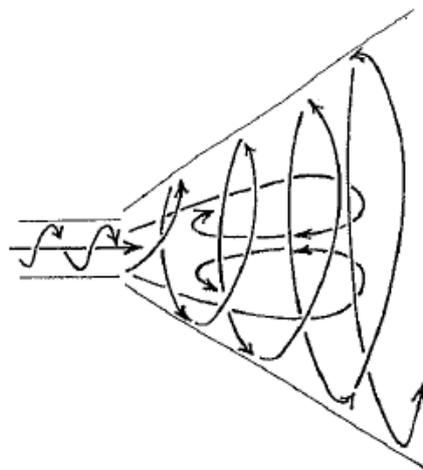


Fig. 3: Swirl Flow with Internal Zone Recirculation

2.2. Boundary conditions

Experimental data obtained from measurements [14] are used in order to approach realistic situation. Entrance conditions are related to the flow rate and the temperature given from the burner exit. The inlet velocities of fuel and primary air are respectively 2.34m/s and 15m/s while their respective temperatures are 120°C and 17°C. The outlet conditions are related to the pressure. The temperature of the nozzle was fixed to 980°C which is the limit of security imposed. According to the secondary air, the inlet velocity is a rotating profile field whose variation depends on the ratio of flow rates injection. About the walls, they are supposed to be adiabatic.

3. Mathematical formulation

3.1. Swirling Flow

Commonly, it is possible to obtain a swirling flow either by using an insert inside the nozzle or by mixing two air flows. The last study on a swirling impinging jet is due to Ward and Mahmood [20]. The swirling jet designed by Ward and Mahmood, which is based on the concept of mixing an axial air flow with a tangential one, is characterized by a radial distribution of the local convective heat transfer, which, respect to the circular jet, is slightly more uniform with a significantly lower heat transfer rate.

3.1.1. Swirl number (S) [9]

The standard method used to stabilize flames in current lean premixed combustors is based on swirling injection. Combustion is established around a hot gas kernel formed by the swirling flow. The lower pressure region created in the central region by swirl generates an internal recirculation zone of burnt gases anchoring the flame. This region designated as the inner recirculation zone (IRZ) is associated with the vortex breakdown process which takes place in swirling flows [3, 4]. The intensity of a swirling flow is given by a dimensionless number S called swirl number. This quantity establishes the relationship between the axial flux of kinetic momentum and the axial flux of axial momentum.

$$S = G_{\theta} / RG_x \quad (1)$$

G_{θ} can also be seen as the flux of tangential momentum and G_x the flux of momentum following the direction of propagation. After developing the previous expression, S becomes:

$$S = \frac{\int_0^R \rho(\overline{uw} + \overline{u'w'})r^2 dr}{R \int_0^R \rho \left(u^2 + \frac{(\overline{w}^2 - w_{\max}^2)}{2} \right) r dr} \quad (2)$$

Where R is the radius of the injector tube and $\vec{V}(u, v, w)$ the velocity decomposed into mean and fluctuating values: $u = \overline{u} + u'$, $v = \overline{v} + v'$, $w = \overline{w} + w'$

3.1.2. Equation of conservation of momentum for a swirling flow [8]

As it has been already mentioned on chap. 2. 2, the entrance of secondary air is characterized by a rotating velocity field. The absence of circumferential gradients in the swirling flow led us to consider the assumption of an axisymmetric configuration in order to establish the initial solution. We have now a 2D model with prediction of the circumferential velocity. Its equation of conservation of tangential momentum is:

$$\frac{\partial}{\partial t}(\rho w) + \frac{1}{r} \frac{\partial}{\partial x}(r \rho u w) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v w) = \frac{1}{r} \frac{\partial}{\partial x} \left[r \mu \frac{\partial w}{\partial x} \right] + \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^3 \mu \frac{\partial}{\partial r} \left(\frac{\partial w}{\partial r} \right) \right] - \rho \frac{vw}{r} \quad (3)$$

3.2. Equations of the combustion

This section presents governing equations of combustion flow in the furnace. Chemical reactions that can be modeled include liquid fuel combustion in which fuel vapor is generated via evaporation of liquid droplets and a combustion reaction occurs in the gas phase.

3.2.1. Governing equations in conservative form [1]

Equations of chemical species reacting flow in conservative form can be written as follows:

$$\frac{\partial \mathbf{w}}{\partial t} + \nabla \cdot \mathbf{F} = \mathbf{s} \quad (4)$$

Where $\mathbf{w} = (\rho u, \rho v, \rho w, \rho E, \rho_k)^T$ with ρ the density, $\vec{V} = (u, v, w)^T$ the vector velocity, E the total energy and

$\rho_k = \rho Y_K$ with Y_K the mass fraction of K species. Flux tensor can be decomposed in two parts:

$$\mathbf{F} = \mathbf{F}(\mathbf{w})^I + \mathbf{F}(\mathbf{w}, \nabla \mathbf{w})^V \quad (5)$$

Where $\mathbf{F}(\mathbf{w})^I$ is non-viscous tensor and has 3 components:

$$f^l = \begin{pmatrix} \rho u^2 + P \\ \rho uv \\ \rho uw \\ (\rho E + P)u \\ \rho_K u \end{pmatrix}, \quad g^l = \begin{pmatrix} \rho uv \\ \rho u^2 + P \\ \rho uw \\ (\rho E + P)v \\ \rho_K v \end{pmatrix}, \quad h^l = \begin{pmatrix} \rho uw \\ \rho vw \\ \rho w^2 + P \\ (\rho E + P)w \\ \rho_K w \end{pmatrix} \quad (6)$$

Where P is the hydrostatic pressure defined in perfect gas state equation.

$F(w, \nabla w)^V$ is viscous tensor and has 3 components:

$$f^V = \begin{pmatrix} -\tau_{xx} \\ -\tau_{xy} \\ -\tau_{xz} \\ -(u\tau_{xx} + v\tau_{xy} + w\tau_{xz}) + q_x \\ J_{x,K} \end{pmatrix}, \quad g^V = \begin{pmatrix} -\tau_{xy} \\ -\tau_{yy} \\ -\tau_{yz} \\ -(u\tau_{xy} + v\tau_{yy} + w\tau_{yz}) + q_y \\ J_{y,K} \end{pmatrix}, \quad h^V = \begin{pmatrix} -\tau_{xz} \\ -\tau_{yz} \\ -\tau_{zz} \\ -(u\tau_{xz} + v\tau_{yz} + w\tau_{zz}) + q_z \\ J_{z,K} \end{pmatrix} \quad (7)$$

Stress tensor is defined as follows:

$$\tau_{ij} = 2\mu \left(S_{ij} - \frac{1}{3} \delta_{ij} S_{ll} \right), \quad i, j = 1, 3 \quad (8)$$

$$S_{ij} = \frac{1}{2} \left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j} \right), \quad i, j = 1, 3 \quad (9)$$

3.2.2. Equation of state of perfect gas

The mixture formed in the furnace is supposed to be mixture of perfect gas:

$$P = \rho \frac{R}{W} T \quad (10)$$

\overline{W} is the mixture molecular weight:

$$\frac{1}{\overline{W}} = \sum_{K=1}^N \frac{Y_K}{W_K} \quad (11)$$

R = 8.3143J/mol.K is the universal constant of perfect gas.

3.2.3. Equation of continuity

Conservation of total mass in multi species flow is satisfied by the following equations:

$$\sum_{K=1}^N Y_K V_i^K = 0 \quad (12)$$

V_i^K represents the velocity component of diffusion of the species k in the i direction (i=1,2,3). The approximation of Hirschfelder-Curtis is used to express this velocity of diffusion:

$$X_K V_i^K = -D_K \frac{\partial X_K}{\partial x_i} \quad (13)$$

$$Y_K V_i^K = -D_K \frac{\overline{W}_K}{W} \frac{\partial X_K}{\partial x_i} \quad (14)$$

Where X_K are gradients of mole fractions.

3.2.4. Heat flux

The total heat flux q_i is the sum of 2 terms: the heat flux by conduction and the heat flux by diffusion of species.

$$q_i = -\lambda \frac{\partial T}{\partial x_i} + \sum_{K=1}^N J_{i,k} h_{S,K} \quad (15)$$

$$q_i = - \underbrace{\lambda \frac{\partial T}{\partial x_i}}_{\text{conduction}} - \underbrace{\rho \sum_{K=1}^N \left(D_K \frac{W_K}{W} \frac{\partial X_K}{\partial x_i} - Y_K \sum_{K=1}^N D_K \frac{W_K}{W} \frac{\partial X_K}{\partial x_i} \right)}_{\text{Diffusion of species}} h_{s,K} \quad (16)$$

4. Implementation in OpenFOAM

The choice of implementation in OpenFOAM as CFD code compared to the commercial counterparts, e.g. Fluent and Ansys CFX etc. is justified by the fact that it offers a free advanced toolbox for solving complex physical problems involving chemical reactions, turbulence and heat transfer with the advantage to be totally open and free, both in terms of source code and in its structure and hierarchical design [10].

4.1. Description of OpenFOAM (Open Field Operation and Manipulation) [12]:

Native development of OpenFOAM is done on a Linux / UNIX platform and specifically in the GCC (Gnu Compiler Collection) C++ compiler. To pre-process and solve cases, two ways are offered: use of the graphical user interface called FoamX or the Terminal. As direct access to OpenFOAM's source code is possible (text files in C), modifications of boundary conditions or input/output control are allowed by editing the files manually. The post-processing is done by ParaView.

4.2. Mesh (Fig. 5):

Although OpenFOAM has a mesh tool called BlockMesh exportation of other meshes is feasible. It is important to note that OpenFOAM is strictly a 3D code. To find an initial solution for a 2D axi-symmetric problem, the mesh must be 3D first with one cell thick and having no solution in Z- direction. We must ensure that the opening of the thickness of both sides between the Y-axis does not exceed the angle whose apex would be on the Y-axis and between 2.5° and 5° [2].

Quadrangular mesh is used due to the simple geometry of the 2D initial solution configuration while tetrahedron/hybrid mesh type is used due to the complex shape of the 3D geometry [19].

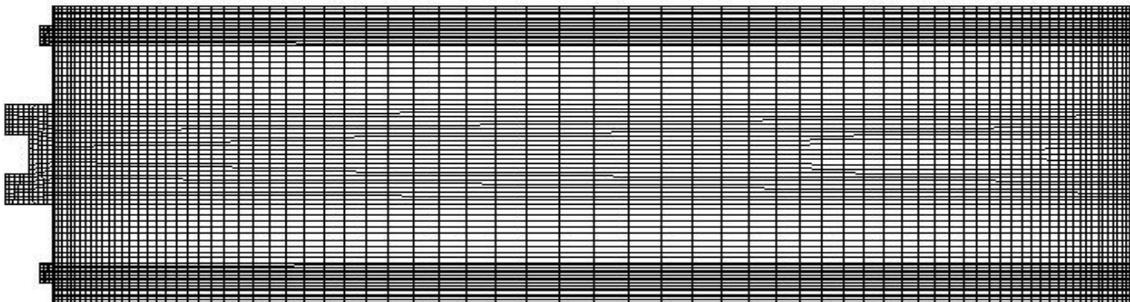


Fig. 4: 2D initial solution mesh view

5. Turbulence and combustion modeling

5.1. Turbulence [8]

The DNS (Direct Numerical Simulation) is effective when simple academic problems occur otherwise it would be numerically expensive as it can be seen in our case because the phenomena and configurations are complex. The advantage of LES (Large Eddy Simulation) is to combine modeling and direct calculation but the computation of its pre-processor and solver costs a lot. Another problem related to LES is that their coupling with other models is not yet mature in OpenFOAM. Taking in account all those points RANS (Reynolds Averaged Navier-Stokes) remains the only one approach more accessible to use in our work. The model chosen is k-ε standard because of its robustness, accuracy, low cost and rich documentation. Its weakness in walls zone can be offset by the Standard Model Wall Function (SWF).

5.2. Combustion [2]

For combustion gas, there are three types of solvers all operating in unsteady state: reactingFoam [10], Xoodles and XiFoam. Only reactingFoam models the non-premixed turbulent combustion. The concept it uses for the chemical species is the PaSR (Partially Stirred Reactor) which is a modified version of the EDC (Eddy Dissipation Concept) where the chemical time scale is handled differently.

- Expression of Air-Fuel reaction mechanism:



- Expression of the reaction rate constant according to Arrhenius kinetic model:

$$k(T) = AT^b \exp\left(-\frac{E_A}{R_u T}\right) \quad (18)$$

Where EA is the activation energy (1.256x10⁸J/kg mol) and A is the pre-exponential factor (2.587x10⁹).

The equation (14) was originally made for laminar case. If we want to adapt it to turbulent case, just define the mixture rate by the ratio:

$$turb_{mix} = \varepsilon / k \quad (19)$$

NB: the reaction mechanism is imported from Chemkin (software tool for solving complex chemical kinetics problems).

- ReactingFoam Code:
 1. Calculate chemical reaction based on turbulent and chemical timescales;
 2. Calculate of the density;
 3. Calculate velocity/pressure fields;
 4. Read species and feed them to the chemistry solver using Chemkin table;
 5. Calculate the temperature from the chemical reactions enthalpy lookup;
 6. Calculate the pressure field using PISO;
 7. Correct the turbulence (pressure-corrector);
 8. Update the density from the temperature;
 9. Return to step 1.

6. Numerical resolution [10]

Discretisation method used by OpenFOAM is Finite Volume (FVM). Only transient models like reactingFoam need temporal discretisation.

6.1. Temporal discretisation

To avoid instability due to the simultaneous calculation of turbulence and combustion, the first solver to use is simpleFoam which is a steady-state incompressible turbulence model. This is done to calculate the cold flow which will be considered as our initial solution. Knowing that reactingFoam is an unsteady solver, the time step depends strongly on the Courant number Cr . This is possible only with the CFL (Courant Friedrichs Lewy) condition:

$$C_r = \frac{\bar{v}\Delta t}{\delta x} \quad (20)$$

If $0 < C_r \leq 0.5$, we have more stability and less speed.

If $1 > C_r \geq 0.5$, we have more speed and less stability.

In this case, $Cr = 0.2$.

6.2. Interpolation schemes:

- Pressure: limitedLinear 1 (second order bounded scheme);
- Velocity: limitedLinearV (TVD scheme recommended for swirl);
- Turbulence: upwind (first order bounded scheme);
- Species: upwind;
- Energy: upwind;
- Pressure-velocity coupling: PISO (Pressure Implicit with Splitting of Operators).

7. Results

7.1. Advantages of swirling secondary air configuration:

7.1.1. Tight confinement

The swirling flow affects the behavior of the flame because we are dealing with a tight confinement, in other words, the dimensions (ie, diameter) of the injector are not negligible compared to the chamber dimensions (Fig. 1 & 2) [13].

7.1.2. Decrease of unburned rate

From thermodynamic studies it has been shown that a small amount of trace species in the combustion products can have a great impact on the CO₂ capture, storage and transportation [18]. Now the mass fraction of fuel (mF_f) at the outlet has almost decreased by an half as can confirm the values and graph below (Fig. 5):

- mF_f (without swirl)= ± 0.1057 ;
- mF_f (with coflow: $S=0$) = ± 0.0478 ;
- mF_f (with swirl: $S=0.3$)= ± 0.0478 .

This will enable a better control of pollution due to CO₂ production.

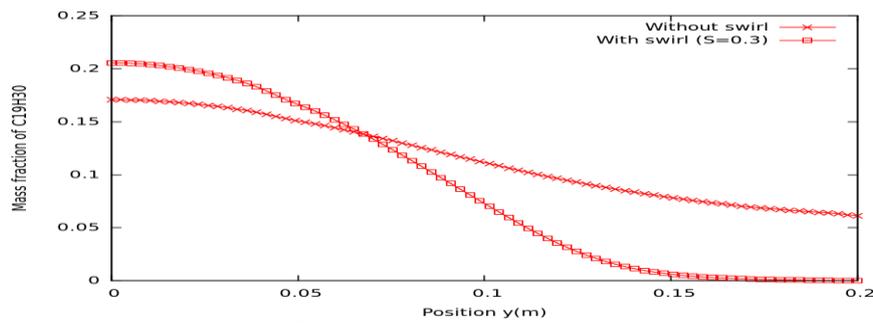


Fig. 5: Outlet mass fraction of $C_{19}H_{30}$

7.1.3. Better protection of walls

Temperature decreases considerably near the walls (Fig. 6). This fall ΔT is inversely proportional to S . This is verified through values of transversal temperature differences observed when $S=0$ or $S=0.3$ (Fig. 7 & 8).

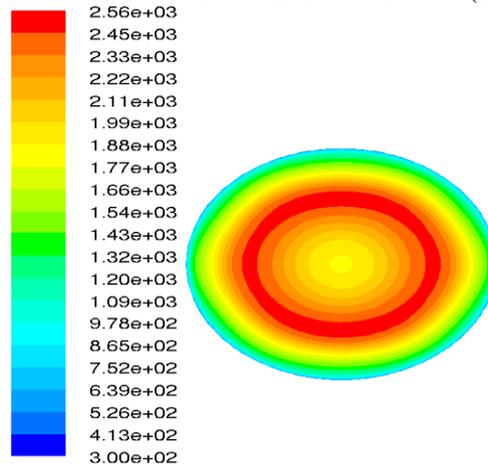


Fig. 6: Radial evolution of temperature contours

- For internal transversal temperature: if $S = 0.3$: $\Delta T_{swirl} = T - T_{swirl} = 1150^\circ C$ (Fig. 7)

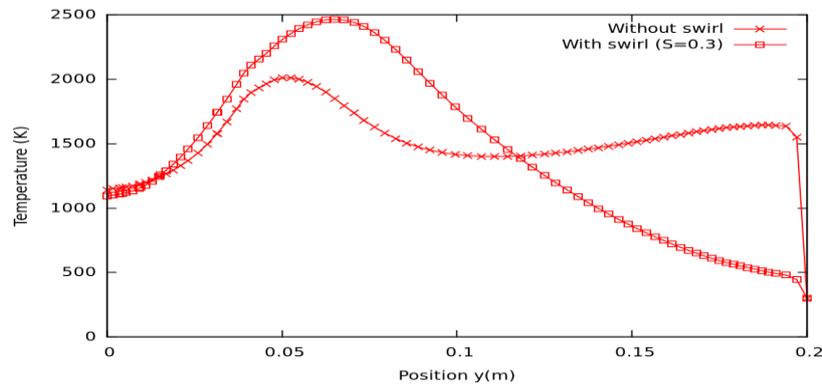


Fig. 7: Internal transversal temperature ($x=0, z=0.52$)

- For outlet transversal temperature: if $S = 0.3$: $\Delta T_{swirl} = T - T_{swirl} = 650^\circ C$ (Fig. 8)

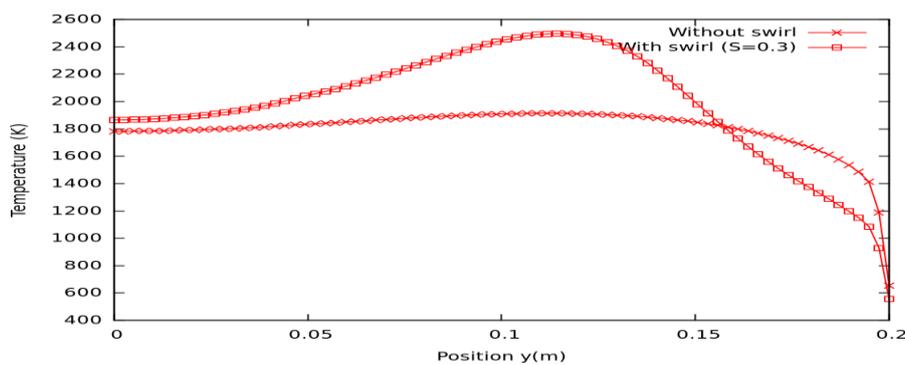


Fig. 8: Outlet transversal temperature

7.1.4. Stabilization of combustion:

In practical combustion chambers the mixing time for the fuel and oxidizer is typically larger than the chemical reaction time [18].

Stabilization of combustion is done by increasing the residence time of the flame and creating recirculation in the reaction zone. The effect produced is to favor the mixture and give to the flame a more compact form [13] (Fig. 9, 10, 11, 12 & 13). The axial evolution of swirling temperature profile is almost linear (Fig. 11). It means that swirl configuration provides better and easiest control of temperature in flow direction.

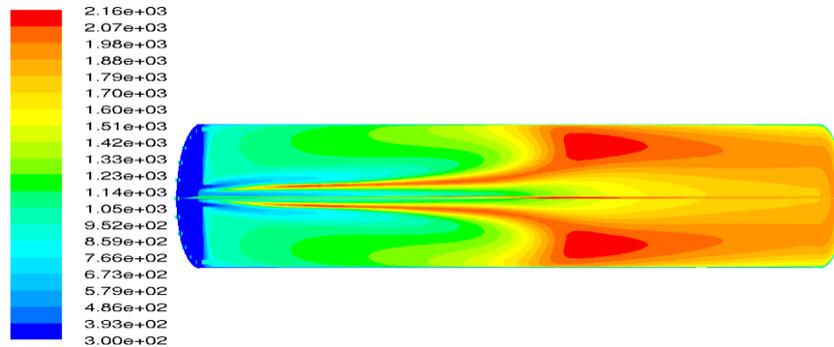


Fig. 9: Temperature contours without swirl

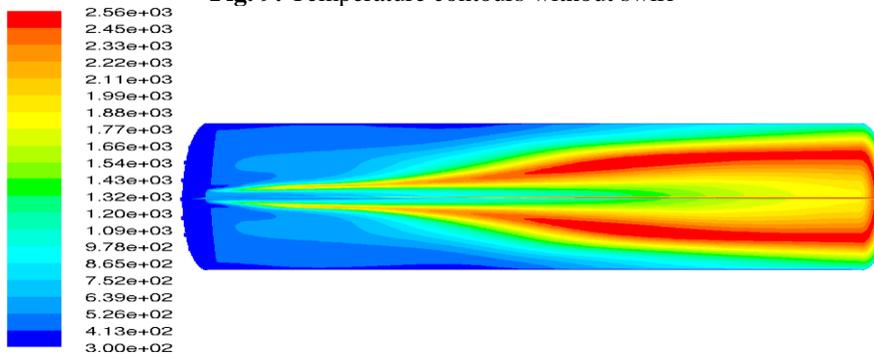


Fig. 10: Temperature contours with swirl

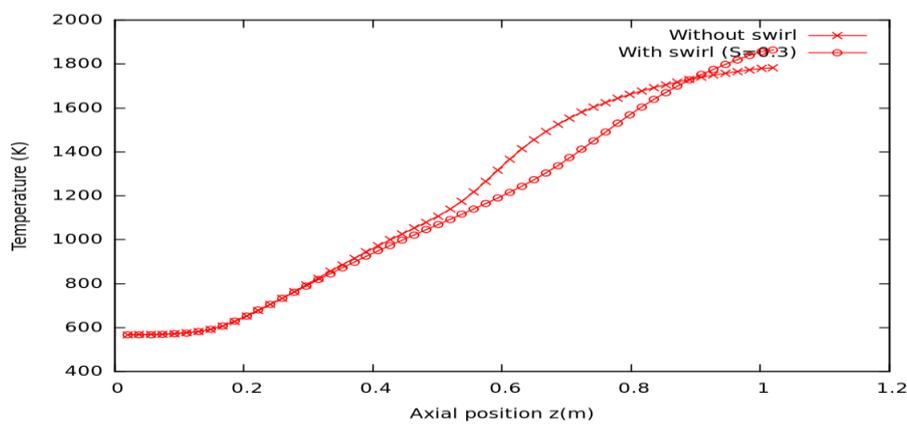


Fig. 11: Axial temperature evolution

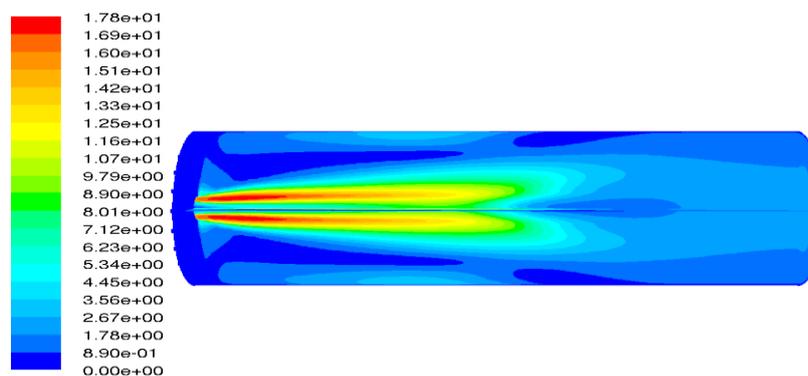


Fig. 12: Velocity contours without swirl

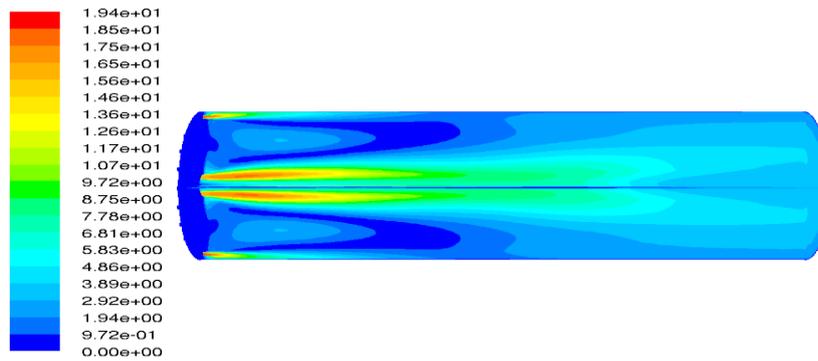


Fig. 13: Velocity contours with swirl

7.1.5. Reduction of the CO₂ formation:

Values and graphs of mF(CO₂) below (Fig. 14 & 15) reveal a decrease of the outlet CO₂ mass fraction in the case of intense swirl:

- Without swirl, mF(CO₂) = ± 0.1893;
- With coflow (S=0), mF(CO₂)_{coflow} = ± 0.1677;
- With swirl (S=0.6), mF(CO₂)_{swirl} = ± 0.1637.

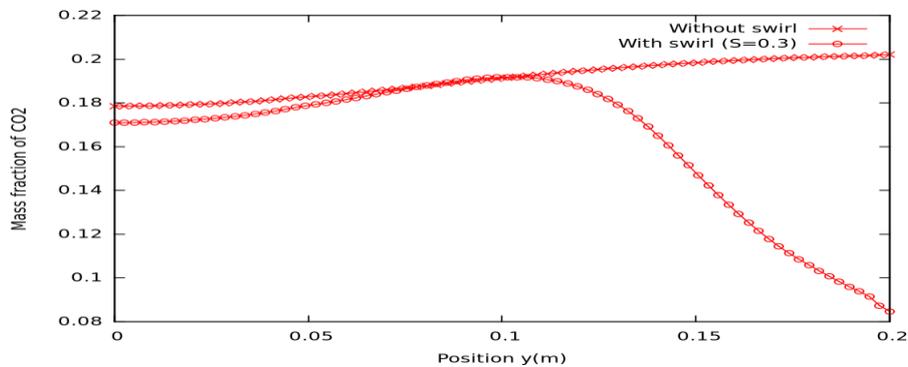


Fig. 14: Outlet mass fraction of CO₂

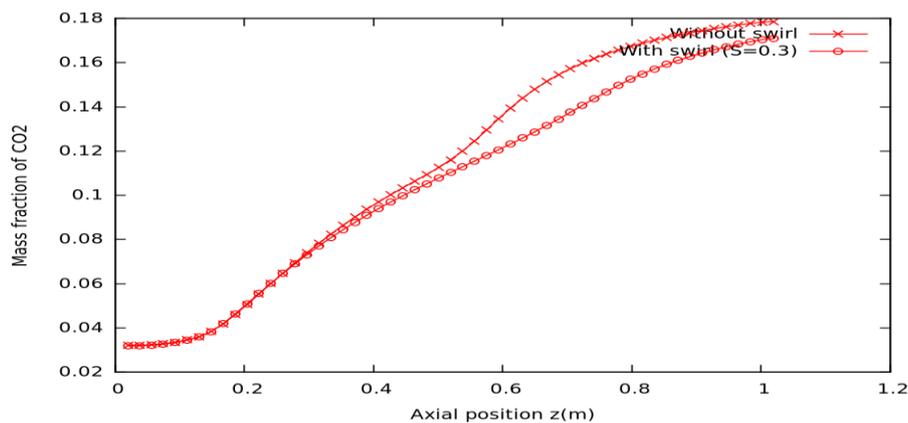


Fig. 15: Axial mass fraction of CO₂

7.2. Disadvantages of swirling secondary air configuration:

7.2.1. Rise of outlet temperature

Calculation of the mass-weighted average value of temperature at outlet gives:

$$\bar{T} = \frac{\int T\rho\vec{V}d\vec{A}}{\int \rho\vec{V}d\vec{A}} \tag{21}$$

- Without swirl, $\bar{T} = \pm 1827\text{K}$;
- With coflow (S=0), $\bar{T}_{coflow} = \pm 1934\text{K}$;
- With swirl (S=0.6), $\bar{T}_{swirl} = \pm 1879\text{K}$.

7.2.2. Amplification of the swirl promotes NO formation

The same observation was made on a small scale by Vauchelles [16]. After 3 test cases ($S=0.3$; $S=0.5$; $S=0.7$) of swirling secondary air injection to a combustion chamber of gas turbine engine, he asserts that the reduction of CO_2 emissions means better combustion and thus a rise of temperature controlling NO.

To reach this conclusion, the following two approaches can be used:

- Calculation of NO concentration (ppm) in the gas mixture:
NO ppm is computed from the following equation:

$$NO_{ppm} = \frac{MF(NO) \times 10^6}{1 - MF(H_2O)} \tag{22}$$

Where:

$$MF(NO) = \frac{mF(NO) \times MW(mixture)}{30} \tag{23}$$

After calculation, it can be seen through the values and graphs of $MW(NO)$ below (Fig. 16 & 17) that the swirl promotes the production of NO.

- Without swirl, $MW(NO) \approx 1019$ ppm;
- With coflow ($S=0$), $MW(NO)_{coflow} \approx 2933$ ppm;
- With swirl ($S=0.6$), $MW(NO)_{swirl} \approx 3152$ ppm.

- Estimation of average mass fraction of outgoing NO:

The mass fraction of outgoing NO experienced a sharp increase with the intensification of the swirl.

- Without swirl, $mF(NO) \approx 0.0011$;
- With coflow ($S=0$), $mF(NO) \approx 0.0092$;
- With swirl ($S=0.6$), $mF(NO) \approx 0.0091$.

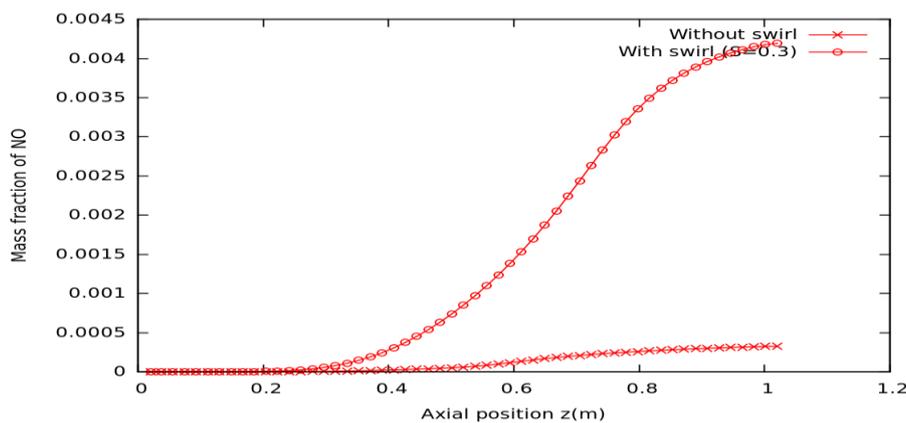


Fig. 16: Axial mass fraction of NO

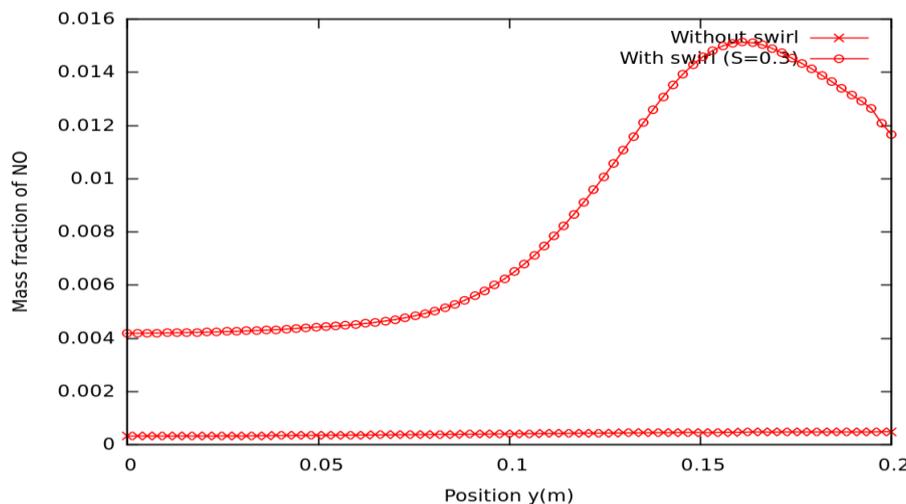


Fig. 17: Outlet mass fraction of NO

8. Conclusions and perspective

To better understand the potential role of a secondary air swirling flow in a turbulent non-premixed combustion chamber two test cases situations are considered : a pure axial jet called "coflow" (with $S=0$) and an weak swirl jet (with $S=0.3$) characterized by the presence of an IRZ (Internal Recirculation Zone). Non-premixed flames can be quenched by lowering the temperatures of the oxidizer stream, or by increasing the flow strain rate (equivalently scalar dissipation rate) [18]. In the first case there is a possible risk of blowing phenomenon for the secondary air flow strain rate is increased. As it is known the blow tends to get rid of the IRZ pushing the flame to hanging or extinction. In the last case [15] a flashback to the injector's tube can occur and lead to a suspension of the flame. Flashback occurs when the gas velocity becomes lower than the burning velocity due to flame propagation within boundary layer, core flow or because of combustion instabilities [22]. Flashback in swirl burners [23] can be caused by a phenomena termed combustion induced vortex breakdown (CIVB) due to rapid expansion at the burner exit creating a recirculation zone which acts as a flame holder: the breakdown of this structure can occur due to flow perturbations and chemical reaction effects causing the CRZ (Central Recirculation Zone) and hence flame to propagate upstream into the premixing zone [23, 24].

Neither of these two phenomena was observed because of the good tight confinement between combustion chamber and swirling air injectors. This led us to introduce confidently discussion about the flame stability based on the influence of number of Swirl (S). Studies will be extended in the next paper to deal with greater values of S .

About the combustion intensity, we noticed simultaneously an increase of NO formation and a decreased of CO₂ production. We know that the two gases are the leading pollutants in combustion sites [7]. This is justified by the fact that the main reaction mechanism involved for NO is that of Zeldovich (1946) which is essentially thermal as it is generated by the rise of the activation temperature in a reaction zone such as an oxidizing environment [15]. This observation is quite normal for in non-premixed condition, fuel is injected in shear region formed near to the zero stream line boundary and recirculation region which provides the low velocity region for flame stabilization with the evolution of high temperatures from the flame. For flames operating in diffusion mode, the reaction zone is stabilized to result in large temperature gradients and hot-spot regions in the entire combustion chamber that result in high NO levels [9] from the combustion of fuels. NO emission increased due to the accelerated chemical kinetics of the combustion process [19]. Concerning CO₂ production, it is possible to reduce more by increasing the inlet air temperature but this will result in increasing the flame temperature which in its turn will increase the production of NO [19]. Another solution to explore about CO₂ production can be the elevation of pressure into the combustor. As said Ahmed E.E. Khalil and al. [19] this will promote the combustion kinetics to enhance the combustion reactions. At low equivalence ratios, increase in pressure diminishes CO by accelerating the rate of conversion of CO into CO₂. At high equivalence ratios, increase in combustion pressure reduces CO emissions, albeit to a lesser extent, by suppressing chemical dissociation. So in one aspect high pressures are beneficial, but on the other hand, high pressure also accelerated NO_x formation leading to higher NO emissions [19].

The future work will be to control temperature so as to expect a good balance between reducing pollution and improving combustion. Moreover, the protection of walls remains a priority. For complete study, it will be better to couple the influence of temperature with the humidity effects near walls by means of transversal evolution of H₂O mass fraction.

To assess the stability of reactingFoam, the role of convection schemes in the solution was investigated. That is the reason why we used successively upwind and limitedLinear to compute the energy and species. The result was the same but instability (represented by "wiggles") occurred in the temperature profile when we tried QUICK scheme. Versteeg and Malalasekera [17] have showed that this is due to the QUICK schemes nature, which has a tendency to produce over-shoots and under-shoots in the results and thus producing the obtained "wiggles".

References

1. Albouze, G., Simulation aux grandes échelles des écoulements réactifs non-prémélangés, Ph.D. thesis, Université de Toulouse, 2009.
2. C. Andersen, E. L. Nielsen Niels, Numerical investigation of a BFR using OpenFOAM, Project of Fluids and Combustion Engineering (2008) p.18.
3. Acta Physicochim, "Thermal NO Mechanism ou Zeldovich Mechanism", 1946, 21, 577.
4. R. W. Bilger, International workshop on measurement and computation of Turbulent Non-premixed Flames: References on Swirl Flames, Updated November 2005.
5. Criner, K., Stabilisation de flammes de diffusion turbulentes assistée par plasma hors-équilibre et par champ électrique, Ph.D. thesis, Université de Rouen, 2008.
6. D. Elkaïm, M. Agouzoul and R. Camarero, "A Numerical Solution for Reacting and Non Reacting Flow", Lectures Notes in Physics, Edited by Dervieux A. and Larouturou B., Springer-Verlag, 1989.
7. Ferrand, L., Modélisation et expérimentation des fours de réchauffage sidérurgiques équipés de brûleurs régénératifs à oxydation sans Flamme, Ph.D. thesis, Ecole des Mines de Paris, 2003.
8. FLUENT Inc., Fluent 5.5 User's guide, October 2000.
9. Gupta, A. K. Lilley, D.G. Syred, N., Swirl Flows, Gordon & Breach Science Pub, 1984.
10. Lundström, A. reactingFoam Tutorial, Chalmers University of Technology, 2008.

11. W. Meier, O. Keck, B. Noll, O. Kunz, & W. Stricker, “Investigations in the TECFLAM Swirling Diffusion Flame: Laser Raman Measurements and CFD Calculations”, *Appl. Phys.* (2000) B71:725-731.
12. OpenCFD, OpenFOAM – User Guide, 1st August 2007, Version 1.4.1.
13. Poireault, B. Mécanisme de combustion dans un brûleur méthane-air de type swirl (40kW): influence de l'intensité de la rotation. Ph.D. thesis, Université de Poitiers, 1997.
14. T. Raffak, M. Agouzoul, M. Mabsate, A. Chik & A. Alouani, “Recent Patent and Modeling of Phosphate Rotary Dryer”, *Recent Patents on Engineering*, June, 2008, Volume 2, Issue 2, pp.132-141.
15. Taupin, B. Etude de la combustion turbulente à faible richesse haute température et haute pression, Ph.D. thesis, Institut national des sciences appliquées de Rouen (2003) pp. 24-27.
16. Vauchelles, D., Etude de la stabilité et des émissions polluantes des flammes turbulentes de prémélange pauvre à haute pression appliquées aux turbines à gaz, Ph.D. thesis, INSA-Rouen, 2004.
17. Versteeg, H. K. Malalasekera, W., *An introduction to computational fluid dynamics, The finite volume method*, second ed. Pearson Education Limited, 2007.
18. C.Y. Liu, G. Chen, N. Sipöcz, M. Assadi, X.S. Bai, Characteristics of oxy-fuel combustion in gas turbines, *Applied Energy* 89 (2012) 387–394.
19. Ahmed E.E. Khalil, Ashwani K. Gupta, Distributed swirl combustion for gas turbine application, *Applied Energy* 88 (2011) 4898–4907.
20. Harun Mohamed Ismail, Hoon Kiat Ng, Suyin Gan, Evaluation of non-premixed combustion and fuel spray models for in-cylinder diesel engine simulation, *Applied Energy* 90 (2012) 271–279.
21. J. Ward, M. Mahmood, Heat transfer from a turbulent, swirling impinging jet, *Proceedings of the 7th International Heat Transfer Conference* 3 (1982) 401-407.
22. Sankaran R, Hawkes ER, Chen JH, Lu T, Law CK. Direct numerical simulations of turbulent lean premixed combustion. *J Phys* 2006; 46:38–42.
23. Fritz J, Kroner M, Sattelmayer T. Flashback in a swirl burner with cylindrical premixing zone. *J Eng Gas Turb Power* 2004; 126(2):276–83.
24. Kroner M, Fritz J, Sattelmayer T. Flashback limits for combustion induced vortex breakdown in a swirl burner. *J Eng Gas Turb Power* 2003; 125(3):693–700.
25. CORIA, Actes de la Sixième Conférence Francophone en Recherche d'Information et Applications, Hyères, France, 2009.

Analysis of Skew Bridges Using Computational Methods

Vikash Khatri¹, P. R. Maiti², P. K. Singh² & Ansuman Kar²

Department of Civil Engineering, Institute of Technology, Banaras Hindu University, Varanasi

Abstract

In spite of increases in computing power, analysis of skew bridge deck has not changed to the same extent. Therefore, there is a need for more research to study the skew bridges using different computational methods. Grillage analyze is a fast and simpler approach compared to the finite element method, and has been used by engineers to analyses bridge deck over a long time. On the other hand the finite element method is thought to be better method for the slab analysis because of its capability to represent the complex geometry of the structure more realistically. In this present study, a bridge deck consists of beam and slab is defined and modeled using grillage and finite element method. The effect of grid spacing on different skew angles on same-span of reinforced concrete bridges using the finite-element method and grillage analogy method is compared. Maximum reactions force, deflection, bending and torsional moments is calculated and compared for both analysis methods. A total of nine different grid sizes (4 divisions to 12 divisions) have been studied on skew angles 30°, 45° and 60° to determine the most appropriate and efficient grid size. It is observed that finite element method (FEM) and Grillage method results are always not similar for every grid size. Bending moment calculated by using FEM overestimates the results obtained by grillage analysis for larger grid sizes. Torsion moment behavior shows reverse of bending moment and difference between reaction values of grid sizes between two methods decreases as skew angle increases. FEM gives lesser variations of bending and torsional moment with the change of grid sizes than Grillage one. Deflection doesn't vary much on the change of the grid sizes. The appropriate grid size is estimated for this narrow and long bridge is seven divisions whose ratio of transverse to longitudinal grid spacing is about 2.

Keywords: Skew slab, FEM, Grillage analysis, Grid size

1. Introduction

Generally, grillage analysis [14] is the most common method used in bridge analysis. In this method the deck is represented by an equivalent grillage of beams. The finer grillage mesh, provide more accurate results. It was found that the results obtained from grillage analysis compared with experiments and more rigorous methods are accurate enough for design purposes. The other method used in modeling the bridges is the finite element method. The finite element method is a well known tool for the solution of complicated structural engineering problems, as it is capable of accommodating many complexities in the solution. In this method, the actual continuum is replaced by an equivalent idealized structure composed of discrete elements, referred to as finite elements, connected together at a number of nodes.

In the skew bridges, the effects of skew on the response of completed structures have been well documented [1, 7, 9], with effects being shown to be more significant for skew angles greater than 30°. Critical values for vertical deflections and bending moments within in-service skewed bridges have been shown to be lower when compared against those in similar right bridges. Conversely, torsional rotations, shears and moments have been shown to be larger for skewed bridges. In addition, studies have also demonstrated that interaction between main support girders and transverse bracing members (diaphragms and cross frames) influences skewed bridge load distribution due to an increase in torsional rotations at certain sections of the longitudinal girders. Additional work has shown that the magnitude of torsional shear rotations at skewed bridge supports is largest at the obtuse corners [2, 3].

While a number of studies [4, 5] dedicated to the response of in-service skewed bridges have been completed, as presented above, there are few studies that focus on the behavior of skewed bridges during construction. There has been a lack of research studying the effects of the disproportionate distribution of dead loads on the superstructure during construction. The reactions at the obtuse angled end of slab support are larger than the other end, the increase in value over the average value ranging from 0 to 50 per cent for skew angles 20 to 50 degrees. The bearing reactions tend to change to uplift in the acute angle corners with increase in skew angle.

For skew angle lesser than 15°, an approximate method [12] of design may be adopted as, it is reasonably correct. According to this method, bending moments are calculated as for a right bridge of span centre to centre of supports measured parallel to the centre line of the roadway. The main reinforcement is provided in a direction parallel to traffic. Cross reinforcement, which is usually taken as 0.2per cent of effective cross section of the slab is placed parallel to the supports. For skew angles greater than 15°, a more rigorous analysis is desirable, but it is complicated. Analytical and experimental methods have been attempted. A number of methods like FEM, grillage analogy method are used to analyses bridge decks. However, grillage analogy method seems to be general, simple, sufficiently accurate, easy to comprehend and convenient to work even on easily available Personal Computers [8]. For the analysis, using STAAD PRO 2007, in

each case the skew bridge deck is discretized in grid of interconnected beams in case of Grillage analogy and grid of interconnected plates in FEM [13].

2. Analysis of Deck Slab Bridge

The effect of different grid spacing on the behavior of skewed bridges under dead & live load (70R) [6] using grillage analogy method and FEM are analyzed and the appropriate grid size for the various skewed slab bridge were found.

The skew slab bridges having 30° , 45° and 60° skew angle of carriage way 7.5m and skew span 12.11 m is considered for present analysis i.e. reactions, bending moment, deflection under dead and live load with different cases of grid spacing (4 to 12 divisions). Slab thickness is of uniform depth 750 mm.

A skew slab bridge is supported on five isolated bearings at each end, adoption of neoprene bearings of spring stiffness of 40t/mm at each support point. The flexibility of support is considered in analysis.

For the analysis, using STAAD PRO 2007, in each case the whole slab is discretized in grid of interconnected beams in case of Grillage analogy and grid of interconnected plates in FEM. For sake of convenience in analysis and also comparison of results, the spacing of transverse grid lines are kept constant and only longitudinal grid lines changes to study the variation of deflection, reaction, bending moment [10, 11] for different mesh size. Rectangle sections of all grillage beams are considered with constant thickness of 750 mm.

2.1 Grid Pattern

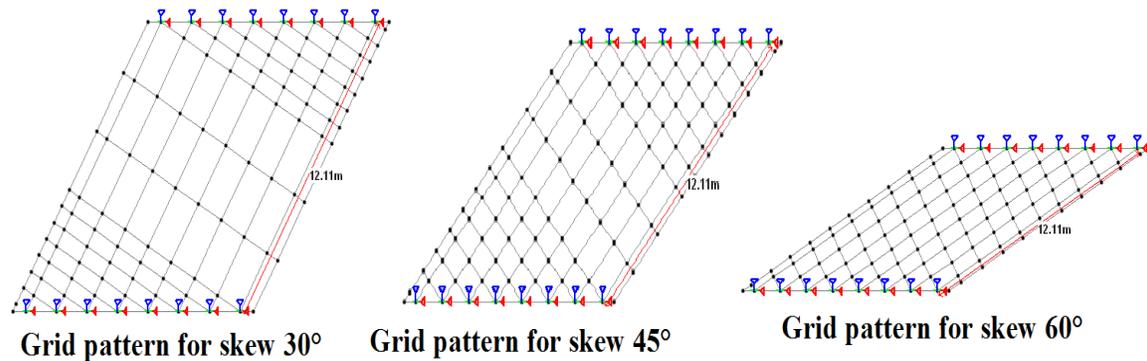


Figure 1: Grid pattern for different skew

All different skewed angle bridges have same grid pattern and same span. STAAD PRO 2007 is the software used for analysis.

- Grillage analogy i.e. given decking system is converted into series of interconnected beams such that given prototype bridge deck and the equivalent grillage of beams are subjected to identical deformations under loading.
- Finite Element Method consists of solving the mathematical model which is obtained by idealizing a structure as an assemblage of various discrete two or three dimensional elements connected to each other at their nodal points, possessing an appropriate number of degrees of freedom.

3. Results and discussions

3.1 Analysis of skew bridge for Skew angle 30°

A skew slab bridge having 30° skew angle supported on five isolated bearings at each end is analyzed for different grid spacing (i.e., 4 to 12 divisions) is analyzed and presented in the fig 2 and 3 and Table 1.

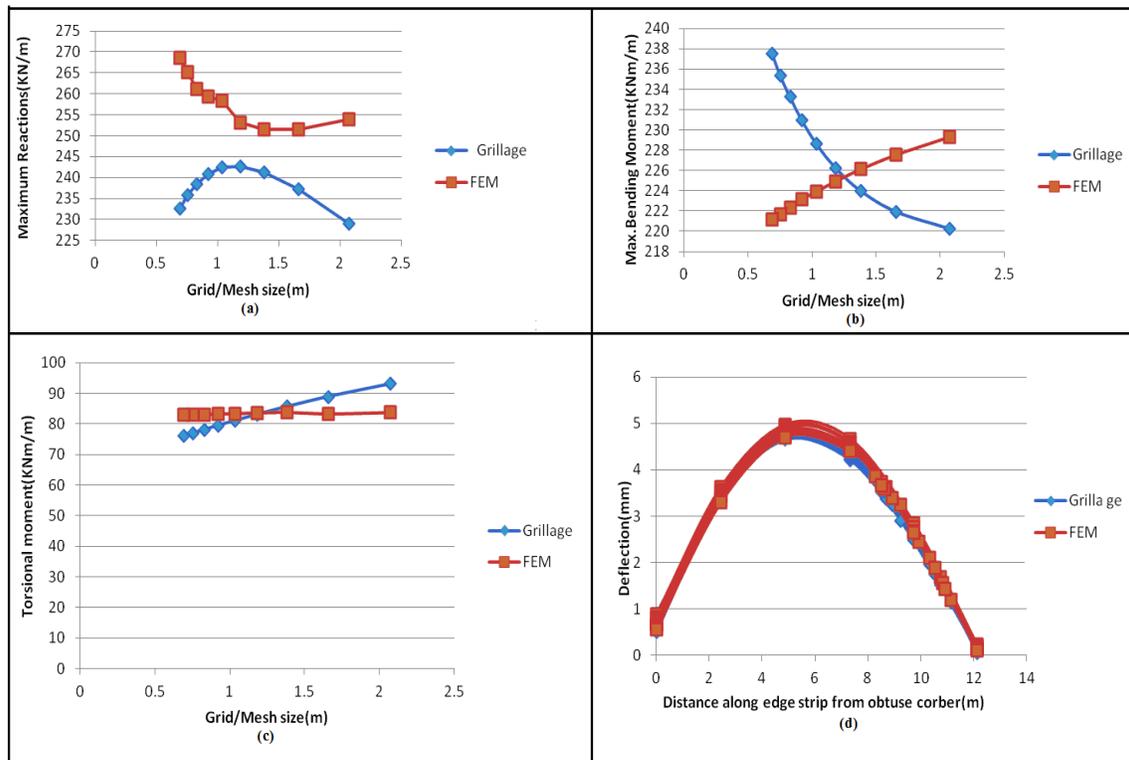


Figure 2: Skew Slab Bridge having 30° skew angle, considering only dead load

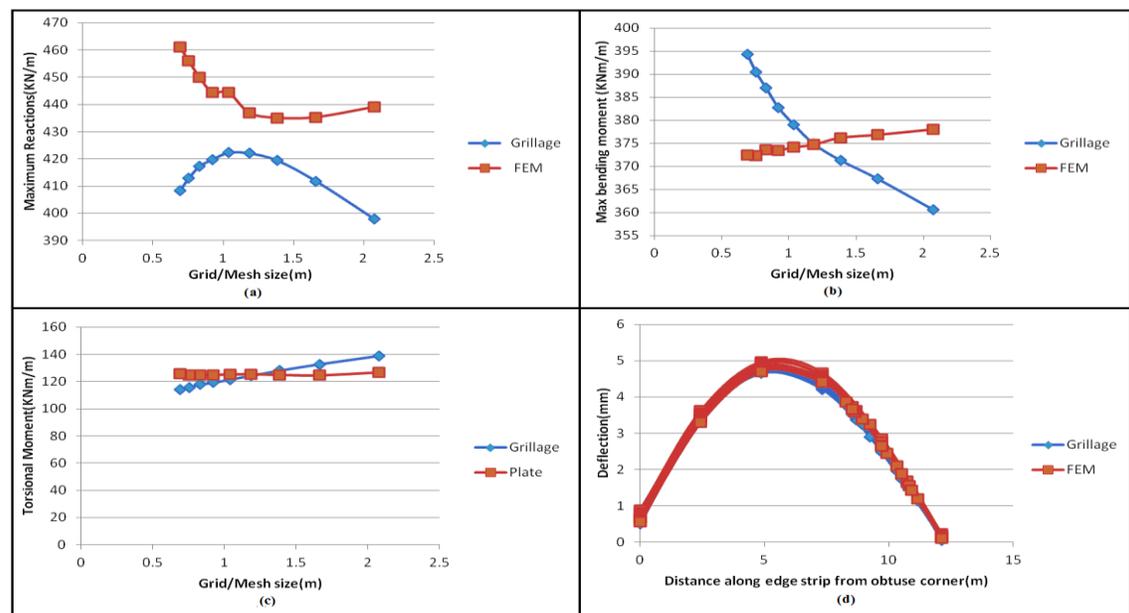


Figure 3: Skew Slab Bridge having 30° skew angle, considering dead load plus live load

It is observed from analysis that, maximum reaction results of FEM overestimate the Grillage results. Grid corresponding to 5, 6 and 7 divisions gives same approx results (<5%). Maximum bending moment results of FEM overestimates grillage results in 4, 5 and 6 divisions. Later, grillage overestimates FEM results and they intersect in between 6 and 7 divisions. So, grid size corresponding to 6 and 7 divisions gives almost same results. Unlike maximum bending moment results, torsional moment results were more in Grillage in 4, 5 and 6 divisions and later it become more in FEM results. Grid corresponding to 6, 7 and 8 divisions gives same (0-2%) results and grid corresponding to 7 divisions gives more accurate result.

Deflection results are almost same in any of grid spacing. i.e., about only 0.2mm changed from 4 divisions to 12 divisions. Although grid corresponding 9 and 10 division gives same results in both methods, but 0.2mm difference corresponding to 6 and 7 divisions also accepted, and it comes closer at 6 and 7 divisions.

As in dead load, in combination of both dead and live load, grid corresponding to 6 and 7 divisions mainly gives closer maximum reactions results in both methods. Grid corresponding to 6, 7 and 8 divisions gives lesser difference in bending moment between two methods. As in dead load case, combination of both dead and live load case grid corresponding to 7 divisions gives same torsional moment in both methods. Deflection results are almost same in any of grid spacing. i.e., about only 0.2mm changed from 4 divisions to 12 divisions. Deflection change is negligible in both methods as in dead load case.

Table 1: Skew Slab Bridge having 30° skew angle

30°Skew Angle	Division	4	5	6	7	8	9	10	11	12	
	Grid size(m)	2.074	1.659	1.382	1.185	1.037	0.921	0.829	0.754	0.691	
Dead Load Only	Reactions (kN/m)	Grillage	228.9	237.2	241.3	242.7	242.3	240.8	238.5	235.7	232.6
		FEM	253.9	251.4	251.6	253	259.2	258.3	261.2	265.1	268.6
		% diff.	9.84	5.67	4.09	4.08	6.51	6.77	8.7	11.07	13.39
	Bending Moment (kNm/m)	Grillage	220.2	221.9	223.9	226.2	228.6	230.9	233.2	235.3	237.5
		FEM	229.3	227.5	226.1	224.9	223.9	223.1	222.3	221.6	221.1
		% diff.	3.95	2.47	0.95	-0.58	-2.09	-3.51	-4.93	-6.17	-7.41
	Torsional Moment (kNm/m)	Grillage	93	88.8	85.6	83	81	79.4	78	76.9	75.9
		FEM	83.7	83.2	83.8	83.6	83.3	83.2	83.1	83	82.9
		% diff.	-11.1	-6.72	-2.14	0.62	2.77	4.53	6.03	7.25	8.4
	Deflection (mm)	Grillage	4.68	4.65	4.65	4.66	4.68	4.71	4.745	4.77	4.80
		FEM	4.97	4.94	4.88	4.84	4.80	4.72	4.75	4.72	4.70
		% diff.	6.23	5.06	4.78	3.57	2.31	0.12	0.02	-1.03	-2.14
Dead Load plus Live Load	Reactions (kN/m)	Grillage	397.8	411.7	419.4	422	422.2	419.7	417.2	412.8	408.3
		FEM	439	435.2	435	436.8	444.4	444.3	449.9	455.9	461.1
		% diff.	9.37	5.39	3.58	3.38	4.99	5.54	7.27	9.44	11.46
	Bending Moment (kNm/m)	Grillage	360.5	367.2	371.2	374.8	379	382.7	387	390.4	394.2
		FEM	378	376.8	376.1	374.6	374.1	373.4	373.5	372.3	372.3
		% diff.	4.61	2.54	1.3	-0.04	-1.31	-2.47	-3.59	-4.87	-5.88
	Torsional Moment (kNm/m)	Grillage	138.5	132.4	127.8	124.2	121.4	119.1	117.8	115.7	114.3
		FEM	126.4	124.5	124.8	125.3	125	124.7	124.6	124.7	125.7
		% diff.	-9.55	-6.35	-2.44	0.869	2.93	4.51	5.48	7.26	9.12
	Deflection (mm)	Grillage	7.62	7.57	7.57	7.57	7.61	7.64	7.69	7.73	7.79
		FEM	8.02	8.07	7.98	7.89	7.82	7.7	7.71	7.67	7.63
		% diff.	5.05	6.22	5.2	3.98	2.72	0.73	0.22	-0.09	-2.05

From the analysis of 30° skew angle bridge (both dead load and combine load cases), we can conclude from the discussions that, mostly in 6 and 7 divisions, the difference of analysis results between them as well as among two methods (Grillage and FEM) are less, and most accurate and closer results given by grid corresponding to 7 grid. Ratios of transverse to longitudinal grid spacing in 7 and 6 divisions are 2.05 & 1.765 respectively.

3.2 Analysis of skew bridge for Skew angle 45°

A skew slab bridge having 45° skew angle supported on five isolated bearings at each end is analyzed for different grid spacing (i.e., 4 to 12 divisions). Ratio of transverse grid spacing to longitudinal grid spacing = 1.83.

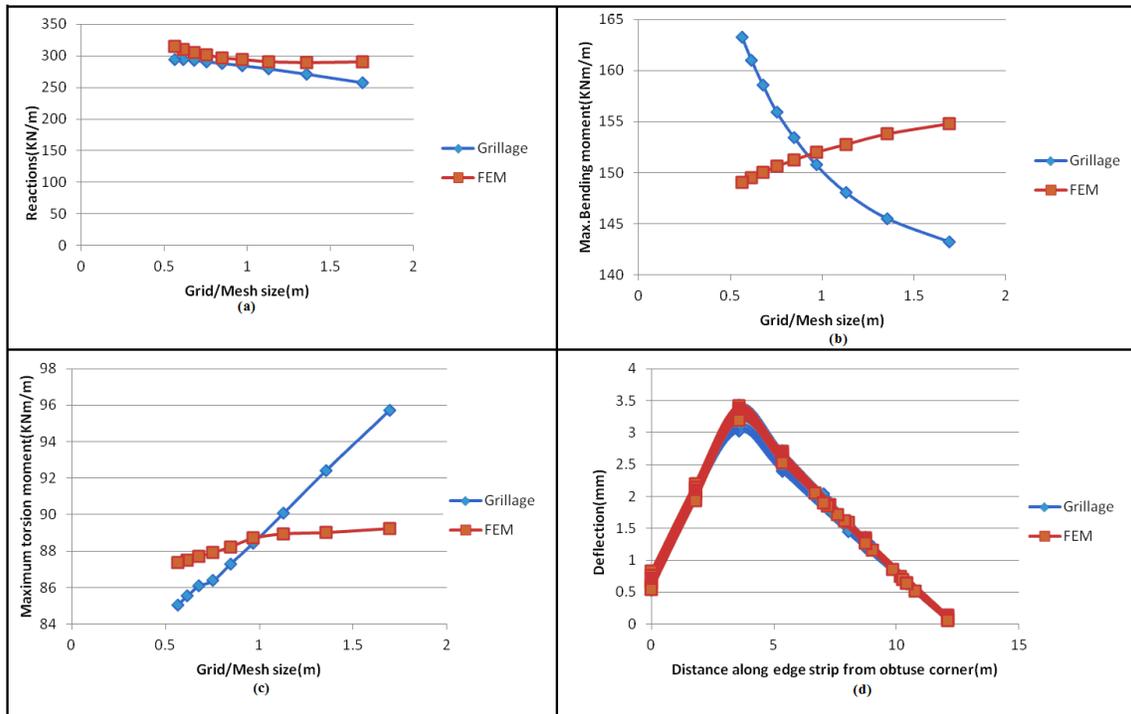


Figure 4: skew Slab Bridge having 45° skew angle, considering only dead load

It is to be noted that difference in maximum reaction between two methods, corresponding to 6, 7, 8 and 9 divisions is less, i.e. about 2-3 %. Deflection results are almost same in any of grid spacing. i.e., deflection change is negligible in both methods. A grid corresponding to 7 and 8 gives only 0.2 mm difference between grillage and FEM results.

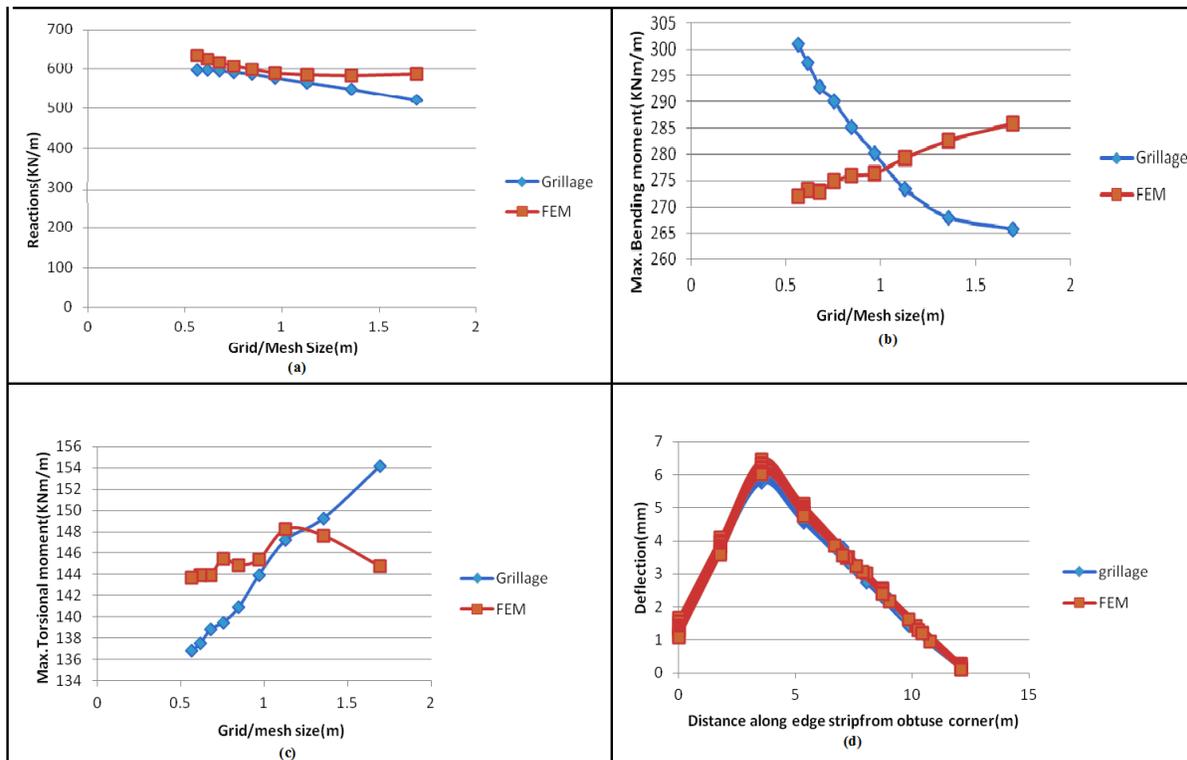


Figure 5: skew Slab Bridge having 45° skew angle, considering dead load plus live load

As in dead load case, combination of dead and live loads, 7, 8 and 9 divisions give closer maximum reaction results between Grillage & FEM method. In dead load plus live load case, FEM overestimates maximum bending moment results over grillage in 4 and 5 divisions, later it underestimates grillage method. It is to be noted that grid corresponding to 6 and 7 gives approximately same results in both method. In combination of dead and live load, grillage over estimates FEM results in 4 and 5 divisions, later it underestimates FEM results. Grid corresponding to 5, 6 and 7 divisions gives lesser difference in results of two methods. We can mark that there is not such variations in torsional moment, in FEM it appears to be constant.

Deflection change due to live load in skew angle 45° is not negligible as that of in dead load case. There is some difference in between two methods compared to other angles. Maximum deflection in both methods becomes same in 9 and 10 divisions. However, previous discussions finds 7 and 8 division appropriate, mainly 7 divisions, which gives 0.4 mm of deflection change between two methods due to combination of dead and live load, it is acceptable. Most of analysis results are same in 6, 7 and 8 but 7 division always gives same results. So, here also 7 divisions we found appropriate. Ratio of transverse grid spacing to longitudinal grid spacing = 1.83

Table 2: skew Slab Bridge having 45° skew angle

45°Skew Angle	Division	4	5	6	7	8	9	10	11	12	
	Grid size(m)	1.693	1.354	1.129	0.967	0.846	0.752	0.677	0.615	0.564	
Dead Load Only	Reactions (kN/m)	Grillage	257.9	270.7	279.1	284.6	288.4	290.8	292.4	293.4	293.5
		FEM	290.5	289.3	290.2	293.5	296.8	301.1	305.4	310.2	314.4
		% diff.	11.20	6.443	3.85	3.01	2.84	3.44	4.27	5.39	6.64
	Bending Moment (kNm/m)	Grillage	143.2	145.5	148.0	150.7	153.4	155.9	158.5	160.9	163.2
		FEM	154.8	153.8	152.7	151.9	151.2	150.6	150.0	149.5	149.0
		% diff.	7.49	5.40	3.08	0.80	-1.46	-3.53	-5.71	-7.66	-9.52
	Torsional Moment (kNm/m)	Grillage	95.7	92.4	90.0	88.4	87.3	86.3	86.0	85.5	85.0
		FEM	89.2	89.0	88.9	88.7	88.2	87.9	87.7	87.5	87.3
		% diff.	-7.25	-3.81	-1.26	0.29	1.05	1.75	1.86	2.25	2.66
	Deflection (mm)	Grillage	3.02	3.03	3.07	3.13	3.19	3.24	3.31	3.37	3.43
		FEM	3.42	3.36	3.33	3.29	3.27	3.24	3.23	3.21	3.19
		% diff.	11.64	9.82	7.62	5.12	2.50	0.03	-2.63	-5.10	-7.47
Dead Load plus Live Load	Reactions (kN/m)	Grillage	522.3	548.3	565.5	575.9	586.2	590.5	594.7	597.0	596.2
		FEM	585.5	581.9	583.8	588.7	597.8	605.3	615.3	624.6	632.8
		% diff.	10.80	5.78	3.12	2.17	1.94	2.43	3.33	4.42	5.78
	Bending Moment (kNm/m)	Grillage	265.6	267.9	273.3	280.1	285.2	290.2	292.8	297.4	300.9
		FEM	285.8	282.6	279.3	276.3	275.9	276.8	272.8	273.1	271.9
		% diff.	7.06	5.19	2.14	-1.35	-3.36	-5.59	-7.33	-8.89	-10.6
	Torsional Moment (kNm/m)	Grillage	154.1	149.2	147.2	143.9	140.8	139.4	138.8	137.4	136.8
		FEM	144.7	147.6	148.2	145.4	144.8	145.4	143.9	143.9	143.7
		% diff.	-6.49	-1.09	0.69	1.00	2.74	4.09	3.53	4.45	4.78
	Deflection (mm)	Grillage	5.75	5.75	5.85	5.89	5.99	6.05	6.16	6.23	6.40
		FEM	6.45	6.34	6.31	6.20	6.17	6.09	6.07	6.01	6.00
		% diff.	10.85	9.16	7.39	5.08	2.89	0.73	-1.56	-3.75	-6.76

3.3 Analysis of skew bridge for Skew angle 60°

A skew slab bridge having 60° skew angle supported on five isolated bearings at each end is analyzed for different grid spacing (i.e., 4 to 12 divisions). It can be marked that as skew angle increases, differences in maximum reaction values between two methods becomes less. In bending moment case, FEM overestimates grillage results in 4, 5 and 6 divisions. Later, it underestimates grillage results and 7, 8 and 9 divisions give almost same results.

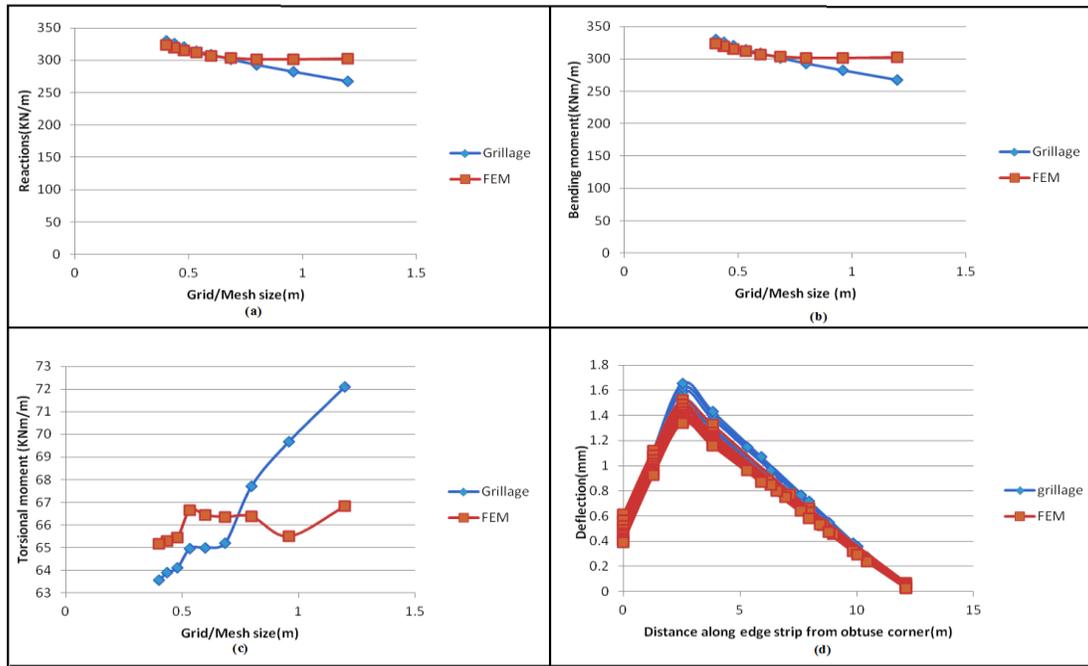


Figure 6: skew Slab Bridge having 60° skew angle, considering only dead load

In bending moment case FEM overestimates grillage results in larger grid sizes and vice versa happens in smaller grid sizes. Grid corresponding to 6 and 7 divisions gives closer results in both methods. It can also be noted that as skew angle increases, more closer is the difference as in reactions above. From above table and graph, we can see that there is not so difference in torsion moment between two methods in any of grid spacing and grillage overestimates FEM results and from 7 divisions it underestimates FEM result. As per above results, 6, 7, 8, 9 and 10 divisions any one can be taken as difference is negligible and also gives almost same value. As similar to other cases, deflection is same everywhere, its change is negligible. Grids corresponding to 6 and 7 divisions give same value in both methods. As in dead load case, same in combinations of dead and live load case, maximum reactions have very less difference between two methods as skew angle increases. Grid corresponding to 7, 8, 9, 10 and 11 divisions gives lesser difference of results between two methods. As in dead load case, same happens in combinations of dead and live load also and grids corresponding to 6 and 7 divisions give almost same value in both methods.

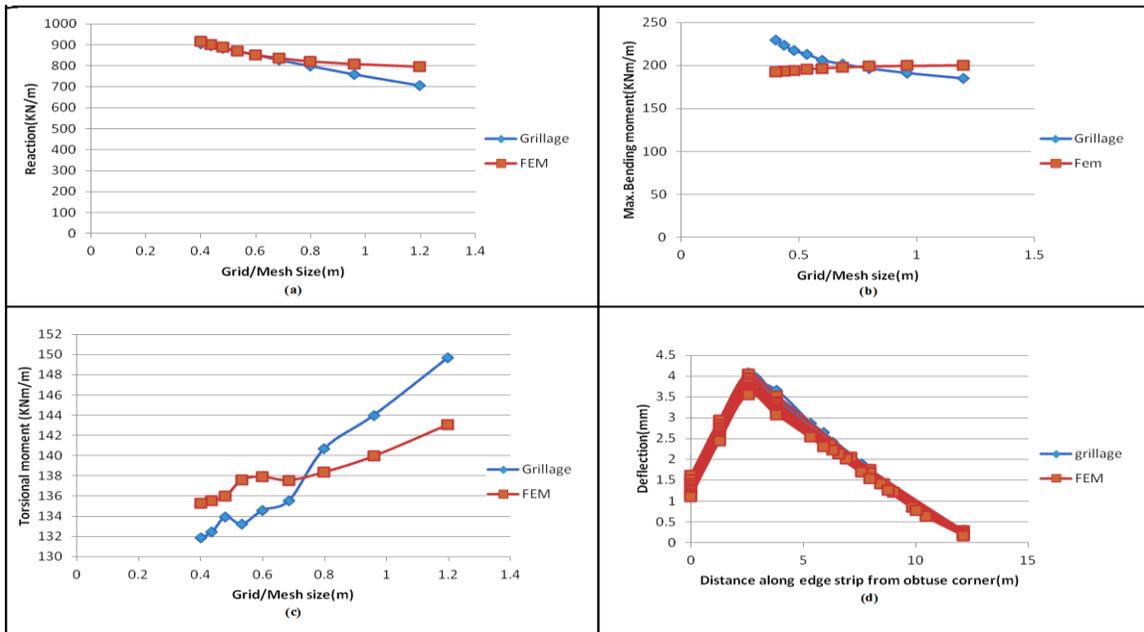


Figure 7: skew Slab Bridge having 60° skew angle, considering dead load plus live load

In torsional moment case grillage overestimates FEM results in larger grid sizes, vice versa happens in lesser grid sizes. Unlike, dead load case, in combination of dead and live load case, considerable difference between results is there. Grids corresponding to 5, 6 and 7 division give lesser difference between two methods.

In this case also deflection change is not so significant. There is not so much difference in any grid size between two methods. Deflection pattern is consistent with bending moment pattern, first FEM overestimates grillage results. Later FEM underestimates grillage results. Grid corresponding to 7 and 8 gives same results.

In skew angle 60° also, all the analysis results are approximately same corresponding to 7 divisions. Ratio of transverse to longitudinal grid spacing =1.858

Table 3: skew Slab Bridge having 60° skew angle

60°Skew Angle	Division	4	5	6	7	8	9	10	11	12	
	Grid size(m)	1.197	0.958	0.798	0.684	0.598	0.532	0.479	0.435	0.399	
Dead Load Only	Reactions (kN/m)	Grillage	267.0	282.0	293.0	301.7	309.0	314.4	320.8	325.6	330.2
		FEM	302.6	301.4	301.9	303.6	307.0	311.9	314.8	319.2	323.7
		% diff.	11.73	6.45	2.95	0.63	-0.63	-0.82	-1.89	-2.00	-2.00
	Bending Moment (kNm/m)	Grillage	77.6	80.2	82.5	84.5	86.4	91.5	93.7	98.4	102.7
		FEM	85.1	84.2	83.6	83.0	82.6	82.1	81.7	81.4	81.0
		% diff.	8.85	4.82	1.40	-1.79	-4.63	-11.4	-14.5	-20.9	-26.8
	Torsional Moment (kNm/m)	Grillage	72.0	69.6	67.7	65.2	64.9	64.9	64.0	63.8	63.5
		FEM	66.8	65.4	66.3	66.3	66.4	66.6	65.4	65.2	65.1
		% diff.	-7.88	-6.39	-2.04	1.73	2.16	2.56	2.04	2.13	2.46
	Deflection (mm)	Grillage	1.42	1.43	1.45	1.48	1.51	1.50	1.58	1.61	1.65
		FEM	1.52	1.48	1.45	1.42	1.40	1.38	1.36	1.35	1.33
		% diff.	6.30	3.36	-0.02	-4.13	-8.19	-8.59	-15.92	-19.74	-23.37
Dead Load plus Live Load	Reactions (kN/m)	Grillage	706.6	760.2	799.3	828.3	851.2	869.5	883.6	895.3	905.3
		FEM	795.9	808.9	821.1	836.3	851.7	870.0	888.8	901.9	915.7
		% diff.	11.22	6.025	2.65	0.959	0.066	0.063	0.581	0.722	1.13
	Bending Moment (kNm/m)	Grillage	184.9	191.3	196.7	201.7	206.3	213.3	217.3	223.6	229.9
		FEM	200.4	199.8	199.1	197.8	196.6	195.7	194.4	193.5	192.4
		% diff.	7.70	4.25	1.22	-1.96	-4.89	-8.99	-11.7	-15.5	-19.5
	Torsional Moment (kNm/m)	Grillage	149.6	143.9	140.6	135.5	134.5	133.2	133.9	132.4	131.8
		FEM	143.0	139.9	138.3	137.5	137.9	137.6	135.9	135.5	135.2
		% diff.	-4.40	-2.79	-1.64	1.49	2.46	3.29	1.52	2.34	2.58
	Deflection (mm)	Grillage	3.79	3.79	3.80	3.84	3.88	3.88	3.98	4.02	4.07
		FEM	4.05	3.93	3.86	3.79	3.72	3.67	3.62	3.58	3.54
		% diff.	6.26	3.65	1.39	-1.34	-4.29	-5.71	-9.85	-12.51	-15.03

4. Conclusion

In the present study comparison of different skew angle bridges is performed using Finite element and Grillage method and results are presented. It is observed from the analysis that mostly seven divisions on gridding is appropriate i.e., ratio of transverse grid lines to longitudinal grid lines is 1.8-2.0. However, six and eight divisions also gives approx similar results as that of seven, so seven divisions can be preferred with ratio within 1.5-2.0 As skew angle increases, difference between reaction values (Grillage and FEM) of grid sizes decreases. In skew angle 30°, there is lot of difference but it decreases, as skew angle increases. In the case of bending moment, FEM results overestimated Grillage results, and after the point of intersection FEM results underestimated Grillage. Deflection is consistent with longitudinal bending moment, it follows the same pattern. Torsion moment follows the reverse pattern of longitudinal moment. Grillage results overestimated the FEM results in larger grid sizes and after intersection it underestimated FEM results. Variation of grid sizes analysis results predicts that, variation in reaction value is same in FEM and Grillage method but variation of bending and torsion moment in FEM is lower than grillage results. So, FEM may be preferred for analysis of skew bridges efficiently with certain limitation.

References

- [1] C. Menassa, M. Mabsout, K. Tarhini and G. Frederick. 2007. Influence of Skew Angle on Reinforced Concrete Slab Bridges. Journal of Bridge Engineering, ASCE. March-April.
- [2] Doug Jenkins, 2004 "Bridge Deck Behaviour Revisited"; MEngSci MIEAust MICE.
- [3] E.J. O'Brien, D.L. Keogh, 1998 "A discussion on neutral axis location in bridge deck cantilevers", Department of Civil Engineering, University College, Dublin, Ireland.
- [4] H. Zeng, J. Kuehn, J. Sun, H. Stalford, 2000 "An Analysis of Skewed Bridge/Vehicle Interaction Using the Grillage Method", University of Oklahoma.
- [5] Huang, P. "The Model Experimental Research on Skew Girder Bridge", J. of Xi'an Highway University. Vol. 19(1), 1998.
- [6] IRC 21-1987, "Specifications and code of Practice for Road Bridges"
- [7] Ibrahim S. I. Harba, "Effect of Skew angle on Behavior of Simply Supported R. C. T-Beam Bridge Decks", ARPN Journal of Engineering and Applied Sciences, VOL. 6, NO. 8, Aug. 2011.
- [8] Jaeger, L. G. and Bakht, B., "Bridge Analysis by Micro-computer", McGraw-Hill, New York, 1990.
- [9] Kar, Ansuman 2011, "Analysis of skew bridges using computational methods", *M. Tech Dissertation*, Department of Civil Engineering, Institute of Technology, Banaras Hindu University, Varanasi
- [10] Mabsout M., Menassa C. and Tarhini K. 2002. Effect of skewness in concrete slab bridges. 9th Int. Conference on Computing in Civil and Building Engineering, Taipei, Taiwan. pp. 663-668.
- [11] Maher Shaker Qaqish, "Distribution of Bending Moments in the Longitudinal T-Beams of Skew Deck-Slab", European Journal of Scientific Research ISSN 1450-216X Vol.13 No.3 (2006), pp. 374-380.
- [12] Maher Kakish, 2007 "Bending moment distribution at the main structural elements of skew deck-slab and their implementation on cost effectiveness", College of Engineering, Al- Balqa Applied University, Salt, Jordan.
- [13] Patrick Théoret, Bruno Massicotte and David Conciatori, "Analysis and Design of Straight and Skewed Slab Bridges", Journal of Bridge Engineering. March 2011.
- [14] Saif Laith Khalid Alomar, "Comparison between grillage model and finite element model for analyzing bridge deck" Master of Engineering Thesis, Faculty of Civil Engineering Universiti Teknologi Malaysia, June-2009.
- [15] Surana, C.S. and Aggarwal, R. "Grillage Analogy in Bridges Deck Analysis", Narosa Publishing House, New Delhi, First Edition 1998.

A New Omni-directional Monopole Antenna for Interference Reduction

T .S. Ghouse Basha ¹, K.Tulasi Krishna ², C.Chandrakala³, V.Kishore⁴, D.Aruna⁵

1 Associate professor, 2 Assistant professor, 3 Assistant Professor, 4 Assistant Professor,
5Assistant Professor

Department of Electronics and Communication Engineering
K.O.R.M.COLLEGE OF ENGINEERING, KADAPA, INDIA.

Abstract- A compact Ultra wideband (UWB) antenna with band- notched characteristic is presented in this paper. It has compact size of 30 mm x 31mm and has ultra wide band operation. A 'C' shaped slot was introduced to achieve band notch function from 4.8 GHz to 6 GHz to avoid interference from WLAN. The proposed antenna has ultra wide band frequency range from 3 GHz to 11.5 GHz for return loss below -10 dB, except for frequency stop band from 5 GHz to 6 GHz. Details of antenna are presented with parametric study for Ultra Wideband Applications. The bandwidth is varied by varying the width W_4 of inner tuning stub and height L_3 of feed and Ultra Wideband width is obtained. The antenna is Omni directional in operating bandwidth and it has good radiation efficiency. The fundamental parameters such as return loss, VSWR, radiation pattern are obtained, which meet standard specifications. Method of moments based IE3D simulator is used to analyze this antenna.

Keywords- Coplanar waveguide, UWB antenna, VSWR, Omni directional radiation pattern

I. INTRODUCTION

Ultra wideband communication systems have received great attraction in wireless world. It is popularly used technology in radar and remote sensing. Ultra Wideband technology provides promising solutions for future communication systems due to excellent immunity to multi path interference, large bandwidth and high speed data rate. A bandwidth from 3.1 GHz to 10.6 GHz is allocated for UWB systems by Federal Communication Commission (FCC) in 2002. From then, the design of UWB antenna became challenging task for engineers in UWB systems. One major issue in UWB is the design of compact size and wide band antenna [1]. Several UWB antennas have been studied for UWB applications [2-3]. Recently slot antennas became popular for UWB applications [4-6]. Bow-tie antenna is one technique for UWB antenna systems [7]. Parasitic elements around the antenna also provide broad bandwidth but increases size of antenna [8]. The planar monopole antennas [9, 10] are better for UWB applications due to small size and stable radiation pattern. UWB antennas can be fed by microstrip line or Coplanar Waveguide (CPW). The CPW antennas are useful for microwave and millimeter applications because they offer low profile and wide bandwidth, ease of integration with circuits. The CPW has low radiation leakage and less dispersion than conventional microstrip lines. They have also low power consumption and high speed data rate for transmitter.

A UWB antenna is also susceptible to interference by narrow band signal of neighboring RF system such as IEEE 802.11a WLAN having operating frequency range 5.125-5.825GHz. So, it is desirable to design UWB antenna with band notch characteristic to avoid interference from the bands [11, 12]. The conventional methods to achieve the notched band are cutting a slot (U-shaped, V-shaped and arc shaped slots) on the patch [13, 14] or embedding a quarter wavelength stub with in a large slot on the patch [15], inserting a slit on the ground [16]. Another way is based on placing parasitic elements near the printed monopole, which play role as filter to reject limited band. A miniaturized UWB antenna with 5 GHz band rejection is also reported [17]. But it has complicated logic. A good UWB antenna should have ultra wide band width, high radiation efficiency, directional or Omni directional radiation pattern.

So, we have proposed CPW-fed antenna for achieving notched band from 5 GHz to 6 GHz which avoids entire WLAN. The antenna provides stable radiation pattern and high radiation efficiency. The antenna geometry is introduced in section II. The antenna parameters return loss, radiation characteristics are discussed in section III. The effect of angle α and L_3 on antenna returns loss and bandwidth also investigated. Conclusions are given in section IV.

II. ANTENNA STRUCTURE

The antenna is fabricated with low cost FR4 substrate with relative permittivity $\epsilon_r = 4.4$ and thickness $h = 1.6$ mm. It has compact area of 30 mm x 31 mm as shown in Figure 1. The antenna is fed by 50 Ω CPW feed line. The gap 'g' between centre conductor and ground plane is 0.4 mm. The various optimized parameters of antenna are $W = 30$ mm, $W_1 = 13.1$ mm, $W_2 = 5$ mm, $W_3 = 3$ mm, $W_4 = 15$ mm, $W_5 = 5.75$ mm, $W_6 = 3.5$ mm, $W_7 = 2.8$ mm, $W_8 = 6$ mm, $L = 31$ mm, $L_1 = 4.75$ mm, $L_2 = 9$ mm, $L_3 = 14$ mm, $L_4 = 8.825$ mm, $L_5 = 1.175$ mm, $L_6 = 1$ mm, $L_7 = 0.5$ mm, $L_8 = 3.6$ mm, $L_9 = 6$ mm. The antenna has two resonating frequencies 3.63 GHz and 7.51 GHz. The obtained resonant frequencies vary depending on the location of inner tuning stub and the gap between CPW-feed line and ground. The antenna has only one layer of substrate with singled sided metallization part. This allows easy manufacturing of antenna with low cost.

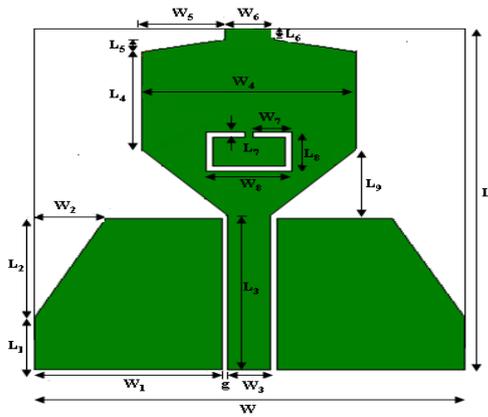


Figure:1 Proposed antenna structure

III. RESULTS AND DISCUSSION

The proposed antenna was simulated and optimized using Zeland IE3D simulator. Figure 2 compares the fabricated antenna with that of simulated antenna. The VSWR curve shows that the antenna achieves an impedance bandwidth of 7.6 GHz ranging from 3 GHz to 11.5 GHz for VSWR below 2 except for notched bandwidth from 5 GHz to 6 GHz. This band is avoided by inserting ‘C’ shaped slot in conductive layer to reduce interference from WLAN. The proposed antenna has good performance to reject unwanted WLAN band. By controlling the size and location of ‘C’ slot, desired notched band can be obtained. The important feature of proposed antenna is the capability of impedance matching at both resonating frequencies 3.62 GHz and 7.54 GHz using CPW feed line.

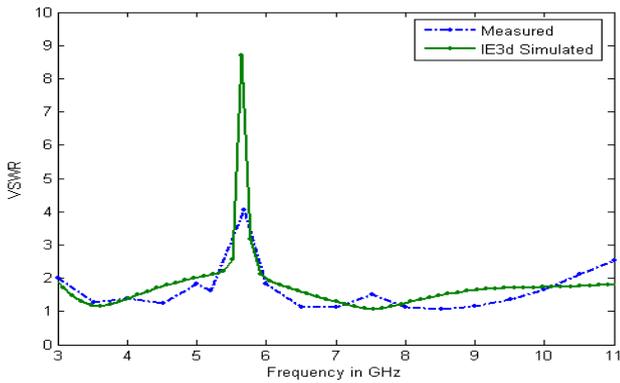


Figure: 2 Simulated and Measured VSWR for proposed antenna

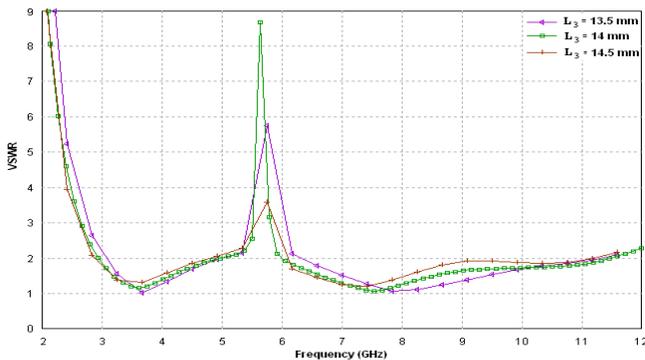


Figure:3 Return loss vs. frequency for the proposed antenna with $L_3 = 13.5, 14, 14.5$ mm

The performance of the antenna has been analyzed by using Method of moments based IE3D simulator. The effect of L_3 and W_4 on the antenna return loss and Band width are studied. These are most sensitive parameters. As L_3 increases from 13.5 mm to 14.5 mm, the band width increases to 7.6 GHz from 7 GHz and then decreases to 7.03 GHz. The higher resonant frequency decreases. When W_4 increases from 14 mm to 16 mm, band width increases to 7.6 GHz and then slightly changes. The higher resonant frequency is slightly shifted down. The effect of L_3, W_4 on resonant frequencies and bandwidth are shown in Figures 3 and 4. They are also represented numerically in table I and table II.

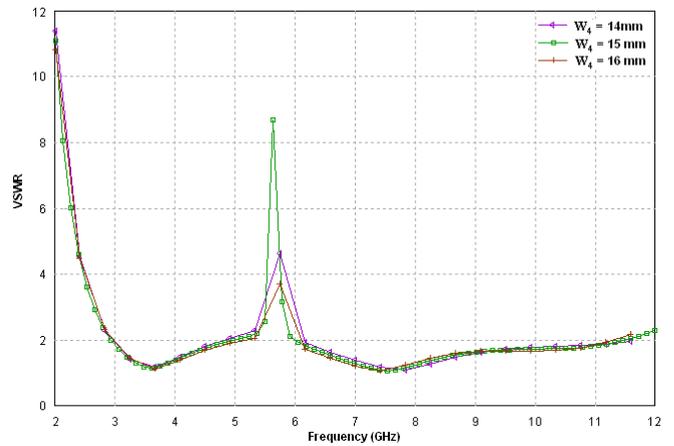


Figure: 4 Return loss vs. frequency for the proposed antenna with $W_4 = 14, 15, 16$ mm Parametric study

TABLE I. EFFECT OF L_3 ON RESONANT FREQUENCIES AND BAND WIDTH

L_3 (mm)	Resonant frequency (GHz)	Band width (GHz)
13.5	3.66, 7.83	7.0
14	3.62, 7.54	7.6
14.5	3.66, 7.41	7.03

TABLE II. EFFECT OF W_4 ON RESONANT FREQUENCIES AND BAND WIDTH

W_4 (in degrees)	Resonant frequency (GHz)	Band width (GHz)
14	3.63, 7.83	7.3
15	3.62, 7.54	7.6
16	3.68, 7.45	7.4

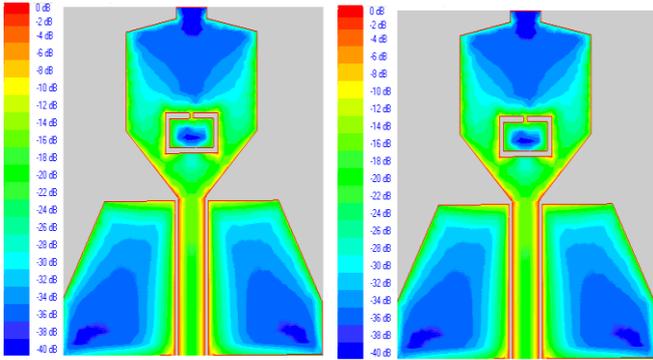


Figure: 5 Current distribution at a) 3.62 GHz b) 7.54 GHz.

The current distribution is more along the transmission line as shown in Figure 5. But, around the C-shaped slot, the current is small. Here, destructive interference takes place for excited currents in the antenna, which makes antenna non responsive in the notched band.

Far field radiation characteristics were also studied. The Figures 6 and 7 present radiation patterns in E and H-planes at 3.62 GHz and 7.54 GHz. The gain patterns are Omni directional in H-plane and monopole like in E-plane. The different frequencies across the operating band width showed similar radiation patterns. So, stable radiation patterns have been obtained for proposed antenna. H-plane patterns show larger cross polarization than E-plane due to strong horizontal surface current components and electric field. The peak antenna gain of antenna varies from 1.5 dBi to 2.8 dBi as shown in Figure 8 and it has maximum value of 2.8 dBi at 4.3 GHz. This antenna has high radiation efficiency of 80% in the operating region.

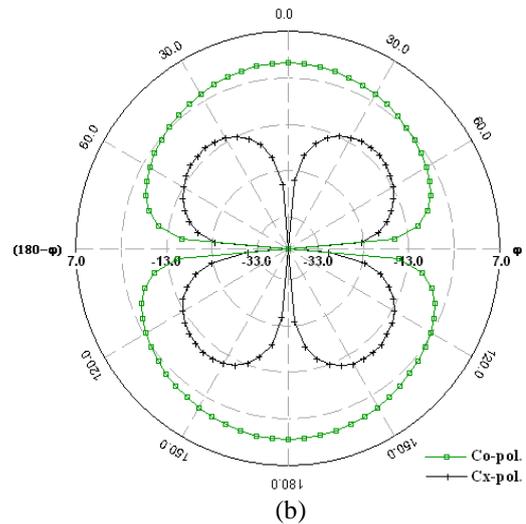


Figure: 6 Radiation pattern in E-plane at a) 3.62 GHz b) 7.54 GHz.

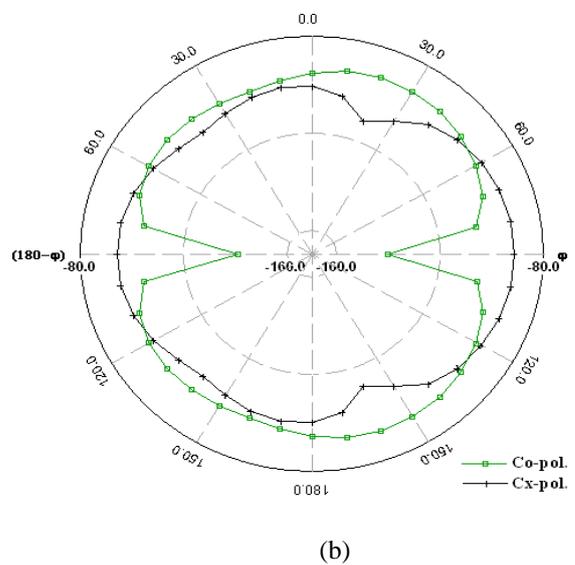
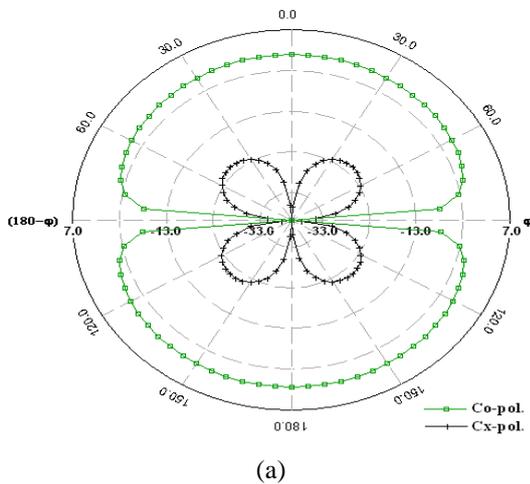
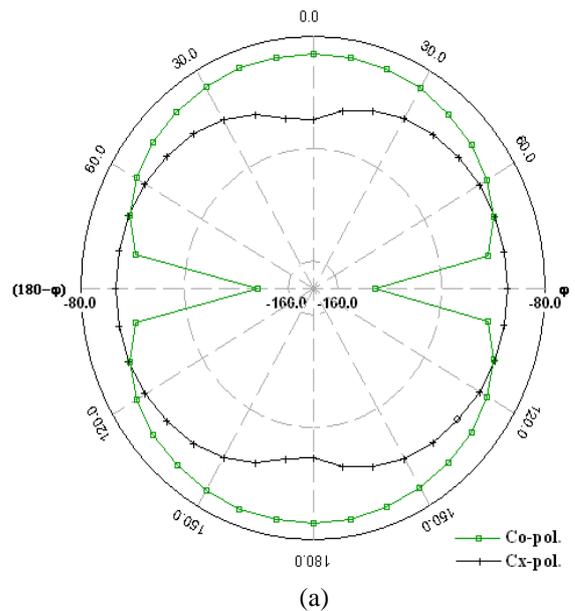


Figure: 7 Radiation pattern in H-plane at a) 3.62 GHz b) 7.54 GHz.

IV. CONCLUSION

A compact CPW fed monopole antenna with band notched characteristic is presented to avoid interference from WLAN. Notch band is achieved by inserting 'C' shaped slot in conductive layer. It has good impedance matching. It has total impedance bandwidth of 7.6 GHz. Parametric study is performed by varying width W_4 of inner tuning stub and height L_3 of conducting feed. The antenna exhibits Omni directional radiation pattern in H-plane with compact size. It has simple structure. The gain and radiation pattern of antenna have been investigated and found to be stable. The antenna has high gain of 2.5 dBi in the operating band width. The antenna has highest radiation efficiency of 80% in the operating region. The antenna is useful for radio communications.

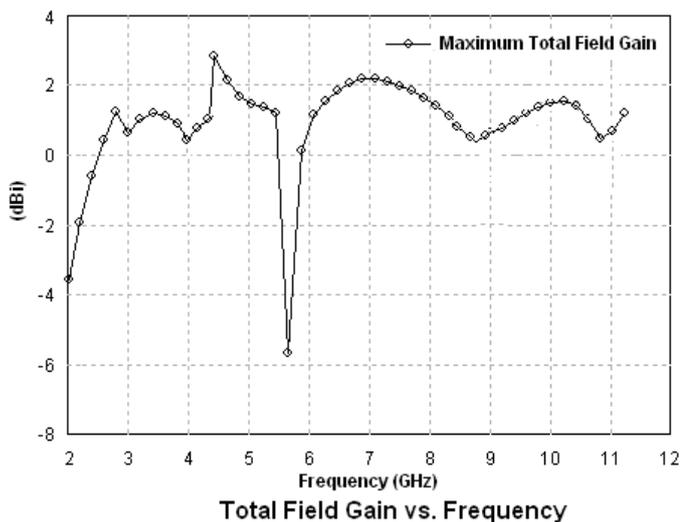


Figure: 8 Gain vs. frequency of proposed antenna

REFERENCES

- [1] Pengcheng Li, Jianxin Liang and Xiaodong Chen, "Study of Printed elliptical/Circular slot Antennas for ultrawideband applications," *IEEE Trans. Antennas Propag.*, vol. 54, no. 6, pp. 1670-1675, 2006.
- [2] G. M. Zhang, J. S. Hong, B. Z. Wang, Q. Y. Qin, J. B. Mo, and D. M. Wan, "A novel multi folded UWB antenna fed by CPW," *J. of Electromagn. Waves and Appl.*, Vol. 21, No. 14, 2109-2119, 2007.
- [3] Shi-Wei qu, Jia-Lin Li, and Quan Xue, "A band notched ultra wide band printed monopole antenna," *IEEE Antennas Wireless Propa. Lett.*, vol. 5, pp. 495-498, 2006.
- [4] T. G. Ma and S. K Jeng, "Planar miniaturized-slot-fed annular slot antennas for ultra-wideband radios," *IEEE trans. Antennas Propag.*, vol. 53, pp. 1194-1202, 2005.
- [5] I. J. Yoon *et al.*, "Ultra-wideband tapered slot antenna with band-cutoff characteristic," *Electron. Lett.*, vol. 41, no. 11, pp. 629-630, 2005
- [6] Y. Gao, B. L. Ooi, and A. P Popov, "Band notched ultra wide band ring monopole antenna," *Microw. Opt. Technol. Lett.*, vol. 48, No. 1, pp. 125-126, 2006.
- [7] C. Y. Huang and D. Y. Lin, "CPW-fed bow-tie slot antenna for ultra-wideband communications," *Electron Lett.*, vol. 42, No. 19, pp. 1073-1074, 2006
- [8] K. H Kim, Y. J .Cho, S.H Hwang. S. O. Park, "Band notched UWB planar monopole antenna with two parasitics," *Electron. Lett.*, vol.. 41, pp. 783-785, 2005.
- [9] K. F. Jacob, M. N. Suma, R. K. Raj, M. Joseph and P. Mohanan, "Planar branched monopole antenna for UWB applications," *Microw. Opt Technol. Lett.*, Vol. 49, No. 1, pp. 45-57, 2007.
- [10] Y. J . Cho, K. H. Kim, D. H choi, S. S. Lee and S. O. Park , "A miniature UWB planar monopole antenna with 5-GHz band rejection filter and time domain characteristics," *IEEE trans Antennas Propag.*, vol.. 54, pp. 1453-1460, 2006.
- [11] Cheol-Bok kim, Jung-Sup Lim, Jae-Sam Jang, Young-Ho Jung, Ho- Sang Lee, Mun So Lee, "Design of wideband notched Compact UWB antenna," in *Proc. IEEE Asia Pacific Microwave conference*, 2007
- [12] Z. F. Yiao, X. Wang, S. G. Zhou, L. Sun and Q. Z. Liu, "A novel dual band-notched ultra wide band slot antenna," in *Proc. IEEE the 8th International Symposium on Antennas and Propagation and EM theory*, pp. 66-69, 2008.
- [13] Wang-Sang Lee, Dong-Zo Kim, Ki-Jin Kim, and Jong-Won Yu, "Wideband Planar Monopole antennas with Dual Band-Notched characteristics," *IEEE Trans. Microw. Theory Tech.*, vol.54, No. 6, Part 2, pp. 2800-2806, 2006
- [14] Xinan Qu, Shun-Shi Zhong, and Wei Wang, "Study of Band-Notched Function for a UWB Circular Disc Monopole Antenna," *Microw. Opt. Technol. Lett.*, vol. 48, No 8, pp. 1667-1670, 2006.
- [15] Y. Gao, B. L. Ooi, and A. P Popov, "Band notched ultra wide band ring monopole antenna," *Microw. Opt. Technol. Lett.*, vol. 48, No. 1, pp. 125-126, 2006.
- [16] Saou-wen Su, and Kin-Lu Wang, "Printed band-notched ultra wideband quasi-dipole antenna," *Microw. Opt. Technol. Lett.*, vol. 48, No 3, pp. 418-420, 2006.
- [17] K. H Kim, Y. J .Cho, S.H Hwang. S. O. Park, "Band notched UWB planar monopole antenna with two parasitic," *Electron. Lett.*, vol. 41, No. 14, pp. 783-785, 2005



T.S. Ghouse Basha is presently working as an Associate Professor and HOD in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa. He carried out his MTech project work in Defence Research and Development Laboratory, Hyderabad and working in teaching field since nine years in different cadres. He received his BTech and MTech from the Department of Electronics and Communication Engineering from JNTU University and Nagarjuna University respectively. He is pursuing his Ph.D in microwave antennas. His areas of interest include microwave antennas, digital signal processing and mobile communications.



D.Aruna is presently working as an Assistant Professor in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa, and pursuing her M.Tech from JNTU Anantapur. She completed her B.Tech in 2008 from JNT University Hyderabad. She working in the field of teaching since 2008. Her areas of interest includes Microprocessor & Interfacing, Probability Theory And communications.



C.Chandrakala is presently working as an Assistant Professor in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa, and pursuing her M.Tech from JNTU Anantapur. She completed her B.Tech in 2006 from JNT University Hyderabad. She working in the field of teaching since 2006. Her areas of interest includes VLSI, Electronic Circuit Analysis, Radar Systems.



V.Kishore is presently working as an Assistant Professor in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa. He completed his B.Tech and M.Tech in 2008 and 2010 from JNT University Hyderabad and JNT University Anantapur. He working in the field of teaching since 2010. His areas of interest includes Embedded Systems, Signals and Systems, Electromagnetic waves.



K.Tulasi Krishna is presently working as an Assistant Professor in the Department of Electronics and Communication Engineering in KORM College of Engineering, Kadapa. He completed his B.Tech and M.E. in 2003 and 2005 from JNT University Hyderabad and Sathya Bhamma University. He is working in the field of teaching since 2009. He worked for 4 years in telecom industry before joining in teaching. His areas of interest include Microprocessor & Interfacing, Data communication Systems, Microwave Antennas.

A Novel Two Stage Binary Image Security System Using (2,2) Visual Cryptography Scheme.

Mr. ROHITH S

Lecturer,
Department of E&C,
NCET, Bangalore,
Karnataka, India-562110

Mr. VINAY G

Student.
Department of E&C,
NCET, Bangalore,
Karnataka, India-562110

Abstract: Visual Cryptography Scheme (VCS) is an encryption method used to encode secret written materials. The idea is to convert the written material into a binary image and encode this image into n shadow image, it is also called as shares of images. The decoding only requires selecting some subset of these n shadow images, making transparencies of them and stacking them on top of each other. Main advantage of this scheme is mathematical computation complexity is reduced compared to conventional cryptographic techniques.

This paper presents design of two stage binary image security scheme using (2,2) Visual Cryptography Technique such that we are not able to get back the original image at the first decoding stage. As the cryptographic algorithm becomes more relevant, it will become inefficient. So, the basic model of Visual Cryptography is not an efficient tool to hide the information anymore. This method aims to improve efficiency of VCS. The performance of proposed algorithm is compared with three different binary images. Result shows that there is no residual information present in the shares.

Key word: Visual Cryptography Scheme, LFSR, Binary Image Security.

I Introduction

With the invention of the Internet, more and more digital data can be accessed via the network. Internet users can transmit and store images with less secured channels. To secure the information one possible technique is cryptography where information is encrypted using key and same key is used to decrypt the information. Here computation complexity of decryption algorithm increases the information security [9]. Visual Cryptography is a technique which provides information security and uses simple decryption algorithm unlike the computationally complex algorithms used in traditional cryptography schemes. This technique allows Visual information such as pictures, text, etc. to be encrypted in such a way that their decryption can be performed by the human visual system. This technique encrypts a secret image into shares such that stacking a sufficient number of shares reveals the secret image [1-4]. In this paper an overview of the emerging Visual Cryptography scheme and design of the proposed scheme is discussed.

An example of (2, 2) Visual Cryptography Scheme is given in figure 1a to 1d. Original binary image is shown in Figure 1a. Figure 1b & 1c are share1 and share2 generated from original image by (2, 2) Visual Cryptography Scheme discussed in section IV. Figure 1d reconstructed image by application of stacking process (BIT OR operation) on share1 and share2. The reconstructed image is visually identical to the original image.

The rest of the paper is organized as follows. The section II gives brief description about earlier work. Applications of Visual Cryptography Schemes are discussed in section III. Design preliminaries are discussed in section IV. Proposed scheme is given section V. In Section VI Results of proposed schemes are discussed. Conclusion is given in section VII.



Figure 1a Original Binary Image

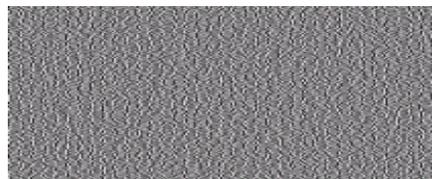


Figure 1b Share 1

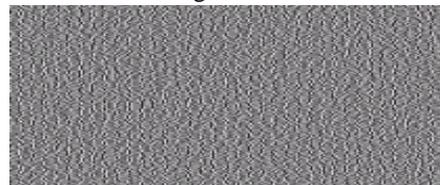


Figure 1c Share 2

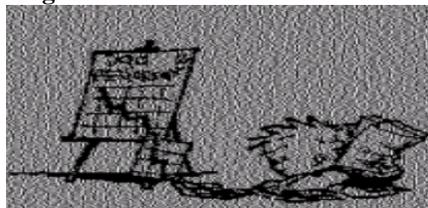


Figure 1d Reconstructed Image by Stacking Process

II Previous Work.

Many authors published different Visual Cryptography Schemes [1-8] for different applications. Each scheme has its own advantages and disadvantages. Few such schemes were given in Table 1.

Table 1 Previous work in Visual cryptography

Authors	Features of the Work
Naor and Shamir [2]	Basic (2,2) Visual Cryptography Scheme was discussed
Ateniese et al[8]	Elegant VCS for general access structures based on the cumulative array method was discussed.
Tzeng and Hu[1]	In this VC scheme secret image can be either darker or lighter than the background
Naor and Pinkas[7]	Methods of authentication and identification between two participants were discussed
Yang et al [6]	Cheating method against some VC schemes were discussed. In their cheating method, the cheater needs to know the exact distribution of black and white sub pixels of the shares of honest participants. Based on this characteristic they proposed a cheat preventing method to prevent the cheater from obtaining the distribution.

III Application of VCS

There are many applications incorporated with the Visual Cryptography Scheme. Two main applications are discussed in this section.

- 1) **Electronic-Balloting System:** Dilemma et al [3] proposed a secret-Ballot Receipts system that is based on (2,2) binary VCS. It generates an encrypted receipt to every voter which allows them to verify the election outcome even if all election computers and records were compromised. At the polling station, the voter will receive a double-layer receipt that prints his/her voting decision. The voter will be asked to give one of the layers to the poll worker who will destroy it immediately with a paper shredder. The remaining one layer will now become unreadable.
- 2) **Encrypting Financial Documents:** The VCS principle can also be applied in transmitting confidential financial documents over Internet. Visual Cryptography is an example of this type of system being proposed by Hawkes et al [4]. Visual Cryptography can encode the original drawing document with a specified (k, n) VCS, then send each of the encoded n shares separately through Emails or FAX to the recipient. The decoding only requires bitwise OR operation on all shares in the specified directory, and needs no extra effort of cryptographic computation. Any malicious attacker who intercepts only m of n shares where $m < k$ will not be able to gain any information about the financial document.

IV Design preliminaries.

A) Visual Cryptography

The basic model of Visual Cryptography was introduced by Naor and Shamir [2] in 1994 accepts binary image $I(x, y)$ as secret image, which is divided into 'n' number of shares. Each pixel of image $I(x, y)$ is represented by 'm' black and white sub pixels in each of the 'n' shared images. It is impossible to get any information about the secret images from individual shares. The other advantage of VCS is that, unlike other cryptography techniques, the secret image recovery does not need difficult computations. i.e the secret information can easily be recovered with enough number of shares through stacking process (human vision) instead of special software or hardware devices.

Naor and Shamir proposed a k out of n scheme and assumed that the image or message is a collection of binary data 0 and 1 displayed as black and white pixels. According to their algorithm, the secret image is turned into n shares and the secret is revealed if any k of them are stacked together. So the image remains hidden if less than k shares are stacked.

The main parameters of Visual Cryptography includes image contrast and the number of sub pixels of recovered image. The contrast of a image is a relative difference between the original and retrieved image. As increased in number of sub pixels while creating the shares, size of the share (column) also increases. This increment in turn affects the quality of retrieved image. Some researchers have focused on contrast degradation and introduced methods to improve the contrast of the reconstructed secret image[5].

2-out-of-2 Secret Image Sharing Scheme

The basic idea of 2-out-of-2 Visual Cryptography sharing scheme is illustrated in Table2. Every secret pixel of the original binary image is converted into four sub pixel of two share images and recovered by simple stacking process. This is equivalent to using the logical OR operation between the shares. As illustrated in Table2 four sub pixels are generated from a pixel of the secret image in a way that two sub pixels are white and two sub pixels are black. The black or white pixel selection is based on random selection. In this paper 8 stages LFSR (Linear Feed Back Shift Register) is considered to generate the random bit and design is discussed in subsection B.

Pseudo code of share generation scheme is given below. Based on pixel bit and random sequence bit, share1 and share2 will be generated. The size of obtained share will be twice the as many columns in the original image.

Table-2 (2, 2) Visual Cryptography Sharing Scheme

```

For i=1 to Size of the image
  If (pixel == 1)
    If (reandom_bit==1)
      Share1=[1 0]
      Share2=[1 0]
    Else
      Share1=[0 1]
      Share2=[0 1]
  Else
    If (reandom_bit==1)
      Share1=[1 0]
      Share2=[0 1]
    Else
      Share1=[0 1]
      Share2=[1 0]
  END
End of the loop
    
```

	White	Black
Pixel		
Prob.	50% 50%	50% 50%
Share 1		
Share 2		
Stack share 1 & 2		

B. LFSR Design

A Linear Feedback Shift Register(LFSR) is a sequential shift register with combinational logic that makes it to pseudo-randomly cycle through a sequence of binary values. In this paper to generate a random bit an 8 bit LFSR is considered with a polynomial $x^8+x^6+x^5+x^4+1$ and corresponding design is given in the figure 2a. Feedback around an shift register comes from a selection of points (D) in the register chain and constitutes XORing these taps to provide tap(s) back into the register as shown in the figure2a. The 8 bit LFSR case D_0, D_4, D_5, D_6 are XORed and feeding back to MSB bit(D_7) and this bit is used as a sequence bit in future. The 8-bit sequence will repeat every after 255 cycle(1 cycle= one shift) and all generated sequence are purely based on the initial 8-bit data present in the register.

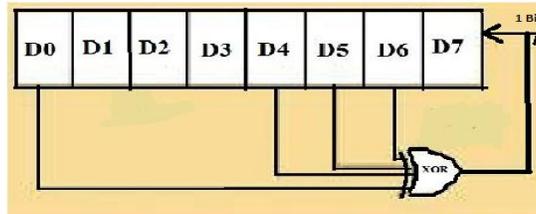


Figure 2a 8 Bit Linear Feed Back Shift Register

V Proposed Scheme

In the Proposed scheme we used (2, 2) Visual Cryptography Scheme. The whole design is divided into two process

- 1) Encryption(Creating shares)
- 2) Decryption(Human Visual System)

Encryption process

Encryption process is partitioned into two stages and shown in figure 3b.

• First stage

An original binary image of size 100X100 has been considered in our design where each pixel is either 0 or 1.

A (2, 2) Visual Sharing technique (as discussed in section IV) is applied on original binary image. From this process two binary share images will be generated.

• Second stage

In the second stage four share images are generated from the two share images obtained from first stage using (2,2) VCS, this is called chain share of order 2.

Further to enhance the security, share encryption is performed on four shares obtained from previous step using 8-bit different LFSR sequence (as discussed in section IV) by XORing the share bit with LFSR bit and it is repeated for whole binary share image. Hence, no two share should generate partial image.

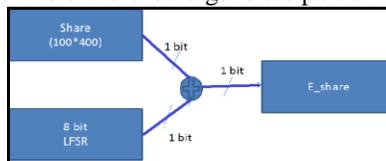


Figure 3a Encryption of shares

Decryption process

Decryption process is reverse process of encryption. It consists of two stages and pictorial representation of decryption process is given in figure3c.

First Stage

To decrypt the four shares from encrypted shares same LFSR sequence must be used which was used in earlier encryption process and bitwise XOR operation performed between encrypted share and LFSR bit sequence. Otherwise original image will not be recovered in future.

From four decrypted share of the previous step, two partial shares were obtained by performing bitwise logical ORing between share 1,2 and share 3,4.

Second Stage

Bitwise logical OR operation performed on two obtained share images from previous stage to get back the retrieved image. The obtained image is identical to original image.

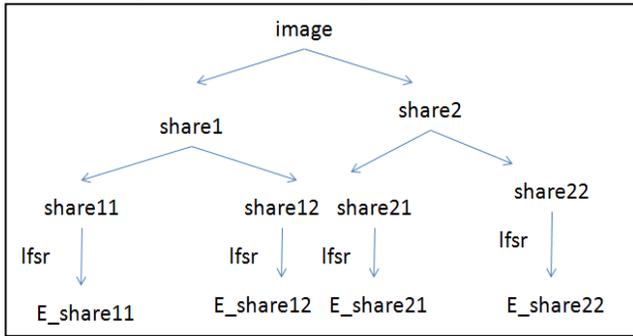


Figure 3b Encryption Process with two stages

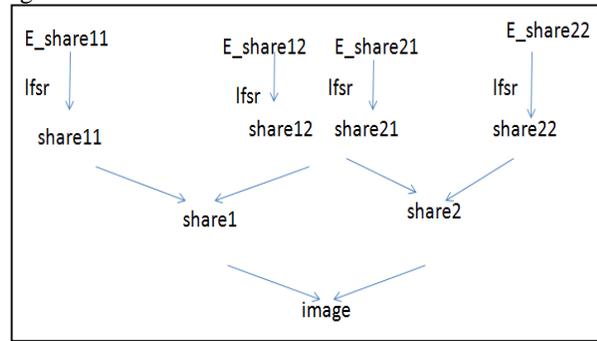


Figure 3c Decryption Process with two stages

VI Results

Matlab 7.1 Software tool is used to analyze the result. To study the performance of proposed scheme three different binary images of size 100X100 were used and shown in figure 5a, 5b, 5c. The stage by stage result of proposed scheme is given in table 3a and 3b.



Figure 5a



Figure 5b

VINAY
ECE
NCET

Figure 5c

Table 3a shows the Encryption process of the original binary image. In the table3a Figure4a shows three different original binary images, Figure4b and 4c are shares generated in the 1st stage. The result shows no residual information of original image perceptually observed among the two shares. In the second stage four different shares are generated and is given in figure 4d to 4g. In this four shares stacking process of any two shares gives residue of original information. So to provide additional security share encryption is performed on all the four shares as discussed in section V. Figure 4h to 4k is scrambled shares using 8 stage LFSR sequence.

Table 3a Encryption process

Original Image 100x100 Fig4a	Figure 5a	Figure 5b	VINAY ECE NCET
Share1 100x400 Fig4b	[Image]	[Image]	[Image]
Share2 100x400 Fig4c	[Image]	[Image]	[Image]
Share11 100x400 Fig4d	[Image]	[Image]	[Image]
Share12 100x400 Fig4e	[Image]	[Image]	[Image]
Share21 100x400 Fig4f	[Image]	[Image]	[Image]
Share22 100x400 Fig4g	[Image]	[Image]	[Image]
E_share11 100x400 Fig4h	[Image]	[Image]	[Image]
E_share12 100x400 Fig4i	[Image]	[Image]	[Image]
E_share21 100x400 Fig4j	[Image]	[Image]	[Image]
E_share22 100x400 Fig4k	[Image]	[Image]	[Image]

Table 3b gives decryption process and it is reverse process of encryption. In the table3b the figure 4l to 4o is same scrambled share image. The same LFSR initial sequence or key need to be used to get back the four shares. The obtained shares are shown in the figure 4p to 4s. It is observed that no residual information is visually seen. Figure 4t and 4u are the shares obtained from BIT ORing the shares shown in Fig 4p and 4q and Fig 4r and 4s respectively. Reconstructed image given in figure 4v obtained by BIT ORing the previous shares i.e.,Fig4t and 4u. From figure4v image can be easily identified but

contrast is poor. In order to obtain the efficient reconstructed image Bitwise XOR is applied between the shares(Fig4t and 4u) and obtained image shown in figure4w.The figure4w image is identical to original image except the size.

Table 3b Decryption Process

V Conclusion

Using the proposed design original image successfully encoded into 4 share images. In This scheme two stages were used it's very difficult to access the original information by unauthorized receiver. In this paper we used only order 2(two stages) because if we increase the order even though we can generate many number of shares, but quality of the reconstructed image will be poor (more number of columns) and needs more storage space for shares.

Difficulties for image reconstruction by unauthorized receiver

1. **Order and orientation of chain share:** If the order of the chain share is unknown and orientation is mismatched we will not get the original image
2. **LFSR sequence:** We have to use same LFSR sequence for encryption as well as decryption to get back the original image
3. **Number of shares:** All the four shares compulsorily necessary to identify the original binary image.

VI Reference

- [1] W.-G. Tzeng and C.-M. Hu, "Anew approach for visual cryptography," *Designs, Codes, Cryptog.* vol. 27, no. 3, pp. 207–227, 2002.
- [2] M. Naor and A. Shamir, "Visual cryptography", in *Eurocrypt'94 Proceeding*, LNCS, vol.950, Spring-Verlag, pp.1-12, 1995.
- [3] D Chaum, Secret-ballot receipts: True voter-verifiable elections, *IEEE Security and Privacy*, 2004,38-47.
- [4] W. Hawkes, A. Yasinsac, C. Cline, An Application of Visual Cryptography to Financial Documents, technical report TR001001, Florida State University (2000).
- [5] M. Naor and A. Shamir "Visual cryptography: Improving the contrast via the cover base" *IACR Eprint archive*, 1996.
- [6] C.-N. Yang and C.-S. Laih, "Some new types of visual secret sharing schemes," in *Proc. Nat. Computer Symp.*, 1999, vol. 3, pp. 260–268.
- [7] M. Naor and B. Pinkas, "Visual authentication and identification," in *Proc. Advances in Cryptology*, 1997, vol. 1294, LNCS, pp. 322–336
- [8] G. Ateniese, C. Blundo, A. De Santis, and D. R. Stinson, "Visual cryptography for general access structures, *Information and computation* 129 (1996), 86-106
- [9] William Stallings," *Cryptography and Network Security*" 4th Edition, Pearson Education India.

AUTHORS



ROHITH S received B.E. degree in Electronics and Communication engineering in 2006 and M.Tech degree in VLSI Design and Embedded systems in 2008 from Visvesvaraya Technological University, Karnataka. He is currently working as a Lecturer at Nagarjuna College of Engineering and Technology, Bangalore,Karnataka-562110. He has total teaching experience of 4years.His main area of interest includes Digital Watermarking, Steganography, Error Control Coding, Cryptography and VLSI Design.



VINAY G pursuing a Bachelor of Engineering degree at Nagarjuna College of Engineering and Technology, affiliated to Visvesvaraya Technological University, Karnataka. He is involved in hobby projects and also a team member of student satellite (STUDSAT-2) project under the guidance of ISRO. His main area of research includes Embedded Systems, Image Processing, Visual Cryptography.

Numerical Investigation of Secondary Flow In An Axial Flow Compressor Cascade

¹T. Suthakar, ²Akash Dhurandhar

¹Associate Professor, ²M.Tech. Scholar,
Department of Mechanical Engineering
National Institute of Technology, Tiruchirappalli – 620015
Tamil Nadu, India

Abstract

A numerical study has been conducted for steady, three dimensional viscous flow fields in an axial flow compressor cascade focusing basically on the secondary flow formation. The study is made on the basis of static pressure distribution along the blades with different inflow angles. The k-ε model was used and a finite volume method was employed to solve the flow governing equations. The present work consists of a compressor cascade analysis with chord length of 5 units and pitch of 0.55 times chord length. For this analysis ideal gas is taken as the fluid with atmospheric pressure and a temperature of 300 K, $\alpha_i = 26.5^\circ, 35^\circ$ and 45° , $Ma = 0.66$, $Re = 0.6 \times 10^6$. Flow analysis is carried out using commercial CFD codes (Fluent 6.3.26). The computational results support the secondary flow occurrence in a compressor cascade for specific conditions.

Key words: secondary flow, turbulence intensity, axial compressor cascade, cascade losses

Nomenclature

Re	Reynolds number
Ma	Mach number
S	Cascade pitch
C	Chord
h	Blade height
Cp	Pressure coefficient

Greek letters

α_i	inlet air angle
α_e	exit air angle
θ	stagger angle

Introduction

Accurate prediction of turbo machine cascade flow and pressure losses remains a challenging task despite a lot of work in this area [4]. Flow in a cascade of turbo machinery suffers various types of losses like annular loss, profile loss and last but not the least secondary losses. The term secondary flows refers to the three- dimensional vortical flow structures that develop in blade passages due to high turning of the flow and non uniform inlet total pressure profiles. [2]. Flow which is transverse to the primary flow direction is termed as secondary flow [2]. Secondary flows are important source of losses in turbo machines, especially in cascades with short height blading and high flow turning. The secondary flows originate from specifically developing end wall boundary layers and are associated with the presence of longitudinal vortices with a dominant stream wise component of the vorticity. They are driven by transverse static pressure gradients and mass forces acting on fluid elements in curvilinear motion through the blade passage. The secondary flows also modify the shape of end wall boundary layers from which they originate [9].

In compressors the classical secondary flow is not so strong because the blade turning is low. However, the end-wall boundary layers are thick and are prone to separate in the corner between the suction surface and the end-wall. This separation is greatly influenced by both secondary flows, which tend to exacerbate it, and tip leakage flows, which tend to prevent separation [6]. The main type of secondary flow is the induced recirculating flow, which leads to the formation of a passage vortex. The source of the induced recirculating flow is the cross flow in the end wall boundary layer that forms as a result of force equilibrium in curvilinear motion.

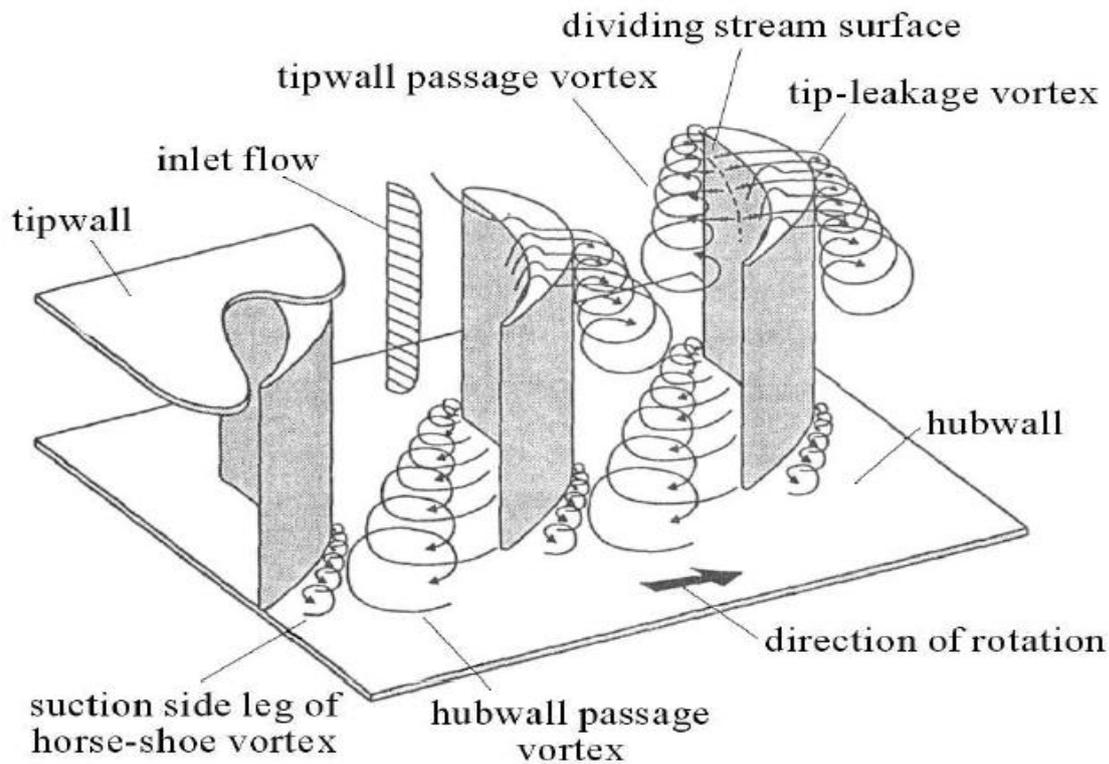


Figure 1: Various types of secondary flows in a cascade of turbo machine.

Mechanisms of secondary flow

Within a blade passage in a cascade, a deflection of the low momentum fluid equal to the mainstream deflection would not be sufficient to obtain the balance between the pressure gradients and the centrifugal forces. As a result the low momentum fluid turns more than the mainstream flow and this deviation from a uniformly deflected stream called secondary flow [11].

The term secondary flows refers to the three-dimensional vortical flow structures that develop in blade passages due to high turning of the flow and non-uniform inlet total pressure profiles. The boundary layer flow along the end wall contains span wise velocity gradients. When the boundary layer flow is turned, transverse velocity components are introduced. These secondary flows, created at the end wall and blade junction, extract energy from the fluid which would otherwise be used to rotate the blades or produce thrust [2].

The secondary flows originate from specifically developing end wall boundary layers and are associated with the presence of longitudinal vortices with a dominant stream wise component of the vorticity. They are driven by transverse static pressure gradients and mass forces acting on fluid elements in curvilinear motion through the blade-to-blade passage [9].

Methodology

Table 1 Details of Investigation

1	Aspect ratio, c/h	1
2	Chord length, c	5 units
3	Pitch of the cascade	0.55 c units
4	Inflow angle (in degree)	26.5 , 35 and 45
5	Mach No.	0.66
6	Reynolds No.	0.6×10^6
7	Viscous model used	k-epsilon
8	Software used	Gambit and Fluent
9	Type of analysis	3D

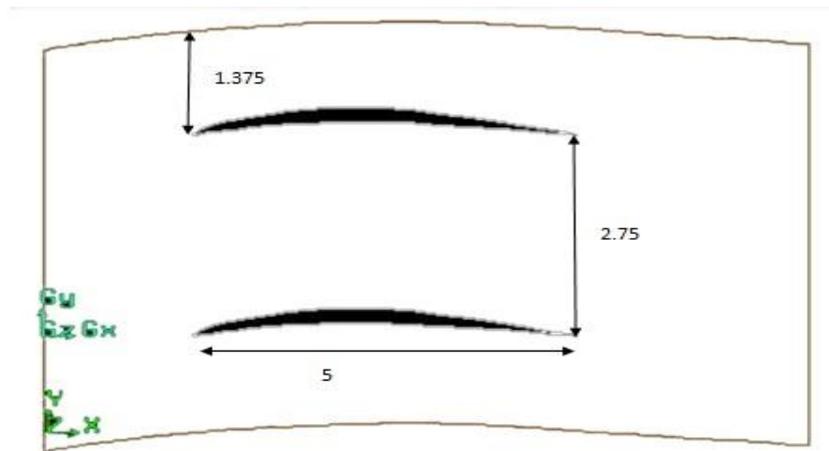


Figure 2: Dimensions of cascade

Using these dimensions three dimensional models are drawn out from this arrangement.

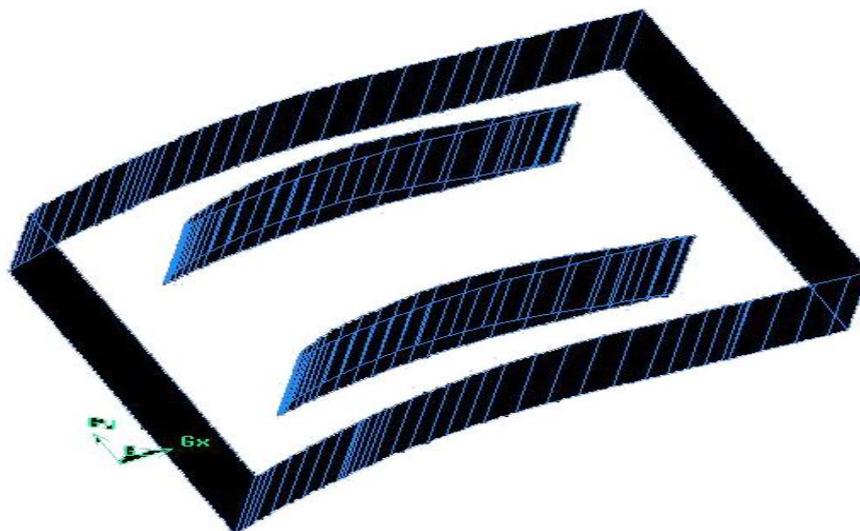


Figure 3: Gambit model of cascade blades with fluid volume (model 1)

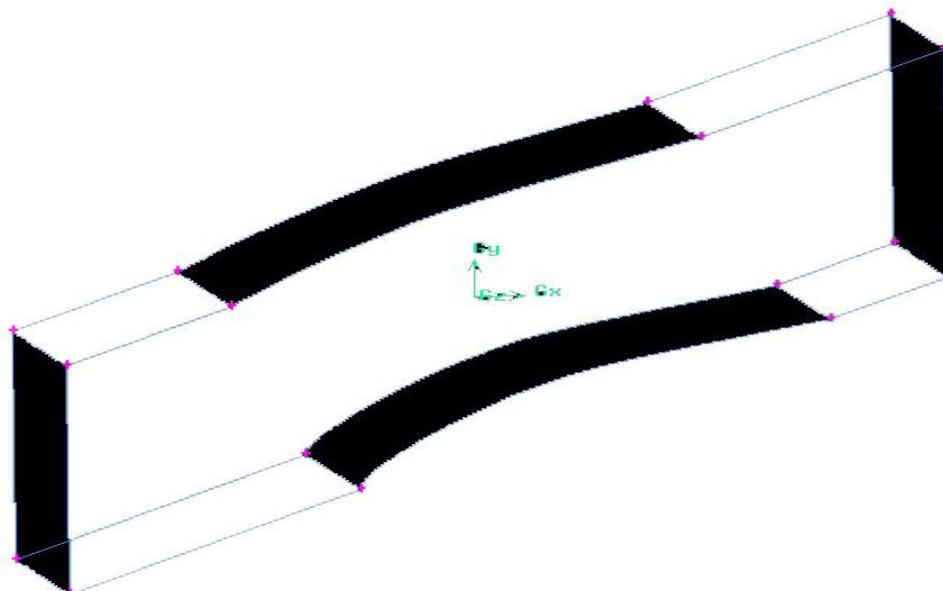


Figure 4: Volume of main flow under consideration (model 2)

This flow is cut out of the main flow stream under consideration from model 1. Since it is hard to numerically calculate for such a bid domain (model 1) in gambit and fluent both.

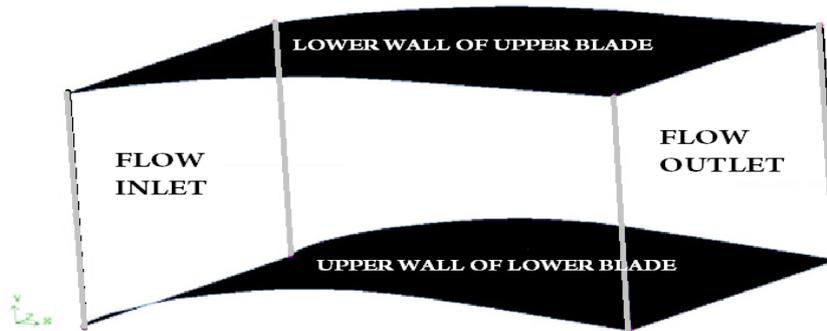


Figure 5: Flow volume between the blades only (model 3)

This figure is again the area between the blades only where we have to check the formation of the secondary flow vortex.

Meshing

Different types of mesh sizes, types, elements are available for edge, face and volume meshing. It must be the optimum size of mesh for getting the accurate results. Gambit software provides a wide range of meshing options for simple as well as complex geometries.

Table 2 Meshing details

1	Elements used for face meshing	Quadrilateral
2	Elements used for volume meshing	Tet/Hyd. & Quad/Tri
3	Interval count	According to length of edge
4	Side	Single sided
5	Successive ratio	1
6	Nodes	33201 (for model 3)

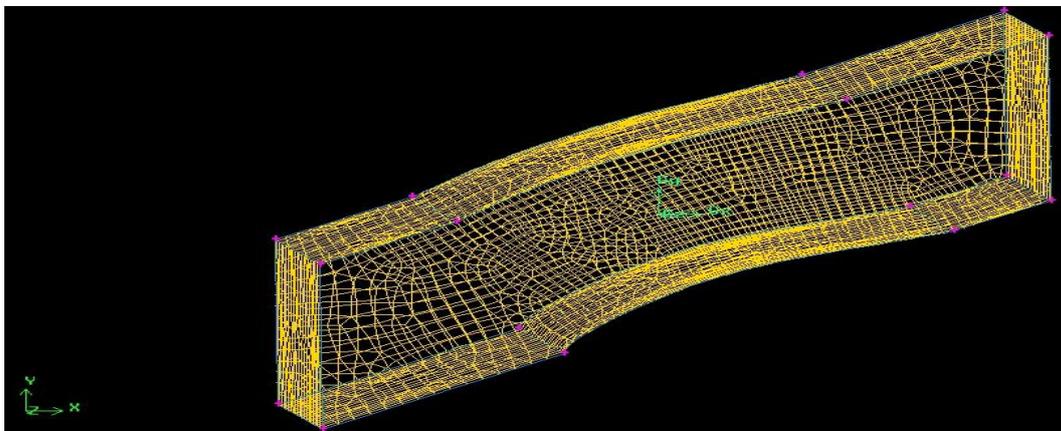


Figure 6: Meshing of model 2

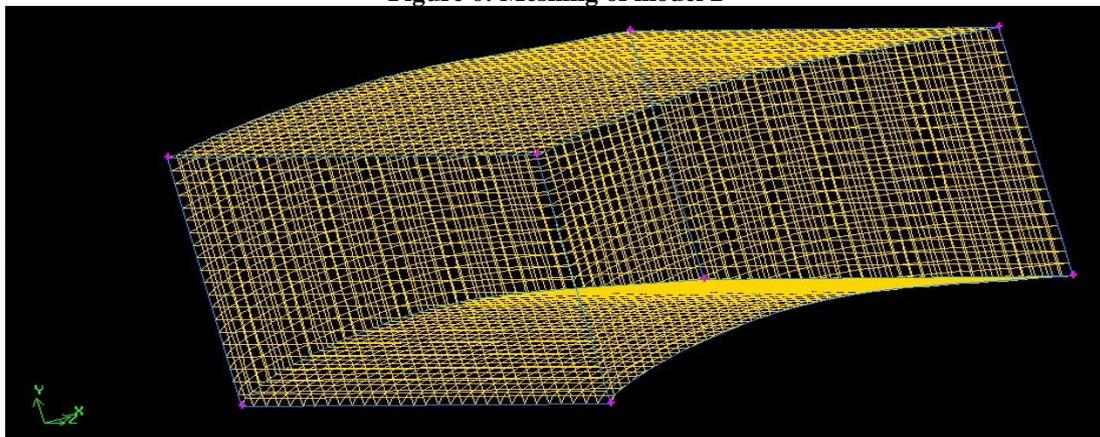


Figure 7: Meshing of model 3

Boundary conditions

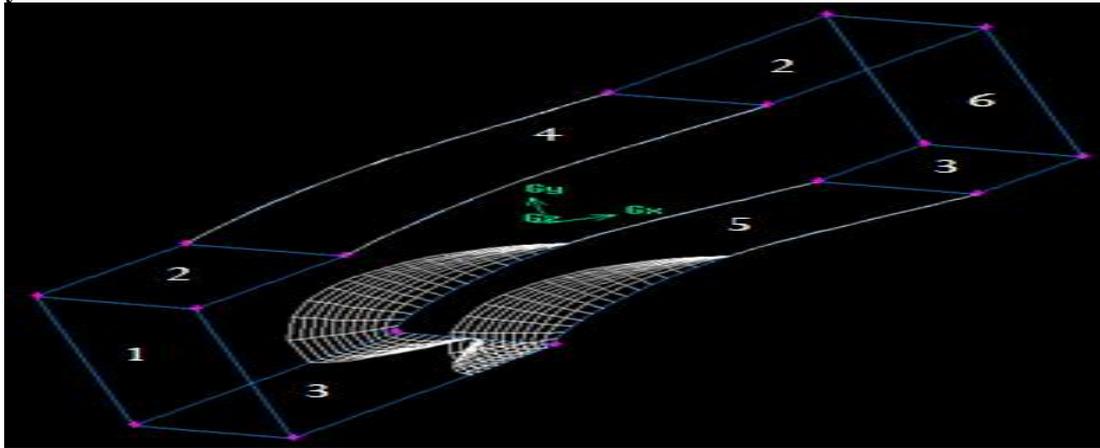


Figure 8: Different faces of model 2

Table 3 Boundary conditions

Face	Name	Boundary Condition
1	Inlet	Velocity
2,3	Side walls	Periodic (Wall)
4	Suction side	Wall
5	Pressure side	Wall
6	Outlet	Pressure

Turbulence modelling

Turbulence Modelling, as turbulence affects the flow behaviour of the turbo machines, its consideration is critical in compressor blade analysis. The turbulence level seems to play a role on both the mixing within, and between the structures. Compressor blade model use k-epsilon turbulence model with realizable option. The k-epsilon model is mainly used for confined flows which usually occurred in turbo machines blade cascades and for high Reynolds number flows such as this problem. The realizable K-epsilon model gives more accurate results than any other turbulence model for the annular cascade. It also correctly predicts the boundary layer flows near the tip wall region in the cascade.

Table 4 Solution Controls for turbulence Modelling

Under relaxation factors	Values
Turbulent Kinetic Energy	0.8
Turbulent Dissipation Energy	0.8
Turbulent Viscosity	1
Discretization	Order
Flow	Second order upwind
Turbulent Kinetic Energy	Second order upwind
Turbulent Dissipation Rate	Second order upwind

Results and discussion

In order to recognize the secondary flow formation between the two blades of compressor cascade, a numerical simulation using CFD method is performed for axial flow compressor cascade. Input conditions are provided in the form of boundary conditions for velocity inlet at inlet face the velocity was set to 250 m/s that is approx. equal to $Ma = 0.66$. The α_i is varied as 26.5° (for model 2 only), 35° and 45° (for model 3). Their intensity is shown to be strongly dependent on the inlet flow angle.

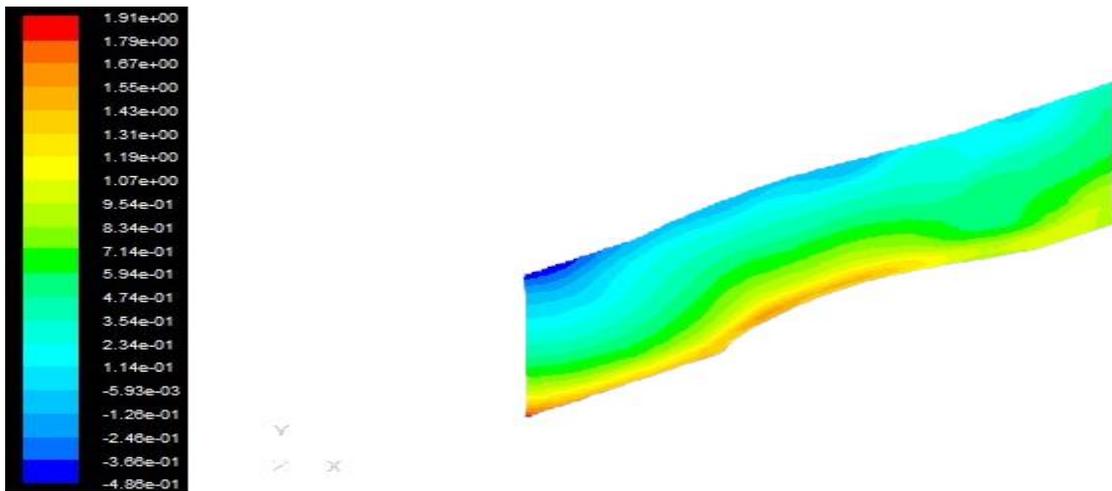


Figure 9: Contours of total pressure for model 2.

In figure 9 we can see the contours of total pressure which is showing a slight increase in the inlet pressure. This is the condition for inclination of 26.5 degree angle. But in this geometry we cannot see the flow behaviour near the blades. So there is a need of geometry of third type.

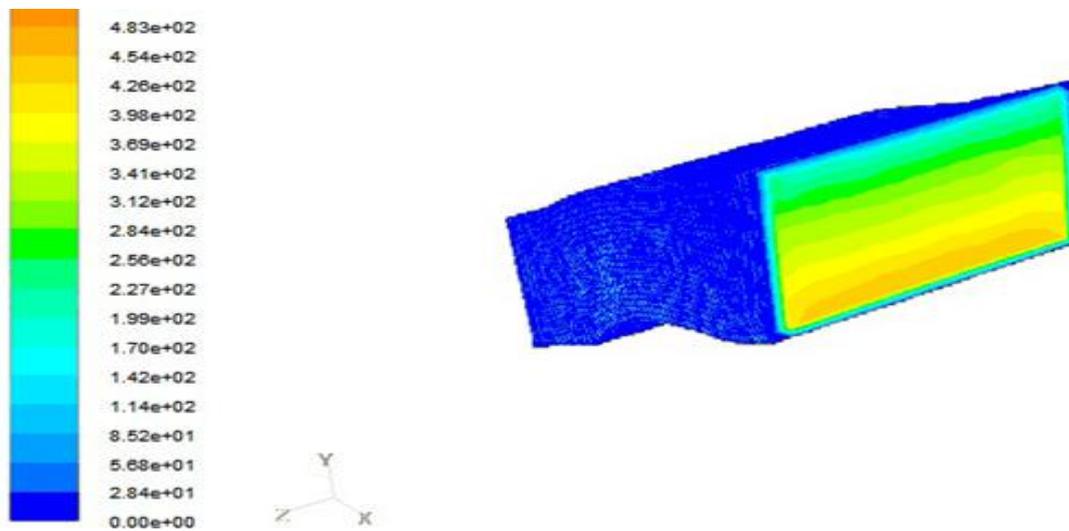


Figure 10: Contours of velocity magnitude for model 2.

Figure 10 tells about the contours of velocity magnitude at the outlet of the flow of model 2. The dark area shows the stagnation of the flow particles next to the wall.

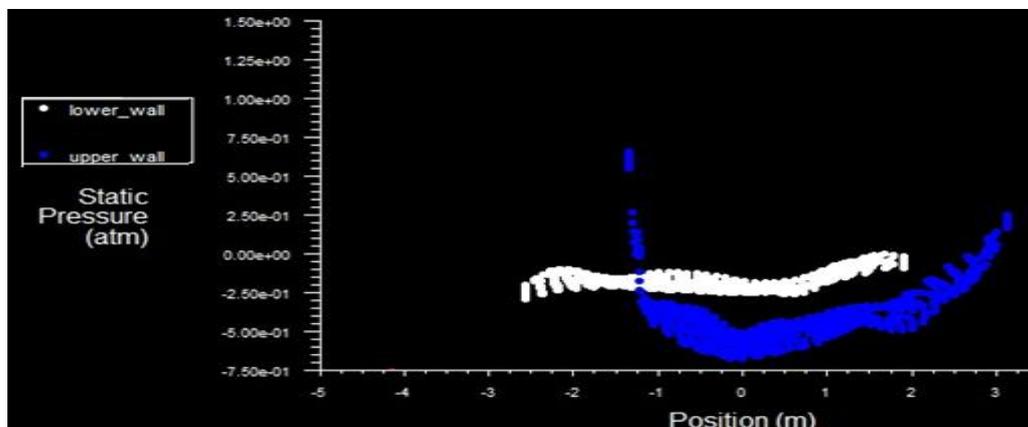


Figure 11: Static pressure for pressure side and suction side for model 2

This figure shows the gauge pressure variation along the length of the blades both in the lower wall that is pressure side and upper wall that is suction side. A little increase in pressure can be seen in both the walls here.

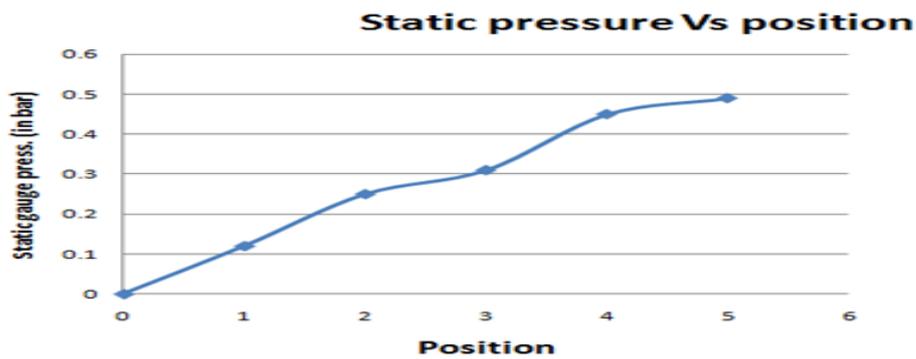


Figure 12: Variation of static pressure with the position in model 2 at 0° incidence

This plot gives the variation of static pressure for the zero degree incident angle. In this we can see that the compressor cascade is giving the rise in the pressure from leading edge to trailing edge.



Figure 13: Variation of Ma with the position in model 2 at 0° incidence

The inlet Mach number taken here is 0.66 but due to the decrease in velocity throughout the cascade flow there is a decrease in Ma can be seen. The outlet Mach number comes around 0.52. All the values taken for plotting these graphs are averaged and approximate.

Model 3 represents the volume of fluid flow between the two blades only. This model consists of 33201 nodes and is not showing behaviour of compressor cascades at high incidence angles since with high incidence angles they act more like a turbine cascade. Here number of iterations done is 500.

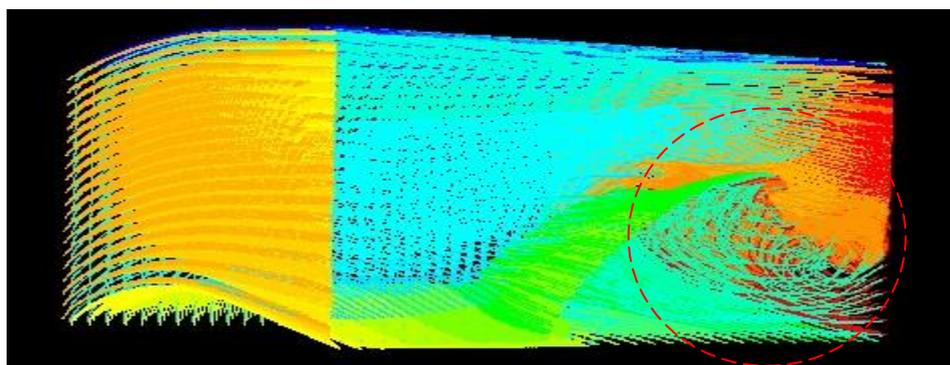


Figure 14: Path lines of flow in model 3

Here the path lines of the fluid flow inside the volume taken under consideration at 35°. The bottom side encircled of upper wall of first blade is showing a vortex kind of structure coming out from the cascade which confirms the formation of secondary flow.

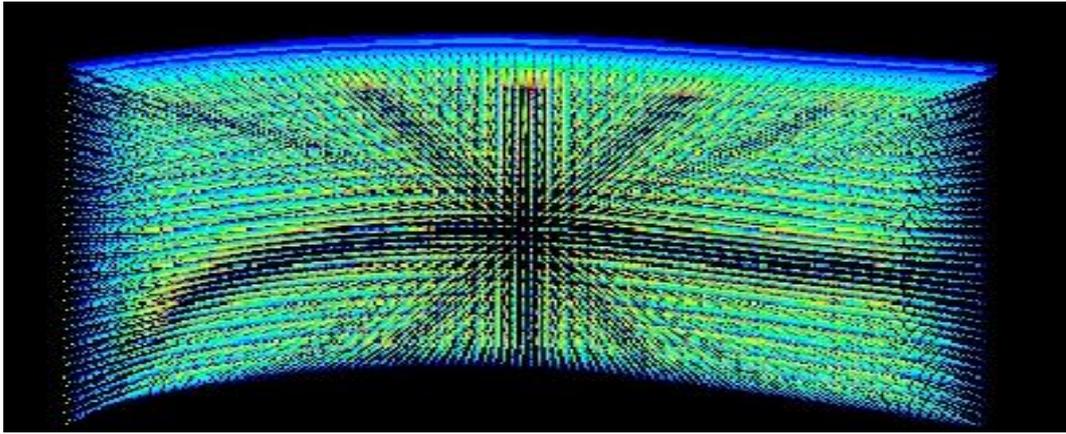


Figure 15: Path lines showing only lower wall and interior flow

This figure is also showing the path lines with respect to the lower wall of second or upper blade only. In that also a main vortex stream can be seen from inlet to outlet of the internal flow through the volume that confirms the formation of passage vortex.

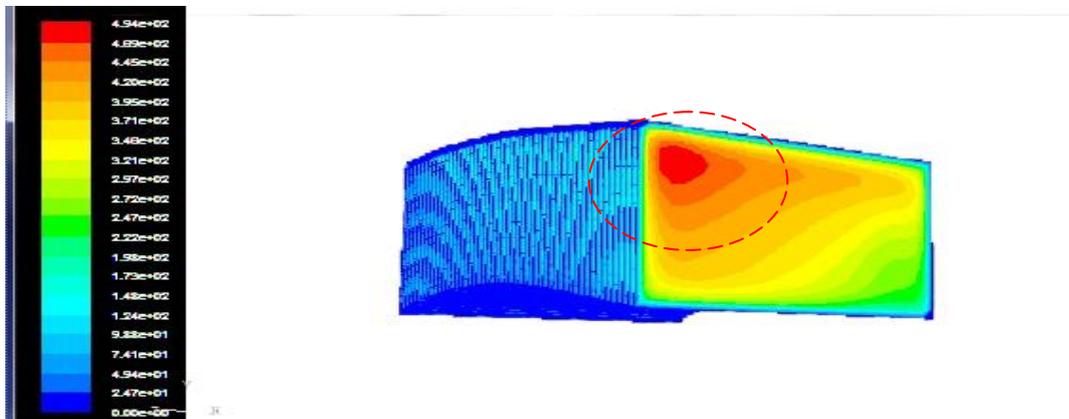


Figure 16: Contours of velocity magnitude (m/s) at outlet for model 3

The above figure shows the counters of velocity at outlet 45° inlet of air. The outside face can be seen here that shows encircled area were the velocity of the air is high. This is confirming the vortex kind of flow taking place at that area in between the blades.

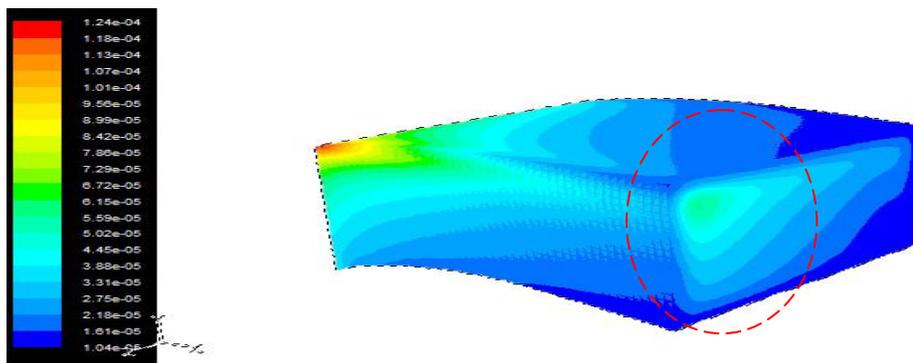


Figure 17: Contours of total pressure (atm) for model 3

This figure also shows the vortex of high pressure region in the outlet face at 45° . So vortex kind of flow exists inside the blade flow volume. The cascade gives us a small pressure rise at most of the place approximately about 0.5 times. The graphs shown below are showing the change in properties of flow while moving from leading edge to trailing edge.

Conclusions

1. By this analysis we can conclude that in axial flow compressors the threat of secondary loss is less because it suffers low inlet angle where secondary flow is hardly formed.
2. In compressors secondary flow effects are less because of less turning or bending while passing through blade passage.
3. Secondary flow effects can be seen in the axial flow compressor with high incidence angles like 35 or 45 degrees.

Appendix (Blade Coordinates)

Suction	Suction	Pressure	Pressure	Mean
5.001	0.000	5.001	0.000	0.000
4.902	0.020	4.901	0.009	0.015
4.703	0.061	4.699	0.028	0.044
4.404	0.119	4.497	0.046	0.083
4.257	0.148	4.245	0.069	0.108
4.010	0.202	3.992	0.092	0.147
3.612	0.285	3.590	0.132	0.208
3.213	0.359	3.189	0.169	0.264
2.811	0.420	2.790	0.198	0.309
2.404	0.456	2.397	0.213	0.335
1.997	0.460	2.004	0.210	0.335
1.744	0.448	1.575	0.201	0.325
1.491	0.427	1.510	0.189	0.308
1.238	0.396	1.263	0.172	0.284
0.986	0.356	1.015	0.150	0.254
0.734	0.304	0.766	0.122	0.213
0.483	0.239	0.517	0.087	0.163
0.383	0.209	0.417	0.071	0.140
0.283	0.174	0.317	0.053	0.114
0.184	0.134	0.216	0.034	0.084
0.086	0.084	0.114	0.013	0.049
0.038	0.053	0.062	0.002	0.028
0.016	0.033	0.340	-0.002	0.015
0.000	0.011	0.010	-0.004	0.004
0.000	0.000	0.000	0.000	0.000

References

1. **A. Hergt, R. Meyer and K. Engel**, Experimental and numerical investigation of secondary flow on compressor blades
2. **Krishna Nandan Kumar and M. Govardhan**, 2010, Secondary Flow Loss Reduction in a Turbine Cascade with a Linearly Varied Height Stream wise End wall Fence. *Hindawi Publishing Corporation, International Journal of Rotating machinery, Volume 2011, Article ID 352819*
3. **Hyon Koon Myong & Seung Yong Yang**, the numerical investigation for getting flow characteristics at blade passage tip clearance in a linear cascade of high performance turbine blade'. *KSME International Journal Vol.17 no. 4 page 606-616,2003*

4. **Nikolay Ivanov, Valery Goriatchev, Evgueni Smirnov and Vladimir RIS**, ‘CFD analysis of secondary flow and pressure losses in a NASA Transonic turbine cascade.’ *Conference of modelling fluid flow, Hungary September 2003*.
5. **J. Dunham** ‘A review of cascade data on secondary losses in turbines’.
6. **J.H. Horlock, J.D. Denton** ‘A review of some early design practise using Computational fluid dynamics and a current perspective.’ *Cambridge University Engineering Department*.
7. **A. R. Howell, M.A.** ‘Flow Dynamics of Axial Compressors.’
8. **Johan Hjarne, Valery Chernoray, Jonas Larsson, Lennart Lofdahl** ‘An Experimental Investigation of secondary flows and loss development downstream of a highly loaded low pressure turbine outlet guide vane cascade.’ *Asme Turbo Expo 2006*.
9. **Piotr Lampart**, ‘Investigation of end wall flows and losses in axial turbines. Part i. formation of end wall flows and losses.’ *Journal of theoretical and applied mechanic 47, 2, pp. 321-342, Warsaw 2009*
10. **Anderson, J. D. Jr.**, 1995: *Computational Fluid Dynamics, The Basics With Applications, McGraw-Hill Inc., Singapore*.
11. **B. Lakshminarayan & J.H. Horlock** , Text book of turbomachinery
12. **Yahya S.M.** 1998, *Turbines, Compressors and Fans*, Tata Mc-Graw-hill publishing company limited.
13. **Dixon S.L.** 1998, *Fluid mechanics and thermodynamics of turbo machines*, Reed Educational and Professional Publishing Ltd.
14. **Ghoshdastidar P**, *Computer simulation of flow and heat transfer* , TMH New Delhi publication
15. **Fluent**, (1998) FLUENT 6 User’s Manual, *Fluent. Inc*
16. **Gambit** user’s manuals.

A MODIFIED MARKER CONTROLLED WATERSHED ALGORITHM WITH LINEAR CONVOLUTION FOR MEDICAL IMAGE SEGMENTATION

Brijesh Shah¹, Jigar Modh² and Satish Shah³

¹ Associate Professor, C S. Patel Institute of Technology, CHARUSAT, Changa, Gujarat, India

² V.T.Patel Departments of Electronics and Communication Engineering,

³ Professor, M. S. University, Baroda, Gujarat, India

ABSTRACT

Image segmentation plays an important role in the medical imaging. Image segmentation is the process to divide the image in the different region which has similar attributes. In this paper we proposed marker controlled watershed segmentation algorithm which works on gradient magnitude and use linear operation. The Linear convolution based image reconstruction algorithm used here is enables us to obtain better than general algorithm and morphological image reconstruction. This also reduces the overall time of algorithm than other marker controlled watershed algorithm. This method can be applied to noisy and degraded images. The use of markers can avoid the problem of over segmentation. The algorithm works on color as well as gray-scale images. The segmentations of X-ray image, MR image and also Mammographic images for breast cancer detection is shown in this paper. This algorithm can also be applied to high resolution satellite images.

KEYWORDS: SEGMENTATION, MARKER CONTROLLED WATERSHED ALGORITHM, MEDICAL IMAGES, LINEAR CONVOLUTION

1. INTRODUCTION

Watershed segmentation [1] is the most powerful segmentation technique in the area of digital image processing. Watershed transformation has widely adopted technique due to its many advantages, including simplicity, speed and complete division of image. Greyscale mathematical morphology of the watershed transform for the image segmentation is first proposed by Digabel and Lantuejoul (1977) [2] and later improved by Li et. al. (2003) [3]. watershed transformation can be called as region based segmentation technique. The proposed algorithm presented here in this paper is more simplified than morphological based image reconstruction. The initiative description of this transform is quite simple: if we consider image as a topographic view, where the height of each point is directly related to its gray level. Now consider the rain is gradually falling on terrain then the watersheds are the lines separate the catchment basins that form. The watershed transform is computed on the gradient of original image. The catchment basin boundaries are located at high gradient point. The watershed transformation is widely used in many fields of image processing like medical image segmentation.

Most important drawback of watershed segmentation is the over segmentation. When the watershed transform is applied to the catchment basins from the gradient of the image, the result of the watershed contains a number of small regions. It makes a resulted segmentation hardly useful. So to use the markers in the image will reduce number of minima in the image. Thus marker controlled segmentation is used for watershed segmentation. [4]

In this paper we explain the marker controlled watershed segmentation for medical images like X-Ray and MR images and Mammographic images. The method overview is explained in section II. The proposed algorithm for this method is explained in section III. Section IV shows the Experimental results of the method and comparison with this algorithm with other algorithm. In this section we use X-Ray Image and MR image and Mammographic Images to see the results. In the end Section V dedicated to conclusion.

2. METHODOLOGY

In watershed segmentation there two basic approaches

Rain fall approach

In this approach local minima is found in all through the image and each local minima assigned an exclusive tag. An intangible water drop placed at each untagged pixel. And it assumes tag value.

Flooding approach

In flooding approach intangible pixel holes are pierced at each local minimum. The water enters in the holes and filling each catchment basin. If the basin is about to overflow, a dam is built on neighbouring ridge line to the height of high altitude ridge point. These dam borders can be called watershed ridge lines. [4]

2.1 Marker –controlled watershed transform

Applying watershed transform directly to the gradient image will results over segmentation. This is due to the irreverent minima, or noise patches or other irregularities. The concept of marker can be used to solve this over segmentation problem whose goal is to detect the presence of homogeneous regions from the image by the set of linear filtering operation. They spatially locate object and background ensuring to keep object interior as a whole [1].The marker controlled watershed algorithm proposed in this paper is robust and flexible method for image segmentation. Where the objects are closed contours and boundaries are ridges. The marker is the binary image used for watershed segmentation is either single marker points or larger marker region. Each marker is related to the specific watershed region. Thus number

of the final number of watershed regions are equals to the number of markers. After segmentation the boundaries of watershed region are arranged to the desired ridges, so each object is separated from its neighbour.

2.2 Linear filtering Operation

Filtering operations reduces or increases the rate of change that occurs in the intensity transition within the image. Areas where there is sudden or rapid change in the intensity appear as hard edge in image, similarly where gradual changes produce soft edges. In the filtering it acts to detect or modify the rate of change at these edges. Filtering techniques are divided into two categories :Convolution filters (Linear filters) and Nonconvolution filters(Nonlinear filters)[14].here in this algorithm we most use convolution filters for image neighbourhoods by multiplying the value within neighbourhood by a matrix of filtering coefficient, however for creating markers we use morphological operation which is based non linear filters.

2.2.1 Convolution

Convolution is the most common to many image processing operators. Convolution is based on simple mathematical operation. Convolution is the way of multiplying two arrays of different sizes to produce a third array of numbers. In the Image processing Convolution is used to implement operators which are the Linear combination of certain input pixel value of image and produce the output pixel value [13]. Convolution is based on class of algorithm which is called spatial filters. This filters uses wide variety of masks or kernels, to calculate different results, depending on desired function.2-D convolution is most important to modern image processing .The basic idea is to scan a window of some finite size over an image. The output pixel value is the weighted sum of input pixels within the window where the weights are the values of the filter assigned to every pixel of the window. The window with its weights is called convolution mask or it can also be called kernels. The mathematical equation of the convolution for image is,

$$convolution(f, k) = \sum_m \sum_n f(x-m, y-n)k(m, n)$$

Where f is the input image and k is the kernel.

From the equation, it can be realized that convolution is similar to dilation in morphological operation with the use of structure element [12].

Steps of convolution in filtering process are,

1. Each pixel in the image neighborhood is multiplied by the contents of the corresponding element in the filtering kernel.
2. The results from the multiplication are summed and divided by the sum of the kernel.
3. The result is scaled and boosted , and used to replace the center pixel in the image neighborhood

2.2.2 Correlation

The correlation operation is closely related to convolution .similar to convolution operation, correlation computes the output pixels as a weighted sum of neighbouring pixels. The difference is that the matrix of weights in this case called correlation kernels, which is 180 degree rotation of convolution kernel. Correlation can be defined as,

$$correlation(f, k) = \sum_m \sum_n f(x+m, y+n)k(m, n)$$

Where f is the input image and k is the kernel,

Both correlation and convolution work as Erosion and Dilation in Morphological operation. In greyscale Images Convolution will increase the brightness of object by taking the neighbourhood maximum when passing with filter, as correlation will reduce the brightness of object by taking the neighbourhood minimum passing with filter.

3. Marker controlled watershed algorithm

In this paper marker controlled watershed algorithm is used which is simple algorithm to create foreground and background markers using linear filtering based image reconstruction. Here Image is converted in to the gradient images so that it represents the edge strength at each pixel [9, 10]. After that we use the correlation followed by convolution of image to traced foreground object. For finding the image convolution we define a circular average filter as mask or kernel. Calculating the extended regional maxima of these reconstructed images is done to get the smooth edge of foreground objects which is better than only regional maxima finding in image. Then we superimpose these markers to the original image. The background markers are created by calculating the Euclidian distance of binary version of above superimposed image. Now the gradient image is modified by the Linear filtering based reconstruction with foreground and background markers. Finally applying watershed transform on it will gives final segmented images of desired objects. The proposed algorithm can be seen in figure 1.

0

Figure 1 Algorithm

4. Experimental Results

Marker controlled watershed algorithm is very fast and advanced technique in image segmentation process. It converts the image in the gradient magnitude. Then Linear filtering operation is applying on gradient image. In the area of Medical Image Segmentation, marker controlled watershed algorithm performs well. It gives better result on X-ray, MR, and Mammographic images. In fig 2 we can see the original medical images. The fig 2.1 is an X-ray Image, in which we can see the stone in the kidney. Fig. 2.2 and 2.3 are MR images, which shows tumour in brain, and fig. 2.4 is the mammographic image, which shows tumour in breast. Fig 3 shows the gradient magnitude of original images. fig 4 is the visualization of the foreground marker, background marker and object boundaries applied on the original image, which was implemented by linear filtering. Finally fig. 5 is final watershed outputs of the original images, where objects like stone in the kidney and tumour in the brain and breast can be extracted from the original images. These results are better than the marker controlled watershed algorithm implemented with the morphological operation, which are shown in fig 6. From fig 6 it is compared that using morphological operation based algorithm, it fails to detect tumour in brain and breast. However it finds stone in kidney but it is more sensitive to other object in image also. Morphological operation is based on dilation, erosion, image opening and closing with reconstruction [5-8]. The modified marker controlled watershed algorithm with liner convolution and correlation is faster than morphological based algorithm. This comparison is shown in table 2, in which average CPU time of both the algorithm are compared .The speed and so that accuracy of new algorithm is higher than old algorithm. Table 1 shows the different CPU time of convolution based algorithm.



Figure 2.1 X-Ray

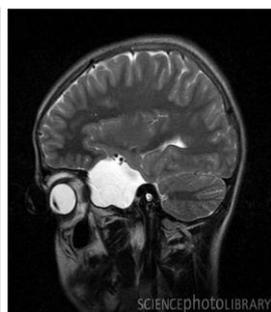


Figure 2.2 MRI-1

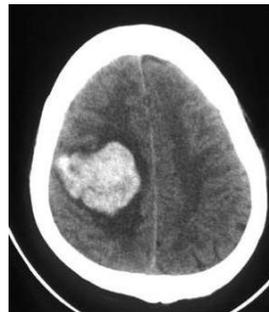


Figure 2.3 MRI-2



Figure 2.4 Mammograph

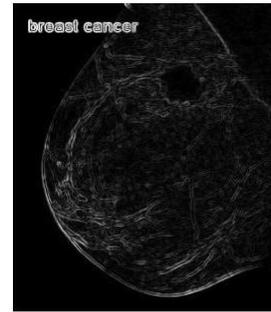
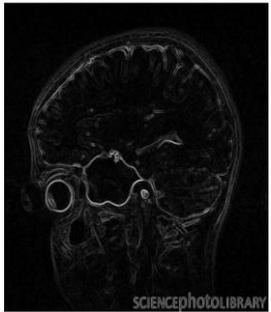


Figure 3 Gradient Magnitudes of Original Images

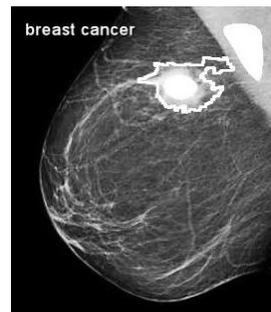
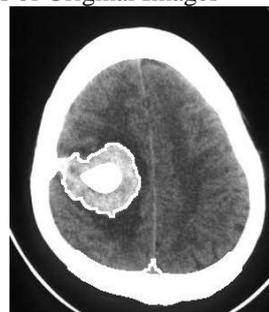


Figure 4 Markers and Objects Applied on Original Images

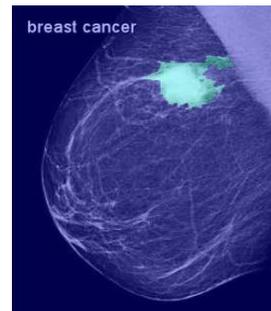
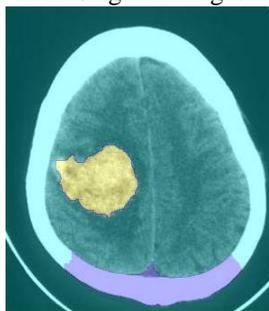
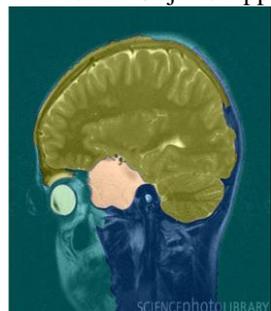
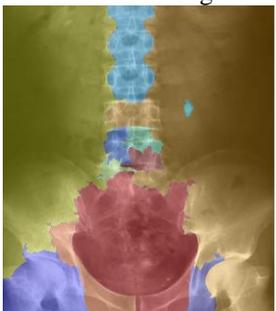


Figure 5 Final Watershed Output

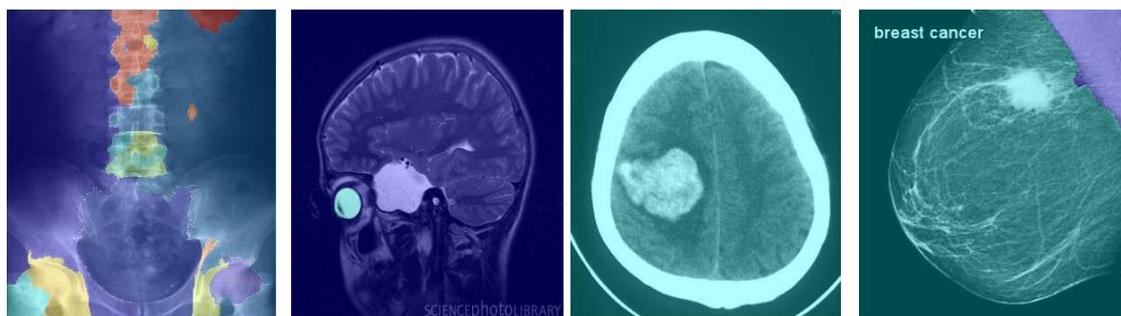


Figure 6 Comparison with Morphological based Watershed Algorithm

Image Type	Image Size	CPU Time-1 (sec.)	CPU Time-2 (sec.)	CPU Time-3 (sec.)	CPU Time-4 (sec.)	Average CPU Time (sec.)
X-ray	1536×2048	19.232	19.462	19.243	19.013	19.237
MR Image-1	530×530	3.370	3.463	3.510	3.426	3.442
MR Image-2	300×319	2.520	2.450	2.440	2.372	2.445
Mammographic Image	272×342	2.589	2.432	2.501	2.697	2.554

Table 1

Image Type	Image Size	Average CPU Time With Morphological Operation based Algorithm (Sec.)	Average CPU Time With Linear Convolution based Algorithm (Sec.)	Accuracy (%)
X-ray	1536×2048	23.199	19.237	17.07 %
MR Image-1	530×530	4.052	3.442	15.05 %
MR Image-2	300×319	3.041	2.445	19.60 %
Mammographic Image	272×342	3.110	2.554	17.88 %

Table 2

5. Conclusion And Future Work

The study shows that proposed Watershed algorithm by foreground marker is able to segment real medical images which containing several irregularities. The formulation is based on markers and simple convolution and correlation, which easily allows a regularization of the watershed, and is a flexible approach to the further optimization parameters. Implementing marker controlled watershed algorithm with morphology will increase computational cost and time consumption, because of several operations. Detection of object in medical images is also better in linear algorithm than morphological operation. So linear operation based marker controlled watershed algorithm is superior than morphology based algorithm. This algorithm can be merged with some advanced techniques like wavelet transform or it can be modify with some standard algorithm like fuzzy-c or k-means segmentation to improve the high resolution images specially in case of satellite images.

REFERENCES

- [1] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transform," in *Mathematical Morphology in Image Processing*, E. R. Dougherty, Ed. New York: Marcel Dekker, 1993, vol. 12, pp. 433-481.
- [2] Digabel, H., & Lantuejoul, C. (1977). Iterative algorithms. In J.-L. Chermant (Ed.), *Actes du Second Symposium Europeen d'Analyse Quantitative des Microstructures en Sciences des Materiaux, Biologie et Medecine*, Caen, 4-7 October 1977, pp. 85-99. Stuttgart: Riederer Verlag.
- [3] Li, H., Elmoataz, A., Fadili, J. & Ruan, S. (2003). "An improved image segmentation approach based on level set and mathematical morphology". In H. Lu & T.Zhang (Eds.), *Proceedings of the Third international Symposium on Multispectral Image Processing and Pattern Recognition*, volume 5286, pp. 851-854.
- [4] J. L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms," *IEEE Trans. Image Processing*, vol. 2, pp. 176-201, 1993.
- [5] J. Serra, *Mathematical Morphology*, vol. 1, Academic Press, London, UK, 1982.
- [6] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, UK, 1982.
- [7] C. R. Giardina and E. R. Dougherty, *Morphological Methods in Image and Signal Processing*, Prentice-Hall, Upper Saddle River, NJ, USA, 1988.

- [8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall, Upper Saddle River, NJ, USA, 2nd edition, 2002.
- [9] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion Simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, 1991.
- [10] Gonzalez, R.C., R.E. Woods and S.L. Eddins, 2004. *Digital Image Processing using MATLAB*. 1st Edn, Pearson Education India, India, ISBN: 9788177588989, pp: 620.
- [11] Imdad Rizvi, B K Mohan, "Wavelet based Marker-Controlled Watershed Segmentation Technique for High Resolution Satellite Images" 2nd International Conference and workshop on Emerging Trends in Technology (ICWET) 2011.
- [12] B.Kisacanin,c.schonfeld,"A Fast threshold Linear Convolution Representation of Morphological Operations",*IEEE Transactions on Image Processing*, ,1994
- [13] Mariusz Jankowski,"Erosion, dilation and related operators" 8th international mathematica symposium, Avignon 2006.
- [14] Steven W.Smith,"The Scientist and Engineer's Guide to Digital Signal Processing",chapter 4.

Smart Lighting and Control using MSP430 & Power Line Communication

Sanjay Belgaonkar¹, E. Elavarasi², Gurjeet Singh³

¹M.Tech Scholar, Dept. of ECE, AMC Engineering College, Bangalore – 560 083

²Assistant Professor, Dept. of ECE, AMC Engineering College, Bangalore – 560 083

³Director, Gill Instruments Pvt. Ltd., Electronic City, Bangalore – 560 100

Abstract

Smart Lighting is a lighting technology designed for energy efficiency. It includes high efficiency fixtures, day lighting and automatic controls that make adjustments based on conditions such as occupancy. This smart lighting system is connected through power line which is also used for communication. Power Line Communication (PLC) is a technology which uses power lines as physical media for data transmission. PLC can offer a “no new wires” solution because the infrastructure has already been established. PLCs are used for transmitting data at rapid speed through a power line in a house, an office, a building and a factory etc. Here, the existing alternating current (AC) power wires serve as a transmission medium by which information is relayed from an AC source. Present paper deals with design and development of a smart lighting system which is controlled by MSP430 microcontroller and power line communication.

Key Words – MSP430, Power Line Communication (PLC), Sensors, Smart Lighting System

Introduction

Since the first caveman learned to control fire, humans have shaped and used light in a constantly expanding array of technologies. Yet lighting – “Smart lighting” – could do much more, according to E. Fred Schubert, Wellfleet Senior Constellation at Rensselaer. Usually lighting consumes a lot of electrical energy every day all around the world. According to the statistics, 20 to 50 percent of total energy consumed in homes and offices are used for lighting. What is surprising is that over 90 percent of the lighting energy expense used for some of the buildings is unnecessary due to the over illumination. The cost of lighting can be very realistic.

PLC utilizes the power line infrastructure in a home, office or other building, both indoor and outdoor, for networking and communication thereby eliminating the expense and inconvenience of new wires or antenna – based networks. The power line can be an extremely difficult and noisy communications medium, characterized by several unpredictable and strong forms of interference. In this paper, we give solutions to overcome the interference and reliability problems that can occur on the power line and offer low cost, robust and superior performance. Reliable, low cost PLC technology enables ubiquitous applications for residential and business markets. PLC opens a whole new world of business opportunities to appliance and electric devices manufacturers, utilities and other service providers, with no need to install new cables or jacks. With PLC, every electrical outlet can become a communication node, part of a PLC network.

I. Smart Lighting System

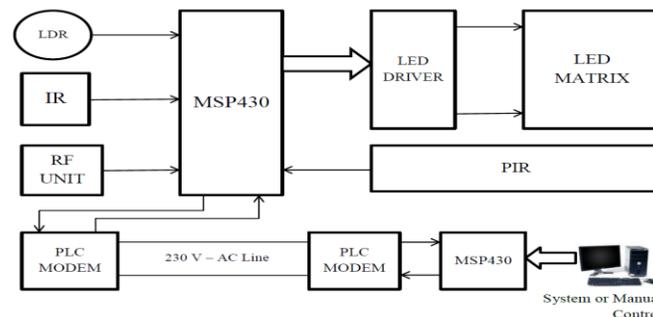


Figure 1: Block diagram of Smart Lighting and Control using MSP430 & PLC modem

The lighting model uses Light Emitting Diodes (LEDs) which are driven by a driver. This lighting model is controlled by MSP430 microcontroller which receives the commands from the master microcontroller or from a manually controlled system. The communication between transmitter and receiver happens through power lines using PLC modem. To this lighting model different sensors are also connected as shown in the diagram. The details about the sensors are given below.

A. SENSORS

1. Passive Infrared Sensor

A Passive Infrared sensor (PIR sensor) is an electronic device that measures infrared (IR) light radiating from objects in its field of view. PIR sensors are often used in the construction of PIR-based motion detectors. Apparent motion is detected

when an infrared source with one temperature, such as a human, passes in front of an infrared source with another temperature, such as a wall. This is not to say that the sensor detects the heat from the object passing in front of it but that the object breaks the field which the sensor has determined as the "normal" state. Any object, even one exactly the same temperature as the surrounding objects will cause the PIR to activate if it moves in the field of the sensors.

2. Infrared Receiver Modules

These IR receiver modules are used here for remote control systems. These modules are miniaturized receivers for infrared remote control systems. Figure 3 gives the block diagram of IR receiver module. PIN diode and preamplifier are assembled on lead frame, the epoxy package is designed as IR filter. The demodulated output signal can directly be decoded by the microcontroller. TSOP348.. is the standard IR remote control receiver series for 3V supply voltage, supporting all major transmission codes.

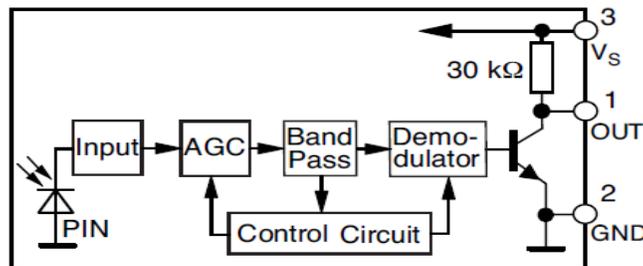


Figure 2: Block diagram of IR receiver module

3. Light Dependent Resistor

Light dependent resistor (LDR) is a resistor whose resistance decreases with increasing incident light intensity; in other words, it exhibits photoconductivity. It can also be referred to as a photoconductor or CdS device, from "cadmium sulfide," which is the material from which the device is made and that actually exhibits the variation in resistance with light level. A photo resistor is made of a high resistance semiconductor. If light falling on the device is of high enough frequency, photons absorbed by the semiconductor give bound electrons enough energy to jump into the conduction band. The resulting free electron (and its hole partner) conduct electricity, thereby lowering resistance.

II. MSP430 MICROCONTROLLER

The MSP430 is a mixed-signal microcontroller family from Texas Instruments. Built around a 16-bit CPU, the MSP430 is designed for low cost and, specifically, low power consumption embedded applications.

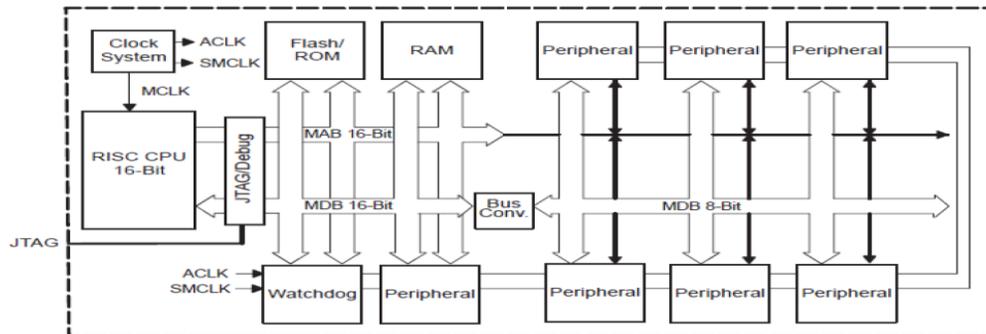


Figure 3: MSP430 Architecture

The MSP430 can be used for low powered embedded devices. The electric current drawn in idle mode can be less than 1 microamp. The top CPU speed is 25 MHz. It can be throttled back for lower power consumption. The MSP430 also utilizes six different Low-Power Modes, which can disable unneeded clocks and CPU. This allows the MSP430 to sleep, while its peripherals continue to work without the need for an energy hungry processor. Additionally, the MSP430 is capable of wake-up times below 1 microsecond, allowing the microcontroller to stay in sleep mode longer, minimizing its average current consumption. Note that MHz is not equivalent to Million instructions per second (MIPS), and different architectures can obtain different MIPS rates at lower CPU clock frequencies, which can result in lower dynamic power consumption for an equivalent amount of digital processing.

The MSP430 incorporates a 16 – bit RISC CPU, peripherals, and a flexible clock system that interconnect using a von-Neumann common memory address bus (MAB) and memory data bus (MDB). Partnering a modest CPU with modular memory – mapped analog and digital peripherals, the MSP430 offers solutions for demanding mixed – signal applications. Figure 3 gives the architecture of MSP430.

Key features of the MSP430x1xx family include:

- Ultralow – power architecture extends battery life
 - 0.1 μ A RAM retention
 - 0.8 μ A real-time clock mode
 - 250 μ A / MIPS active
- High – performance analog ideal for precision measurement
 - 12 – bit or 10 – bit ADC --- 200 ksp/s, temperature sensor
 - 12 – bit dual – DAC
 - Comparator – gated timers for measuring resistive elements
 - Supply voltage supervisor
- 16 – bit RISC CPU enables new applications at a fraction of the code size.
 - Large register file eliminates working file bottleneck
 - Compact core design reduces power consumption and cost
 - Optimized for modern high – level programming
 - Only 27 core instructions and seven addressing modes
 - Extensive vectored – interrupt capability
- In – system programmable Flash permits flexible code changes, field upgrades and data logging.

III. Power Line Communication

Power Line Communication is a communication technology that enables sending data over existing power cables. This means that, with just power cables running to an electronic device (for example) one can both power it up and at the same time control / retrieve data from it in a half – duplex manner. PLC is the usage of electrical power supply networks for communication purposes. In this case, electrical distribution grids are additionally used as a transmission medium for the transfer of various telecommunications services. The main idea behind PLC is the reduction of cost and expenditure in the realization of new telecommunication networks.

The application of electrical supply networks in telecommunications has been known since the beginning of the twentieth century. The first Carrier Frequency Systems (CFS) had been operated in high-voltage electrical networks that were able to span distances over 500 km using 10-W signal transmission power. Such systems have been used for internal communications of electrical utilities and realization of remote measuring and control tasks. Also, the communications over medium- and low-voltage electrical networks has been realized. Ripple Carrier Signaling (RCS) systems have been applied to medium- and low-voltage networks for the realization of load management in electrical supply systems.

A. STANDARDS

The communications over the electrical power supply networks is specified in a European standard CENELEC EN 50065, providing a frequency spectrum from 9 to 140 kHz for powerline communications (Tab. 1). CENELEC norm significantly differs from American and Japanese standards, which specify a frequency range up to 500 kHz for the application of PLC services.

Band	Frequency Range (in KHz)	Max. Transmission Amplitude (in V)	User Dedication
A	9 – 95	10	Utilities
B	95 – 125	1.2	Home
C	125 – 140	1.2	Home

Table 1: CENELEC bands for power line communication

CENELEC norm makes possible data rates up to several thousand bits per second, which are sufficient only for some metering functions (load management for an electrical network, remote meter reading, etc.), data transmission with very low bit rates and the realization of few numbers of transmission channels for voice connections. However, for application in modern telecommunications networks, PLC systems have to provide much higher data rates (beyond 2Mbps). Only in this case, PLC networks are able to compete with other communications technologies, especially in the access area.

B. TECHNOLOGY BENCHMARK

The following table compares the different technologies depending on the required baud rate, target price and system performance.

	Price	Baud	Distance	Network	Power	Deployment	Operational
		rate		coverage	consumption	cost	cost
PLC	✓	✓	✓	✓	✓	✓	✓
BPL	✗	✓	✗	✗	✗	✓	✓
ZIGBEE	✓	✓	✗	✗	✓	✗	✓
RFISM	✓	✗	✓	✗	✗	✗	✓
GSM/GPRS	✗	✓	✓	✓	✗	✗	✗

Table 2: Comparison of PLC with other technologies

PLC: Power Line Communications

BPL: Broadband PL

RF ISM: Radio Frequency in Industrial, Scientific and Medical radio band

IV. Conclusions

Thus we provide a smart lighting system using MSP430 microcontroller which helps to reduce energy usage and cost by eliminating over – illumination and unnecessary waste. This system provides centralized control of all lighting within a home or commercial building, allowing easy implementation of scheduling, occupancy control, daylight harvesting and more through power line communication. This system also supports demand response and will automatically dim or turn off lights to take advantage of demand response incentives and cost savings.

References

- [1] Smart Lighting http://www.rpi.edu/research/magazine/summer04/futurechips_1.html
- [2] Texas Instruments User’s Guide for MSP430x1xx Family – 2006
- [3] IR Receiver Modules for Remote Control Systems, Datasheet, Vishay Semiconductors, Rev. 5, 23 – Jun – 2003.
- [4] Pyro electric Infrared Sensors, Datasheet, Murata Manufacturing Co., Ltd. – 2011.
- [5] Lighting - Occupancy Sensors, <http://www.doi.gov/greening/energy/occupy.html>
- [6] E. Fred Schubert (2003). "1". *Light-Emitting Diodes*. Cambridge University Press. ISBN 0819439568.
- [7] F.N.Pavlidou, A.J.Han Vinck, J.yagdni, B.Honary, “ Power Line Communications State of the Art and Future Trends”, IEEE Communications Magazine, pp 34 – 40, Vol. 41 – Issue 4, April – 2003.
- [8] Research Paper: A Purroy, A. Sanz, J.I. Garcia Nicolas, I Urriza, “ Research Areas for Efficient Power Line Communication Modems”, Dept. of Electronics & Communication Engg., University of Zaragoza, Spain.
- [9] HomePlug Power line Alliance: <http://www.homeplug.com>
- [10] Research Paper: Yu – Ju Lin, Haniph A. Latchman, Minkyu Lee and Srinivas Kata, “A Power Line Communication Network Infrastructure for the Smart Home”, Electrical & Computer Engg. Department, University of Florida.
- [11] Jae – Jo Lee, Choong Seon, Hong Joon – Myung Kang and James Won – Ki Hong, “Power Line Communication Network Management in Korea”, International Journal of Network Management – 2006; 16.
- [12] Book: “Braodband Power Line Communications Network Design” by Halid Hrasnica, Abdlfatteh Haidue, Ralf Lehnert, John Wiley & Sons Ltd., Edn. – 2004.

Author's Information



Mr. Sanjay Belgaonkar received his **B.E** degree in **Electronics & Communication Engineering** from Visvesvaraya Technological University, Belgaum, Karnataka State, India in the year 2009. Presently he is pursuing his **M.Tech** in **Digital Electronics and Communication** from AMC Engineering College, Bangalore. Affiliated to VTU, Belgaum. His interest areas are Communication, Signal Processing etc.



Prof. E Elavarasi received her **B.E** degree in **Electronics & Communication Engineering** from Visvesvaraya Technological University, Belgaum, Karnataka State, India in the year 2004 and **M.Tech** in **Digital Electronics and Communication** from the same University in the year 2007. She has also received **MBA** degree in **HR & Finance** from Bangalore University in the year 2011. She has 5 years of teaching experience and presently she is working as Assistant Professor in the Dept. of ECE, AMC Engineering College, Bangalore. She has published several research papers in International and National conferences. Her areas of interest include Signal Processing, Communication, Multirate, Error Control Coding, Digital Circuits and Logic Design etc.



Mr. Gurjeet Singh is a **Founder Director and Design Head** of **Gill Instruments Pvt. Ltd.**, Electronic City, Bangalore. Gill Instruments is a Third party Company of Texas Instruments. Gill Instruments provides design solutions in the area of Embedded Systems and Manufacture Development tools for MSP430 family. He worked with General Electric Bangalore for 1 year. He holds a Bachelor's degree in **Electronics & Telecommunication** from J.N.E.C Aurangabad. He has been working closely with Texas Instruments for university programs. He has taken the initiative to set up embedded system labs across the country in various universities and Engineering colleges. He contributed some of the best development tools available today for embedded system designs. He also conducts 3 day Embedded System design workshop, where he has trained more than Five thousand engineers in the past 6 years.

The e-Health scenario with latest trends in EMR applications: A Review of EMR techniques with healthcare framework

Onkar S Kemkar¹ Dr P B Dahikar²

¹PCD ICSR, VMV College Campus, Wardhaman Nagar, Nagpur – 440008.

²Kamla Nehru Mahavidyalaya, Sakkardara Square, Nagpur – 440009.

Abstract:

Over the past few years, information systems have become increasingly important in healthcare delivery. The use of computers in a wide range of medical applications and healthcare management is one potential alternative to reducing the overall costs of healthcare delivery. Further, the use of sophisticated decision support systems is envisaged to improve quality of clinical decision making.

The paper discusses what medical informatics is, the definition of EMR, efforts required for extraction of patient data from heterogeneous EHRs. Gain new knowledge within secure distributed systems and software agents. Based on experimental work and some of the pilot studies a system has been designed and developed in the field of EMR & EHR

Keywords— medical informatics, ehealth, e-records, health informatics, EMR, EPR

I. INTRODUCTION

The health of a nation is the product of many factors and forces that combine and interact. Information on health care and nutrition, access to safe drinking water, public and private health care infrastructure, access to preventive health and medical care and the health insurance are among the contributing factors.

The advances in medical science and biomedical engineering on one side and Information and Communication Technology (ICT) on the other are offering wide opportunities for improved health care. [1]

In the healthcare industry there are some areas where leading-edge ICT developments are employed and developments at the leading-edge will be important. In other areas the health system is some way behind other industries in the adoption and application of information management and information systems. In general, the situation seems to be one of relatively slow progress through the evolving computing paradigms of functional computing, enterprise computing and network computing. Some functions are highly automated, but integrated enterprise computing in hospitals and clinics is still rather rare. As a result, linking and integrating the healthcare system (network computing) remains a major challenge. [2]

The Internet has a greater potential to fundamentally transform both the structure and the core processes of medicine than any new technology we have seen in the past

fifty years. Professional resistance to adoption of the technology and political problems associated with protecting the confidentiality of patient records pose the two biggest hurdles to fully realizing this potential.

So, vision of the health care information infrastructure is possible using technologies that support the sharing of medical **e-records** while maintaining patient privacy. Revolutionizing Health Care through Information Technology that contained comprehensive findings regarding the potential of information technology (IT) to reduce medical errors, lower costs, and improve patient care. It also recommended a technological framework for transitioning from manual, paper-based health records to a modern, computerized electronic records infrastructure. [3]

II. WHAT IS MEDICAL INFORMATICS?

Health Information systems have nationalities. All the HIS in the world have a common goal: to support healthcare professionals in improving overall efficiency, cost effectiveness and ultimately the quality of patient care. However, they differ widely in the implementation. The growing trend towards shared care requires that these systems be able to share their data. The objective is to propose a generic system that will provide integrated access to all the information and knowledge necessary to treat patients. [4]

Medical informatics is the field concerned with "the cognitive, information processing, and communication tasks of medical practice, education, and research, including the information science and technology to support these tasks (as reported in Greenes and Shortliffe, (1990)". [1]

III. WHAT IS ELECTRONIC MEDICAL RECORD (EMR)?

According to the Medical Records Institute, the electronic patient record is a "computer-stored collection of health information about one person linked by a person identifier." The institute further identifies five distinct levels of computerization for patient information systems:

- Level 1: in this level, health organizations are still depending on paper-based medical records
- Level 2: Here, the goal is to digitize the medical record and create an electronically available record;
- Level 3: the electronic medical record is the upgraded version of the computerized medical record
- Level 4: here, the electronic patient record combines several enterprise-based electronic medical records concerning one patient and assembles a record that goes beyond the enterprise-based retention period; and
- Level 5: the more comprehensive collection of an individual's health information is the electronic health record [5, 6].

A. Features

The universal features that **EPRs** should have lifelong records for every person, online access to patient records

for practitioners and genuinely seamless care resulting from the above two features.

- Integrates information from many sources; from blood pressure monitors to complex imaging systems
- Provides a single access point for relevant, concise, accurate and active data about a patient to authorized users in different locations. These records may be available for:

1. Patient care
2. Administration
3. Clinical Audit
4. Financial Audit
5. Research
6. Education

- The UI features are easily configurable to meet the varying demands of each medical practitioner. [4]

B. Evaluation

We try and to answer the following questions while evaluating our system:

- Will the evaluation provide a comprehensive and objective assessment of the achievement of goals?
- Will it form the basis for continued improvement in the application and use of information?
- Will the evaluation be realistic?
- Will the findings of the evaluation be accurate?
- Will it convey technically adequate information about the features that determine the worth or merit of the system being evaluated? [9]

IV. OBJECTIVES OF EMR

“ehealth is not an evolutionary concept but a revolutionary concept and at the heart of every revolution, there is the need for a sudden massive change, at the core of which is the human mind.”

By taking into consideration “Health for All,” we’ll discuss some objectives for which e-record has originated.

- To develop a generic approach to integrate the various components to provide decision support, using the existing health information systems, imaging systems, case bases, clinical guidelines and protocols, and modeling and simulation tools as shown in figure 1.[6,7]

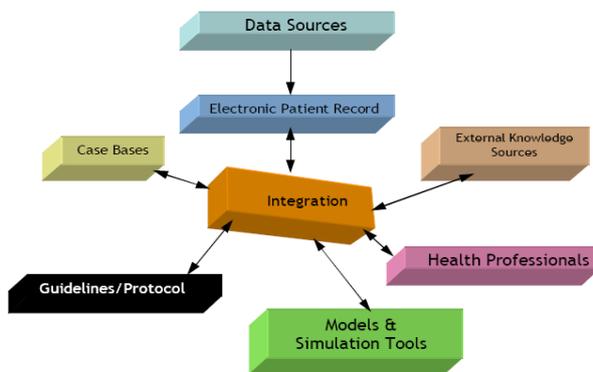


Figure1: overview of the Information Architecture of the System.[4]

- EPR is not homogenous. It is taken from a multitude of sources. The data which feeds or populates the EPR already resides in a variety of highly heterogeneous and autonomous information systems and simply integrating this data does not necessarily result in a valid EPR. Clinical

data may be entered using a multiplicity of methods such as keyboard, voice recognition software, touch screens, scanners or from medical instrumentation such as blood pressure monitors, ECG machines, laboratory analyzers or imaging devices. We will have to develop a middle-ware client-server component-based approach, acting as an intermediary as shown in figure 2.[10,11]

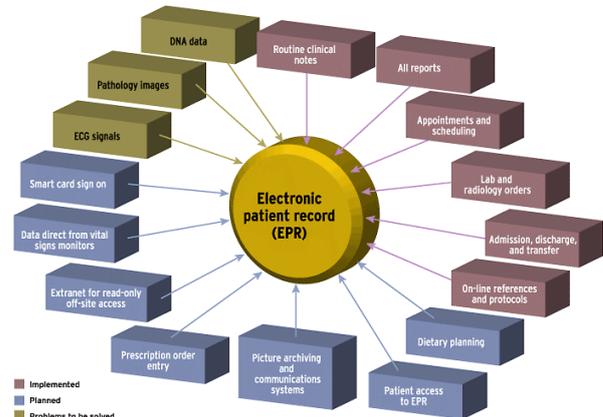


Figure 2: EPR with functions that have been incorporated & those that have not in the current crop of HIS.[4]

A. What are advantages of EMR?

- Efficiency isn't the only benefit. For individual patients, access to good care becomes easier and safer when records can easily be shared. Important information -- such as blood type, prescribed drugs, medical conditions and other aspects of our medical history -- can be accounted for much more quickly. At the very least, an existing electronic medical record (EMR) can save time at the doctor's office. At most, quick access to our records can be lifesaving if an emergency occurs and answers to those questions are needed during the emergency decision-making process.

- Second advantage is safety. In the past, the way a doctor obtained your health history was by asking you. Each time you visited a new doctor's office, you filled out forms about your history, including previous surgeries, or the drugs you take on a regular basis. If you forgot a piece of information, or if you didn't write it down because it seemed unimportant to you, then your doctor didn't have that piece of your medical puzzle to work with. However, when doctors share records electronically, your new doctor only needs to ask your name, birthdates, and possibly another piece of identifying information. He can then pull up your records from their electronic storage space. All of the information he needs to see will be there in full.

- In the past, when a doctor closed his practice, retired, moved, or even died, patient records could easily get lost or relocated, making it impossible for patients to get the records they needed to take to a new doctor. Keeping these records electronically, especially in the cases where patients can also gain access to them, means the patient won't be left without the records he may need.

V. eHealth & TODAY'S SCENARIO

A. eHealth

eHealth is an overarching term used today to describe the application of information and communications technologies in the health sector. It encompasses a whole

range of purposes from purely administrative through to health care delivery. For example: within the hospital care setting, eHealth refers to electronic patient administration systems; laboratory and radiology information systems; electronic messaging systems; and, telemedicine -- teleconsults, telepathology, and teledermatology.

A fundamental building block of all these applications is the Electronic Health Record, which allows the sharing of necessary information between care providers across medical disciplines and institutions. Other important uses of eHealth are found in the areas of continuous medical education and public health awareness and education.

B. Today's Scenario

Let's now consider the different options available as far as the healthcare service provider software & healthcare tertiary caregivers are concerned:

- The patient is mobile and its health travels with him. Healthcare services should be available pervasively, integrated into the patient's environment.
- Mobile Grids can provide an infrastructure for an efficient development, provision and maintenance of complex e-health applications.
- Recent advances in information and communication technologies (ICT), miniaturization of sensor devices and computers, as well as the wider availability of connectivity and wireless networks gives rise to a number of services and applications involving wirelessly connected sensors and actuators in a body area network (BAN), personal area network (PAN) or in the environment.
- Connectivity between medical sensors/devices, gateways/hubs and healthcare services is present.
- Health professionals rely on accurate data coming from certified devices. Patient safety is most important, but privacy is also taken care of, as required by legislation (EU Directive 46/95, HIPAA, etc.), e.g. personal health data can be used based on patient consent, only for certain purposes, on a need-to-know basis by certain healthcare providers (e.g. a doctor that has a care relationship with the patient).

Keeping with the pace of cutting edge technology, the major drift has occurred in terms of business logic as eHealth scenario is concerned. Now cloud computing & mobile computing are the most hap things. These technologies are in, so we will have to reconsider the traditional ways of using EHR or EMR, the process of remolding has began & we are currently doing a exp. on the data modeling & the results are quite interesting. As the Indian Govt. is also promoting to use the HL7 standards, the stiffness in the solution is reducing to a greater extent. We will be discussing the Indian scenario later when the article progresses.

VI. EMR & ITS IMPLICATIONS

Although the concept of EMR is relatively old, still it has to recoup & the scope for research & area of activity is wide as compared to the other allied healthcare apps. The aspect of implication lies in the implementation; this is universal as far as the software utilization is concerned.

Considering this point we will be elaborating the Indian case scenario:

The Government of India in February 2009 set up the Unique Identification Database Authority of India (UIDAI) with the aim of providing a Unique Identification number to all Indian citizens.

These positive developments impelled us to start a discussion on the use of UID as a healthcare identifier.

Some suggestions were given for the software vendors for the change in their software.

- UIDAI may decide policies for encryption, security and sharing with external applications and showing the UID on screens of the application etc. and specification to be able to process a UID as an UHID and set a standard.
- Each software vendor should add a field in patient record in which the UID may be stored compliant to standards set by government.
- Software Vendors should declare that their product is compliant to the standard and register themselves with UIDAI and or some other government authority and obtain compliance certification.
- This will help during transfer of patients data when patient moves from
 1. One doctor to another doctor
 2. One hospital to another hospital
 3. One city to another city
 4. One state to another state
 5. Database of one software to another software

These are some of the recommendations. Now if we look into the deeper perspective of this, we want to have software which is scalable, expandable & e records should be such that it can be migrated over the database with a click or button. Our research aims for that. We will be converting a universal query into EDI so that it is scalable & the values in the distributed databases will be transferred from one database to another. This implies that with extensive research & database modeling the level of EMR can be aggravated.

VII. FEATURES VS. USABILITY

It is essential to use a Unique Healthcare Identifier (UHID) for each patient/citizen of India and the proposed UID can be used for this purpose. This assumes great significance as:

1. The individual patients in India are free to access healthcare provider of his/her choice.
2. The providers may refer the patient to another healthcare provider.
3. The patients may change the city, district or state for various reasons as well as change the provider.
4. In each healthcare establishment, all necessary facilities are not available and hence several diagnostic and therapeutic services are taken from other agencies.
5. More than one healthcare provider records healthcare information of each patient.
6. The information is maintained in unconnected and unrelated databases, electronic or otherwise, under disparate identification numbers.

7. Benefits of using electronic medical records for public health cannot be overemphasized.
8. If each electronic database maintains the information of each patient with their UID, it will help integration of medical information of each patient collected by various healthcare providers.
9. Integrated medical information of all patients will be helpful for both the experts deciding public health policies and the public health informaticians.

VIII. BENIFITS

- The patient saves traveling time and money for availing expert opinion.
- The doctors maintain the records in an organized form.
- Statistics generation is online /immediate for future planning.
- Disease prevalence /epidemics will be detected immediately.
- Doctors get Continuing Medical Education through the network.
- Mass Education can be distributed through the network.
- Dais and ANMs get online training as well as advise during difficulties.
- Fringe benefits of computerization like stocks maintenance, accounting etc. shall be availed. Immunization records, maintenance of Child Health, AIDS and other public Health Programs can be followed up.

IX. CONCLUSION

A meeting of minds has taken place. After all, a computer is required for sending the records and a doctor who keeps his records on a computer just needs to send the same record for a consultation as and when required for a specialist's opinion.

Realizing that the inputs required for both are similar and the efforts required for use can be complementary, both have now come together to bring out a comprehensive solution which solves the need of the health care industry. If all records of patient can be kept on the computer, the same quality of health care will be available in all the corners of our country saving on travel and time off work.

The challenge for ongoing national health record projects around the world is to take into account all the different types of EHRs and the needs and requirements of different health care professionals and consumers in the development of EHRs.

A further challenge is the use of international terminologies in order to achieve semantic interoperability.

ACKNOWLEDGEMENT

The author wish to acknowledge the help received from Dr D A Deshpande, Director, PCD Institute of Computer Studies and Research, VMV College campus for various inputs & methodological techniques in telemedicine.

REFERENCES

- [1] Telemedicine Concepts, Modalities and Implementation : Onkar S Kemkar, Publication : i4 Journal Volume 5 Issue 1/2 ISSN # 0975-2757:2011

- [2] Information Technology and the Revolution in Healthcare: John Houghton, Article taken from csesinfo website, June 2002
- [3] Enabling the 21st Century Health Care Information Technology Revolution: Communication of the ACM vol. 50, No. 2, February 2007
- [4] Medical Informatics: M. Habibullah Pagarkar Spring 2004
- [5] The EMR: An Electronic Bridge Between Medicine and Science: white paper of faster curves:2005
- [6] Think Research Using Electronic Medical Records: Faster Curves Medical white papers:2005
- [7] Hallvard Lærum Doctors Use of Electronic Medical Records Systems in Hospitals: Cross Sectional Survey British Journal of Medicine, vol 323, pp 13441348, 2001
- [8] Robin Beaumont Managing your Clinical Information Session 8: The Electronic Patient / Healthcare Record British Journal of Medicine, Vol. 38, pp. 303310, sept. 2001
- [9] Electronic Health Records Overview. Key Components of Electronic Health Records Journal of the Health Information Management, vol 20, Issue 2,
- [10] Definition, structure, content, use and impacts of Electronic health records: a review of the research literature: international journal of medical informatics vol 77, pp 291-304, July 2008
- [11] EPR (Electronic Patient Record) Laboratory – Simulated Environment to Learn about a Hospital EPR System “ Knowledge Management & E-Learning: An International Journal, Vol.3, No.1.
- [12] Healthcare Informatics Online: Electronic Patient Records EMRs & EHRs: by Dave Garets and Mike Davis October 2005- Healthcare Informatics
- [13] Problems with the Québec Electronic Health Records Strategy: A May 5 article in The Gazette by Kevin Dougherty
- [14] EHR 120 - Understanding Features and Functions <http://www.centerforhit.org/online/chit/home/cme-learn/tutorials/ehrcourses/ehr120.html>
- [15] EHR 301 - Selecting the Right EHR: <http://www.centerforhit.org/online/chit/home/cme-learn/tutorials/ehrcourses/ehr301.html> , Accessed on 15/05/2010

E-GOVERNANCE INITIATIVE IN TECHNICAL INSTITUTE – A CASE STUDY OF DR.B.R.AMBEDKAR INSTITUTE OF TECHNOLOGY, ANDAMAN & NICOBAR ISLANDS (INDIA)

UTPAL SHARMA

Principal

Dr. B.R. Ambedkar Institute of Technology
Andaman and Nicobar Islands

SUNIL KUMAR CHAKRABORTY

Faculty Member, Department of Computer Science
Dr. B.R. Ambedkar Institute of Technology
Andaman and Nicobar Islands

ABSTRACT

The process transformation and e-governance in technical institutes is one of the important components for improving the efficiency. This also improves the student's satisfaction index apart from making them more IT savvy. This paper deals with the case study showing how e-Governance initiative with in-house capacity has transformed the institute process without incurring any expenditure.

1. INTRODUCTION

Higher education in general and technical education in particular continues to expand in response to changing student's demand to meet the needs of diversity of markets. It is a widely accepted fact that technical education is a driver for our country's emergence as a key force in global economy. In other words, improvement in the process and governance of technical institutes will have direct bearing on the quality of the student rolled out, thereby giving impetus to the market, which is likely to absorb them.

Process Improvement Approach In Management Of Technical Institute

It is well accepted fact that administrators of most of the technical institute spend maximum time in attending routine matters apart from concentrating on expansion of academic activities of the institute. Considerably less time is spent on structured process mapping, identification of redundant & critical processes, process re-engineering followed by process transformation. There is a distinct difference between a manufacturing process and the process of academic institute, since the measurement criteria of process parameter is well defined in the former and its effect on quality can also be quantified. While the measurement of academic process parameter and quantifying its effect on the product quality is a difficult task.

The approach suggested for process improvement of an educational institute can be seen in Figure 1.

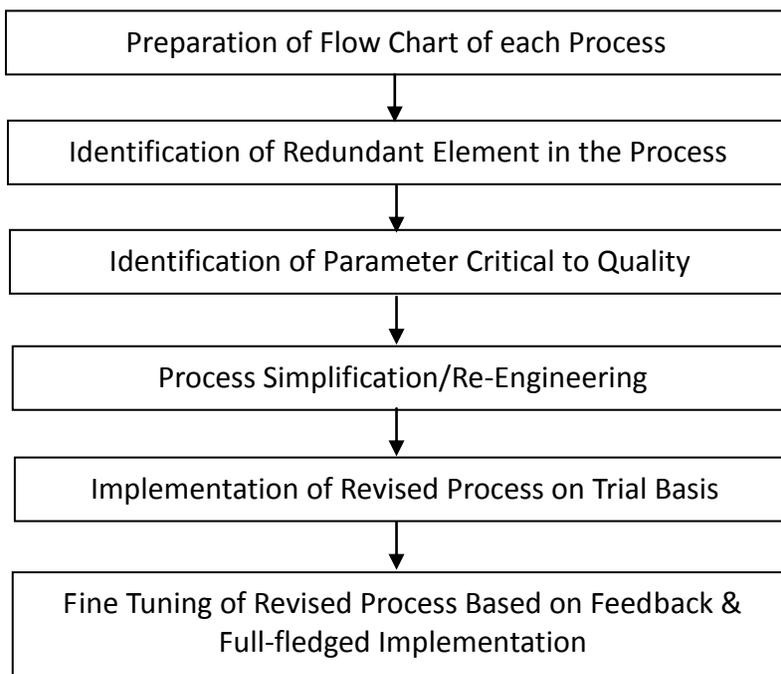
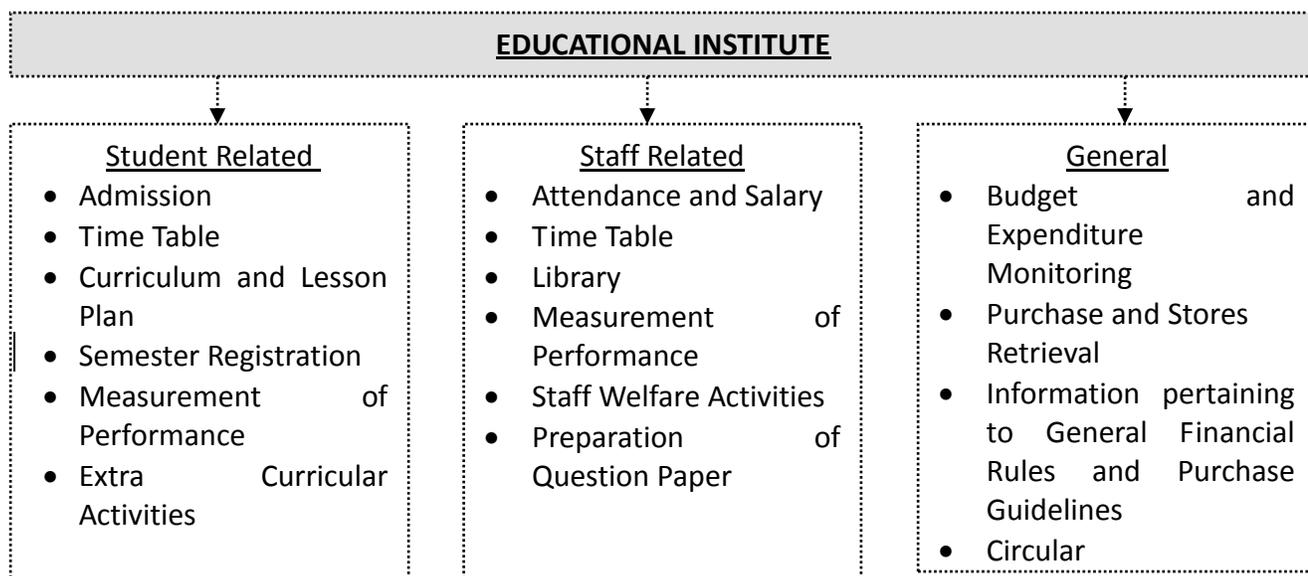


Figure 1: Approach for Process Improvement

The process of educational institute can be broadly divided into three major components viz. student related, -staff related and General as mentioned in Figure 2.



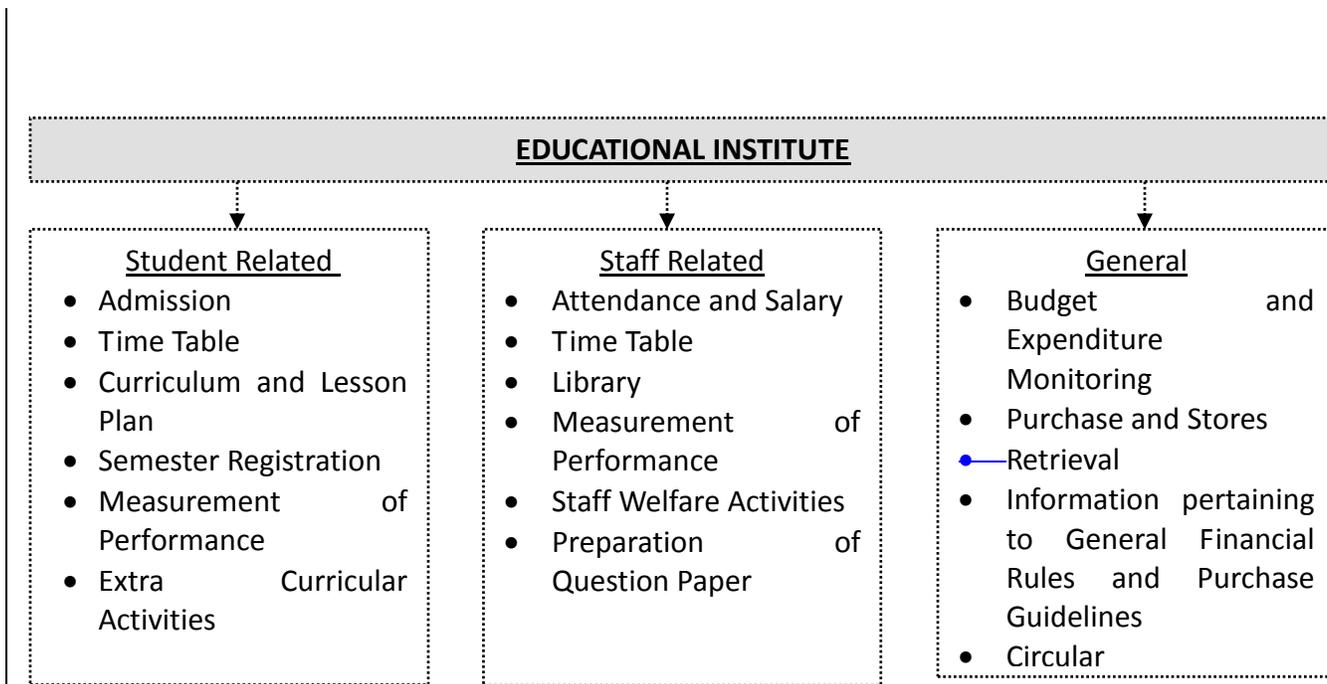


Figure 2: Major Process of Educational Institute

Most of the progressive institutes have gone for e-Gov initiatives by computerization of most of the above processes. The general model followed for computerization is outsourcing the development of applications including its maintenance, while some institutes have developed some of the applications in-house and implemented the same. It has been observed that the second option though time consuming and tedious, but helps in the long run to sustain and improve, apart from facilitating in-house capacity building

2. DR. B.R. AMBEDKAR INSTITUTE OF TECHNOLOGY – A BRIEF

Dr. B.R. Ambedkar Institute of Technology (BRAIT) earlier known as Dr. B.R. Ambedkar Govt. Polytechnic was established in 1984 at Port Blair (Andaman and Nicobar Islands) by A & N Administration. In 1999, the institute has been accredited with ISO 9000 and is perhaps one of the first technical institute in Govt. sector to get ISO 9000 accreditation. In 2009, the institute has been upgraded to Engineering College offering Engineering Degree Programmes in addition to Diploma Programmes. In last 10 years, the various department of the institute have taken up real time projects. The contribution of Computer Science Department of the institute is relatively higher as various applications developed for different departments of A & N Administration have been successfully implemented.

The model followed for development of the application by Department of Computer Science, BRAIT is as under –

- Identification of platform as per market demand.
- Imparting training to the staff and students on the selected platform by the industry experts.
- Taking up real time projects of different departments of Administration.
- Issuance of certificates, linked with the successful trial and client approval.
- Making the guide (faculty member) responsible for supporting the application on continuous basis

Some of the applications developed for various departments of A & N Administration by Computer Science Department of BRAIT in last 10 years are given below in Table 1.

S.No.	Name of the Project	Platform Used	Client Department	Year
1	Electricity Consumer Billing	Visual FoxPro	Electricity Department	2000
2	Vehicle Registration &	Oracle Developer 2000 &	Andaman & Nicobar Police	2001

S.No.	Name of the Project	Platform Used	Client Department	Year
	Issuance of License	Oracle 8i		
3	Loan Monitoring System of DIC & Khadi & Village Industry Board	Ms-Access	Directorate of Industries	2001
4	Inventory Management of Sagarika Emporium	Ms-Access	Directorate of Industries	2001
5	Account Management System	MS Access & SQL Server 2000	All Department of A & N Administration	2003
6	Production Disposal of Sawn Timber	ASP.Net 2.0 & SQL Server 2000	Forest Department	2004
7	Biometric Finger Print Attendance System	VB6.0 & SQL Server 2000	Directorate of Industries & ITI	2004
8	Employment Exchange Registration and Renewal System	ASP.Net 2.0 & SQL Server 2000	Employment Exchange, Port Blair	2005
9	Library Management System	Ms-Access & SQL Server 2000	State Library	2005
10	Material Management System	ASP.Net 2.0 & SQL Server 2000	Directorate of Shipping Service	2008
11	Store and Vehicle Management System	ASP.Net 2.0 & SQL Server 2000	Police Motor Transport	2010

Table 1: Various Application Developed by BRAIT

2.1 e – GOV INITIATIVES IN DR. B.R. AMBEDKAR INSTITUTE OF TECHNOLOGY

The Computer Science department of the institute has developed a portal which provides a link to various applications which have been developed in various platforms such as VB6.0, ASP.NET 2.0, JSP, and PHP as front end and SQL Server 2000, Oracle 8i as back end. Different platforms have deliberately been selected to improve the faculty skill set in handling different packages. The portal is available in Local Area Network (LAN) and is deployed in the institute server with proper security. The screen shot of the portal can be seen in Figure 3



Figure 3: Institute Intra Portal

STUDENT RELATED

• **Admission**

- The complete admission process right from the submission of application, scrutiny and preparation of merit list is carried out using this package. The statutory guidelines for admission issued by CBSE and A & N Administration are followed while admitting the students and counseling is carried out in online mode bringing greater transparency in the admission process. The package has a provision to generate a report on the preference of seat for different discipline by the students which is compared with the previous year data for assessment of trend. The personal and academic data of the student admitted is migrated through a different package (Student Information System) for further analysis.

• **Student Information System**

- Student Information System (SIS) is a comprehensive application developed to capture all the inputs of the student's right from its admission to his monthly/semester performance for the complete academic period of their association with the institute. Some of the modules of SIS is as under in Table 2:

Module	Features	User
Semester Registration (Multiple Point Entry & Credit System)	Captures Subject Registered by the student for the semester and generate Bar-coded Challan form for payment of semester fee	Students with the guidance of Academic Member of Department
Attendance	Captures student attendance for the month	Concerned Subject Teacher, access given only for three working days of succeeding month
Marks	Captures student marks (Internal/External)	Internal: Subject Teacher, access given only for nine working days after the conduction of last examination External: Academic Member of Concerned Department
Examination Registration (Regular/Backlog)	Captures Exam Registered by the student for the semester and generate Bar-coded Challan form for payment of semester fee	Students with the guidance of Academic Member of Department
Curriculum	Provides access to curriculum	Student and Staff
Lesson Plan	Provides Lesson Plan of all the subject along with the week map and learning resource used	Staff and Student, (Concerned Subject Teacher uploads the lesson plan)
Skill Map	Captures the skill set of student and make a matrix for identifying the grey area	Student in consultation with Academic Member
Student Notes	Hand written notes on various topics/subjects are made available	Student and Staff (Students are motivated to upload the notes)
Question Paper	Board/University Question Paper of various subjects are made available	Staff and Student
*Co-Curricular	Captures option of various students for participation in various co-curricular activities	<ul style="list-style-type: none"> • Students: For selecting option for participation. • Student Secretary: For preparation of report for conduction of programme. • Cultural Convener: For identifying the student who have not shown interest in

Module	Features	User
		any events

* Trial Stage

Table 2: Modules in Student Information System

The above application is developed in ASP.Net 2.0 as Front End and Oracle 8i as Back End and has been designed to generate various reports which are used to monitor student performance, class performance, the academic performance of the same student throughout the study in the institute, the trend of the class till final year, etc. The students have access to the relevant reports using smart card/enrollment number. The data is also used to carry out result analysis for finding out the root cause of poor result. It also generates reports which are required to be sending to various Govt. Agency, University and Board. The details of some of the reports are as under:

- Compilation of Attendance & Marks and Generation of Defaulter followed by Letter to Parents
- Internal Marks Statement as per University/Board Format
- GPA and CGPA Calculation based on Final Results Sheet
- Result Analysis Report for the Semester, Staff, Subject, Subject Type(Analytical/Descriptive) and Marks Range to identify probable cause of poor results.
- Individual Student/Class Trend
- Component wise fees compilation for transfer into different account heads.

Screen shot of few reports generated is as under in Figure 4 (a) & 4 (b):

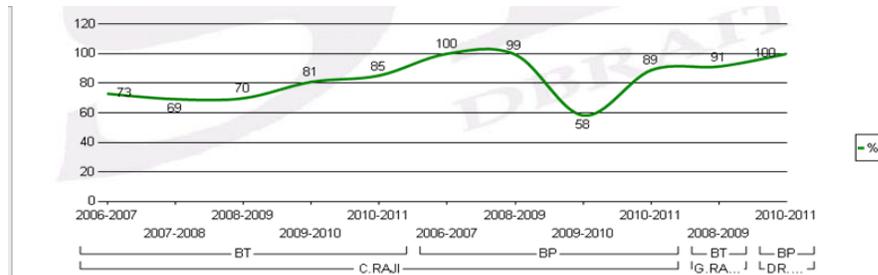


Figure 4 (a): Subject Wise Result Analysis

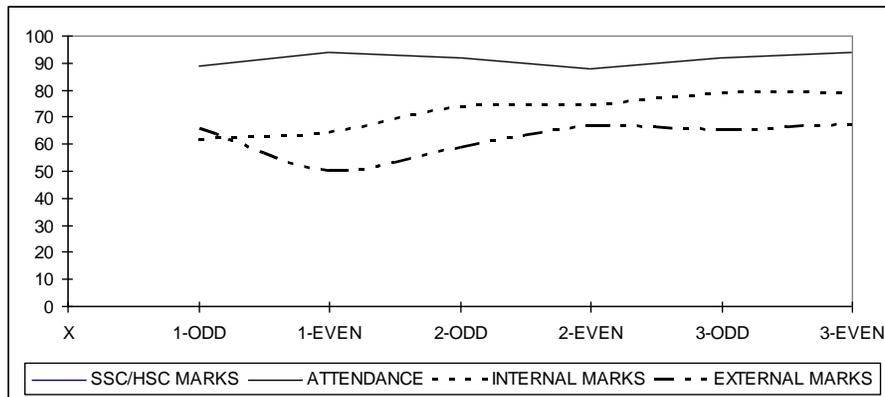
RESULT ANALYSIS OF INDIVIDUAL STUDENT

Enroll No. 08050004

Name of the Student **P.KARTHICK**

Deptt. **DCE**

Attendance/Marks	X	1-ODD	1-EVEN	2-ODD	2-EVEN	3-ODD	3-EVEN
SSCIHSC MARKS	47.06						
ATTENDANCE		89.00	94.00	92.00	88.00	92.00	94.00
INTERNAL MARKS		61.56	63.93	73.59	74.38	78.64	78.82
EXTERNAL MARKS		65.53	50.21	58.75	66.60	65.31	67.00



• Examination Time Table

- The Examination Timetable package is developed in ASP.Net 2.0 as Front End and Oracle 8i as Back End. The main objective of this package is to prepare a timetable both for internal examination and for external examination based on the Subject Registration of the student (Regular/Arrears) paper. The students have an option to use the package and give the choice of the subject where holiday is required before the examination.-

2.2 STAFF RELATED

• Attendance & Pay Roll

This module captures Biometric Finger Print of all the staff for daily attendance and the reports generated showing the monthly attendance both incoming and outgoing time is used for preparation of salary. The application also has a provision to capture incoming and outgoing during the working hours. The application also updates the leave; the software is being further modified for applying various types of leave. The screen shot for monthly attendance statement is as under in Figure 5:

DR.B.RAMBEDKAR INSTITUTE OF TECHNOLOGY
PAHARGAON, PORT BLAIR

STATEMENT OF MONTHLY ATTENDANCE FOR THE MONTH OF : October- November 2011

DEPARTMENT: A.V.CELL

S.No.	Name	Designation	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	01	02	03	04	05	06	07	08	09
1	Sanjeeb Roy	Phy Edn Instructor	0727	0729	0731	0701	0729	0701	S	0727	0722	0707	0707	0709	SAT	S	0727	0715	H	0737	0708	0711	S	0712	0711	0726	CL	0706	0722	S	H	0702	0651
			1618	1533	1542	1553	1534	1718	S	1535	1611	1534	1538	1540	SAT	S	1535	1526	H	1538	1535	1611	S	1539	1530	1526		1544		S	H	1524	1554
2	Pal Swamy	Reprography Attd	0736	EL	EL	EL	EL	S	1023	0748	0747	0822	0726	SAT	S	0738	CL	H	RH	0741	0727	S	0731	0725	0727	0729	0805	0807	S	H	0735	0729	
			1532	1528				S	1529	1528	1529	1529	1528	SAT	S	1526		H		1530	1526	S	1531	1528	1528	1527	1527	1527	S	H	1531	1620	

Figure 5:- Monthly Attendance Statement

• Time Table

The portal also has the linkage with the timetable software which has been developed by the external agency and used by the institute to finalize the semester timetable. The package has built in feature to avoid clash of classroom/lab and teachers. Various reports are generated to get the loading of individual teacher/labs. The screen shot of timetable application is given in Figure 6.

The screenshot shows a web-based timetable application. On the left is a sidebar with filters for 'Classes' (CO -1 EVEN), 'Teachers', 'Classrooms', and 'Summary timetables'. The main area displays a grid for 'CO -I EVEN' classes. The grid has columns for days (1-8) and rows for days of the week (Mo-Sa). Each cell contains a course code and section number, such as 'DT(P)', 'CAAM(P)', 'AM(L)', 'LIB', 'PRG IN C(L)', 'CAAM(L)', 'CSK(L)', 'OS(L)', 'DP-I', 'PSE-(P)', 'AV', 'PE', 'PROG IN C(P)', and 'CULTURAL'. The grid also includes time slots for each day.

Figure 6: Screenshot of Timetable Application

• Question Paper Generation

One of the most time consuming activities of teaching learning process is preparation of question papers and evaluation of answer script. In order to reduce the time in preparation of question paper and maintain required confidentiality, an application has been developed in JSP as Front End and Oracle 8i as Back End, which picks up questions on random basis from the question bank and generates question paper as per the desired format specifying number of questions with 02, 04, 06, 08 marks. The software also supports various figures such as circuit diagram, symbols, mathematical expression, network models, etc. The screen shot of specification table used to select number of question of different marks and for finalizing the format is shown in Figure 7. The question banks are

updated on continuous basis by the subject teacher.

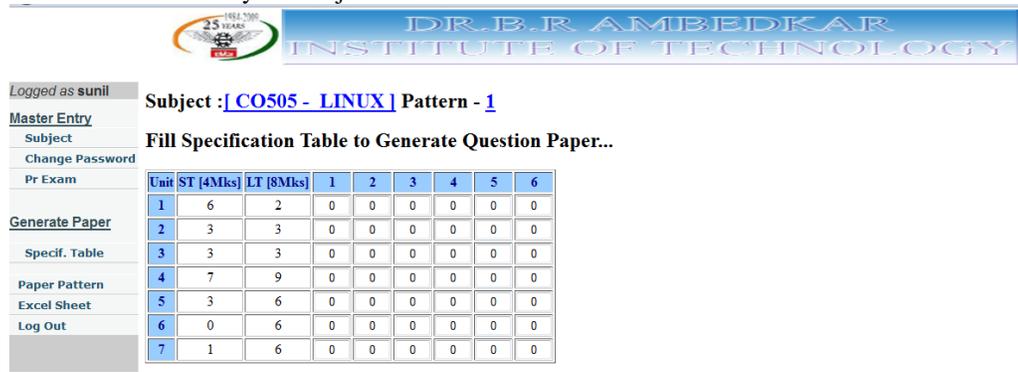


Figure 7: Specification Table Format

• **Staff Welfare Fund**

In order to provide financial support to the supporting staff, all the staff of institute joined together and started contributing fixed amount per month which is now accumulated to Rs. 25 Lakhs. A package has been developed to facilitate the processing of loan request including disbursement of loan. It also monitors the recovery and generates MIS.

2.2 **GENERAL**

All Government Institutes are required to follow certain rules and regulations as laid down by Central/State Government. There are procedures for purchase of goods, service rules governing the service conditions of employees. Similarly private institutes have also framed certain business rules which are followed by the institute. Timely retrieval of important document is also one of the major factor required in decision making by the administrators. This module addresses all these issues.

• **Budget & Expenditure Monitoring**

This application was developed by Computer Science & Engineering Department for A & N Administration way back in 2003. After detailed system study the package was developed to prepare all the bills which are sent to Pay & Accounts Office (Govt. Treasury). This has been implemented by all the DDO's (Drawing & Disbursing Officer) of A & N Islands except Forest Department and APWD since they follow different accounting system. It captures all input data to generate following bills:

- a. Salary Bill
- b. LTC/TA/Leave Salary Advance Bills
- c. Children Education Subsidy/Stipend Bill
- d. Fully Vouched Contingent Bills (All Purchase Bills)

The application developed has a provision to block the fund for which the bill is under process and post the expenditure head wise, scheme wise for all the departments of A & N Administration. It also generates an MIS to keep track on the number of days the bills were held-up at various stages thereby bringing the sense of accountability and transparency. The expenditure report of the institute as on 18th of Nov, 2011 may be seen in Table 3.

2203 **Technical Education**

00105 **Polytechnics**

NON PLAN

	Budget Estimate	Blocked	%	Balance (1-2)	Expenditure	Balance (1-4)	Target Achieved	Bill Process

07.00.01	Salaries	28000000	23184646	82.80	4815354	19797997	8202003	70.71	%	3386649
07.00.03	OTA	50000	19654	39.31	30346	17753	32247	35.51	%	1901
07.00.06	Medical Treatment	50000	17810	35.62	32190	6697	43303	13.39	%	11113
07.00.11	Domestic Travel Expenses	200000	204790	102.40	-4790	189290	10710	94.65	%	15500
07.00.13	Office Expenses	700000	99744	14.25	600256	99744	600256	14.25	%	0
		29000000	23526644	81.13	5473356	20111481	8888519	69.35	%	3415163

PLAN

		Budget Estimate	Blocked	%	Balance (1-2)	Expenditure	Balance (1-4)	Target Achieved	Bill Process	
07.00.01	Salaries	36150000	29653537	82.03	6496463	26043109	10106891	72.04	%	3610428
07.00.02	Wages	7000000	5065750	72.37	1934250	5060012	1939988	72.29	%	5738
07.00.03	OTA	150000	125297	83.53	24703	123797	26203	82.53	%	1500
07.00.06	Medical Treatment	400000	375030	93.76	24970	273558	126442	68.39	%	101472
07.00.11	Domestic Travel Expenses	600000	315059	52.51	284941	191375	408625	31.90	%	123684
07.00.13	Office Expenses	10000000	4608291	46.08	5391709	3697674	6302326	36.98	%	910617
07.00.20	Other Administration Expenses	1500000	1156439	77.10	343561	727213	772787	48.48	%	429226
07.00.28	Professional Services	10000000	6258363	62.58	3741637	6258363	3741637	62.58	%	0
07.00.31	Grant - in - Aid	100000	0	0.00	100000	0	100000	0.00	%	0
07.00.34	Scholarship/ Stipends	2300000	1596099	69.40	703901	1266127	1033873	55.05	%	329972
07.00.50	Other Charges	1400000	949792	67.84	450208	806532	593468	57.61	%	143260
		69600000	50103657	71.99	19496343	44447760	25152240	63.86	%	5655897

4202 ——— Technical Education

02.104 ——— Polytechnics

PLAN

04.00.52	Machinery & Equipments	17000000	7639739	44.94	9360261	6610781	10389219	38.89	%	1028958
		17000000	7639739	44.94	9360261	6610781	10389219	38.89	%	1028958
TOTAL		115600000	81270040	70.30	34329960	71170022	44429978	61.57	%	10100018

Table 3: Expenditure Report of the Institute

• **Purchase and Stores**

This module facilitates the indenting department to enter their requirement in a format (Purchase Requisition Form- PRS) giving detailed specification and approximate rate. The other columns of PRS such as date and quantity of last drawn, Stock Reference are filled from the database. The indenting department submits the hard copy which is signed by all other members and sent to purchase section. Purchase section uses the information available in the soft format and processes further. The software generates Tender Notice by incorporating enlisted vendors if it is a limited tender. In case of Open Tender, the notice is generated in a different format. After receiving a tender, the name of the firm, the rates and other details are fed into the system to generate Techno Commercial Evaluation (TCE) Report. After the scrutiny by the purchase team, system generated purchase order is issued to the vendor. After the goods are received further processing is done in Expenditure Monitoring Module

- **Important Document Finder**

A module has been developed to facilitate the decision maker to retrieve required document using a search engine. The search engine is flexible and it responds to one of the criteria such as a keyword, subject, date etc. The document comprises of .pdf, .doc, or .txt format.

- **e-Circular**

In order to ensure timely communication to all the team members, a module has been developed which has a provision of uploading all the circulars/inter departmental notes. This is accessed by the concerned; there is a provision of searching a circular from the archive using keyword, date, subject or circular number.

3. CONCLUSION

The process of e-Governance in BRAIT was started 10 years ago following a structured approach of process mapping, process re-engineering followed by computerization. The package has undergone continuous revision with respect to platform as well as features. The benefits accrued by the above e-Governance initiatives are as under: -

- Improving the work culture of the institute and self discipline among the staffs
- Capacity building of staffs
- Institute became a resource centre for supporting IT initiatives of A & N Administration
- Enhanced employment opportunities for the students due to better skill sets
- System driven with no scope for adhoc decisions in routine academic matters

Andaman & Nicobar Islands being geographically isolated from mainland India poses a challenge in adopting best practices which is otherwise prevalent in other institutes of our country. The BRAIT team has successfully transformed the challenge into opportunity and implemented the e-GOV initiatives in the institute with no additional expenditure.

Optimal Power Flow Using Differential Evolution Algorithm With Conventional Weighted Sum Method

Rohit Kumar Verma¹, Himmat Singh² and Laxmi Srivastava³

^{1, 2, 3} Department Of Electrical Engineering, Madhav Institute Of Technology And Science, Gwalior

Abstract

Optimal reactive power dispatch is one of most important task in the today's power system operation. This paper present optimal reactive power dispatch with the help of differential evolution algorithm. The optimal reactive power dispatch is a non linear constraints multi objective optimization problem where the real power loss, voltage deviation and fuel cost are to be minimized under control and dependent variables.. Reactive power optimization is a mixed integer nonlinear optimization problem which includes both continuous and discrete control variables. The suggested algorithm is used to find the setting of control variables, such as voltage, transformer tap position and reactive compensation devices to optimize a certain objective. A Differential Evolution Algorithm based approach is proposed to handle the problem as a true multi-objective optimization problem. The standard IEEE 30-bus test system is used and the results show the effectiveness of Differential Evolution Algorithm and confirm its potential to solve the multi-objective optimal reactive power dispatch problem. The results obtained by Differential Evolution Algorithm are compared and validated with conventional weighted sum method to show the effectiveness of the proposed algorithm.

Keywords: - Differential evolution algorithm, Power loss minimization, Voltage deviation, Multi-objective weighted sum method.

I. Introduction

Optimization concerns to the analyze of problems where an objective function is minimized or maximized by consistently choosing the values of real and/or integer variables from within an permitted set. Many real world and theoretical problems may be simulated in the general model of an optimization problem. Power

System is one of the composite fields in electrical engineering, where optimization plays a significant role [1].

Reactive power optimization [2] is one of the difficult optimization problems in power system operation and control. To ameliorate the voltage profile and to decrement the active power losses along the transmission lines beneath various operating conditions, power system operator can choose a number of control tools such as switching reactive power sources, charging generator voltages and adjusting transformer tap settings. Therefore, the trouble of the reactive power bump off can be optimized to ameliorate the voltage profile and minimize the system losses as well [3].

It is a non- linear optimization trouble and various numerical proficiencies have been assumed to solve this optimal reactive power dispatch trouble [5]. These include the gradient method [6-7], Newton method [8] and linear programming [9].The gradient and Newton methods endure from the difficulty in

handling inequality constraints. To employ linear programming, the input output function is to be conveyed as a set of linear functions, which may lead to loss of accuracy. Recently, worldwide optimization proficiencies such as genetic algorithms have been suggested to figure out the reactive power optimization problem [10-11]. Genetic algorithm is a random search technique based on the mechanics of natural selection [12].In GA-based RPD problem it starts with the randomly generated population of points, ameliorates the fitness as generation continues through the application of the three operators-selection, crossover and mutation. But in the recent research some inadequacies are identified in the GA execution. This debasement in efficiency is apparent in applications with highly hypostasis objective functions i.e. where the parameters being optimized are highly correlated. In addition, the untimely convergence of GA degrades its performance and abbreviates its search capability. In addition to this, these algorithms are found to take more time to reach the optimal solution.

The reactive power planning is one of the most significant and ambitious problems because it has lots of objective functions like Voltage Deviation, Real power loss and installation cost of the reactive power sources is to be minimized at the same time [13]. The number of variables and parameters of the objective functions are optimized for figuring out the reactive power compensation problem. To solve the reactive power compensation troubles various numerical troubles are formulated.

When an optimization trouble calls for more than one objective function, the project of determining one or more optimum solutions is known as multi-objective optimization [2].

One major trouble consociated with suggesting a multiobjective problem as single objective trouble is that an optimal solution may be extremely contingent on how the weights are set [14-23]. This can be of great business in cases where weights are randomly assigned. To overcome this problem weighted sum methods are to be proposed.

In this paper, the Differential Evolution Algorithm (DE) combine with weighted sum method based approach is proposed for solving the multiobjective VAR dispatch optimization problem [15]. The problem is formulated as a nonlinear constrained multiobjective optimization problem where the real power loss and the bus voltage deviations are treated as contending objectives. A hierarchical bunching technique is carried out to allow for the power system operator with a representative and accomplishable Pareto optimal set. The strength and potential of the proposed approach to figure out the multiobjective VAR dispatch problem are presented.

The new approach introduced in this paper to solve the reactive power compensation problem is based on Differential Evolution Algorithm, proposed by Storn and Price in 1995, a variant of the Differential Evolution Algorithm (DE) that has

tested to be a very competitory algorithm solving several optimization problems [16-17].

II. Problem Formulation

A. Minimization Of System Power Losses

The RPD problem aims at minimizing the real power loss in a power system while satisfying the unit and system constraints. This goal is achieved by proper adjustment [18-19] of reactive power variables like generator voltage magnitudes (V_{Gi}), reactive power generation of capacitor banks (Q_{Ci}) and transformer tap settings (T_k).

This is mathematically stated as:

$$\text{Min } F_1 = P_{Loss} = \sum_{k \in N_l} \sum_{i,j=(i,j)} g_k (V_i^2 + V_j^2 - 2V_i V_j \cos(\theta_i - \theta_j))$$

The real power loss given by (1) is a non-linear function of bus voltages and phase angles which are a function of control variables.

B. Voltage Profile Improvement (Voltage Deviation)

Bus voltage is one of the most important security and service quality indices. Improving voltage profile can be obtained by minimizing the load bus voltage deviations from 1.0 per unit. The objective function can be expressed as:

$$\text{Min } F_2 = VD = |V_i - 1.0| \quad (2)$$

C. Minimization Of Fuel Cost

The objective of the ELD is to minimize the total system cost by adjusting the power output of each of the generators connected to the grid. The total system cost is modeled as the sum of the cost function of each generator (1). The generator cost curves are modeled with smooth quadratic functions, given by:

$$\text{Min } F_3 = F_{cost} = \sum_{i=1}^{NG} a_i + b_i P_{Gi} + c_i P_{Gi}^2 \quad (\$/h) \quad (3)$$

Where NG is the number of online thermal units, P_{Gi} is the active power generation at unit i and a_i , b_i and c_i are the cost coefficients of the i^{th} generator.

iii: System Constraints

The real power loss given by (1) is a non-linear function of bus voltages and phase angles which are a function of control variables. The minimization problem is subjected to the following equality and inequality constraints:

A. Equality Constraints (Load flow constraints)

$$P_{Gi} - P_{Di} - V_i \sum_{j=1}^{N_B} V_j (G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)) = 0, \quad (4)$$

$i = 1, 2, \dots, N_B$

$$Q_{Gi} - Q_{Di} - V_i \sum_{j=1}^{N_B} V_j (G_{ij} \sin(\theta_i - \theta_j) - B_{ij} \cos(\theta_i - \theta_j)) = 0, \quad (5)$$

$i = 1, 2, \dots, N_{PQ}$

B. Inequality constraints includes

• Voltage constraints:

$$V_i^{\min} \leq V_i \leq V_i^{\max} ; i \in N_B \quad (6)$$

• Generator reactive power capability limit:

$$Q_{Gi}^{\min} \leq Q_{Gi} \leq Q_{Gi}^{\max} ; i \in N_G \quad (7)$$

• Reactive power generation limit of capacitor banks:

$$Q_{Ci}^{\min} \leq Q_{Ci} \leq Q_{Ci}^{\max} ; i \in N_C \quad (8)$$

• Transformer tap setting limit:

$$T_k^{\min} \leq T_k \leq T_k^{\max} ; i \in N_T \quad (9)$$

• Transmission line flow limit

$$S_l \leq S_l^{\max} ; l \in N_l \quad (10)$$

• Generation capacity constraint

$$P_{Gi}^{\min} \leq P_{Gi} \leq P_{Gi}^{\max} ; i \in N_B \quad (11)$$

• Power balance constraint

(1) The total power generation must cover the total demand P_D and the real power loss in transmission lines P_L . This relation can be expressed as:

$$\sum_{i=1}^{NB} P_{Gi} = P_D + P_L \quad (12)$$

III. DIFFERENTIAL EVOLUTION ALGORITHM

In 1995, Storn and Price proposed a new floating point encoded evolutionary algorithm for global optimization and named it differential evolution (DE) algorithm owing to a special kind of differential operator, which they invoked to create new off-spring from parent chromosomes instead of classical crossover or mutation [20]. DE algorithm is a population based algorithm using three operators; crossover, mutation and selection. Several optimization parameters must also be tuned. These parameters have joined together under the common name control parameters. In fact, there are only three real control parameters in the algorithm, which are differentiation (or mutation) constant F , crossover constant CR , and size of population NP . The rest of the parameters are dimension of problem D that scales the difficulty of the optimization task; maximum number of generations (or iterations) GEN , which may serve as a stopping condition; and low and high boundary constraints of variables that limit the feasible area [20]. DE works through a simple cycle of stages, presented in Fig. 1.

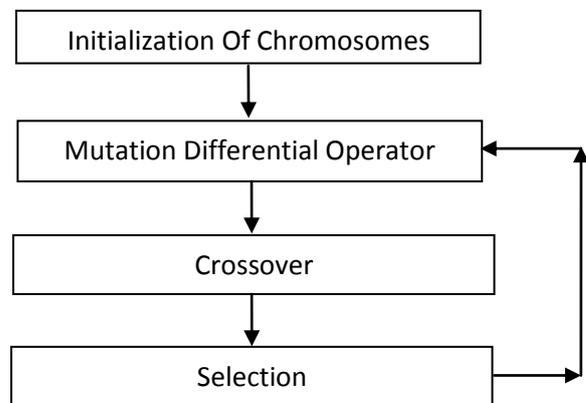


Fig-1 DE Process cycle

These stages can be cleared as follow:

A. Initialization

At the very beginning of a DE run, problem independent variables are initialized in their feasible numerical range. Therefore, if the j^{th} variable of the given problem has its lower and upper bound as x_j^L and x_j^U , respectively, then the j th component of the i^{th} population members may be initialized as:

$$x_{ij}(0) = x_j^L + \text{rand}(0,1) \cdot (x_j^U - x_j^L) \quad (13)$$

Where $\text{rand}(0, 1)$ is a uniformly distributed random number between 0 and 1.

B. Mutation

In each generation to change each population member, a donor vector is created. It is the method of creating this donor vector, which demarcates between the various DE schemes. However, in this paper, one such specific mutation strategy known as DE/rand/1 is discussed. To create a donor vector for each i^{th} member, three parameter vectors and are chosen randomly from the current population and not coinciding with the current x_i . Next, a scalar number F scales the difference of any two of the three vectors and the scaled difference is added to the third one whence the donor vector is obtained. The usual choice for F is a number between 0.4 and 1.0. So, the process for the j th component of each vector can be expressed as:

$$v_{ij}(t + 1) = x_{r1j}(t) + F \cdot (x_{r2j}(t) - x_{r3j}(t)) \quad (14)$$

C. Crossover

To increase the diversity of the population, crossover operator is carried out in which the donor vector exchanges its components with those of the current member $x_i(t)$. Two types of crossover schemes can be used with DE technique. These are exponential crossover and binomial crossover. Although the exponential crossover was proposed in the original work of Storn and Price [20], the binomial variant was much more used in recent applications [21]. In this paper, binomial crossover scheme is used which is performed on all D variables and can be expressed as:

$$u_{ij}(t) = \begin{cases} v_{ij}(t) & \text{if } \text{rand}(0,1) < CR \\ x_{ij}(t) & \text{else} \end{cases} \quad (15)$$

Where $u_{ij}(t)$ represents the child that will compete with the parent $x_{ij}(t)$.

D. Selection

To keep the population size constant over subsequent generations, the selection process is carried out to determine which one of the child and the parent will survive in the next generation, i.e., at time $t = t + 1$. DE actually involves the Survival of the fittest principle in its selection process. The selection process can be expressed as:

$$\bar{x}_i(t + 1) = \begin{cases} \bar{u}_i(t) & \text{if } f(\bar{u}_i(t)) \leq f(\bar{x}_i(t)) \\ \bar{x}_i(t) & \text{if } f(\bar{x}_i(t)) < f(\bar{u}_i(t)) \end{cases} \quad (16)$$

Where $f()$ is the function to be minimized. So, if the child yields a better value of the fitness function, it replaces its parent in the next generation; otherwise, the parent is retained in the population. Hence the population either gets better in terms of the fitness function or remains constant but never deteriorates

Flow chart for steps follows for calculation in Differential Evolution Algorithm

IV. RESULTS AND DISCUSSION

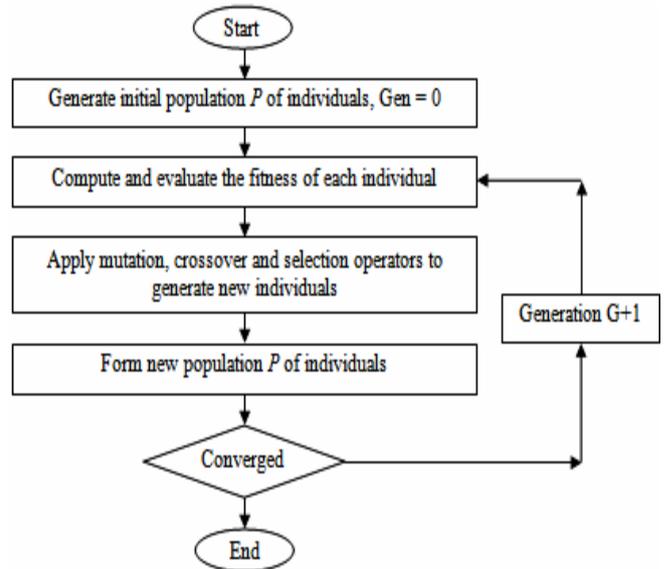


Fig: - 2 Flow Chart Of Algorithm

All the coding for the proposed algorithm using DE was done in MATLAB 7.8.0.347 (2009A) running on Pentium(R) Dual-Core CPU E5200, 2.50GHz, 1.00 GB. Simulation is carried out on IEEE 30 bus test system. In order to compare the results obtained from DE, the optimization problem is also solved using a conventional technique. The all data is and initial value is taken from [22]

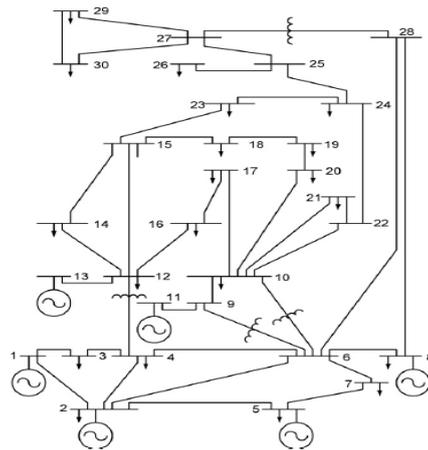


Fig:- 3 One Line Diagram Of Ieee-30 Bus System

Table 1:- Parameter used in calculation

No. Of Population	06
No. Of Iteration	200
Crossover	0.8
Mutation	0.8

Table 2:- Optimal result when objective function taken as multi-objective. See Fig: - 4

P_{loss} vs VD	
W₁= 0.8184	
P _{loss}	VD
5.3378	0.2483

Table 3:- Best results of individually run of Ploss and Voltage Deviation as main function

	Min	Max	Initial	Ploss	VD
V ₁	0.95	1.1	1.04	1.10	1.00
V ₂	0.95	1.1	1.01	1.09	1.02
V ₅	0.95	1.1	1.01	1.07	1.01
V ₈	0.95	1.1	1.05	1.08	1.01
V ₁₁	0.95	1.1	1.05	1.04	1.09
V ₁₃	0.95	1.1	1.07	1.10	1.06
T ₁₁	0.9	1.1	1.06	1.03	0.98
T ₁₂	0.9	1.1	1.03	0.95	0.95
T ₁₅	0.9	1.1	1.06	1.06	1.03
T ₃₆	0.9	1.1	0.0	1.00	1.00
QC ₁₀	0.0	5.0	0.0	5.0	2.67
QC ₁₂	0.0	5.0	0.0	5.0	5.00
QC ₁₅	0.0	5.0	0.0	5.0	5.00
QC ₁₇	0.0	5.0	0.0	1.10	1.38
QC ₂₀	0.0	5.0	0.0	2.30	3.38
QC ₂₁	0.0	5.0	0.0	1.41	5.00
QC ₂₄	0.0	5.0	0.0	4.45	4.22
QC ₂₉	0.0	5.0	0.0	2.49	5.00
P loss (MW)			5.85	4.76	6.40
VD			1.16	1.01	0.19
@=See Fig :- 5	\$= See Fig :- 6		@	\$	

To demonstrate the effectiveness of the proposed algorithm, these different cases have been considered as follows:

Case 1:- Minimization of power loss and voltage deviation using weighted sum method considered as multi-objective.

Case 2:- Individual minimization of system power loss and voltage deviation

Case 1:- Minimization of power loss and voltage deviation using weighted sum method considered as multi-objective

The problem was handled as a multi objective optimization problem where both power loss and voltage deviations were optimized simultaneously with the proposed approach. For completeness and comparison purposes, the problem was also treated as a single objective optimization problem by linear combination of PL and VD is considered:-

$$\text{Minimize } f = W_1 * f_1 + (1 - W_1) * f_2 \quad (17)$$

W₁ is a weighting factor. To generate 41 nondominated solutions, the algorithm was applied 41 times with varying W₁ as a random number W₁= rand [0, 1]. The best P_{loss} and best VD solutions are given in Table 2.

Case2:- Individual minimization of system power loss and voltage deviation

At first, the P_{loss} and VD objectives are optimized individually in this case we can take power loss and voltage deviation as a main function simultaneously. The best results of P_{loss} and VD functions when optimized individually are given in Table 3.

Table 2:- Optimal result when objective function taken as multi-objective. See Fig: - 4

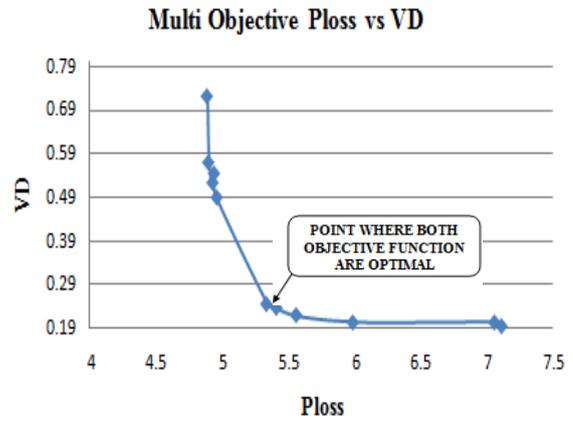


Fig 4 Pareto optimal graph between VD

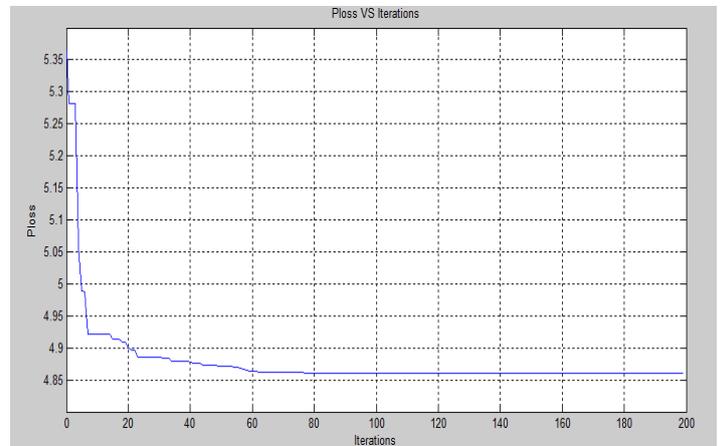


Fig 5 Graph between Ploss and No. of iterations

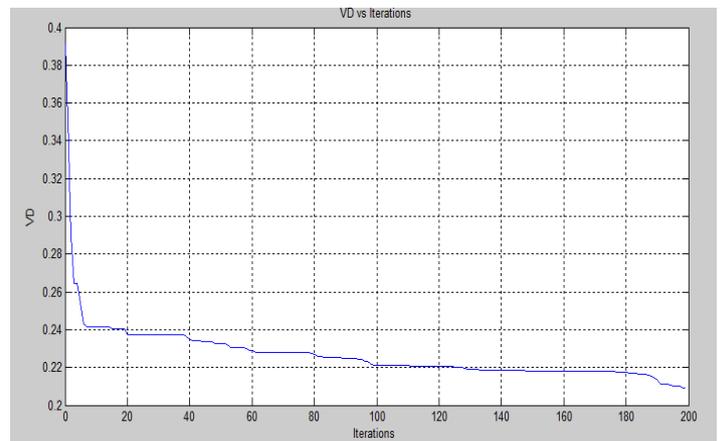


Fig 6 Graph between VD and No. of iterations

Conclusion

In this paper differential evolution algorithm has been proposed and successfully applied to solve the optimal power flow problem for multiobjective. For solving the optimal power flow problem we can consider two objective functions these are Ploss and voltage deviation. Result shows that differential evolution based optimal power flow algorithm is adequate to minimize the P_{loss} and VD in the system. The proposed approach has been tested on standard IEEE-30 bus system. The Pareto graph proves the effectiveness of this proposed algorithm.

Acknowledgement

The authors are thankful to Director, HOD (Electrical Department) Madhav Institute of Technology and Science, Gwalior (M.P.) India for providing support and facilities to carry out this research work

REFERENCES

- [1]. M. Varadarajan Optimal “Power Flow Solution Using Differential Evolution”, In November 2007
- [2]. P. Aruna Jeyanthi, D. Devaraj, “Multi-Objective Genetic Algorithm For Reactive Power Optimization Including Voltage Stability” In:2010 International Journal of Engineering Science and Technology Vol. 2 (7), 2010, 2715-2729
- [3]. Ramesh Subramanian, Kannan Subramanian and Baskar Subramanian, “Application of a Fast and Elitist Multi-Objective Genetic Algorithm to Reactive Power Dispatch” , in:2009 Serbian Journal Of Electrical Engineering vol. 6, no. 1, May 2009, 119-133
- [4]. K.R.C. Mamandur, R.D. Chenoweth: Optimal Control of Reactive Power Flow for Improvements in Voltage Profiles and for Real Power Loss Minimization, IEEE Transactions on Power Apparatus and Systems, Vol. PAS-100, No. 7, 1981, pp. 3185-3194.
- [5]. P.Aruna Jeyanthi and Dr.D.Devaraj, “Hybrid Particle Swarm Optimization fo Multi-objective Reactive Power Optimization with Voltage Stability Enhancement”, in:2010 Control, Communication and Power Engineering 2010
- [6]. Lee K Y ,Paru Y M ,Ortiz J L –A united approach to optimal real and reactive power dispatch , IEEE Transactions on power Apparatus and systems 1985: PAS-104 : 1147-1153
- [7]. A.Monticelli , M .V.F Pereira ,and S. Granville , “Security constrained optimal power flow with post contingency corrective rescheduling” , IEEE Transactions on Power Systems :PWRS-2, No. 1, pp.175-182.,1987.
- [8]. Deeb N, Shahidehpur S.M, Linear reactive power optimization in a large power network using the decomposition approach. IEEE Transactions on power system 1990: 5(2) : 428- 435
- [9]. Devaraj, and B. Yegnanarayana, “Genetic algorithm based optimal power flow for security enhancement”, IEEE proc- Generation. Transmission and Distribution; 152, 6 November 2005.
- [10]. Deb, K. (201): Multi – objective optimization using evolutionary algorithms 1st ed. (John Wiley & Sons, Ltd.).
- [11]. Kwang Y. Lee and Frank F.Yang, “Optimal Reactive Power Planning Using evolutionary Algorithms: A Comparative study for Evolutionary Strategy, Genetic Algorithm and Linear Programming”, IEEE Trans. on Power Systems, Vol. 13, No. 1, pp. 101-108, February 1998.
- [12]. D Goldberg, “Genetic algorithms in search, optimization and machine learning”, Addison-Wesley,1989.
- [13]. S. Jaganathan, S. Palaniswami and G. Maharaja Vignesh, “Applications of Multi Objective Optimization to Reactive Power Planning Problem Using Ant Colony Algorithm”, in:2011 European Journal of Scientific Research ISSN 1450-216X Vol.51 No.2 (2011), pp.241-253
- [14]. Steven M. Small, and Benjamin Jeyasurya, Senior Member, IEEE, “Multi-Objective Reactive Power Planning: A Pareto Optimization Approach”, in:2007 The 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007
- [15]. M. A. Abido J. M. Bakhshwain, “A Novel Multiobjective Evolutionary Algorithm For Optimal Reactive Power Dispatch Problem”, in:2003 ICECS-2003 0-7803-8163-7/03/\$17.00 © 2003 IEEE
- [16]. P. Gardel, B Barán, H Estigarribia, U Fernández, S. Duarte, “Multiobjective Reactive Power Compensation with an Ant Colony Optimization Algorithm”, National University of Asuncion, Paraguay
- [17]. Dorigo M. and Di Caro G., “The Ant Colony Optimization Meta-heuristic” Ei David Corne, Marco Dorigo and Fred Glover, editors, New Ideas in Optimization, pages 11-32. McGraw-Hill, London, 1999.
- [18]. R. Storn, K. Price, “Differential evolution—a simple and efficient adaptive scheme for global Optimization over continuous spaces”, in: Technical Report TR-95-012, ICSI, 1995.
- [19]. Zhenyu Yang, Ke Tang, Xin Yao,” Differential evolution for high-dimensional function optimization”, in: IEEE Congress on Evolutionary Computation (CEC2007), 2007, pp. 3523–3530.
- [20]. R. Storn, K. Price, “Differential Evolution – A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces”, Technical Report TR- 95-012, ICSI, 1995.
- [21]. J. Vesterstrøm, R. Thomsen,” A comparative study of differential evolution, particle swarm optimization, and evolutionary algorithms on numerical benchmark problems”, in: IEEE Congress on Evolutionary Computation, 2004, pp. 980–987.
- [22]. M.A. Abido, “Multiobjective optimal VAR dispatch using strength pareto evolutionary algorithm”, in:2006 IEEE Congress on Evolutionary Computation, Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada, July 16–21, 2006.
- [23]. B. Fogel, Z. Michalewicz, How to solve it: Modern Heuristics, NewYork, NY, Springer, 2004.

Minimization of Reactive Power Using Particle Swarm Optimization

¹Vivek Kumar Jain, ²Himmat Singh, ³Laxmi Srivastava

^{1,2,3}Department of Electrical Engineering, Madhav Institute of Technology and Science, Gwalior (M.P.), India

Abstract— This paper presents an efficient and reliable Particle Swarm Optimization (PSO) algorithm for solving Reactive power optimization including voltage deviation in Power System. Voltage deviation is the capability of a power system to maintain up to standard voltages at all buses in the system under standard conditions and under being subjected to a disturbance. Reactive power optimization is a complex combinatorial programming problem that reduces power losses and improves voltage profiles in a power system. To overcome this shortcoming, a multi-objective particle swarm optimization is proposed and applied in reactive power optimization on IEEE-30 bus. Here the RPO problem has been formulated as a constrained multi-objective optimization problem by combining of two objective functions (real power loss and voltage profile improvement) linearly shows that the particle swarm optimization more effectively solve the reactive power optimization problem in power system.

Keywords — Reactive power optimization, multi-objective particle swarm optimization, voltage deviation and loss minimization.

I. INTRODUCTION

The difficulty of reactive power optimization is in a straight line concerned not only with service excellence and reliability of supply, but also with financial system and safety of the power systems. It is of large importance to preserve suitable voltage levels at all power system buses, since every part of current day equipments which extend electric power such as illumination; thermal appliances, electronic appliances and motors, are designed for use surrounded by a definite workstation voltage, the nameplate voltage. If the voltage deviates from this value, the helpfulness, life suspense, and the superiority of performance of the equipment will suffer. Some electrical equipment is more sensitive to voltage variations than others such as motors. Even though the fact that more than a few voltage-deviation techniques are accessible to electric power system operational staff, power systems are silent subjected to voltage instabilities and in some belongings to voltage collapses that could lead to unexpected system breakdowns. Reactive power organize has become an important aspect for many reasons.

The need for most efficient operation of power systems has increased with price of fuel. For a given distribution of power, the losses in the system can be reduced by minimizing the flow of reactive power. In many case power transmitted through older circuit has

been increased, requiring the application of reactive power control measures to restore stability margins.

Different techniques have been descriptions [18 - 19] to assess voltage deviation of power systems to find the possible ways to improve the voltage stability boundary. Voltage is considered as one of the most important parameters of the quality of power supply. Its deviation from the normal value may be damaging and luxurious. Therefore, the RPO problem is large-scale extremely constrained nonlinear non-convex optimization difficulty [1]. In this paper, the main concerns are proper planning and organization of control variables which are whichever transformer tap changers, shunt capacitors, generators reactive Vars in an interconnected power system such that real power becomes least. [4-9]. Here the RPO problem has been formulated as a constrained multi-objective optimization problem by combining of two objective functions (real power loss and voltage profile improvement) linearly. Usually, PSO has a more global searching ability at the found of the sprint and a local search in close proximity to the end of the sprint. Therefore, while solving problems with more some degree of optima, there are more potential for the PSO to find out local optima at the finish of sprint. However, the reactive power optimization problem does have these properties itself. For these reasons, a reliable global advance to power system optimization problems would be of significant value to power engineering civilization. The problem of reactive power planning in a power system can be exposed to be a combinatorial optimization problem through a number of methods have been proposed to solve the problem using the particle swarm optimization algorithm. Compared with other optimization techniques, particle swarm optimization has comparable or even superior search performance for some hard optimization problems in real power systems [2-3].

PSO was applied to different areas of power systems. It was used to optimize the reactive power flow in the power system network to minimize real power system losses [10]. Several evolutionary algorithms such at the same time as genetic algorithms (GA) [13-16]; ant colony optimization (ACO) [11-12] have been used for optimization of structures. These optimization algorithms are all the rage and widely used due to their high possible for modeling engineering problems and simple programming in computers. These optimization algorithms have many similarities. All of them discover the aim freedom by a population of potential designs using some simulation development operators with accidental environment. A power system needs to be with sufficient reactive coffers to meet the improved reactive power demand under seriously

loaded conditions and to avoid voltage instability problems.

The proposed come within reach of has been examined and tested on the standard IEEE 30-bus test system with different objectives that reproduce voltage profile improvement, and voltage deviation improvement. Leaving presented an A Reactive Power Optimization Solution Based on multi-objective Particle Swarm Optimization in Power Systems [17]. The optimality of the proposed approach has been tested by comparing the results obtained by other evolutionary algorithms.

II. PROBLEM FORMULATION

The aim of reactive power optimization problem is to minimize the power losses in the transmission network and develop voltage quality. The control variables are generators bus voltages, transformer tap positions and switch-able shunt capacitor banks.

The equality constraints are power/reactive power equalities, the inequality constraints include bus voltage constraints, generator reactive power constraints, reactive source reactive power capacity constraints and the transformer tap position constraints, etc [12]. The equality constraints can be automatically satisfied by load flow calculation, while the lower/upper limit of control variables corresponds to the coding on the Particle Swarm optimization so the inequality constraints of the control variables are satisfied which can be described as follows:

Objective Function:

$$F = F_1 + F_2 = P_{Loss} + VD \quad (1)$$

Where,

P_L = network real power loss

nl = Number of line

VD = Voltage deviation

This is mathematically stated as:

$$F_1 = P_{Loss} = \sum_{k=1}^{nl} G_k [V_i^2 + V_j^2 - 2V_i V_j \cos \theta_{ij}] \quad (2)$$

Constraints:

1. Real Power Constraints:

$$P_{Gi} - P_{Di} - V_i \sum_{j=1}^{N_B} V_j (G_{ij} \cos(\theta_i - \theta_j) + B_{ij} \sin(\theta_i - \theta_j)) = 0, \quad i = 1, 2, \dots, N_B \quad (3)$$

2. Reactive Power

$$Q_{Gi} - Q_{Di} - V_i \sum_{j=1}^{N_B} V_j (G_{ij} \sin \theta_{ij} - B_{ij} \cos \theta_{ij}) = 0,$$

Constraints:

$$i = 1, 2, \dots, N_B \quad (4)$$

Where,

V_i = Voltage magnitude at bus I

V_j = Voltage magnitude at bus j

P_i, Q_i = Real and reactive powers injected into network at bus i

G_{ij}, B_{ij} = Mutual conductance and susceptance between bus i and bus j

Q_{gi} = Reactive power generation at bus i

N_B = Total number of buses excluding slack bus

N_{PQ} = Number of PQ buses

θ_{ij} = Voltage angle difference between bus i and bus j

3. Bus Voltage magnitude constraints :

$$V_i^{\min} \leq V_i \leq V_i^{\max} ; i \in N_B \quad (5)$$

4. Transformer Tap position constraints :

$$t_k^{\min} \leq t_k \leq t_k^{\max} ; i \in N_T \quad (6)$$

5. Generator bus reactive power constraint:

$$Q_{gi}^{\min} \leq Q_{gi} \leq Q_{gi}^{\max} ; i \in N_g \quad (7)$$

6. Reactive power source capacity constraints

$$Q_{ci}^{\min} \leq Q_{ci} \leq Q_{ci}^{\max} ; i \in N_c \quad (8)$$

7. Transmission line flow constraints:

$$|s_l| \leq |s_l^{\max}|$$

$$S_l \leq S_l^{\max} ; l \in N_l \quad (9)$$

The symbol used is follows:

t_k = Tap setting of transformer at branch k

Q_{ci} = Reactive power generated by i^{th} capacitor bank

Q_{gi} = Reactive power generation at bus i

S_l = Apparent power flow through the i^{th} branch

N_B = Total number of buses

G_k = Conductance of buses

N_T = Number of tap-setting transformer branches

N_c = Number of capacitor banks

N_g = Number of generator buses

State variable are restricted by adding them as a quadratic penalty terms to the objective function. Therefore the equation (1) is changed to the following form:

$$\text{Min. } F_T = f + K_v \sum_{i=1}^{N_{PQ}} (V_i - V_i^{\text{lim}})^2 + K_q \sum_{i=1}^{N_g} (Q_{gi} - Q_{gi}^{\text{lim}})^2 + K_f \sum_{i=1}^{N_b} (S_i - S_i^{\text{lim}})^2 \quad (10)$$

Where K_v , K_q and K_f are the penalty factors for the bus voltage limit violations, generator reactive power limit violations and line flow violations respectively.

$$X_i^{\text{lim}} = X_i^{\text{max}} \text{ if } X_i > X_i^{\text{max}} \quad (11)$$

$$X_i^{\text{lim}} = X_i^{\text{min}} \text{ if } X_i < X_i^{\text{min}} \quad (12)$$

$$F = \frac{k}{F_T} \quad (13)$$

Where k is a large constant this is used to amplify $1/F_T$ the value of which is usually small, so that the fitness value of the chromosome will span a wider range.

III. REACTIVE POWER OPTIMIZATION WITH VOLTAGE PROFILE

The significance of reactive power supervision increases regularly, in the direction of increasing reactive power require. Here the RPO problem has been formulated as a constrained particular objective optimization problem by combining of two objective functions (real power loss and voltage profile improvement) linearly. Voltage is very significant in power management; as it must be high sufficient to support loads and must be in the neighborhood of to the position plenty not to cause any responsibility of equipment. Hence, voltage must be controlled from every position and should be maintained. This can be recognized to a great amount by controlling reactive power utilization and resources.

The controllable devices such as generator, capacitor, and reactor devices are used for decreasing the loss and increasing the voltage control in reactive power optimization. At the same time, these devices consist of constraints for the optimization problem. In this study, there are three object function optimization called RPOVD. These functions are: active power loss, voltage profile of load buses and penalty function of reactive power sources. Generally V_{ref} . Is taken as 1.0 p.u.

$$F_2 = V_{dev} = \frac{\sum_{i \in N_{PQ}} |V_i - V_{ref}|}{N_{PQ}} \quad (14)$$

Where,

V_{dev} = Voltage deviation

V_{ref} = load bus reference voltage value.

V_i = load bus voltage

N_{PQ} = load bus number

IV. PARTICLE SWARM OPTIMIZATION APPROACH

PSO algorithm, originally introduced by Kennedy and Eberhart (1995). Similar to evolutionary algorithm, the PSO technique conducts searches using a population of particles, corresponding to individuals. Each particle represents a candidate solution to the reactive power problem. In a PSO system, particles change their positions by flying around in a multidimensional search space until a relatively unchanged position has been encountered, or until computational limitations are exceeded. In social science context, a PSO system combines a social-only model and a cognition-only model. These particles are randomly initialized and freely fly diagonally the multidimensional investigate space. Throughout flight, each particle updates its own velocity and position based on the best experience of its own and the complete population. The social-only component suggests that individuals ignore their own experience and adjust their behavior according to the successful beliefs of the individual in the neighborhood. On the other hand, the cognition-only component treats individuals as isolated beings. A particle changes its position using these models. [21-22] The PSO system simulates the knowledge evolution of a Social organism, in which N individuals, a potential Solution to a problem is represented as a particle flying in D -dimensional search space, with the position vector $X_i = (X_{i1}, X_{i2}, X_{i3}, \dots, X_{iD})$ and velocity $V_i = (V_{i1}, V_{i2}, V_{i3}, \dots, V_{iD})$ Each particle records its best previous position (the position giving the best fitness value) as $P_{besti} = P_{besti1}, P_{besti2}, \dots, P_{bestid}$ called personal best position. The global version of the PSO keeps track of the overall best value (g^{best}), and its location, obtained thus far by any particle in the population. At each iteration, each particle competes with the others in the neighborhood or in the whole population for the best particle (with best fitness value among neighborhood or the population) with best position $g^{besti} = g^{besti1}, g^{besti2}, \dots, g^{bestid}$ called global best position. Each particle tries to modify its position using the following information:

The velocity and position of all particles are randomly set to within pre-defined ranges. And velocity updating (fig 1). At each iteration, the velocities of all particles are updated according to:

$$V_{id} = w_i v_{id}^k + C_1 \text{rand}_1 (P_{id} - X_{id}) + C_2 \text{rand}_2 (P_{gd} - X_{id}) \quad (15)$$

And

$$w_i = w_{\text{max}} - \frac{w_{\text{max}} - w_{\text{min}}}{\text{iter}_{\text{max}}} \times \text{iter} \quad (16)$$

i = Index of particle, $i \in \{1, \dots, n\}$, n = Population size

d = Dimension, $d \in \{1, \dots, N\}$

X_{id} = present position of particle i on dimension d

C_1 = Self confidence factor

C_2 = Swarm confidence factor

P_{id} = Personal best

P_{gd} = Global best

W= Inertia weight

w_{max} =initial weight

w_{min} = final weight

I_{ter} = current iteration

$Iter_{max}$ = maximum iteration

The use of the inertia weight w has provided improved performance in a number of applications. As originally developed, w often is decreased linearly from about 0.9 and 0.4 during a run. Suitable collection of the inertia weight provides a balance between global and local examination, and results in less iteration on average to find a sufficiently optimal solution. C_1 and C_2 are known as learning factors. They represent the weighting of stochastic acceleration terms that pull every one particle towards the p^{best} and g^{best} positions. Commonly C_1 and C_2 are set to 2.0 which will make the search cover neighboring sections centered at p^{best} and g^{best} .

$$X_{id} = X_{id} + V_{id} \quad (17)$$

The current position that is the searching point in the solution breathing space can be modified by the equation and as shown in Fig. 1.

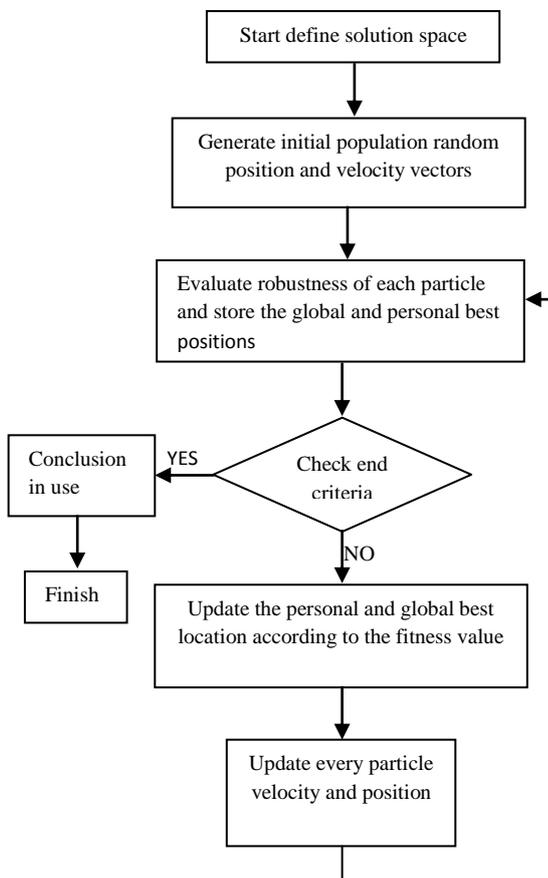


Fig-1 Flow Chart for PSO

The PSO algorithm simple in thought, trouble-free to implement and computational efficient. The unique process for establishing PSO is as follows:

1. Initialize population of particle with random position and velocities and N dimensions in the difficulty space.
2. Identify the particle in the swarm through the best achievement so far, and assign its index to the changeable g .
3. For every particle, assess the desired optimization fitness function in N variables. Evaluate particle's robustness evaluation with its p^{best} . If present value is better than p^{best} , then set p^{best} equal to the current value, and p^{best} equals to the current location x in N-dimensional space.
4. Modify the velocity and position of particle according to equations (15) and (17).
5. Loop to stop 3 until a principle is met, typically a productively superior fitness or a maximum number of iterations.

V. MULTI-OBJECTIVE RPO WITH PARTICLE SWARM OPTIMIZATION INCLUDING VD

Multi-objective Reactive power optimization (RPO) is a significant optimization procedure in terms of voltage deviation and active power loss. The reactive power optimization is realized on IEEE30 bus test system with particle swarm optimization and voltage deviation. There are 6 generator bus, 24 load bus and 41 transmission lines with tap setting transformers.

Bus voltage is one of the largest part significant security and service excellence indices. Improving voltage profile can be obtained by minimizing the load bus voltage deviations from 1.0 per unit. The most favorable settings of the PSO were obtained by the following parameters are given below:

Maximum no. of iteration	:	250
Population size	:	30
w_{max}	:	0.9
$w_{min.}$:	0.4
C_1	:	2
C_2	:	2

The effectiveness of the PSO algorithm has been demonstrated through solution of multi-objective reactive power optimization with voltage deviation problem in, IEEE 30-bus test system.

VI. RESULT AND DISCUSSION

The voltage deviation and reactive power optimization applied on different particle swarm optimization in power system with loss coefficient. The software programs were written in MATLAB 7.8 Language for Reactive Power Optimization, Voltage Deviation and implemented on PSO (particle swarm optimization) with system configuration HP Core 2 Second Gen. i3 Processor and 2 GB RAM. The system data is given in Table2 and Table3 [20, 22]. These system minimum and maximum limits for the control variables along with the initial settings are given in the Table1. The problem was handled as a multi-objective optimization problem where both power loss and voltage deviations

shown in fig.2 (Pareto optimal graph between v d and ploss) were optimized simultaneously with the proposed approach.

TABLE 1. Control Variables Setting And Best Results Of Ploss And Vd As A Main Function

	Min.	Max.	initial (Base case)	Proposed PSO algorithm
V1	1.0	1.1	1.05	1.0824
V2	1.0	1.1	1.04	1.0470
V5	1.0	1.1	1.01	1.0347
V8	1.0	1.1	1.01	1.0209
V11	1.0	1.1	1.05	1.0376
V13	1.0	1.1	1.05	1.0402
T11	1.0	1.1	1.078	1.0196
T12	1.0	1.1	1.069	1.0783
T15	1.0	1.1	1.032	1.0573
T36	1.0	1.1	1.068	1.0963
Qc10	0.0	5.0	0.0	1.2677
Qc12	0.0	5.0	0.0	1.0610
Qc15	0.0	5.0	0.0	0.8607
Qc17	0.0	5.0	0.0	0
Qc20	0.0	5.0	0.0	2.5792
Qc21	0.0	5.0	0.0	1.7678
Qc23	0.0	5.0	0.0	1.6902
Qc24	0.0	5.0	0.0	0.5076
Qc29	0.0	5.0	0.0	0.6881
Power loss(MW)			5.8708	5.3714
Voltage deviation			1.4888	0.6190

The proposed particle swarm optimization based RPO algorithm has been applied to standard IEEE 30-bus test system with system line data and bus data as given in table 2 and 3.

TABLE 2: IEEE 30 BUS SYSTEM DATA

Bus no.	Load		Bus no.	Load	
	P(p.u.)	Q(p.u.)		P(p.u.)	Q(p.u.)
1	0.000	0.000	16	0.035	0.018
2	0.217	0.217	17	0.090	0.058
3	0.024	0.012	18	0.032	0.009
4	0.076	0.016	19	0.095	0.034
5	0.942	0.190	20	0.022	0.007
6	0.000	0.000	21	0.175	0.112
7	0.228	0.109	22	0.000	0.000
8	0.300	0.300	23	0.032	0.016
9	0.000	0.000	24	0.087	0.067
10	0.058	0.020	25	0.000	0.000
11	0.000	0.000	26	0.035	0.023
12	0.112	0.075	27	0.000	0.000
13	0.000	0.000	28	0.000	0.000
14	0.062	0.016	29	0.024	0.009
15	0.082	0.025	30	0.106	0.019

Helpfulness of PSO algorithm is established for solving RPO difficulty with linear arrangement of following

two objective functions treating as a single objective function:

- (1) Minimization of system power losses (P_{Loss})
- (2) Voltage Profile Improvement (VD)

The parameters of the particle swarm optimization algorithm for solving reactive power optimization (RPO) problem in proposed PSO algorithm for best control variable settings shown in the Table the real power loss and voltage deviation are 5.6252 MW and 0.4825 p.u. PSO algorithm gives the best results which illustrate the effectiveness of proposed algorithm.

TABLE 3: IEEE 30 BUS LINE DATA

Line no.	From bus	To bus	Line Impedance	
			R(p.u.)	X(p.u.)
1	1	2	0.0192	0.0575
2	1	3	0.0452	0.1852
3	2	4	0.0570	0.1737
4	3	4	0.0132	0.0379
5	2	5	0.0472	0.1983
6	2	6	0.0581	0.1763
7	4	6	0.0119	0.0414
8	5	7	0.0460	0.1160
9	6	7	0.0267	0.0820
10	6	8	0.0120	0.0420
11	6	9	0.0000	0.2080
12	6	10	0.0000	0.5560
13	9	11	0.0000	0.2080
14	9	10	0.0000	0.1100
15	4	12	0.0000	0.2560
16	12	13	0.0000	0.1400
17	12	14	0.1231	0.2559
18	12	15	0.0662	0.1304
19	12	16	0.0945	0.1987
20	14	15	0.2210	0.1997
21	16	17	0.0824	0.1932
22	15	18	0.1070	0.2185
23	18	19	0.0639	0.1292
24	19	20	0.0340	0.0680
25	10	20	0.0936	0.2090
26	10	17	0.0324	0.0845
27	10	21	0.0348	0.0749
28	10	22	0.0727	0.1499
29	21	22	0.0116	0.0236
30	15	23	0.1000	0.2020
31	22	24	0.1150	0.1790
32	23	24	0.1320	0.2700
33	24	25	0.1885	0.3292
34	25	26	0.2544	0.3800
35	25	27	0.1093	0.2087
36	28	27	0.0000	0.3960
37	27	29	0.2198	0.4153
38	27	30	0.3202	0.6027
39	29	30	0.2399	0.4533
40	8	28	0.6360	0.2000
41	6	28	0.0169	0.0599

Conclusion

In this paper, minimization of reactive power using particle swarm optimization (PSO) algorithm. In order to prove the usefulness of algorithm it is applied to standard reactive power with voltage deviation problem by combining of two objective functions (real power loss and voltage profile improvement) linearly. There are 6 generator bus, 24 load bus and 41 transmission lines with tap setting transformers. At the minimization of the voltage deviation, the more optimum result is taken as the active power loss. The proposed approach is analyzed and demonstrated on the standard IEEE-30 bus test system. The results obtained by proposed algorithm demonstrate its robustness and helpfulness.

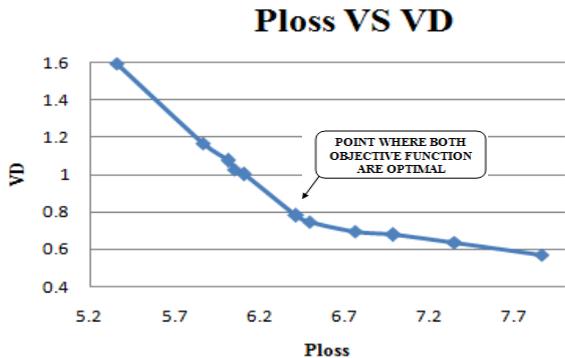


Fig 2 Pareto optimal graph between VD and Ploss

Acknowledgement

The authors are thankful to Director, Madhav Institute of Technology and Science, Gwalior (M.P.) India for providing support and facilities to carry out this research work. vivekkumar.jn@gmail.com

References

- [1] Momoh JA, Zhu JZ. Improved interior point method for OPF problems, IEEE Trans Power Syst. 1999; 14(3): 1114-20.
- [2] M. Huneault, F. D. Galiana, (1991). A survey of the optimal power flow literature, IEEE Trans. Power Syst., Vol. 6 (2) (pp.762-770).
- [3] A. Momoh, M. E. El-Hawary and R. Adapa, A (1999). Review of Selected Optimal Power Flow Literature to 1993 Part I: Nonlinear and Quadratic Programming Approaches, IEEE Trans. Power Syst., Vol. 14 (1) (pp. 96-104).
- [4] K. Aoki, W. Fan, and A. Nishikori, "Optimal var planning by approximation method for recursive mixed- integer programming", IEEE Trans. Power Systems, Vol. 3, No. 4, pp. 1741-1747, Nov. 1988.
- [5] N. Deeb and S.M. Shadidepur, "An efficient technique for reactive power dispatch using a revised linear programming approach", ElectricPowerSystemsResearch, Vol.15, No.2, pp.121-134, 1988.
- [6] S.M. Shaidehpour and N. Deeb, "An overview of the reactive power allocation in electric power systems", Electric Machines and Power Systems, Vol. 18, pp. 495-518, 1990.
- [7] R. Yokoyama, T. Niimura, and Y. Nakanishi, "A coordinated control of voltage and reactive power by heuristic modelling and approximate reasoning", IEEE Trans. Power Systems, Vol. 8, No. 2, pp. 636-670, May 1993.
- [8] K. Iba, H. Suzuki, K. Suzuki, and K. Suzuki, "Practical reactive power allocation/operation planning using successive linear programming", IEEE Trans. Power Systems, Vol. 3, No. 2, pp. 558-566, May 1998.
- [9] R.C. Bansal, "Optimization methods for electric power systems: an overview", Int. Journal of Emerging Electric Power Systems, Vol. 2, No. 1, article No. 1021, 2005.
- [10] H. Yoshida, K. Kawata, Y. Fukuyama, S. Takayama, and Y. Nakanishi, (2000). A particle swarm optimization for reactive power and voltage control considering voltage security assessment, IEEE Trans. Power Syst., Vol. 15 (4) (pp. 1232-1239).
- [11] Camp CV, Bichon BJ, Stovall SP. Design of steel frames using ant colony optimization, Journal of Structural Engineering 131 (2005)369-79.
- [12] Haibo Zhang, Lizi Zhang, Fanling Meng, "Reactive Power Optimization Based on Genetic Algorithm", Proceedings of International Conference on Power System Technology, POWERCON '98. 1998, Vol. 2, pp. 1448 – 1453.
- [13] Salajegheh E, Gholizadeh S. Optimum design of structures by an improved genetic algorithm using neural networks, Advances in Engineering Software, 36(2005) 757-67.
- [14] Kaveh A, Kalatjari V. Genetic algorithm for discrete sizing optimal design of trusse using the force method, International Journal for Numerical Methods in Engineering 55 (2002)55-72.
- [15] Erbatur F, Hasancebi O, Tutuncu I, Kitic H. Optimal design of planar and space structures with genetic algorithms, Computers and Structures, 75(2000) 209-24.
- [16] Rajeev S, Krishnamoorthy CS. Discrete optimization of structures using genetic algorithms, Journal of Structural Engineering, 118(1992)1233-50.
- [17] M. R. Alrashidi and M. E. El-Hawary, (2006). A Survey of Particle Swarm Optimization Applications in Power System Operations, Electric Power Components and Systems, 34:12, (pp. 1349-1357).
- [18] Kessel and H. Glavitsch, 1986, "Estimating the Voltage Stability of Power System," IEEE Trans. Power delivery, Vol. PWRD-1, No. 3, pp. 346- 354.
- [19] H. Wan, J. D. Mc Calley, V. Vittal, 2000, "Risk Based Voltage Security Assessment", IEEE Transaction on Power Systems, Vol. 15, No. 4, pp. 1247-1.
- [20] A.A. Abou El Ela, M.A. Abido, S.R. Spea , "Differential Evolution algorithm for optimal reactive power dispatch ", Electric Power Systems Research 81 (2011) 458-464.
- [21] W.N.W Abdullah, H. Saibon A.A.M Zain, "Genetic Algorithm for Optimal Reactive Power Dispatch", Proceedings of International Conference on Energy Management and Power Delivery, 1998. Vol. 1, 3-5 Mar 1998, pp. 160-164.
- [22] H.D. Chiang, J.C. Wang, O. Cockings, and H.D. Shin, "Optimal capacitor placements in distribution systems: part 2: solution algorithms and numerical results", IEEE Trans. Power Delivery, Vol. 5, No. 2, pp. 634-641, April 1990.

Speaker Recognition System Using Combined Vector Quantization and Discrete Hidden Markov model

Ameen Khan A[#], N V Uma Reddy[#] and Madhusudana Rao.*

[#] Student AMC Engineering college, VTU Karnataka, India, associate professor AMCEC, VTU Karnataka, India associate professor, NGIT Bangalore, India.,

Abstract-- This paper presents a speaker verification system using a combination of Vector Quantization (VQ) and Hidden Markov Model (HMM) to improve the HMM performance. A Malay spoken digit database which contains 100 speakers is used for the testing and validation modules. It is shown that, by using the proposed combination technique, a total success rate (TSR) of 99.97% is achieved and it is an improvement of 11.24% in performance compared to HMM. For speaker verification, true speaker rejection rate, impostor acceptance rate and equal error rate (EER) are also improved significantly compared to HMM.

Index Terms-- Speaker recognition, speaker verification, hidden Markov model, vector quantization

I. Introduction

Speaker recognition is the process of automatically recognizing who is speaking on the basis of individual information included in speech waves. This technique makes it possible to use the speaker's voice to verify their identity and control access to services such as voice dialing, banking by telephone, telephone shopping, data base access services, information services, voice mail, security control for confidential information areas, and remote access to computers.

Speaker recognition or verification is a biometric modality that uses an individual's voice for recognition or verification purpose. It is a different technology from speech recognition, which recognizes words as they are articulated, which is not biometrics. Speech contains many characteristics that are specific to each individual. For this reason, listeners are often able to recognize the speaker's identity fairly quickly even without looking at the speaker.

Speaker recognition methods can be divided into text-independent and text-dependent methods. In a text-independent system, speaker models capture characteristics of somebody's speech which show up irrespective of what one is saying. In a text-dependent system, on the other hand, the recognition of the speaker's identity is based on his or her speaking one or more specific phrases, like passwords, card numbers, PIN codes, etc.

At the highest level, all speaker recognition systems contain two main modules feature extraction and feature matching. Feature extraction is the process that extracts a small amount of data from the voice signal that can later be used to represent each speaker. Feature matching involves the actual procedure to identify the unknown speaker by comparing extracted features from his/her voice input with the ones from a set of known speakers.

All speaker recognition systems have to serve two distinguished phases. The first one is referred to the enrollment sessions or training phase while the second one is referred to as the operation sessions or testing phase. In the training phase, each registered speaker has to provide samples of their speech so that the system can build or train a reference model for that speaker. In case of speaker verification systems, in addition, a speaker-specific threshold is also computed from the training samples. During the testing (operational) phase the input speech is matched with stored reference model(s) and recognition decision is made. This paper presents a text dependent speaker verification using a combination approach of VQ and DHMM. The objective is to improve the performance of DHMM in a speaker verification system.

ii. Feature Extraction

The acoustic speech signal contains different kind of information about speaker. This includes "high-level" properties such as dialect, context, speaking style, emotional state of speaker and many others [3].

The speech wave is usually analyzed based on spectral features. There are two reasons for it. First is that the speech wave is reproducible by summing the sinusoidal waves with slowly changing amplitudes and phases. Second is that the critical features for perceiving speech by humans ear are mainly included in the magnitude information and the phase information is not usually playing a key role [4].

A. Short term analysis

Because of its nature, the speech signal is a slowly varying signal or quasi-stationary. It means that when speech is examined over a sufficiently short period of time (20-30 milliseconds) it has quite stable acoustic characteristics [5]. It leads to the useful concept of describing human speech signal, called "short-term analysis", where only a

portion of the signal is used to extract signal features at one time. It works in the following way: predefined length window (usually 20-30 milliseconds) is moved along the signal with an overlapping (usually 30-50% of the window length) between the adjacent frames.

Overlapping is needed to avoid losing of information. Parts of the signal formed in such a way are called frames. In order to prevent an abrupt change at the end points of the frame, it is usually multiplied by a window function. The operation of dividing signal into short intervals is called windowing and such segments are called windowed frames (or sometime just frames). There are several window functions used in speaker recognition area [4], but the most popular is Hamming window function, which is described by the following equation:

$$w(n)=0.54-0.46\cos(2n\pi/N-1) \quad (1)$$

where N is the size of the window or frame. A set of features extracted from one frame is called feature vector. Overall overview of the short-term analysis approach is represented in Figure 1

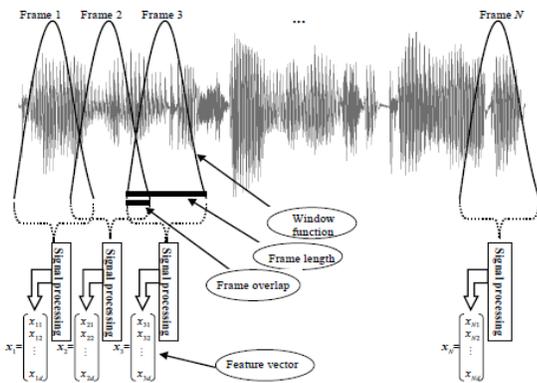


Fig.1 short term analysis

B. Cepstrum

The speech signal $s(n)$ can be represented as a “quickly varying” source signal $e(n)$ convolved with the “slowly varying” impulse response $h(n)$ of the vocal tract represented as a linear filter. Separation of the source and the filter parameters from the mixed output is in general difficult problem when these components are combined using not linear operation, but there are various techniques appropriate for components combined linearly. The cepstrum is representation of the signal where these two components are resolved into two additive parts. It is computed by taking the inverse DFT of the logarithm of the magnitude spectrum of the frame. This is represented in the following equation:

$$\text{Cepstrum}(\text{frame})=\text{IDFT}(\log(|\text{DFT}(\text{frame})|)) \quad (2)$$

By moving to the frequency domain we are changing from the convolution to the multiplication. Then by taking logarithm we are moving from the multiplication to the addition. That is desired division into additive components. Then we can apply linear operator inverse DFT, knowing that the transform will operate individually on these two parts and knowing what Fourier transform will do with quickly varying and slowly varying parts.

C. Mel-Frequency Cepstrum Coefficients

MFCC extraction is similar to the cepstrum calculation except that one special step is inserted, namely the frequency axis is warped according to the Mel-scale. Summing up, the process of extracting MFCC from continuous speech is illustrated in Figure 2

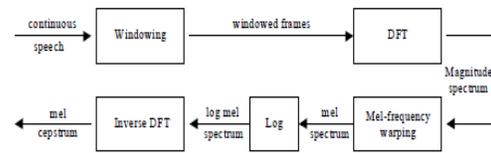


Fig.2 Computing of Mel-cepstrum

A “Mel” is a unit of special measure or scale of perceived pitch of a tone [5]. It does not correspond linearly to the normal frequency indeed it is approximately linear below 1 kHz and logarithmic above [5]. One useful way to create Mel-spectrum is to use a filter bank, one filter for each desired Mel-frequency component. Every filter in this bank has triangular band pass frequency response. Such filters compute the average spectrum around each center frequency with increasing bandwidths, as displayed in Figure 3.

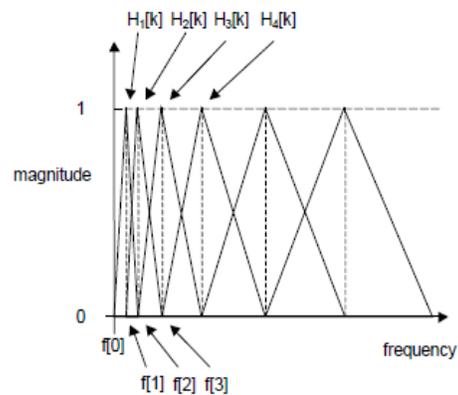


Fig.3 Triangular filters used to compute Mel-cepstrum

This filter bank is applied in frequency domain and therefore, it simply amounts to taking these triangular filters on the spectrum. In practice the last step of taking

inverse DFT is replaced by taking discrete cosine transform (DCT) for computational efficiency.

iii. Feature Matching And Speaker Modeling

The modeling is a process of enrolling speaker to the Identification system by constructing a model of his/her voice based on the features extracted from his/her speech sample. The matching is a process of computing a matching score, which is a measure of the similarity of the features extracted from the unknown speech sample and speaker model.

A. Vector Quantization

Vector quantization (VQ) is a process of mapping vectors from a vector space to a finite number of regions in that space. These regions are called clusters and represented by their central vectors or centroids. A set of centroids, which represents the whole vector space, is called a codebook. In Speaker identification, VQ is applied on the set of feature vectors extracted from the speech sample and as a result, the speaker codebook is generated.

Mathematically a VQ task is defined as follows: given a set of feature vectors, find a partitioning of the feature vector space into the predefined 30 number of regions, which do not overlap with each other and added together form the whole feature vector space. Every vector inside such region is represented by the corresponding centroid [6]. The process of VQ for two speakers is represented in Figure 4.

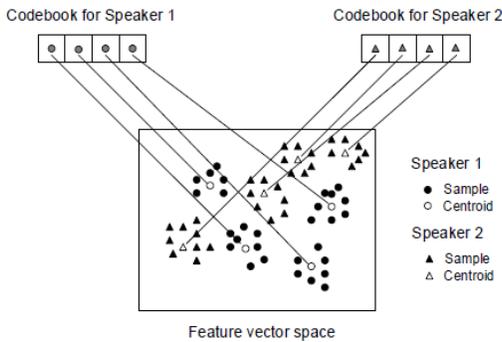


Fig.4 vector Quantization for two speakers

B. Clustering the training vectors (LBG algorithm)

For clustering a set of L training vectors into a set of M codebook vectors. The algorithm is formally implemented by the following recursive procedure:

1. Design a 1-vector codebook: this is the centroid of the entire set of training vectors (hence, no iteration is required here).
2. Double the size of the codebook by splitting each current codebook Y_n according to the rule

$$Y_n^+ = Y_n(1+\epsilon)$$

$$Y_n^- = Y_n(1-\epsilon)$$

where n varies from 1 to the current size of the codebook, and is ϵ a splitting parameter (we choose $\epsilon=0.01$).

3. Nearest-Neighbor Search: for each training vector, find the codeword in the current codebook that is closest (in terms of similarity measurement), and assign that vector to the corresponding cell (associated with the closest codeword).
4. Centroid Update: update the codeword in each cell using the centroid of the training vectors assigned to that cell.
5. Iteration 1: repeat steps 3 and 4 until the average distance falls below a preset threshold
6. Iteration 2: repeat steps 2, 3 and 4 until a codebook size of M is designed.

Figure 6 shows, in a flow diagram, the detailed steps of the LBG algorithm. “Cluster vectors” is the nearest-neighbor search procedure which assigns each training vector to a cluster associated with the closest codeword. “Find centroids” is the centroid update procedure. “Compute D (distortion)” sums the distances of all training vectors in the nearest-neighbor search so as to determine whether the procedure has converged.

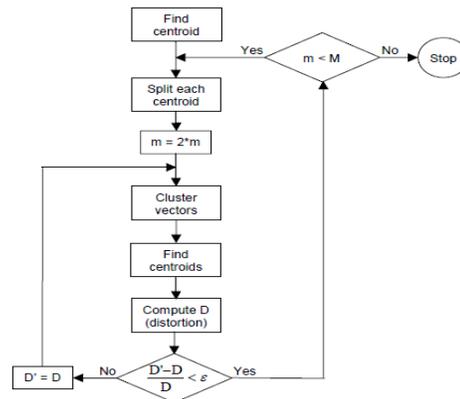


Fig.5 Flow diagram of LBG algorithm

C. Discrete Hidden Markov Model(DHMM)

A speaker verification system consists of two phases which is the training phase and the verification phase. In the training phase, the speaker voices are recorded and processed in order to generate the model to store in the database. While, in the verification phase, the existing reference templates are compared with the unknown voice input. In this project, we use the Discrete Hidden Markov Model (DHMM) method as the training/recognition algorithm[10].

The most flexible and successful approach to speech recognition so far has been HMM. The goal of HMM parameter estimation is to maximize the likelihood of the

data under the given parameter setting. General theory of HMM has been given in [1,7,8]. There are 3 basic parameters in HMM which is:

1. π - The initial state distribution
2. a - The state-transition probability matrix. Probability $A_{ij} = P(\text{state at time } t = S_j / \text{state at time } t-1 = S_i)$.
3. b - Observation probability distribution. probability $b_{jk} = P(\text{observation at time } t, o_t = v_k / \text{state at time } t = S_j)$.

In the training phase, a HMM model for each speaker is generated. Each model is an optimized model for the word it represents. For example, a model for the word 'satu' (number one), has its a , b , and π parameters adjusted so as to give the highest probability score whenever the word 'satu' is uttered, and lower scores for other words. Thus, to build a model for each speaker, a training set is needed. This training set consists of sequences of discrete symbols, such as the codebook indices obtained from the Vector Quantization stage.

Here, an example is given of how HMM is used to build models for a given training set. Assuming that N speakers are to be verified, first we must have a training set of L token words, and an independent testing set. To do speaker verification, the following steps are needed:

1. First we build an HMM for each speaker. The L training set of tokens for each speaker will be used to find the optimum parameters for each word model. This is done using the re-estimation formula.
2. Then, for each unknown speaker in the testing set, first characterize the speech utterance into an observation sequence. This means using an analysis method for the speech utterance so that we get the feature vector, and then the vector is quantized using Vector Quantization. Thus, we will get a sequence of symbols, with each symbol representing the speech feature for every discrete time step.
3. We calculate a , b and π parameters for the observation sequence using one of the speaker models in the vocabulary. Then repeat for every speaker model in the database.

After N models have been created, the HMM engine is then ready for speaker verification. A test observation sequences from an unknown speech utterance (produced after vector quantization of cepstral coefficient vectors), will be evaluated using the Viterbi algorithm.

The log-Viterbi algorithm is used to avoid precision underflow. For each speaker model, probability score for the unknown observation sequence is computed. The speaker whose model produces the highest probability

score and matches the ID claimed is then selected as the client speaker[10].

D. Decision

The next step after computing of matching scores for every speaker model enrolled in the system is the process of assigning the exact classification mark for the input speech and it is based on the computed probabilities. This process is represented in Figure 6.

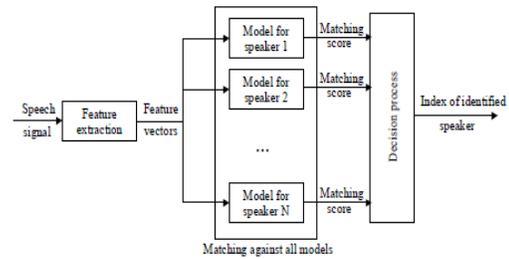


Fig 6 Decision process

The model with highest probability is selected. Practically, decision process is not so simple and for example for so called open-set identification problem the answer might be that input speech signal does not belong to any of the enrolled speaker models.

E. Performance Criteria

The basic error measures of a verification system are false acceptance rate (FAR) and false rejection rate (FRR), as defined below

$$FAR = \frac{\text{Number of accepted imposter claims}}{\text{total number of imposter accesses}} \times 100$$

$$FRR = \frac{\text{Number of rejected genuine claims}}{\text{total number of genuine accesses}} \times 100$$

Overall performance can be obtained by combining these two errors into total success rate (TSR) where:

Speaker verification threshold or equal error rate (EER) is calculated as .

$$T = \frac{\mu_1 \sigma_2 + \mu_2 \sigma_1}{\sigma_1 + \sigma_2} \quad (3)$$

Iv. Results And Discussion

Table 1 shows a summary of the verification results for the experiments performed. A total success rate (TSR) of 99.97% was achieved using this combination technique compared to stand alone HMM which is 89.87%. The TSR performance improve significantly by 11.24%. Using the combination technique, true speaker rejection rate achieved was 0.06% while impostor acceptance rate was

0.03% and equal error rate (EER) of 11.72% was achieved.

Table 1
Spoken Digit Data Base

Method	FAR	FRR	TSR	EER
VQ	25.3%	9.4%	76.8%	12%
HMM	6.2%	2.1%	89.8%	4%
VQ-HMM	0.08%	0.03%	98.6%	2%

V. Conclusion

This paper has shown that the combination approach of VQ and HMM can improve the HMM performance in a noise-free environment. The Malay spoken digit database which contains 100 speakers has been used to test and validate the system. It is shown that, a total success rate (TSR) of 99.97% was achieved using this combination technique compared to HMM which was 89.87%. The TSRs performance improves significantly by 11.24%. For FRR, FAR, and EER, the combination technique also shows improvement. Further work will concentrate on noisy environment to evaluate the robustness of the system.

Vi.References

[1] C. Wheddon and R. Linggard, *Speech and Language Processing*, Chapman and Hall, UK, 1990, pp. 209-230.

[2] Evgeny Karpov “Real-Time Speaker Identification” University of Joensuu Department of Computer Science.

[3] J. M.Naik, “Speaker Verification: A Tutorial”, *IEEE Communications Magazine*, January 1990, pp.42-48.

[4] S. Furui, *Digital Speech Processing, Synthesis and Recognition*, New York, Marcel Dekker, 2001.

[5] J. R. Deller, J. H. L. Hansen, J. G. Proakis, *Discrete-Time Processing of Speech Signals*, Piscataway (N.J.), IEEE Press, 2000.

[6] F. K. Soong, A. E. Rosenberg, L. R. Rabiner and B. H. Juang, “A Vector Quantization Approach to the Speaker Recognition”, *AT&T Technical Journal*, Vol. 66, pp. 14-26, Mar/Apr 1987.

[7] L.R. Rabiner, “A Tutorial on Hidden Markov Models and Selected Application in Speech

Recognition”, *Proceeding of The IEEE*, Vol.77, No.2, February 1989.

[8] L.R. Rabiner and B.H. Juang, *Fundamental of Speech Recognition*, Prentice Hall, New Jersey, 1993.

[9] Tomi Kinnunen “Spectral Features for Automatic Text-Independent Speaker Recognition” *University of Joensuu Department of Computer Science*. December, 2003.

[10] Mohd Zaizu Ilyas, *Member, IEEE*, Salina Abdul Samad, *Senior Member, IEEE*, Aini Hussain, *Member, IEEE* and Khairul Anuar Ishak, *Member, IEEE* “Speaker Verification using Vector Quantization and Hidden Markov Model” 2007-11-12 December 2007, Malaysia

Design, Implementation and Performance Analysis of an Integrated Vedic Multiplier Architecture

Ramachandran.S*, Kirti.S.Pande

Department of Electronics and Communication
Amrita School of Engineering
Kasavanahalli, Bangalore- 560035, India

Abstract

Fundamental and the core of all the Digital Signal Processors (DSPs) are its multipliers and speed of the DSPs is mainly determined by the speed of its multipliers. Multiplication is the most fundamental operation with intensive arithmetic computations. Two important parameters associated with multiplication algorithms performed in DSP applications are latency and throughput. Latency is the “real delay of computing a function”. Throughput is a measure of “how many computations can be performed in a given period of time”. The execution time of most DSP algorithms is dependent on its multipliers, and hence need for high speed multiplier arises.

Urdhva tiryakbhyam sutra performs faster for small inputs and Nikhilam sutra for larger inputs. Here a novel Integrated Vedic multiplier architecture, which by itself selects the appropriate multiplication sutra based on the inputs, is proposed. So depending on inputs, whichever sutra is faster, that sutra is selected by the proposed integrated Vedic multiplier architecture. In the simulation results, it can be seen that Urdhva performs faster for small inputs, but Nikhilam performs better for large inputs (more than twice as much for 64 bit multiplicands).

Keywords— Integrated Vedic multiplier architecture, Nikhilam sutra architecture, Urdhva tiryakbhyam sutra.

I. INTRODUCTION

Vedic mathematics is an extract from four Vedas (books of wisdom). It is actually a sub module of “Sthapatya-veda” (book on civil engineering and architecture), which is an upya-veda (supplement) of Atharva Veda. Owing to its simplicity and regularity, it finds its utility and applications in the fields of geometry, trigonometry, quadratic equations, factorization and calculus.

His holiness Jagadguru Shankaracharya Bharati Krishna Teerthaji Maharaja (1884-1960) comprised all this work. He did an extensive research in the Vedas and came up with the simplified form of calculations which are yet so powerful. He came up with complete explanations and presented them in the form sutras. He constructed 16 sutras (formulae) and 16 Upya sutras (sub formulae) after extensive research in Atharva Veda. As all these sutras were extracts from swamiji’s own findings and research, these are not found explicitly in the Veda. ‘Vedic’ is a term derived from the word ‘veda’ which means the store-house of all knowledge.

The power of Vedic mathematics is not only confined to its simplicity, regularity, but also it is logical. Its high degree of eminence is attributed to the aforementioned facts. It is these phenomenal characteristics, which made Vedic mathematics, become so popular and thus it has become one of the leading topics of research not only in India but abroad as well.

Vedic mathematics’ logics and steps can be directly applied to problems involving trigonometric functions, plane and sphere geometry, conics, differential calculus, integral calculus and applied mathematics of various kind.

The awe striking and the mind boggling feature of Vedic mathematics lies in the fact that it simplifies the complicated looking calculations in conventional mathematics to a simple one in a much faster and efficient manner. This is attributed to the fact that the Vedic formulae are claimed to be based on the “natural principles on which the human mind works”. Hence this presents some effective algorithms which can be applied to various branches of engineering.

The architecture of Multipliers can be generally classified into three categories. First is the “serial multiplier” which emphasizes on hardware optimization of chip area. Second is “parallel multiplier” which performs high speed mathematical operations, the drawback being relatively larger chip area consumption. The final one is “serial-parallel multiplier” which is a trade-off between time consuming serial multipliers and the area consuming parallel multipliers.

II. RELATED WORKS

A high speed multiplier design (ASIC) using Vedic mathematics was presented in [1]. The idea for designing the multiplier and adder unit was adopted from ancient Indian mathematics “Vedas”. Based on those formulae, the partial products and sums are generated in single step which reduces the carry propagation from LSB to MSB. The implementation of the Vedic mathematics and their application to the complex multiplier ensured substantial reduction of propagation delay in comparison with Distributed Array (DA) based architecture and parallel adder based implementation which are most commonly used architectures.

The implementation of the Vedic algorithms in DSP is highlighted in [2]. In this, multiplication process based on Vedic mathematics and its implementation on 8085 and 8086 microprocessors was shown. A comparative study of processing time of conventional multipliers for 8085 and 8086 was done. It was shown that there is an appreciable saving in the processing time of the Vedic multiplier as when compared to that of a conventional multiplier.

A time, area, power efficient multiplier architecture using Vedic mathematics was shown in [3]. In this a comparative study of the array multiplier, Carry save multiplier, Wallace tree multiplier, Booth multiplier and Vedic multiplier was done in detail. The study clearly showed that though array and booth multipliers are faster among the conventional multipliers, they are so because of some trade-off with complexity and high power consumption respectively.

A fast, low power multiplier architecture based on Vedic mathematics was shown in [4]. This paper presented a new architecture for multiplication which uses the modified binary tree network (MBT). This architecture focuses on generating all partial products in one step. The generated partial products are added by the MBT network. This also showed evidence of increase in speed.

Reduced bit multiplication algorithm for digital arithmetic was shown in [5]. It mainly consisted of the in depth explanation of Urdhva tiryakbhyam sutra and the Nikhilam sutra. These sutras are the extracts from the Vedas which are the store house of knowledge. The former was suggested for smaller numbers and the latter was suggested for larger numbers. This paper showed that multiplication of two 8 bit numbers can be effected by reducing it further into two 4 bit numbers and likewise.

VHDL implementation of a NXN multiplier based on the Vedic mathematics was shown in [6]. This proposed a way to implement the design of the Urdhva sutra based multiplier as a bottom up design methodology.

A novel design for square and cube architecture was shown in [7]. It was very clearly evident from the explanations that, the Vedic square and cube architecture were faster than the conventional square and cube calculations.

All the Vedic based calculations, sutra explanations and complete discussions were made in [8]. This does all explanations right from scratch. This contains the complete essence of the Vedic techniques, calculations and all the associated sutras with the explanations in detail. This forms the base and is the fundamental of any Vedic mathematics based calculations.

III. DESIGN APPROACH

A. Urdhva tiryakbhyam sutra illustration

Multiplication is based on an algorithm called Urdhva Tiryakbhyam (Vertical & Crosswise) of ancient Indian Vedic Mathematics. Urdhva Tiryakbhyam Sutra is a general multiplication formula applicable to all cases of multiplication. The Sanskrit term means “Vertically and crosswise”. The idea here is based on a concept which results in the generation of all partial products along with the concurrent addition of these partial products in parallel [5]. The parallelism in generation of partial products and their summation is obtained using Urdhva Tiryakbhyam explained in Fig. 1.1. Since there is a parallel generation of the partial products and their sums, the processor becomes independent of the clock frequency. Thus the multiplier will require the same amount of time to calculate the product and hence is independent of the clock frequency. The advantage here is that parallelism reduces the need of processors to operate at increasingly high clock frequencies. A higher clock frequency will result in increased processing power, and its demerit is that it will lead to increased power dissipation resulting in higher device operating temperatures. By employing the Vedic multiplier, all the demerits associated with the increase in power dissipation can be negotiated. Since it is quite faster and efficient its layout has a quite regular structure. Owing to its regular structure, its layout can be done easily on a silicon chip [4]. The Vedic multiplier has the advantage that as the number of bits increases, gate delay and area increases very slowly as compared to other multipliers, thereby making it time, space and power efficient. It is demonstrated that this architecture is quite efficient in terms of silicon area/speed [4].

To illustrate this multiplication scheme, let us consider the multiplication of two decimal numbers ($325 * 738$). Line diagram for the multiplication is shown in Fig.1.2. Initially the LSB digits on the both sides of the line are multiplied and added with the carry from the previous step. This generates one of the bits of the result and a carry. This carry is added in the next step and the process goes on likewise. If more than one line are there in one step, all the results are added to the previous carry. In each step, least significant bit act as the result digit and all other digits act as carry for the next step. Initially the carry is taken to be zero.

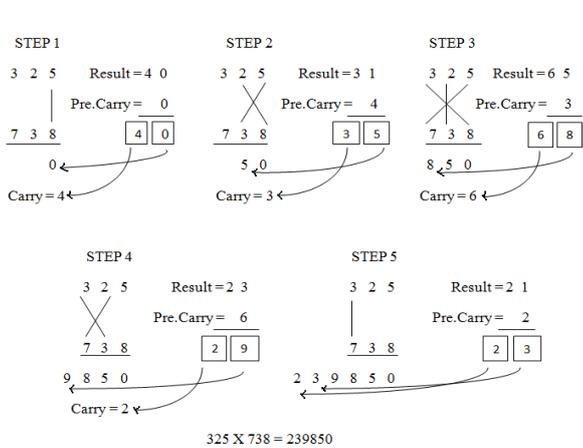


Fig. 1.1 Multiplication of two numbers using Urdhva sutra

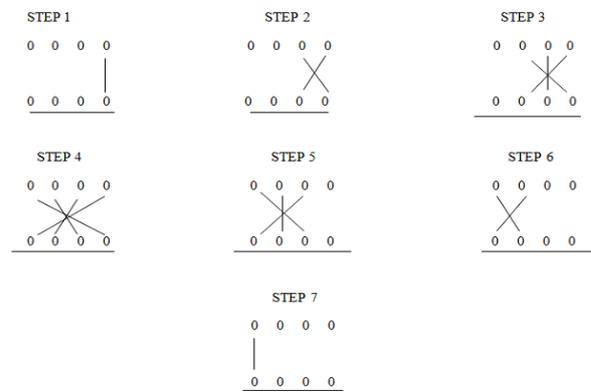


Fig. 1.2 Line diagram of the multiplication

To make the methodology more clear, an alternate illustration is given with the help of line diagrams in Fig. 1.2, where the dots represent bit “0” or “1” [4]. Here in order to illustrate the multiplication algorithm, we consider the multiplication of two binary numbers $a_3a_2a_1a_0$ and $b_3b_2b_1b_0$. As the result of this multiplication would be more than 4 bits, it is expressed as $\dots r_3r_2r_1r_0$. Line diagram for multiplication of two 4-bit numbers is shown in Fig. 1.2. This just maps the illustration in Fig.1.1 in binary system. Least significant bit r_0 is obtained by multiplying the least significant bits of the multiplicand and the multiplier. The process is followed as per the steps shown in Fig. 1.1. The same method is used and the design for multiplication of two 64 bit numbers is done using the bottom up methodology. The architecture is shown in Fig. 1.3 which is followed as per the concept of [9][10].

A31-A0 and B31-B0 generates Q31-Q0 of the output and the remaining bits are forwarded as carry to next. A31-A0, B63-B32 and A63-A32, B31-B0 and the carry from previous step together forms Q63-Q32 of the output and the remaining bits are forwarded to the next. Finally A63-A32 and B63-B32 and the carry from the previous step together form the Q127-Q64 of the output. Thus a 64 bit multiplication results by multiplying four 32 bit numbers in parallel.

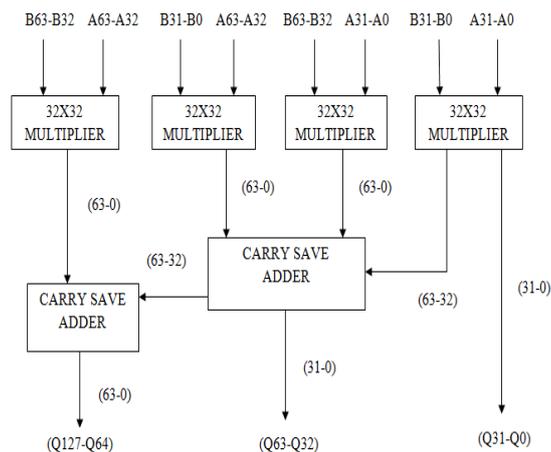


Fig. 1.3. Architecture of 64 bit Urdhva sutra [9][10]

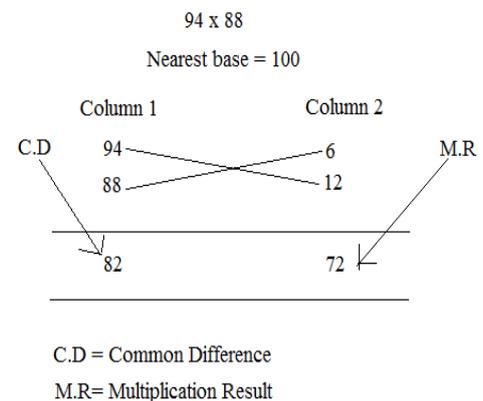


Fig. 1.4. Nikhilam sutra illustration [4]

B. Nikhilam sutra illustration

The Sanskrit term Nikhilam means “all from 9 and last from 10”. It is also applicable to all cases of multiplication, but it tends to be more efficient when the numbers involved are large. This is because, it just finds out the compliment of the large number from its nearest base to perform the multiplication operation on it [4]. Larger the original number, lesser the complexity of the multiplication. This sutra is illustrated by considering the multiplication of two decimal numbers ($94 * 88$) where the chosen base is 100 which is nearest to and greater than both these two numbers.

The right hand side (RHS) of the product can be obtained by simply multiplying the numbers of the Column 2 ($6 * 12 = 72$). The left hand side (LHS) of the product can be found by cross subtracting the second number of Column 2 from the first number of Column 1 or vice versa, i.e., $94 - 12 = 82$ or $88 - 6 = 82$. The final result is obtained by concatenating RHS and LHS (Answer = 8272).

The proposed Nikhilam sutra architecture is shown in Fig. 1.5 and is based on the above illustration of the sutra.

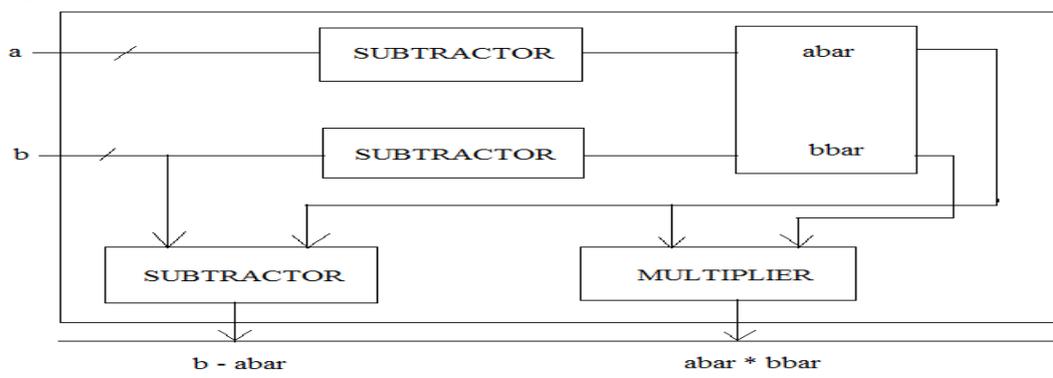


Fig. 1.5. Proposed architecture of Nikhilam sutra multiplier

It is known from literature that Urdhva based multiplier is expected to work faster for small inputs and Nikhilam sutra based multiplier for large inputs. Hence, Integrated Vedic Multiplier Architecture is proposed in this paper, which is capable

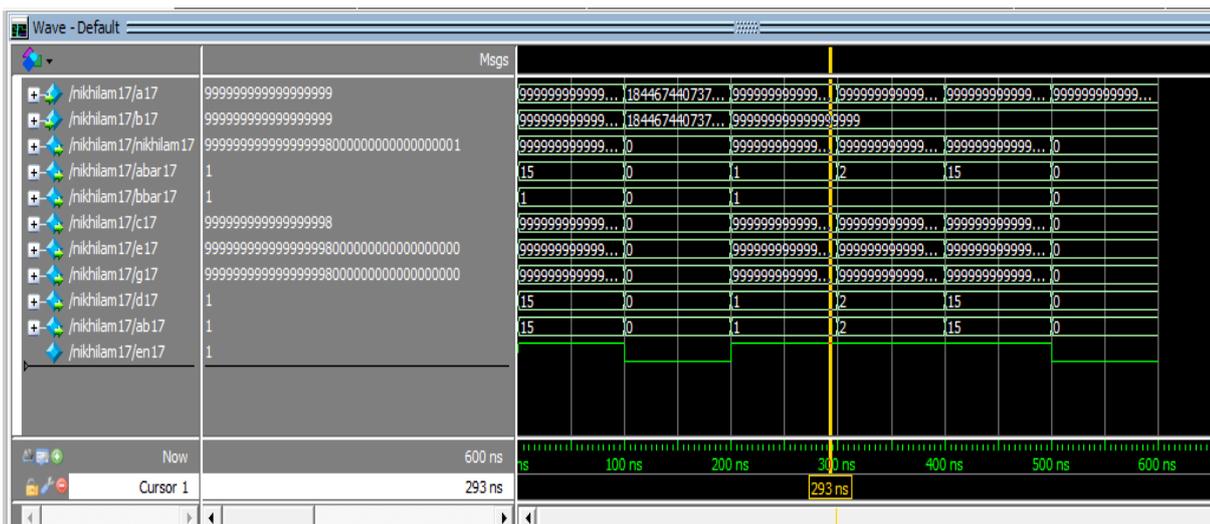
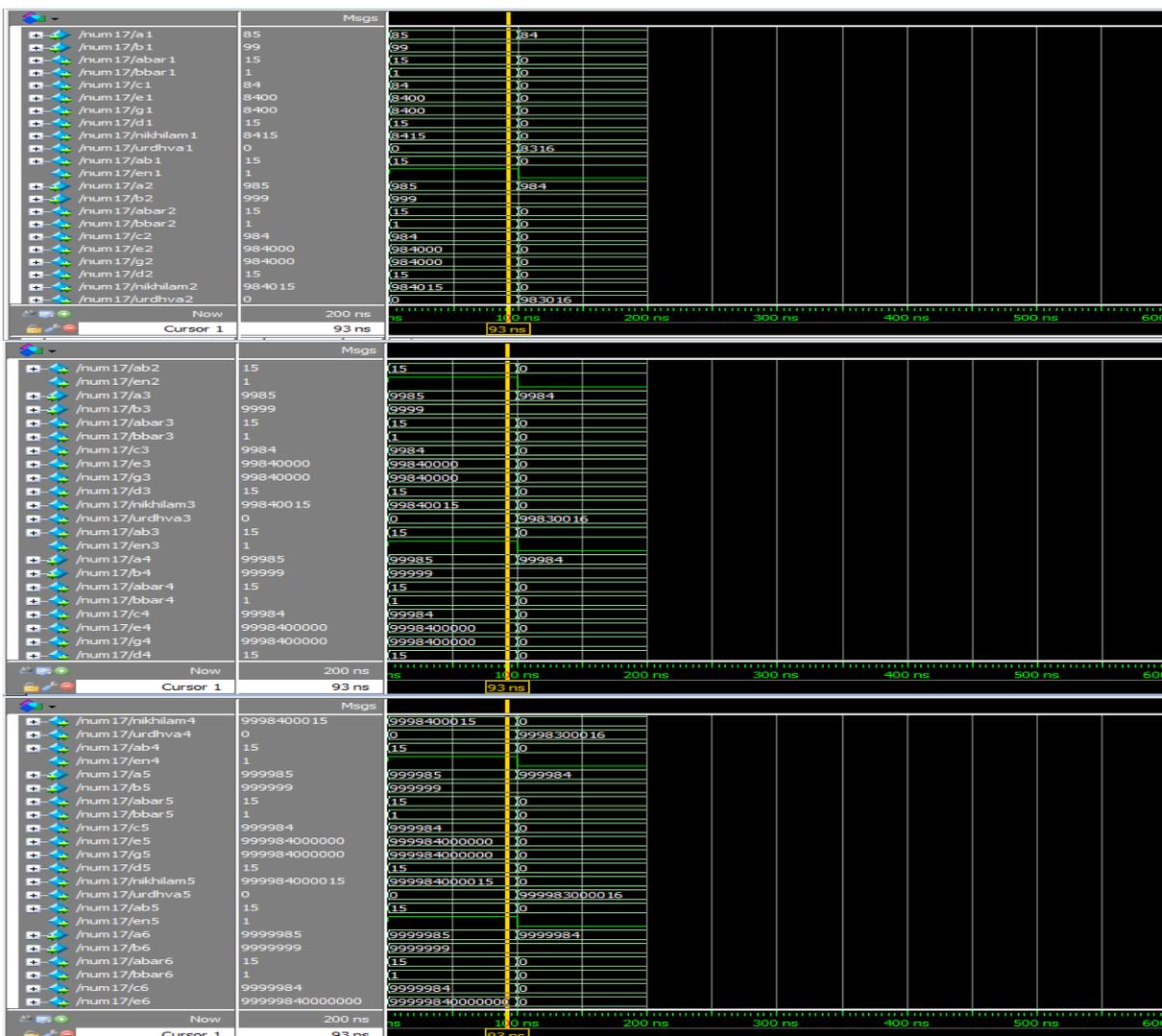


Fig. 1.8.Simulation result of Proposed 64x64 Nikhilam multiplier

Here if both ‘a’ and ‘b’ are within the permissible limits of the Nikhilam condition, ‘en’ becomes ‘1’ and if ‘a’ and ‘b’ are outside Nikhilam limit, ‘en’ becomes ‘0’ and Nikhilam stops functioning (i.e.) “OFF”.



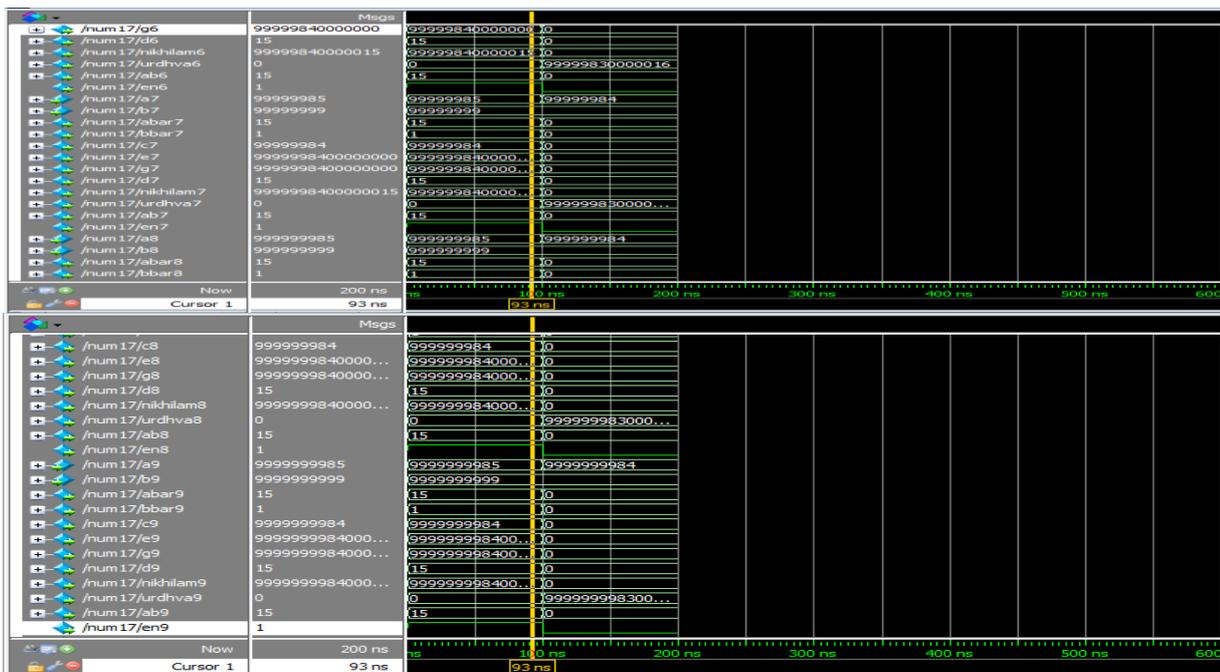


Fig. 1.9.Simulation result of Proposed 64x64 Integrated Multiplier Architecture

Fig.1.9. shows the output of the proposed Integrated Multiplier Architecture. This clearly shows that based on the conditions, only one multiplier sutra performs the multiplication at any given time. When the inputs are large and close to the base such as 100, 1000, 10000, etc., Nikhilam sutra does the multiplication and saves time. When any other normal input is given (i.e.) when the inputs are small, Nikhilam sutra stops working and Urdhva multiplier performs the multiplication. Thus this proposed paper becomes an Integrated Vedic Multiplier Architecture which ensures that the best multiplier is performing the multiplication and saves time depending on the given inputs.

B. Speed Analysis Report

The designs of 8x8 bits, 16x16 bits, 32x32 bits and 64x64 bits Vedic multiplier have been implemented on *Xilinx ISE 11 series* for a series of ten multiplicands each, which fall in the Nikhilam range. It is therefore seen from Table.1.that, on an average, in case of 8 bit multipliers, Urdhva performs better than Nikhilam because of the small size of the multiplicands. However, as the size of the multiplicands increase, Nikhilam performs much faster than Urdhva & achieves an increase of 135.04% (more than twice as fast) in speed for 64 bit multiplicands. The speed analysis is as shown in Table.1.

Size \ Type	8x8	16x16	32x32	64x64
Urdhva	13.455 ns	25.083 ns	44.667 ns	77.132 ns
Nikhilam	18.699 ns	20.094 ns	24.075 ns	32.816 ns
% increase in speed	Urdhva faster by 38.97%	Nikhilam faster by 24.82%	Nikhilam faster by 85.53%	Nikhilam faster by 135.04%

Table.1. Synthesis Report

V. CONCLUSION AND FUTURE SCOPE

The time taken for multiplication operation is reduced by employing the Vedic algorithms. Here integrated Vedic multiplier architecture is proposed for further reduction in time. Depending on the inputs, the better sutra is selected by the architecture itself. Future work includes the integration of the divider block, multiply and accumulate (MAC) unit, thereby making it into a Vedic Arithmetic and Logical unit (ALU).

REFERENCES

- [1] Prabir Saha, Arindham Banerjee, Partha Battacharyya, Anup Dhandapat, "High speed design of complex multiplier using Vedic mathematics", *Proceedings of the 2011 IEEE students technology symposium, IIT Kharagpur*, pp. 237-241, Jan. 2011.
- [2] Purushottam D. Chidgupkar and Mangesh T. Karad, "The Implementation of Vedic Algorithms in Digital Signal Processing", *UICEE, Global J. of Engng. Educ., Vol.8, No.2*, pp. 153-158, 2004.
- [3] Himanshu Thapliyal and Hamid R. Arabnia, "A Time-Area- Power Efficient Multiplier and Square Architecture Based On Ancient Indian Vedic Mathematics", *Department of Computer Science, The University of Georgia, 415 Graduate Studies Research Center Athens, Georgia 30602-7404, U.S.A.*
- [4] E. Abu-Shama, M. B. Maaz, M. A. Bayoumi, "A Fast and Low Power Multiplier Architecture", *The Center for Advanced Computer Studies, The University of Southwestern Louisiana Lafayette, LA 70504.*
- [5] Harpreet Singh Dhillon and Abhijit Mitra, "A Reduced- Bit Multiplication Algorithm for Digital Arithmetics", *International Journal of Computational and Mathematical Sciences*, 2008.
- [6] Shamim Akhter, "VHDL Implementation of Fast NXN Multiplier Based on Vedic Mathematics", *Jaypee Institute of Information Technology University, Noida, 201307 UP, INDIA, 2007 IEEE.*
- [7] Himanshu Thapliyal, Saurabh Kotiyal and M. B Srinivas, "Design and Analysis of A Novel Parallel Square and Cube Architecture Based On Ancient Indian Vedic Mathematics", *Centre for VLSI and Embedded System Technologies, International Institute of Information Technology, Hyderabad, 500019, India, 2005 IEEE.*
- [8] Jagadguru Swami Sri Bharati Krishna Tirthji Maharaja, "Vedic Mathematics", *Motilal Banarsidas, Varanasi, India, 1986.*
- [9] Himanshu Thapliyal and M.B Srinivas, "VLSI Implementation of RSA Encryption System Using Ancient Indian Vedic Mathematics", *Center for VLSI and Embedded System Technologies, International Institute of Information Technology Hyderabad-500019, India.*
- [10] Abhijeet Kumar, Dilip Kumar, Siddhi, "Hardware Implementation of 16*16 bit Multiplier and Square using Vedic Mathematics", *Design Engineer, CDAC, Mohali.*

ABOUT THE AUTHORS



Ramachandran.S received his Bachelor of Engineering in Electronics and Communication Engineering from Anna University, Chennai in 2009. He is currently Pursuing Master of Technology in VLSI Design in Amrita University, Bangalore.

His main research interests include VLSI design, Digital design.



Kirti.S.Pande was born in Maharashtra, India on March 16, 1980. She graduated from Nagpur University in 2001. She completed post graduation from Amrita Vishwa VidyaPeetham in 2010. She Worked as a lecturer at various universities like Nanded University and Visveshwaraya Technological University. Presently she is working as an Asst. Professor at Amrita Vishwa VidyaPeetham and has been engaged with research in VLSI design, Processor based design and FSM based Digital design.

Performance Comparison Study of AODV, OLSR and TORA Routing Protocols for MANETS

Manjeet Gupta¹ and *Sonam Kaushik²

¹ Assistant Professor, C.S.E, Kurukshetra University,
JMIT(Radaur), Haryana, India

²C.S.E, Kurukshetra University,
Kurukshetra, Haryana, Pin-136118, India

Abstract

Mobile Ad hoc Networks (MANETS) are the special type of wireless network, where mobile nodes are connected through wireless interfaces forming a temporary network. They don't need fixed infrastructure. Due to higher mobility in nodes and dynamic infrastructure of MANETS, Routing is important issue in ad hoc networks. There are many routing protocol in MANETS like AODV, TORA, DSDV, OLSR, DSR etc. MANETS is classified in three routing protocols. This research paper make a comparison of these routing protocol based on the performance metrics like packet delivery fraction, end-to-end delay and throughput. Simulation is used to compare the performance of AODV, OLSR and TORA. NS2 (Network Simulator version2) is used as simulator. With the help of ns-2, result shows that AODV's performance in PDF and throughput metrics is better than OLSR and TORA. For end-to-end delay metrics TORA perform better than OLSR and AODV.

Keywords: . AODV, MANET, OLSR, Routing Protocols, , TORA.

1. Introduction

Mobile Ad-Hoc Networks are autonomous and self-configuring wireless systems. MANETS consist of mobile nodes that are free to move in and out of the network. These node can be mobile phone, system etc. Mobility affects the power indulgence of the nodes in a MANET. This is because of the high overhead incurred in Route Discovery and Route Maintenance in mobile nodes. Due to higher mobility of nodes they form random topologies depending on their connectivity with each other in the network. The dynamic topology makes the routing protocol design complex.

APPLICATIONS OF MOBILE AD HOC NETWORKS

Table 1: Application of MANETS

<i>Applications</i>	<i>scenarios/services</i>
Educational Applications	<ul style="list-style-type: none">• In Universities and campus.• To Setup virtual classrooms, conference rooms.
Tactical networks	<ul style="list-style-type: none">• In military , battlefield• Rescue operations
Entertainment	<ul style="list-style-type: none">• Outdoor Internet access.• Multi-user games.• Robotic pets.• Wireless P2P networking.
Enterprise Networking	<ul style="list-style-type: none">• Personal Area Network .• Conferences, meeting rooms.• Networks at construction sites.

Emergency services	<ul style="list-style-type: none"> • Search and rescue operations • Disaster recovery • In hospitals • Policing and fire fighting
Sensor networks	<ul style="list-style-type: none"> • Home applications: smart sensor nodes. • Environmental applications

1.1 MANETs routing protocols

Ad-Hoc network is called as Mobile Ad-Hoc Network (MANET) because of mobility of nodes in network. They are IBSS (Independent Basic Service Set), because they does not need AP(Access Point) for communication in nodes. MANETs is a self-configuring network and form an uninformed topology. These nodes behave like routers in network to route the packet. MANETs are used in those areas where wire and wireless infrastructures are unreachable. Due to rapid change of topology in MANETs, MANETs routing protocols are required. The routing protocol is required whenever the source needs to communicates with destination. Routing protocols are classified as Proactive (Table Driven Routing Protocol), Reactive (On Demand Routing Protocol) and Hybrid (having the advantages of both proactive and Reactive routing protocols) routing protocols.

MANETs routing protocols are classified as:-

- A. Reactive protocols
- B. Proactive protocols
- C. Hybrid protocols

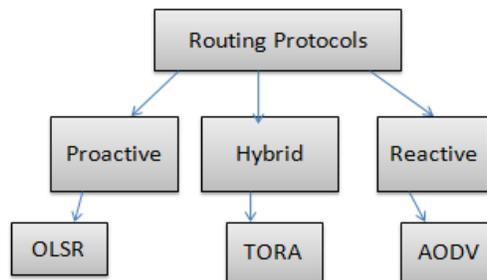


Fig. 1 MANETs Routing Protocols

A. Reactive Protocols:

Reactive protocols are also called as on demand driven reactive protocols. It is mainly used to find the route between source and destination as needed. As per the demand of source this routing protocol initiate route discovery, to find the route to the destination. Then this route is used for further communication [1, 2] e.g. AODV.

B. Proactive Protocols:

Proactive protocols also called as Table driven routing protocols. Each node maintains routing tables which are consistent and up-to-date containing routing information for every node in the network. Whenever new node is entered in the network or removes from the network, control messages are sends to neighboring nodes then they update their routing tables. This routing protocol uses link-state routing algorithms which frequently flood the link information about its neighbors. Proactive routing protocols are OSPF and OLSR [2].

C. Hybrid Routing Protocol:

Hybrid routing protocol have advantages of both proactive and reactive routing protocols. Firstly it behave like proactive routing protocol, because in starting nodes have tables. Then whenever nodes finds that they does not have route to destination, they start route discovery and behave like reactive routing protocols .Hybrid protocols are TORA and ZRP.

1.2 Overview of Protocols

A. Ad-Hoc on Demand Distance Vector Protocol (AODV):

AODV [3] is reactive protocol, when a source wants to initiate transmission with another node as destination in the network, AODV use control messages to find a route to the destination node in the network. AODV will provide topology information (like route) for the node. Fig.2 shows the message routing for AODV protocol. Node “A” wants to send messages to another node “F”. It will generate a Route Request message (RREQ) and forwarded to the neighbors, and those node forward the control message to their neighbors’ nodes. Whenever the route to destination node is located or an intermediate node have route to destination. They generate route reply message (RREP) and send to source node. When the route is established between “A” and “F”, node then they communicate with each other.

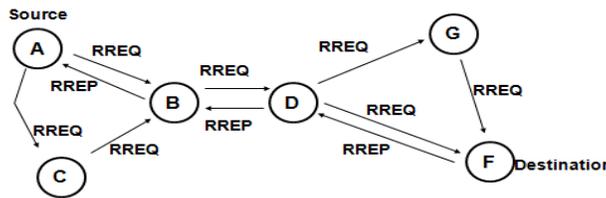


Fig 2Message routing in AODV

B. Optimized Link State Routing Protocol (OLSR) :

OLSR is proactive routing protocol or table driven protocol. Initially nodes have routing tables and they update their routing tables time to time. It is based on the link-state algorithm. Each node maintains the topology information of network and sending this information from time to time to neighbors. The uniqueness of OLSR is that it minimizes the size of control messages and rebroadcasting by using the MRP (Multipoint Relaying). The basic concept of MRP is to reduce the loops of retransmissions of the packets. Only MRP nodes broadcast route packets. The nodes within the network maintain a list of MRP nodes. MRP nodes are selected within the environs of the source node. The selection of MRP is done by the neighbor nodes in the network, with the help of HELLO messages.

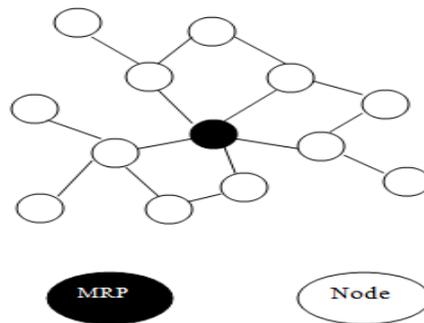


Fig. 3.3 Flooding Packets using MRP

C. Temporally Ordered Routing Algorithm (TORA):

TORA is a hybrid protocol, based on a “link reversal” algorithm. It discovers multiple routes to a destination, create routes quickly, and diminish communication overhead. Nodes have routing tables, so it helps the sending node to find the route to destination with the help of given tables. Routing tables also maintains the longer routes to avoid discovering newer routes. When a node finds that a route to a destination is no longer valid, it adjusts its height so that it is a local maximum with respect to its neighbors and transmits an UPDATE packet. If the node has no neighbors of finite height with respect to this destination, then the node discover a new route. When a node detects a network partition, it generates a CLEAR packet which resets routing tables and removes invalid routes which does not exist from the network.

1.3 Related Work

AODV, OLSR and TORA have lot of attention in recent times. B.M Sobral et. al. [5] compared the performance of AODV and OLSR using a self-configuration mechanism for Heterogeneous Wireless Mesh Networks. As relevant results they got some improvements related to the original OLSR and AODV protocols, by applying the self-configuration capacity, such as, the increased throughput of the overall network, improvement of the delay of discovery local neighboring routes by

reducing the HELLO traffic messages and the MPR Count metric and finally the improvement of the delivery of packets due to smaller dropped packets.

Dong-Won Kum, Jin-su-park et. al. [6] they propose an Mobility aware Hybrid Routing (MHR) approach for WMNs, which varies its routing between reactive and proactive to adapt to node mobility using ns-2. Applied MHR approach to AODV, called AODV-MHR and compared its performance with that of AODV and OLSR in terms of routing overhead, total throughput, and the average end-to-end delay by simulation. Simulation results showed that AODV-MHR's performance was enhanced by the advantages attained by using both reactive and proactive routing approaches.

Jing Xie, Yuming Jiang [7], they propose a threshold-based hybrid routing protocol that supports a mobile node to selectively run the routing protocol based on its velocity. They Theoretically analyzed THRP, OLSR, AOD and ZRP and found THRP can show better scalability than both OLSR and AODV since it partly shortens the AODV route discovery process and constrains the local control overhead of OLSR to flood in the whole network.

Julian Hsu, Sameer Bhatia Mineo Takai [8], They compare the Performance of AODV, DSR, OLSR, OLSR v2 and ZRP in REALISTIC SCENARIOS. AODV protocol performed best in this type of scenario, with a slight edge in overall throughput.

2. Methodology

2.1 Simulation Environment:

Simulations are done to compare these routing protocols. Simulator ns-2 is used for performance comparison. The network simulator ns-2 developed by the VINT research group at University of California at Berkeley in 1995. The network simulator NS2 is a discrete event network simulation. NS2 is used to simulate the proposed algorithm. It work on network layer and inform about link breakage. The implementation of the protocol has been done using C++ language in the backend and TCL language in the frontend. TCL (Tool Command Language) is compatible with C++ programming language.

During interpretation two files trace files and nam files are to be generated. Network Animator (.nam) file, records all the visual events that happened during the simulation. Trace files (.tr), records the entire network event that occur during the simulation. And file is post analyzed with the help of awk scripts.

Table 2 : Simulation Parameter

<i>Parameter</i>	<i>value</i>
Simulation Time	50 Sec
No. of Nodes	50
Traffic Type	CBR
Pause Time	10 Sec
Maximum X-Y coordinate value	1000 M
Packet Size	512 byte
MAC Protocol	802.11
Mobility Model	Random Waypoint
Routing Protocols	AODV, OLSR, TORA
Observation Parameters	EED, Throughput, PDF

2.2 Performance Metrics:

The estimation of performance of AODV, OLSR and TORA is done on the basis of following Performance metrics:

- **Packet Delivery Ratio:** It is the ratio of the packets received by destination to those generated by the sources. CBR traffic type is used by source. It specifies the packet loss rate, which limits the maximum throughput of the network. The routing protocol which have better PDR, the more complete and correct. This reflects the usefulness of the protocol. And provide good performance.

$$\text{Packet Delivery Ratio} = (\text{Received Packets} / \text{Sent Packets})$$

- **End to End Delay:** Average end-to-end delay is the average time it taken by the packet to reach to destination in seconds.
- **Throughput:** No. of packet passing through the network in a unit of time. It is measure in kbps.

3. Result and Discussion

3.1 Packet Delivery Ratio Graph:

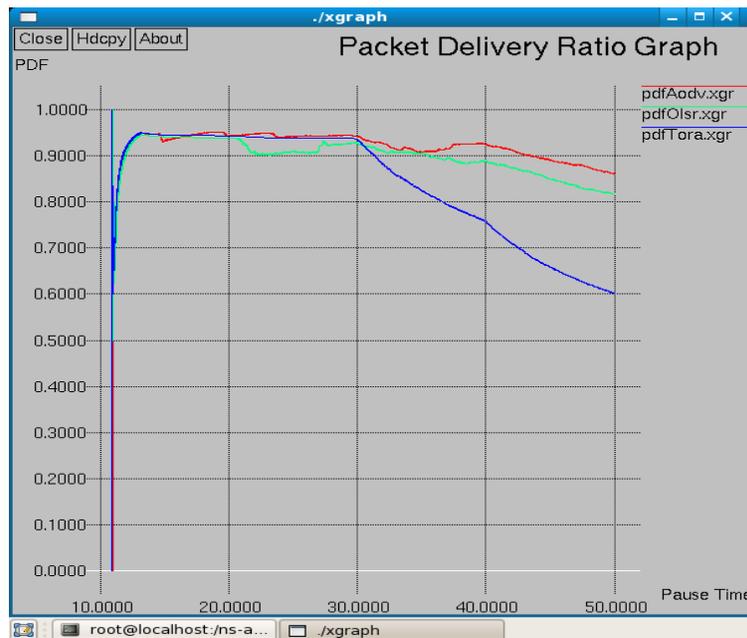


Figure 3: Packet Delivery Ratio vs. Pause Time

Figure 3 shows the PDR of the three protocols AODV, TORA and OLSR. The PDR of AODV is greater than OLSR and TORA. At the high mobility all three protocols behave same but with less mobility AODV has maximum PDR, OLSR lies in between AODV and TORA. TORA has minimum PDR. Routing protocol which has higher PDR, it performed better as good the routing protocol performed. Because it states that maximum packets are received by destination node.

3.2 End to End delay Graph

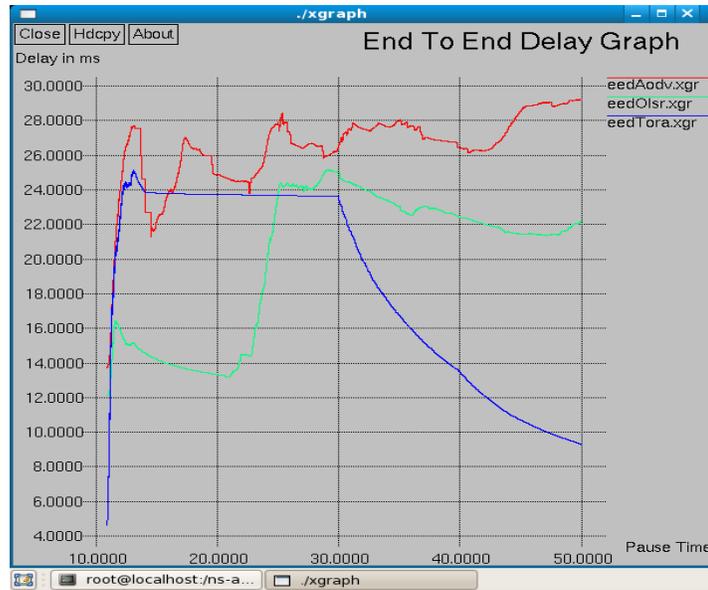


Figure 4: End to End delay vs. Pause Time

Figure 4 compares the average End-To-End packet delay between AODV, TORA and OLSR. In this figure, TORA exhibited the lowest average end-to-end delay, while AODV had the highest delay. The end-to-end delay of OLSR is lowest at a higher mobility, but with higher mobility it is increases and become greater than that of TORA. TORA had lowest average end-to-end delay because it has routing tables need not to rediscover the route for the same destination.

3.3 Throughput Graph:

Below Fig. 5 shows the total throughput of AODV, TORA and OLSR. Total throughput is the amount of packet transferred through the network per unit time. At a higher mobility, OLSR exhibit the lowest total throughput, while TORA had the highest throughput at a higher mobility. The average throughput of AODV is higher than OLSR, TORA. Higher the throughput, better the routing protocol performed.

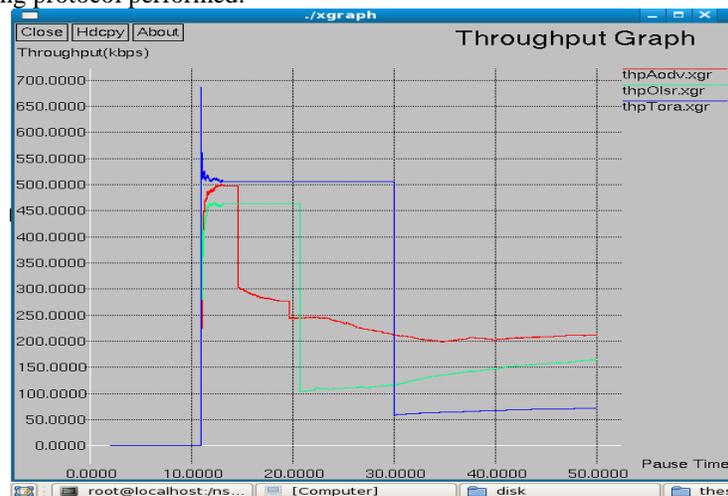


Figure 5: Throughput (kbps) vs. Pause Time

Table 3: Performance Summary of AODV, OLSR, AND TORA

Performance metrics	Routing Schemes	High Mobility	Low Mobility
Packet Delivery Ratio	AODV	High	High
	OLSR	High	Medium
	TORA	High	Low
Total throughput	AODV	Medium	High
	OLSR	Low	Medium
	TORA	High	Low
Average end-to-end delay	AODV	High	High
	OLSR	Low	Medium
	TORA	Medium	Low

Table 4: Performance Comparison of AODV, OLSR, AND TORA

Performance metrics		AODV	OLSR	TORA
Packet Delivery Function	CBR	s:3934 r:3383	s:2993 r:2445	s:1595 r:958
	r/s	.8599	.8169	.6006
	F	4465	4013	0
Total throughput(in kbps)	Avg. Thp.	211.26	164.71	72.07
	Start Time	10.91	10.95	10.90
	Stop Time	49.98	49.99	49.58
Average end-to-end Delay(in sec)		29.2420	22.175877	9.30079

4 Conclusion:

Mobile Ad-Hoc Networks has the ability to deploy a network where a traditional network infrastructure environment cannot possibly be deployed. With the importance of MANET comparative to its vast potential it has still many challenges left in order to overcome. Performance comparison of routing protocol in MANET is one of the important aspects. In this paper, I have analyzed the behavior and different performance matrices for MANETs using different protocols. (AODV, OLSR and TORA) and compared their performance matrices, like End to end delay, Packet delivery Fraction and Throughput . In Table 3 and 4 performance comparisons of routing protocols AODV, OLSR, TORA is shown using ns2 simulator. For Throughput and PDF, AODV behaving the best and for End to End delay is concern TORA is taking less delay.

4 ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my guide Mrs. Manjeet Gupta, Assistant Professor, Computer Science and Engineering Department for immense help, guidance, stimulating suggestions, and encouragement all the time with this thesis work. This work would have not been possible without her support. She always provided a motivating and enthusiastic atmosphere to work with; it was a great pleasure to do this thesis under her supervision.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction, to help me stay calm in the oddest of times and keep moving at times when there was no hope.

Sonam Kaushik

References

- [1] C.E.Perkins and E.M.Royer, "Ad-Hoc on Demand Distance Vector Routing", Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications, pp.90-100, Feb, 1999.
- [2] C.M barushimana, A.Shahrabi, "Comparative Study of Reactive and Proactive Routing Protocols Performance in Mobile Ad-Hoc Networks", Workshop on Advance Information Networking and Application, Vol. 2, pp. 679-684, May, 2003.
- [3] C.Parkins, E.B.Royer, S.Das, A hoc On-Demand Distance Vector (AODV) Routing. July 2003, [Online]. Available: <http://www.faqs.org/rfcs/rfc3561.html>. [Accessed: April. 10, 2010]
- [4] T.Clausen, P.Jacquet , "Optimized Link State Routing Protocol (OLSR)", RFC 3626 October, 2003.
- [5] Lucas Guardalben, Joao B.M. Sobral,"A Performance Evaluation of OLSR and AODV Routing Protocols Using a Self-Configuration Mechanism for Heterogeneous Wireless Mesh Networks" guardalben,bosco@inf.ufsc.br 978-1-4244-2413-9/08/©2008 IEEE
- [6] Dong-Won Kum,Jin-su-park," Mobility aware Hybrid Routing (MHR) approach for WMNs" {80kumsy, yzcho}@ee.knu.ac.kr,2010
- [7] Jing Xie,Yuming Jiang," Threshold-based hybrid routing protocol for Manets" ymjia@ieee.org, 2007
- [8] Julian Hsu, Sameer Bhatia Mineo Takai," compare the Performance of AODV,DSR, OLSR, OLSR v2 and ZRP in REALISTIC SCENARIOS"

EIA for Ramapada Sagar (Polavaram) Irrigation Project using the Model of RS and GIS

Sreeramulu. Y^{1*}, Murali Krishna.I.V²

1. Associate Professor, Dept.of Civil Engg, KSRM College of Engineering, Kadapa-516003, A.P., India

2. Adjunct Professor, Asian Institute of Technology, Bangkok, Mobile No 09848049624

Abstract

At Global level, Asian countries like India and China have experienced untold environmental degradation and ecological deterioration in the past century, with little or no real solution to alleviate many of these concerns. Poorly planned human interference has been the major cause. Adequate information and appropriate technology are limiting factors for effective environmental management. Hence, efforts to improve, conserve and protect the environment will include not only the resolution of political policies but also the application of a state-of-the-art scientific approach to planning and implementation. The process of Environmental Impact Assessment (EIA) was developed as an effective planning tool. The genuine conduct of this process will go a long way in reducing environmental deterioration. Because of the dynamic characteristics and multivariate nature of the environment, it has often been difficult to collate, analyze and interpret its data sets. However, this great complexity can be overcome with the present research of engineering management system model of Remote sensing and geographical information system and related technology with the ground truth verification.

This Research study deals with the Environmental impact assessment for an irrigation project i.e.Ramapada Sagar (Polavaram) Irrigation project, which has been carried out in parts of West Godavari, East Godavari and Khammam districts of Andhra Pradesh. About 14,400 Sq Kilometer area of remote sensing data have been collected and analyzed for environmental impact assessment by using emerging GIS technology. The prime objective is to study the environmental impact of the project on land use land cover environment, water environment, , submergence area and formulate suitable environmental management plan for minimizing expected adverse impacts during as well as after the implementation of project.

From the land use land cover environment, the crop land covers 79.8% with an area of 3,92,766 ha, plantations 49,505ha, built up land occupies 23,262ha (4.72%), forest occupies a mere 0.38%, tanks and streams occupies 3,819 ha (0.82%) and scrub land occupies 19,680ha (3.37%). The slopes of the catchment mostly range between 1% to 5% (i.e., nearly level to gently sloping). The agriculture is practiced generally in the nearly level to gently sloping areas. In the catchment about 91.55% of the area comes under nearly level to gently sloping category (i.e., 0% - 5% slope). Only 0.45% of the catchment area comes under moderately sloping to strongly sloping (i.e., between 5% - 15% slope). The nearly level, very gently sloping, gently sloping areas are 77.25%, 21.44%, and 0.86% respectively. From the slope map, erosion intensity can be estimated and suitable measures that are required to restrict the siltation of the proposed Reservoir by biotic treatment, engineering treatment and gully control works can be suggested.

It has been observed during water environment analysis; an area of about 50 sq.km is under inundation at low water level +135ft (41.15m), 200 sq.km is under inundation at +140ft (42.67m), and an area of 360 sq.km is under Inundation at +150ft (45.72m) FRL.

Key words: Environmental impact assessment, Ramapada sagar (Polavaram) Project, Remote sensing (RS), Geographic Information system (GIS), Catchment area, land use land cover , water environment , submergence / Inundation

1.0 Introduction

Environmental Impact Assessment (EIA) is a planning and management tool for sustainable development, aimed at providing decision-makers with information on the likely consequences of their actions. Thus EIA can be considered as being anticipatory in nature. EIAs define and assess the potential physical, biological, socio-economic and health effects of the proposed project in a manner that allows for a logical and rational decision to be made about the proposed action (Glasson et al., 1995; Wathern, 1988; Wood, 2003) [32].

Asian countries like India and China have experienced untold environmental degradation and ecological deterioration in the past century, with little or no real solution to alleviate many of these concerns. Poorly planned human interference has been the major cause. Adequate information and appropriate technology are limiting factors for effective environmental

management. Hence, efforts to improve, conserve and protect the environment will include not only the resolution of political policies but also the application of a state-of-the-art scientific approach to planning and implementation. The genuine conduct of this process will go a long way in reducing environmental deterioration. Because of the dynamic characteristics and multivariate nature of the environment, it has often been difficult to collate, analyze and interpret its data sets. However, this great complexity can be overcome with the present research of Engineering Management system Model of Remote sensing and Geographic Information System (GIS) and related technology with the ground truth verification. The prime objective is to study the environmental impact of the project on land use land cover environment, water environment and submergence area and formulate suitable environmental management plan for minimizing expected adverse impacts during as well as after the implementation of project.

1.1 Study Area

This study deals with Environmental impact assessment for developmental project i.e. Ramapada Sagar (Polavaram) Irrigation project, which has been carried out in parts of West Godavari, East Godavari and Khammam districts of Andhra Pradesh. The Godavari River originates in the Nasik district of Maharashtra, India and flows through West Godavari district of Andhra Pradesh, India and discharges directly into the Bay of Bengal Sea. The study area is represented in Figure 1.1, located between $81^{\circ}-46^{\circ}$ E longitude and $17^{\circ}-13^{\circ}$ N latitude, and covers parts of the Survey of India topographic sheet numbers 65G/7,8,11,12.

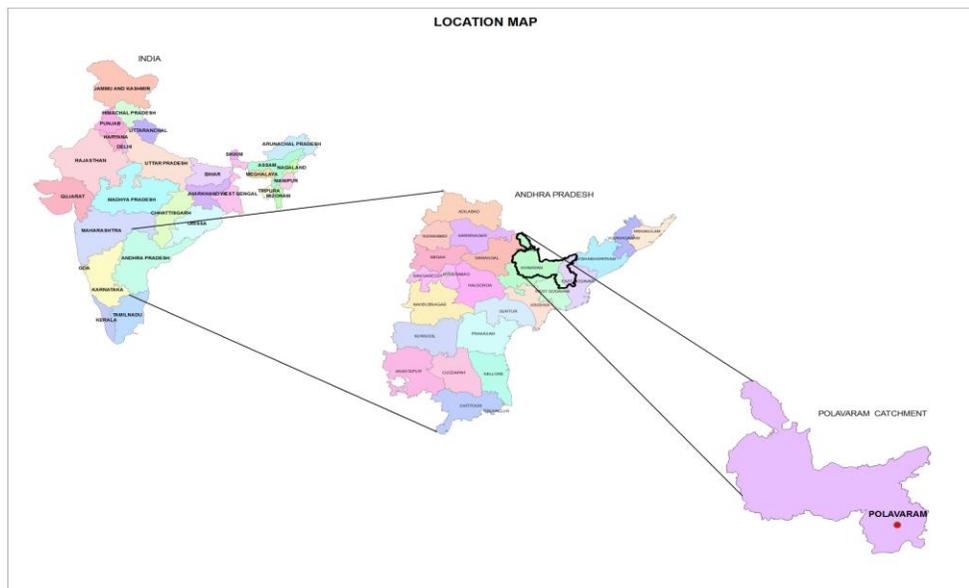


Figure 1.1: Location map of Ramapada Sagar (Polavaram) Project Study area

The proposed project is a multipurpose project benefiting the upland areas of Visakhapatnam, East Godavari, West Godavari, Krishna and Khammam districts. It also supplies drinking water to Visakhapatnam township and villages enroute. The general climate of the command is characterized by hot summer and general dryness which gives copious rainfall. May is the hottest month with a maximum day temperature 44°C and the minimum temperature is about 22°C . The south-west monsoon season during mid of June and ends by mid October. The annual average rainfall is about 858.65 mm. In the study area the humidity is very high during peak rainy season. Mean humidity generally varies between 62% and 80% as per the data collected for 27 years. Highest humidity of 80% is during June to September. Average wind velocity varies between 4.2 kmph in April and 8.47 kmph in November. The average rainfall in the command area is 859 mm. The finalized project site consists of a zoned Earth-cum-Rockfill dam with an impervious core across the existing river course. The spillway is located in the right flank saddle and power dam on the left flank saddle. The FRL of the reservoir is proposed as +150ft (45.72m).

2.0 Research work Model proposed

Model Name	Description	Major Features	Data Requirements	Out Put
<p>CE – EIPS – RS & GIS</p> <p>[Civil engineering Environmental impact assessment for an irrigation project interms of submergence using Remotesensing and Geographic information system]</p>	<p>2D – Longitudinal, Vertical Reservoir water quantity and quality model for Environmental Impact Assessment using Remote sensing and GIS</p>	<ul style="list-style-type: none"> - Description of Land use / Land cover - Assessment of Water Environment - Assessment of submergence area under various discharge flows - Allow simulation of most major physical , chemical and biological process and associated water quality constituents 	<p>Primary data :</p> <ul style="list-style-type: none"> - Water Quantity , Quality data and coefficients - Physical data , cross section geometry , elevations and locations of nodes; Lateral inflows and tributaries ; control structures <p>Secondary data :</p> <ul style="list-style-type: none"> - River Flows, Depths, spot heights and velocities - Water Quantity and Quality targets at system control points. 	<ul style="list-style-type: none"> - Classification of Land use / Land Cover area and suggestive crops - Submergence / Inundation area due to the Reservoir water quantity (Printed and / or plotted) - Vertical profiles and outflow values for constituents over time (Printed and / or plotted) - Reports of water quality parameter values for drinking and Irrigation purposes - Catchment area treatments - Site selection for Resettlement & Rehabilitation purposes

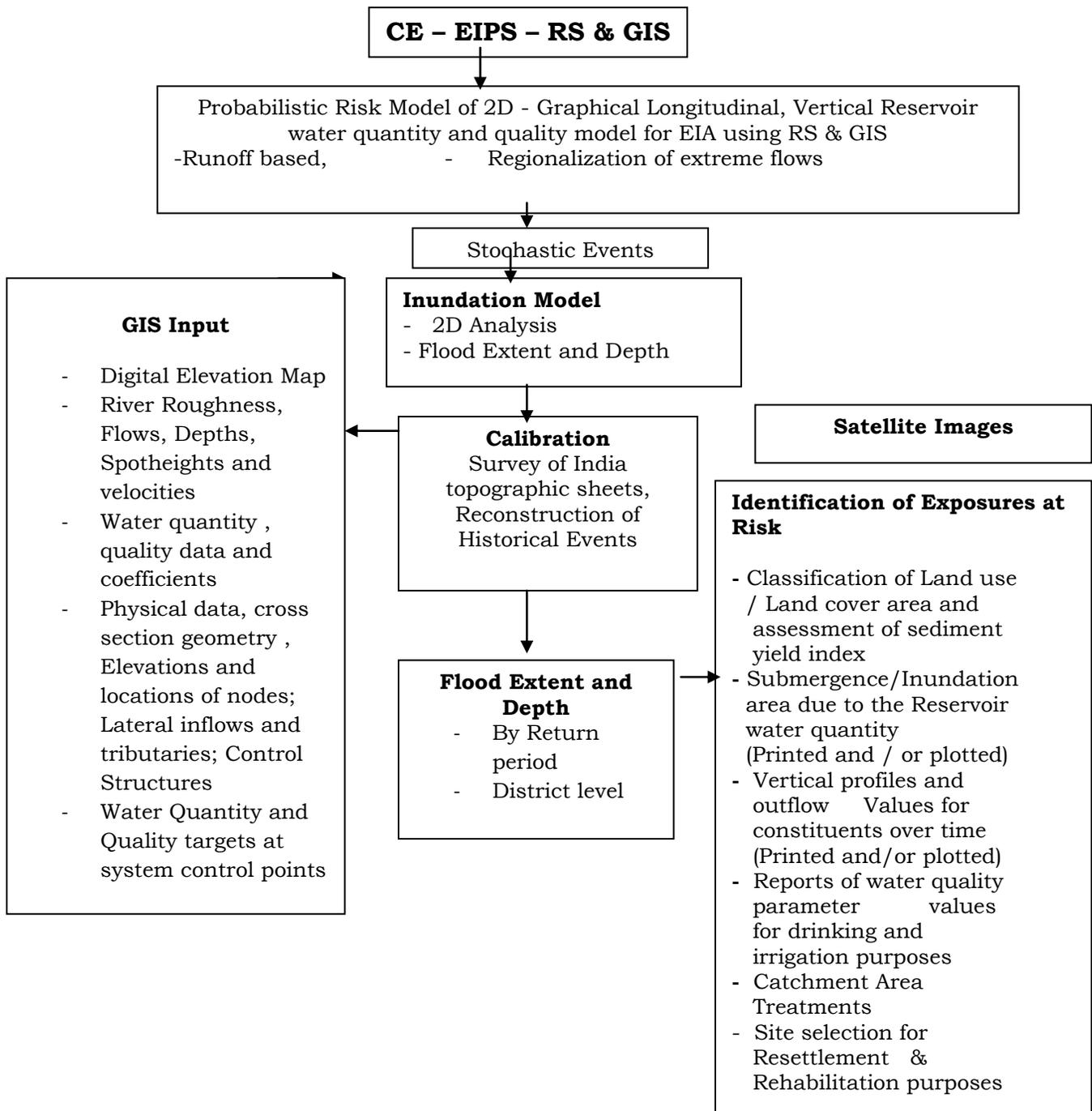


Figure.2.1 Flow chart of Research work Graphical model proposed on Environmental Impact assessment for Ramapada sagar (Polavaram) Irrigation Project using Remote sensing and Geographic Information system with ground truth verification

2.1 Mathematical Expressions involved in the Research work model proposed

Probability Distributions:-

Generalized Pareto Distribution (GPD)

Cumulative Distribution Function when $P [X \leq x]$, $CDF F(x) = 1 - e^{-y}$ where
 $Y = -K^{-1} \log [1 - k(x - \xi) / \alpha]$

$$Y = (x - \xi) / \alpha \quad \text{for } k = 0$$

Generalized Extreme value distribution [GEV]

CDF , $F_x(x) = \exp \{ - [1 - (kx - \xi) / \alpha]^{1/k} \}$ for $k \neq 0$

$$F_x(x) = \exp \{ - \exp[-(x - \xi) / \alpha] \}$$
 for $k = 0$

Runoff Quantity Density Function =

Derivative $f(x) = d/dx [F(x)]$

$$\begin{aligned} f(x) &= d/dx (1 - e^{-y}) \\ &= -d/dx \{ 1 - [k(x - \xi) / \alpha] \}^{1/k} \\ &= 1 / \alpha \{ 1 - [k(x - \xi) / \alpha] \}^{(1-k)/k} \end{aligned}$$

- X = Rainfall
- f(x) = Runoff
- K = Slope factor / Velocity
- α = Time
- ξ = Infiltration
- Y = Runoff function coefficient

3.0 Materials and methods

The study utilised the Survey of India toposheets of 1:50,000 scale for studying the catchment area, command area and also the submergence area for identification of land use land cover, slope and soils, etc. The satellite based remote sensing imageries are procured from the NRSA for image analysis and arriving at the land use land cover, slope, soils, surface drainage, etc. and also for preparation of catchment area treatment plan based on Sediment Yield Index method. But in this paper land use / land cover, slope aspect and submergence / inundation are only represented.

The methodology comprises access to Remote Sensing Data and analysis of the same, ground truth verification, collection of primary and secondary data, group discussions etc. The study covers land environment, water environment and submergence area. The data were processed through computers with suitably designed software. The following tasks have been undertaken to meet the study objectives

- Generate thematic maps of various natural resources
- Integrate the thematic maps
- Define the plan of implementation

The input data from all of the above diverse sources are translated into the thematic maps by the methods of: interpretation, classification, manipulation, integration, editing and analysis. The data translation into thematic maps employed the GIS software Arc/Info, Arc View, Arc Map and the remote sensing software ERDAS. The Impact is determined by taking the Spot heights as reference.

4.0 Results

The following **Figure 4.1, Figure 4.2 and Figure 4.3** are the output of Research work model proposed:

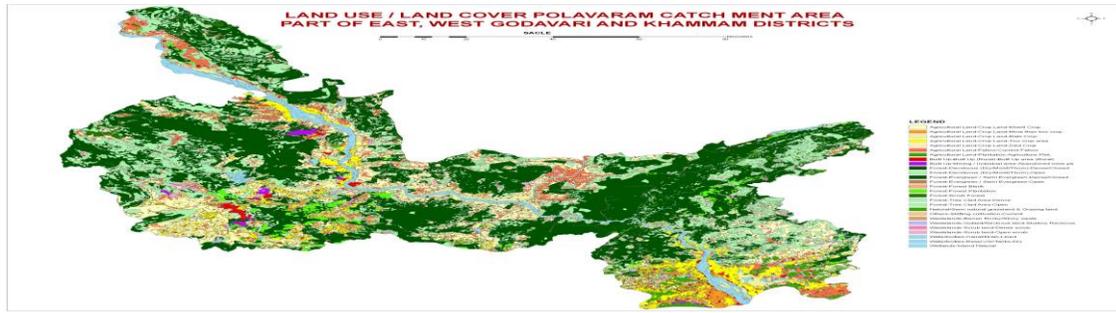


Figure 4.1 shows the Land use / Land cover of Ramapada Sagar (Polavaram) Catchment Area , year 2008-09 (Satellite data: 05.11.2008)

Digital interpretation of IRS LISS-III FCC on 1:50,000 scale is worked out in ERDAS for identification of different land use and land cover classes based on the image characteristics. The multitemporal imagery are interpreted for the details of the crop land in the two harvest seasons known as the kharif and Rabi seasons. Based on ground truth verification the boundaries are finalized which synchronizes well with the physiography, slope and soil of the area. The analysis of remote sensing data provided the area under different land use and land cover under different categories of the catchment and this area has been represented in Figure 4.1. The crop land covers 79.8% with an area of 3,92,766 ha, plantations 49,505ha, built up land occupies 23,262ha (4.72%), forest occupies a mere 0.38%, tanks and streams occupies 3,819 ha (0.82%) and scrub land occupies 19,680ha (3.37%). The major soils in the command area are moderately deep gravelly sandy loams and lateritic upland (1,92,308 ha); followed by deep fine sandy loam (1,39,403 ha). The other types of soils are deep sandy loam on undulating land, moderately deep sandy loams on rolling lands, deep clayey soils on gently sloping, very deep fine loamy / clayey soils in valley, etc. The majority of the soil types fall in land irrigability/ classification 2 & 3.

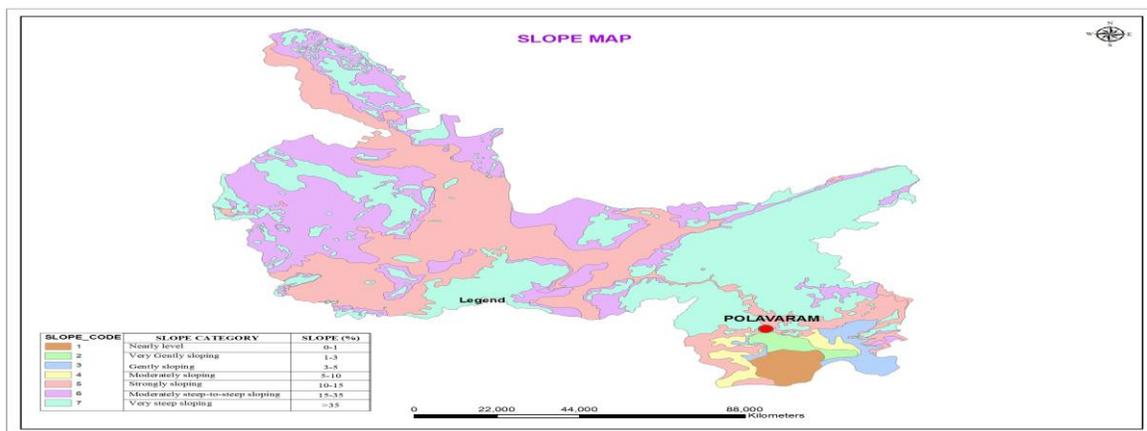


Figure 4.2 Slope Map

The slope map (Figure 4.2) has been prepared for the catchment areas using 1:50,000 scale topographical maps of Survey of India showing contours of 20 m interval using tan method. The different classes of slopes have been categorized as per the guidelines suggested by All India Soil and Land Use Survey (AIS&LUS). The vertical drop is measured from the contour interval and the horizontal distance between the contours is measured by multiplying the map distance with the scale factor. Finally the slope percentage is calculated. From the figure 4.2, the slopes of the catchment mostly range between 1% to 5% (i.e., nearly level to gently sloping). The agriculture is practiced generally in the nearly level to gently sloping areas. In the catchment about 91.55% of the area comes under nearly level to gently sloping category (i.e., 0% - 5% slope). Only 0.45% of the catchment area comes under moderately sloping to strongly sloping (i.e., between 5% - 15% slope). The nearly level, very gently sloping, gently sloping areas are 77.25%, 21.44%, and 0.86% respectively. From the slope map, erosion intensity can be estimated and suitable measures that are required to restrict the siltation of the proposed Reservoir by biotic treatment, engineering treatment and gully control works can be suggested.

Analysis of Water Environment

Surface Water Quality

100 Water samples were collected during the year 2007-08 at the proposed site and got analysed. It is concluded that the Godavari River water at the project site is found to be chemically suitable for irrigation and drinking water purposes.

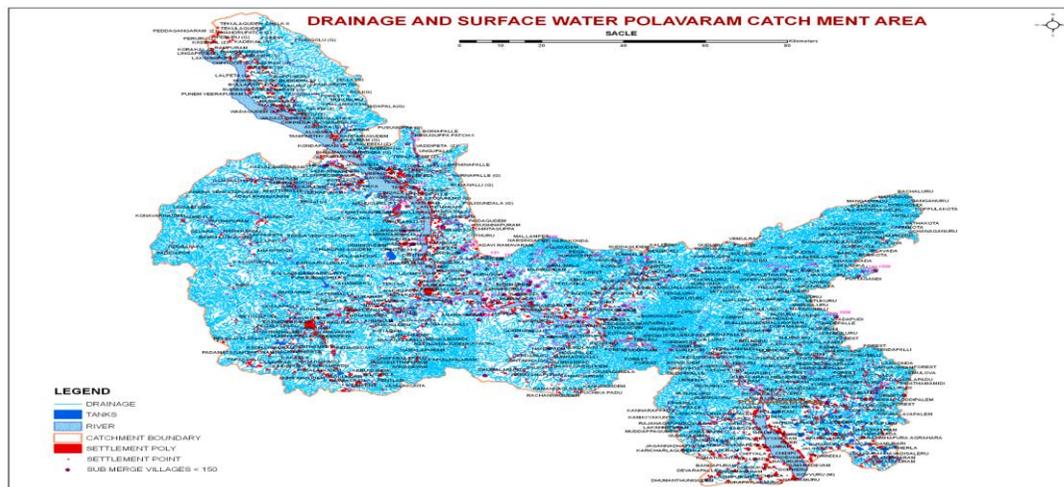


Figure 4.3 Drainage and Surface water of Ramapada Sagar (Polavaram) Catchment Area

From the Figure 4.3 , due to the construction of the Ramapada Sagar (Polavaram) on the Godavari River, an area of about **50 sq.km** is under inundation at low water level +135ft (41.15m), **200 sq.km** is under inundation at +140ft (42.67m), and an area of **360 sq.km** is under Inundation at +150ft (45.72m) FRL.

5.0 Conclusions

The model of Civil Engineering Environmental impact assessment for an irrigation project interms of submergence using Remotesensing and GIS (CE-EIPS – RS&GIS) is developed and successfully worked out. The analysis of remote sensing data provided the area under different land use and land cover under different categories of the catchment and this area has been studied. The model also helps for the preparation of water environment with the representation of submergence / inundation area under different water levels.

The final conclusion is, that this Civil Engineering Management system model using Remote sensing and Geographical information system with ground truth verification is a unique model for assessing results for Environmental impact assessment process interms of, Land use and Land cover environment, Water environment and submergence area etc., . This model help contribute to the resolution of the many environmental problems plaguing in Asian countries.

Acknowledgment

My sincere thanks to the JNTUH higher authorities for providing research facilities. I am thankful to Prof.G.L.Sivakumar babu and Prof.Sivapullaiah, IISC, Bangalore for advises.

6.0 REFERENCES

- [1]. Alwashe.M.A. and A.Y.Bokhari, Monitoring vegetation changes in AL Madinah. Saudi Arabia, using Thematic Mapper data.Int.J.Remote Sensing, 14(2), pp 191-197, 1993.
- [2]. Anselin, L.1998, exploratory Spatial Data Analysis in a Geocomputational Environment, pp.77-94 in Longley, P.A.Brooks, S.M.Mc Donnell.R, and Macmillan, B.,eds., Geocomputation: A Primer, New York: Wiley.
- [3]. Anjaneyulu.Y,Environmental Impact Assessment Methodologies, 2002, 30-72
- [4]. Ade Abiodu, A.2000.Development and Utilization of Remote Sensing Technology in Africa. Photogrammetric Engineering and Remote Sensing, 66, no.6,674-686.
- [5]. Banham, W. and D. Brew (1996). "A review of the development of Environment Impact Assessment in India" Project Appraisal 11(3) : 195-202.
- [6]. Bethel, J/S.1995. Photogrammetry and Remote Sensing . In Chen W.F.(ed.), The Civil Engineering Handbook, 1964-2001, New York; CRC Press, Inc
- [7]. Biksham Gujja, Ramakrishna Sangem, Vinod Goud, Sivarama Krishna, (2006), Perspectives on Polavaram, 1, pp.19-40.
- [8]. Canter, L.W. (1996) Environmental Impact Assessment (New York, Mcgrahill)
- [9]. Dee, N.Drobny, N.L. and Duke, K.(1973) An environmental evaluation system for water resource planning. Water resources Research, 9, 523-553.

- [10]. Eidenshink.J.C. The 1990 conterminous U.S.AVRR data set. Photogrammetric Engineering and Remote Sensing, 58(6), 809-813, 1992.
- [11]. Estes, J.E.MeGuire,K.E.Fletcher, G.A and Foresman, T.W.(1987) Coordinating hazardous waste management activities using geographical information systems. International Journal of Geographical Information Systems 1(4), 359-377.
- [12]. Fisher, P.F.1991 Spatial Data Sources and Data Problems. In Maguire, D.J.Good child, M.F., and Rhind, D.W.(eds.) Geographical Information Systems: Volume 1:Principles, 175-189, Essex, England: Longman.
- [13]. Gautam N.C. and Narayana L.R.A., (1982): Suggested Land Use / Land Cover classification system for India using remote sensing techniques, Pink publishing house, Mathura.
- [14]. Goodchild, M.F.1993a. The State of GIS for Environmental Problem – Solving Pp.8-15 in Goodchild, Parks, and Steyaert, 1993,op.cit.
- [15]. Harmancioglu, N.B.,Singh, V.P. and Aplaslan, M.N.(eds) 1998. Environmental Data management. Dordrecht: Kluwer Academic Publishers.
- [16]. Hinoton,J.C.1996. GIS and Remote Sensing Integration for Environmental Applications. International journal of Geographical Information Systems, 10 no.7,877-891.
- [17]. International Journal of Sediment Research, Vol.23, No.2,2008, pp.167-173.
- [18]. Karanth, K.R (1987), Ground Water Assessment Development and Management, Tata McGraw-Hill Publishing Company Limited, New Delhi.
- [19]. Larsen, L.,1999. GIS in Environmental monitoring and Assessment. In Longley. P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W.(eds.,) Geographical Information Systems, Volume 2: Management Issues and Applications. 999-1007. New York: Wiley.
- [20]. Levinsohn, A.2000 GIS Moves from Computer – Aided Mapping to Spatial Knowledge Representation. Geoworld. 13 no.2
- [21]. Lillesand.T.M., and Kiefer, R.W.1994 Remote Sensing and Image Interpretation. New York: Wiley.
- [22]. Environmental Assessment - Andhra Pradesh Irrigation Department Book-I
- [23]. Environmental Management Plan - Andhra Pradesh Irrigation Department Book– II.
- [24]. Journal of American society of Civil Engineers, vol.26, 2008, pp.105-110.
- [25]. MOEF Guide Lines , Ministry of Environment , New delhi
- [26]. John Glasson, Riki Therivel, Andrew Chadwick 1995 , Introduction to Environmental impact assessment, pp 47-60
- [27]. All India soil and land use planning, 1991; Methodology for priority delineation survey, Ministry of Agriculture, Govt.of India, New Delhi, India.

Speaker Features And Recognition Techniques: A Review

1. Dr. Mahesh S. Chavan

Professor and Head of Electronics Engg. Dept.
KIT's College of Engineering, Kolhapur,
Maharashtra, India

2. Mrs. Sharada V. Chougule.

Assistant Professor,
Electronics and Telecommunication Engg. Dept.
Finolex Academy of Management and Technology, Ratnagiri,
Maharashtra, India

Abstract

This paper gives an overview of various methods and techniques used for feature extraction and modeling in speaker recognition. The research in speaker recognition have been evolved starting from short time features reflecting spectral properties of speech (low-level or physical traits) to the high level features (behavioural traits) such as prosody, phonetic information, conversational patterns etc. Low level acoustic information such as cepstral features has been dominated as these features gives very low error rates (especially in quiet conditions). But they are more prone to error in noisy conditions. In this paper various features along with modeling techniques used in speaker recognition are discussed.

Introduction

Speech is the product of a complex behaviour conveying different speaker-specific traits that are potential sources of complementary information. Human speech production can be modeled by so-called *source-filter model* featured. As the name suggests, the model considers the voice production mechanism as a combination of two components: the *voice source* and the *acoustic filter*. The "source" refers to the airstream generated by the larynx and the "filter" refers to the vocal tract. Both of the components are inherently time-varying and assumed to be independent of each other [1], [2]. In this model we separate the source $G(w)$ from the articulation and radiation $H(w)$ (Furui S. 2001). The speech signal is then represented by the cascade connection of $G(w)$ and $H(w)$ giving: $S(w) = G(w)H(w)$. For voiced speech, the source is modeled as an impulse train with period T . For unvoiced speech, the source is modeled as Gaussian white noise. In each case, the source is amplified by the gain factor G in proportion to the volume of speech. The signal is then passed through a time-varying digital filter to represent the articulation and radiation applied by the vocal track.

Historically, all speaker recognition systems have been mainly based on acoustic cues that is nothing but physical traits extracted from spectral characteristics of speech signals. So far the features derived from the speech spectrum have proven to be the most effective in automatic systems, because the spectrum reflects the geometry of system that generate the signal. Therefore the variability in the dimensions of the vocal track is reflected in the variability of the spectra between the speakers. However, studies [3] have proved that there is a large amount of information suitable for speaker recognition being the top part related to learned traits and the bottom part to physical traits.

Speaker Recognition

One objective in automatic speaker recognition is to decide which voice model from a known set of voice models best characterizes a speaker; this task is referred to as speaker identification. In the different task of speaker verification, the goal is to decide whether a speaker corresponds to a particular known voice or to some other unknown voice.

There are two modes of operation for speaker identification: in the closed-set mode, the system assumes that the unknown voice must come from the set of known voices; in open-set mode, the speakers that do not belong to the set of known voices are referred to as impostors. An important application of speaker identification technology is forensics, identifying the suspects among a set of known criminals. Automatic speaker recognition systems can be further classified according to the speech modality: text-dependent or text-independent. In text-dependent recognition, the user must speak a phrase known to the system, which can be fixed or prompted. The knowledge of a spoken phrase can provide better recognition results. In text-independent recognition, the system does not know the phrase spoken by the user. Despite the unconstrained phrase selection, this makes the system to be more complex. However, text-independent speaker recognition systems have more applications than text-dependent ones in real life.

There are generally two phases [1] in building or using a speaker recognition system. The first phase is called enrolment or training phase, in which a user enrolls by providing voice samples to the system. The system extracts speaker-specific information from the voice samples to build a voice model of the enrolled speaker. The second phase is called the classification or recognition phase, in which a test voice sample is used by the system to measure the similarity of the user's voice to the previously enrolled speaker models to make a decision. In a speaker identification task, the system measures the similarity of the test sample to all stored voice models. In speaker verification task, the similarity is measured only to the model of the claimed identity. The decision also differs across systems. For example, a closed-set identification task outputs the identity of the recognized user; besides the identity, an open-set identification task can also choose to reject the user in case the test sample do not belong to any of the stored voice models; a verification task chooses to accept or reject the identity claim.

The effectiveness of speaker recognition system is in measures differently for different tasks. Since the output of a closed-set speaker identification system is a speaker identity from a set of known speakers, the identification accuracy is used to measure the performance. For open-set systems there are two types of errors: false acceptance of an impostor and false rejection of a known speaker. The performance measure can also incorporate the cost associated with each error.

Like most pattern recognition problems, a speaker recognition system can be partitioned into four modules: feature analysis and extraction, speaker modelling, pattern matching and decision logic.

Feature Analysis and Extraction

From human speech production mechanism, it is possible to identify individual using the speech data. Speech contains speaker specific information due to vocal track and excitation source. Larynx is the major excitation source, whereas vocal track is the major resonant structure. Speaker information is due to particular shape, size and dynamics of vocal track and also the excitation source. These features related to physiological nature of human speech production are called physical traits, which are used in state-of-art systems. However human speaker recognition relies on other sources of information like speaking style, pronunciation etc. Such features are referred to as behavioral traits. Further, the behavioral traits like how the vocal tract and excitation source are controlled during speech production are also unique for each speaker. The information about the behavioral trait is also embedded into the speech signal and can be used for speaker recognition. Thus the information present in speech signal carries the identity of speaker at different levels. To properly represent speech data, it is necessary to analyse it using suitable analysis techniques. The analysis techniques aims at selecting proper frame size and shift for analysis and also at extracting the relevant features in the feature extraction stage [4].

The information about the audio category is contained in the excitation source (sub-segmental), system/physiological (segmental) and behavioral (suprasegmental) characteristics of the speech data.

Sub-segmental analysis is one in speech signal is analysed using the frame size and shift of very small duration (3-5ms). This technique is used mainly to analyse and extract the characteristics of the excitation source. Since the excitation source information is relatively fast varying compared to the vocal tract information, small frame size and shift are required to best capture the speaker-specific information, which is the reason for the choice of 3-5 ms for frame size and shift [5].

In *segmental analysis* (used for extraction of short-term spectral features), features are computed from short frames of about 20-30 milliseconds in duration. They are usually descriptors of the short-term spectral envelope which is an acoustic correlate of timbre, i.e. the "colour" of sound, as well as the resonance properties of the supra-laryngeal vocal tract.

High-level features are generally related to a speaker's learned habits and style, such as particular word usage or idiolect. For humans, the information about the audio category is perceived by listening to a longer segment of audio signal. This other level of information contained in the audio signal is the *suprasegmental information* that is the variation of the signal over long duration. In this case, speech is analysed using the frame size and shift in the range of 50-200 ms. Studies made in [5],[6],[7] shows the significance of suprasegmental features in speaker recognition systems. These features are useful as their structure is not affected by the frequency characteristics of the transmission systems. Each of the four basic acoustic features of speech signal, i.e. pitch, intensity, duration and speech quality, is a carrier of a variety of types of linguistic, paralinguistic and non-linguistic information [8][9].

Feature Extraction Techniques

Feature extraction is said to be the heart of speaker recognition system.. Feature extraction is one where the input speech is processed to obtain those features of the input speech which are useful in speaker identification. The function of feature extraction is to convert speech waveform to some type of parametric representation called *feature vectors* for further analysis and processing by the classifiers. This stage extract the speaker-specific information in the form of feature vectors at reduced data rate. The feature vectors represent the speaker-specific information due to one or more of the following: Vocal tract,

excitation source (Physical traits) and behavioral traits. For speaker recognition, features that exhibit high speaker discrimination power, high inter-speaker variability, and low intra-speaker variability are desired. Although there are no exclusive features conveying speaker identity in the speech signal, from the source-filter theory of speech production it is known that the speech spectrum shape encodes information about the speaker's vocal tract shape via resonances (formants) and glottal source via pitch harmonics.

Feature extraction is necessary for several reasons. First, speech is a highly complex signal which carries several features mixed together. In speaker recognition we are interested in the features that correlate with the physiological and behavioral characteristics of the speaker. Other information sources are considered as undesirable noise whose effect must be minimized. The second reason is a mathematical one, and relates to the phenomenon known as *curse of dimensionality* [10], which implies that the number of needed training vectors increases exponentially with the dimensionality. Furthermore, low-dimensional representations lead to computational and storage savings.

The ideal feature should have [10]:

- large between-speaker and small within-speaker variability
- be difficult to impersonate/mimic
- not be affected by the speaker's health or long-term variations in voice
- occur frequently and naturally in speech
- be robust against noises and distortions

It is unlikely that a single feature would fulfill all the listed requirements. Fortunately, due to the complexity of speech signals, a large number of complementary features can be extracted and combined to improve accuracy. The selection of features depends largely on the application (co-operative/non co-operative speakers, desired security/convenience balance, database size, amount of environmental noise).

Types of Features

A vast number of features have been proposed for speaker recognition. We divide them into the following classes:

- Spectral features
- Dynamic features
- Source features
- Suprasegmental features
- High-level features

Spectral features are descriptors of the short-term speech spectrum, and they reflect more or less the physical characteristics of the vocal tract. *Dynamic features* relate to time evolution of spectral (and other) features. *Source features* refer to the features of the glottal voice source. *Suprasegmental* features span over several segments. Finally, *high-level features* refer to symbolic type of information, such as characteristic word usage.

The anatomical structure of vocal apparatus is easy to extract in automatic fashion. For same sound the location and magnitude of peaks observed in spectra is different. Early text-dependent speaker recognition systems utilized information from short-time spectrum to provide unique features for speaker [8]. These features consisted of energy measurements from the outputs of a bank of a filter. LPC is one of the feature extraction method based on the source-filter model of speech production. The basic problem in LPC analysis is to determine prediction coefficients from the speech frame. B.S. Atal in 1976 [9] uses linear prediction model for parametric representation of speech derived features. The predictor coefficients and other speech parameters derived from them, such as the impulse response function, the auto-correlation function, the area function, and the cepstrum function were used as input to an automatic speaker recognition system, and found the cepstrum to provide the best results for speaker recognition. Joseph P. Campbell in [1], uses all-pole LP (linear prediction) to model a signal by a linear combination of its past values and a scaled present input. Reynolds in 1994 [12] compared different - features useful for speaker recognition, such as Mel frequency cepstral coefficients (MFCCs), linear frequency cepstral

coefficients (LFCCs), LPCC (linear predictive cepstral coefficients) and perceptual linear prediction cepstral coefficients (PLPCCs). From the experiments conducted, he had concluded that , of these features, MFCCs and LPCCs give better performance than the other features. Both MFCC and LPCC coefficients are used to extract vocal track information, but uses different technique to extract the features . MFCC extraction is similar to the cepstrum calculation except that one special step is inserted, namely the frequency axis is warped according to the Mel-scale using mel filter bank. The filter bank outputs are then converted to cepstral coefficients by applying the inverse discrete cosine transform (IDCT). In case of LPCCs, first, LPCs are obtained for each frame using Durbin's recursive method, and then these coefficients are converted to cepstral coefficients. The predictor coefficients themselves are rarely used as features, but they can be transformed into robust and less correlated features such as LPCC, line spectral frequencies (LSFs) and *perceptual linear prediction cepstrum coefficients (PLPCC)* [12,13] or *eigen-MLLR coefficients* [14]. Experimental evaluation of recognition accuracy of the MFCC, LPCC and PLPCC was made in [13] and result of this report is that all features perform poorly without some form of channel compensation, however, with channel compensation MFCC slightly outperform other types. Cepstrum representation of the speech signal has shown to be useful in practice. The features discussed above are called short-term (spectral) or low-level features. These features are used in most state-of speaker recognition systems as these easy to compute and yield good performance (Reynolds et al.,2003). However, it is not without drawbacks. The main disadvantage of the cepstrum is that it is quite sensitive to the environment and noise [15].

Of all the various spectral features, MFCC, LPCC,LSF and PLP are the most recommended features which carry information about the resonance properties of vocal track [17]. Most of the current implementations use some kind of spectral envelope features to parameterize the voice (MFCC, LPCC...), achieving a great performance. But recent researches are trying to include long term information into the system, in order to reduce error rates.

Unlike short-term spectral information, long-term information is being used which convey supra-segmental information, such as prosodic and speaking style. Andre G. Adami and Douglas A. Reynolds in 2003 [16], presented two new approaches that demonstrated effective ways to model and apply prosodic contours for text independent speaker verification tasks. In first approach the relation between dynamics of fundamental frequency (f_0) and energy trajectories were used to characterize the speaker's identity. In this, global distribution of energy and f_0 features were such as $\log f_0$, \log energy and their first order derivatives were created. In second accent and intonation information from a known set of frequently and naturally occurring words found in conversational speech. They had used n-grams to model the sequence.

From source-filter model, it was shown that speech signal can be decomposed into two parts: the source part and the system part. The system part consists of the smooth envelope of the power spectrum and is represented in the form of cepstrum coefficients, which can be computed by using either the linear prediction analysis or the mel-filter-bank analysis. Most of the automatic speaker recognition systems reported in the literature utilise the system information in the form of cepstral coefficients. These systems perform reasonably well. The source contains information about pitch and voicing. This information is also important for humans to identify a person from his/her voice. Hassan Euaidi and Jean Rouaf in 2004 [18] had proposed an approach which jointly exploits the information of the vocal tract and the glottis source. The approach synchronously takes into account the correlation between the two sources of information. The fundamental frequency and the MFCC coefficients were used to represent the information of the source and the vocal tract, respectively. Experiments that integrate the a-posteriori probability of observing a MFCC vector given the knowledge of the pitch frequency have been reported. It was also shown that systems based on voiced segments yield good scores. However, when the dependence of the source and vocal tract is taken into account, the best results were observed for durations T lower than 500 ms. Speech prosody refers to the intonation, energy and rate of the speech. It is well known that these features are characteristics of each person, so that they carry information about the speaker. Furthermore, prosody is uncorrelated with the spectral envelope shape. Therefore, supposedly adding prosodic features to the already used spectral features may lead to an improvement in the system's performance. Najim Dehak, Pierre Dumouchel in their work [19] , introduced the use of continuous prosodic features for speaker recognition and showed how they can be modeled using joint factor analysis. Similar features have been successfully used in language identification. These prosodic features were pitch and energy contours spanning a syllable-like unit. They were extracted using a basis consisting of Legendre polynomials. Tharmarajah Thiruvaran [20] used frequency modulation (FM) features for improving accuracy of speaker identification . Due to the similarity between amplitude modulation (AM) feature and the conventional Mel frequency cepstrum coefficients (MFCC), FM features were used. It was shown that, the correlation between FM feature components was observed to be very small compared with that of Mel filter bank log energies, thus reducing the need for decorrelation. FM feature components were shown to be very nearly Gaussian distributed, Digital Energy Separation Algorithm (DESA) was used as a front-end in speaker identification system.

E. Shriberg, L. Ferrer , S. Kajarekar in [21] described a new approach to modeling idiosyncratic prosodic behaviour for automatic speaker recognition. The approach computes various duration, pitch, and energy features for each estimated syllable in speech recognition output, quantizes the features, forms N-grams of the quantized values, and models normalized

counts for each feature N-gram using support vector machines (SVMs) referred to as SNERF-grams (N-grams of Syllable based Nonuniform Extraction Region Features). S.R. Mahadeva Prasanna, Jinu Mariam Zachariah and B. Yegnanarayana (2004) used the features from spectral, duration and pitch. The substantiation from the different sources were combined using a multilayer perceptron neural network. It was shown that not only that the performance of verification improved, but also the non-spectral features such as duration and pitch were found to be robust for variations due to channel [22]. In order to improve the speaker recognition accuracy, in [23], pitch was applied to GMM-based speaker recognition (SR). The circular average magnitude difference function (CAMDF) method was used to extract the pitch. An endpoint detection method based on the pitch was proposed. In this work, mel-frequency cepstral coefficient (MFCC) based on the pitch, the pitch contour, the pitch first-order difference and the pitch changed rate features were selected as the features of the SR. Experimental results showed improvement in recognition rate using proposed endpoint detection method, than that of the speaker recognition system using the MFCC parameters only.

The last decade has seen increased interest in exploring such higher-level features in automatic speaker recognition. High level features are based on voice timbre and accent/ pronunciation of speaker and also on lexicon - the kind of words the speakers tend to use in their conversations. The work on such "high-level" conversational features was initiated in [24] where a speaker's characteristic vocabulary, the so-called *idiolect*, was used to characterize speakers. The idea in "high-level" modeling is to convert each utterance into a sequence of *tokens* where the co-occurrence patterns of tokens characterize speaker differences. Elizabeth Shriberg in her article [25] demonstrated how higher-level features can contribute to performance in a state-of-the-art system. Various features such as cepstral and cepstral-derived, phonetic (acoustic tokenization), prosodic, lexical features along with their performance was discussed. It was shown that, systems based on frame-level cepstral or cepstral derived features show higher accuracy than longer-range systems. Within the set of cepstral-based systems, the MLLR system had best performance, because it takes advantage of linguistic information from ASR. Of the longer-range systems, the conditioned syllable-based prosody sequence system was shown to be the most successful. Recently a project titled *SuperSid* is undertaken to explore the effectiveness of high level information for speaker recognition [47]. Specifically, methods to extract and model speaker specific patterns in acoustic cues (the sound of the person's voice), speech prosody, word and phone pronunciations (idiosyncratic or dialectical distinctions), word usage (characteristic phrases or word selection), and interactions with conversational partners (taciturn or dominating in conversations) were examined. The fusion of features and classifiers to improve the recognition performance is proposed in this project.

In above section, we have discussed various features and features extraction techniques which is the front end of any speaker recognition system. The features discussed are from low-level spectral features such as MFCC, LPCC, PLP etc., representing the vocal track dynamics as well as features such as LP residues, pitch, pitch contours representing vocal fold or excitation source features. These features are related to physical traits of speaker. We have also discussed the prosodic and high level features which represent behavioural characteristics of speaker. The choice of feature is based on applications, accuracy demand and robustness in various operating conditions and environments, channel parameters, complexity of computation etc.

Feature Modeling

In the previous section, we have discussed so called *measurement* step in the speaker identification where a set of speaker discriminative characteristics is extracted from the speech signal. In this section, we go through the next step called *classification*, which is a decision making process of determining the author of a given speech signal based on the previously stored or learned information [1]. This step is usually divided into two parts, namely *matching* and *modeling*. The modeling is a process of enrolling speaker to the identification system by constructing a model of his/her voice, based on the features extracted from his/her speech sample. The matching is a process of computing a *matching score*, which is a measure of the similarity of the features extracted from the unknown speech sample and speaker model [26]. Once the feature vectors corresponding to the "speech" frames have been extracted, the associated speech data also known as training/enrolment data is used to build a speaker specific model. During the verification phase, the trained model is used to authenticate a sequence of feature vectors extracted from utterances of unknown speakers.

The statistical approaches for constructing the relevant models can be divided into two distinct categories: generative and discriminative. Training of generative models typically involves data specific to the target speakers where the objective is that the models can faithfully capture the statistical properties of the speaker specific speech signal. Training of discriminative models involves data corresponding to the target and imposter speakers and the objective is to faithfully estimate the parameters of the manifold which distinguishes the features for the target speakers from the features for the imposter speakers. An example of a popular generative model used in speaker verification is *Gaussian Mixture Models* (GMMs) and an example of a popular discriminative model is *Support Vector Machines* (SVMs).

Classical speaker models can be also categorized into nonparametric and parametric models. They are also called template models and stochastic models, respectively. Vector quantization (VQ) [30] and dynamic time warping (DTW) [31] are representative examples of template models for text-independent and text-dependent recognition, respectively. In stochastic

models, each speaker is modeled as a probabilistic source with an unknown but fixed probability density function. The training phase is to estimate the parameters of the probability density function from the training data. The likelihood of the test utterance with respect to the model is used for pattern matching. The Gaussian mixture model (GMM) [32,33] and the hidden Markov model (HMM) [34, 35] are the most popular stochastic models for text-independent and text-dependent speaker recognition, respectively. Speaker models can also be classified into generative and discriminative models. The generative models such as GMM and VQ estimate the feature distribution within each speaker independently. While the discriminative models such as artificial neural networks (ANNs) [36, 37] and support vector machines (SVMs) [38, 39] model the boundary between speakers.

Vector quantization (VQ) is a lossy data compression method based on the principle of block coding. It is a fixed-to-fixed length algorithm. In the earlier days, the design of a vector quantizer (VQ) is considered to be a challenging problem due to the need for multi-dimensional integration. In 1980, Linde, Buzo, and Gray (LBG) proposed a VQ design algorithm based on a training sequence. The use of a training sequence bypasses the need for multi-dimensional integration. A VQ that is designed using this algorithm are referred to in the literature as an LBG-VQ [40]. In 1985, Soong *et al.* used the LBG algorithm for generating speaker-based vector quantization (VQ) codebooks for speaker recognition [41]. VQ is often used for computational speed-up techniques and lightweight practical implementations. It also provides competitive accuracy when combined with background model adaption (Kinnuen et al. 2009). Zhong-Xuan [42] Yuan presented a new approach to vector quantization in which feature vector is represented by a binary vector called binary quantization (BQ). The performance criterion of vector quantization, distortion measure, was employed for investigating the effectiveness of BQ. The results shown very good performance in terms of memory space and computation required. Also the identification system had shown strong robustness in additive White Gaussian noise.

In 1995, Reynolds proposed Gaussian mixture modeling (GMM) classifier for speaker recognition task. GMM [33], is a stochastic model which has become the *de-facto* reference method in speaker recognition. The GMM needs sufficient data to model the speaker, and hence good performance. It can be considered as an extension of the VQ model, in which the clusters are overlapping. That is, a feature vector is not assigned to the nearest cluster as in VQ, but it has a nonzero probability of originating from each cluster. GMM is composed of multivariate Gaussian components [17]. A GMM super-vector characterizes a speaker's voice by the GMM parameters such as the mean vectors, covariance matrices and mixture weights. The parameters of the model are typically estimated by maximum likelihood estimation, using the *Expected-Maximization* algorithm. The matching function in GMM is defined in terms of *likelihood* [32,33]. It was shown that GMM outperformed the other modeling techniques. Therefore, state-of-the-art speaker recognition systems use GMM as classifier due to the better performance, probabilistic framework and training methods scalable to large data sets [43]. As GMM needs sufficient data to model the speaker, Reynolds in [44], introduced GMM-UBM (universal background model), in which UBM is trained from speech data collected from large number of speakers, which acts as a speaker independent model. In the GMM approach, speaker models are obtained from the adaption of a universal background model (UBM) through the maximum *a posteriori* (MAP) criterion. The UBM is usually trained by means of expectation-maximization (EM) algorithm from a background dataset, which includes a wide range of speakers, languages, communication channels, recording devices and environments. The GMM-UBM becomes a standard technique for text-independent speaker recognition due to its reliable performance. The discriminant classifier based on support vector machine (SVM) were of great interest in speech field. In speaker recognition, an important revolution was proposed, mainly by [45]. SVMs are typically trained in binary mode to discriminate between the speaker's data and impostor data. The impostor data consists of several speakers and can coincide with the data used to train the SI GMM. The resulting SVM is a hyperplane separating the two classes in the predefined kernel space. During testing, the same kernel is used to compute a signed distance between the test sample and the hyperplane. This distance is used as a similarity measure or score, with positive values indicating that the sample is on the target speaker side of the hyperplane (but note that the decision threshold may be set to a nonzero value to bias the outcome in accordance with a given decision cost model). One disadvantage of the GMM-based approach is that it models the features as a bag of frames ignoring sequence information. Researchers have explored other modeling techniques, such as hidden Markov models (HMMs), to model sequence information (Newman et al. 1996). HMM-based approaches have been shown to outperform the GMM-based approach given enough training data. Another approach has been to model blocks of features, preserving the temporal information (Gillick et al. 1995). It uses a mixed approach, associating the robustness of the statistical modeling provided by the GMM-UBM paradigm with the discriminating power of SVMs. This approach, denoted GMM super-vector SVM with linear kernel (GSL), uses the GMM-UBM to model the training or testing data. A super-vector is extracted from the corresponding GMM (obtained from UBM by MAP procedure), composed by concatenation of the mean coefficients of all the GMM components. The super-vectors are then used as inputs of the SVM classifier [46].

Pattern matching and decision logic

The next step after computing of matching scores for every speaker model enrolled in the system is the process of assigning the exact classification mark for the input speech. Matching gives a score which represents how well the test feature vectors are close to the reference models. This process depends on the selected matching and modeling algorithms. The feature extraction and pattern matching are same for different speaker recognition tasks, but the decision depends on the nature of task. In closed-set identification task, the decision is simply the speaker index that yields the maximum score. In template matching, decision is based on the computed distances, whereas in stochastic matching it is based on the computed probabilities. In template matching, the speaker model with smallest matching score is selected, whereas in stochastic matching, the model with highest probability is selected. Practically, decision process is not so simple and for example for so called open-set identification problem the answer might be that input speech signal does not belong to any of the enrolled speaker models. It is quite difficult to characterize the performance of speaker verification systems in all applications due to the complexities and differences in the enrolment/testing scenarios. Having computed a match score between the input speech-feature vector and a model of the claimed speaker's voice, a verification decision is made whether to accept or reject the speaker or request another utterance (or, without a claimed identity, an identification decision is made). If a verification system accepts an impostor, it makes a false acceptance (FA) error. If the system rejects a valid user, it makes a false rejection (FR) error. The FA and FR errors can be traded off by adjusting the decision threshold, as shown by a Receiver Operating Characteristic (ROC) curve. The operating point where the FA and FR are equal corresponds to the equal error rate. The accept or reject decision process can be an accept, continue, time-out, or reject hypothesis-testing problem. In this case, the decision making, or classification, procedure is a sequential hypothesis-testing problem. On the other hand, the computation of speaker identification is measured as a ratio of the number of correctly identified examples to the total number of examples considered for the testing.

Conclusion

In this paper, we have presented an overview of the various features, the extraction methods and modeling techniques of speaker recognition. The low level features such as cepstral features work well in ideal conditions, but their performance is degraded in real time situations. Use of high level information can add complementary knowledge to improve the performance of recognition system. In practical situations many negative factors are encountered including mismatched handsets for training and testing, limited training data, unbalanced text, background noise and non-cooperative users. The techniques of robust feature extraction, feature normalization, model-domain compensation and score normalization methods are necessary. There are number of research problems that can be taken up, such as human-related error sources, real-time implementation, and forensic interpretation of speaker recognition scores. For this it is important to explore stable features that remain insensitive to variation of speakers voice over time and are robust against variation in voice quality due to physical states or disguises. The problem of distortion in the channels and background noise also requires being resolved with better techniques.

References

1. J.P. Campbell, Jr., "Speaker Recognition: A Tutorial," Proceedings of IEEE, vol. 85, no. 9, Sept. 1997 .
2. Tomas F. Quatieri, "Discrete Time Speech Signal Processing, Principles and Practice", Pearson Education 2006.
3. Marcos Faundez-Zanuy and Enric Monte-Moreno, "State-of-the-art in Speaker Recognition ", IEEE A&E Systems Magazine, May 2005
4. Jayanna, H.S. and S.R.M. Prasanna, 2009. Analysis, feature extraction, modeling and testing techniques for speaker recognition. IETE Technical Review, 2009, Volume 26, issue 3
5. B. Yegnanarayana, S.R.M. Prasanna, J.M. Zachariah, and C.S. Gupta,, "Combining evidence from source, suprasegmental and spectral features for a fixed-text speaker verification system," IEEE Trans. Speech Audio Process. , vol. 13(4), pp. 575-82, July 2005.
6. Grazyna Demenko, "Analysis of suprasegmental features for speaker verification", 8th Australian International Conference on Speech Science and Technology, 2000
7. B. Yegnanarayana, K. Sharat Reddy, and S.P. Kishore, "Source and system features for speaker recognition using AANN models," in proc. Int. Conf. Acoust., Speech, Signal Process. , Utah, USA, Apr. 2001.
8. P.D. Bricker, "Statistical techniques for talker identification", Bell Svst. Techn. Jour. Vol.50 April 1971 B. S. Atal "Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification "J. Acoust. Soc. Am. Volume 55, Issue 6, pp. 1304-1312 (1974)
9. Tommy Kinnunen, "Spectral Features for Automatic Text-Independent Speaker Recognition " thesis *University of Joensuu* ,Dec.2003
10. Douglas A. Reynolds and Richard Rose , 'Robust Text Independent Speaker Identification using Gaussian Mixture Speaker Models ', IEEE transaction on Speech and Audio Processing, Vol.3, No.1, January 1995

11. D.A. Reynolds, "Experimental evaluation of features for robust speaker identification," IEEE Trans. Speech Audio Process. , vol. 2(4), pp. 639-43, Oct. 1994
12. X. Huang, A. Acero and H.-W. Hon, Spoken language processing, Upper Saddle River, New Jersey, Prentice Hall PTR, 2001.
13. N. J.-C. Wang, W.-H. Tsai, L.-S. Lee, "Eigen-MLLR Coefficients as New Feature Parameters for Speaker Identification", Eurospeech 2001 – Scandinavia
14. J. R. Deller, J. H. L. Hansen, J. G. Proakis, "Discrete-Time Processing of Speech Signals", Piscataway (N.J.), IEEE Press, 2000.
15. Andre G. Adami and Douglas A. Reynolds, "Modeling prosodic dynamics for Speaker Recognition ", IEEE ,ICASSP 2003
16. Tomi Kinnunen and Haizhou Li, "An overview of text independent speaker recognition: From features to supervectors ", ScienceDirect, Speech Communication 2010
17. Hassan Euaidi and Jean Rouaf, "Pitch and MFCC dependent GMM models for speaker identification systems " , CCECE IEEE, 2004
18. Najim Dehak, Pierre Dumouchel, and Patrick Kenny, "Modeling Prosodic Features with Joint Factor Analysis for Speaker Verification" , IEEE Trans. Speech and Language Proces 2004.
19. Tharmarajah Thiruvaran, Eliathamby Ambikairajah, "Speaker Identification using FM Features " , ATP Research Laboratory, National ICT Australia
20. E. Shriberg, L. Ferrer , S. Kajarekar, A. Venkataraman, "Modeling prosodic feature equences for speaker recognition " , Science Direct, Speech Communication 2005
21. S.R. Mahadeva Prasanna, Jinu Mariam Zachariah and B. Yegnanarayana. " Neural Network Models for Combining Evidence from Spectral and Suprasegmental Features for Text-Dependent Speaker Verification" IEEE 2004
22. ZHU Jian-wei, SUN Shui-fa," Pitch in Speaker Recognition", IEEE 2009
23. Doddington, G. "Speaker recognition based on idiolectal differences between speakers", Proc. 7th European Conference on Speech Communication and Technology (Eurospeech 2001)
24. Elizabeth Shriberg, " Higher-Level Features in Speaker Recognition " SRI International, International Computer Science Institute, Berke
25. B. S. Atal, "Automatic Recognition of Speakers from their Voices", Proceedings of the IEEE, vol 64, 1976
26. H. Gish and M. Schmidt, "Text Independent Speaker Identification", IEEE Signal Processing Magazine, Vol. 11, No. 4, 1994
27. J. M. Naik, "Speaker Verification: A Tutorial", IEEE Communications Magazine, January 1990
28. D. A. Reynolds, "An Overview of Automatic Speaker Recognition Technology", ICASSP 2002
29. L. R. Rabiner and B. H. Juang, "Fundamentals of Speech Recognition," Prentice-Hall, NJ, 1993
30. S. Furui, "Cepstral analysis technique for automatic speaker verification," IEEE Transactions on Acoustics, Speech and Signal Processing ,April 1981
31. D. Reynolds, T. Quatieri, and R. Dunn, "Speaker verification using adapted Gaussian mixture models," Digital Signal Processing , January 2000
32. D. Reynolds, and R. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," IEEE Trans. on Speech and Audio Processing 3, January 1995
33. J. Naik, L. Netsch, and G. Doddington, "Speaker verification over long distance telephone lines," In Proc. ICASSP, Glasgow, May 1989
34. M. BenZeghiba, and H. Bourland, "User-customized password speaker verification using multiple reference and background models," Speech Communication, September 2006
35. L. Heck, Y. Konig, M. Sonmez, and M. Weintraub, "Robustness to telephone handset distortion in speaker recognition by discriminative feature design," Speech Communication , June 2000.
36. B. Yegnanarayana, and S. Kishore, "AANN: an alternative to GMM for pattern recognition," Neural Networks , April 2002
37. W. Campbell, J. Campbell, D. Reynolds, E. Singer, and P. Torres-Carrasquillo, "Support vector machines for speaker and language recognition," Computer Speech and Language, April 2006
39. V. Wan and W. Campbell, "Support vector machines for speaker verification and identification," Proceedings of the 2000 IEEE Signal Processing Society Workshop, vol.2, 2000
40. Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design," IEEE Transactions on Communications, January 1980.

- 41 F.K. Soong, A.E. Rosenberg, L.R. Rabiner, and B.H. Juang, "A Vector quantization approach to speaker recognition," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. , vol. 10, Apr. 1985
- 42 Zhong-Xuan Yuan and Chong-Yu, "Binary quantization of feature vectors for robust text independent speaker identification", IEEE Trans. Speech Audio Process. , vol. 7(1), January 1999
- 43 W.M. Campbell, J.P. Campbell, D.A. Reynolds, E. Singer, and P.A. Torres-Carrasquillo, "Support vector machines for speaker and language recognition," Computer Speech and Language , vol. 20, 2006
- 44 D.A. Reynolds, T.F. Quateri, and R.B. Dunn, "Speaker verification using adapted Gaussian mixture models," Digital Signal Processing , vol. 10, 2000.
- 45 W. Campbell, D. Sturim, D. Reynolds and A. Solomonoff, "SVM based speaker verification using GMM supervector kernel and NAP variability compensation", Proc. Int. Conf. Acoustics, Speech and Signal Processing, 2006
- 46 Joseph P. Campbell, W.M. Campbell, Wade Shen, " Forensic speaker recognition-A need for caution ", IEEE Signal Processing Magazine, March 2009
- 47 Douglas Reynolds, Walter Andrews, Joseph Campbell, Jiri Navratil, Barbara Peskin, Andre Adami, Qin Jin, Daviekc Klusacek, Joy Abramson, Radu Mihaescu, John Godfrey, Douglas Jones, Bing Xiang, "Exploiting High-Level Information for High-Performance Speaker Recognition" , 2002

Computing Over a Multi Cloud for MTC Applications

Challa Vanitha Reddy

C.V.S.R College Of Engineering, JNTU

Battula Sudheer Kumar

C.V.S.R College Of Engineering, JNTU

Abstract

IT organizations can now outsource computer hardware by leasing CPU time through cloud computing services. The problem here is the effectiveness is becoming less due to burden on single cloud while working with MTC applications. This paper deals with defining the feasible solutions with MTC applications using the programming models for computing over a Multi clouds instead of single cloud for effectiveness.

Keywords: Cloud Computing, MTC applications, HTC

1. Introduction

This decade is marked by a paradigm shift of the industrial information technology towards a subscription based or pay-per-use service business model known as cloud computing. This paradigm provides users with a long list of advantages, such as provision computing capabilities; broad, heterogeneous network access; resource pooling and rapid elasticity with measured services [1].

The single cloud will not be more effective for MTC paradigm (Many-Task Computing (MTC) paradigm [2] embraces different types of high-performance applications involving many different tasks, and requiring large number of computational resources over short periods of time. These tasks can be of very different nature, with sizes from small to large, loosely coupled or tightly coupled, or compute-intensive or data-intensive). As the single cloud will not be efficient for MTC paradigm we can also go with Computing Clusters.

Computing clusters have been one of the most popular platforms for solving MTC problems, specially in the case of loosely coupled tasks (e.g. high-throughput computing applications). However, building and managing physical clusters exhibits several drawbacks: 1) major investments in hardware, specialized installations (cooling, power, etc.), and qualified personal; 2) long periods of cluster under-utilization; 3) cluster overloading and insufficient computational resources during peak demand periods. To overcome the above drawbacks in cluster computing we go for computing clusters over a multi cloud.

The simultaneous use of different cloud providers to deploy a computing cluster spanning different clouds can provide several benefits:

- High-availability and fault tolerance, the cluster worker nodes can be spread on different cloud sites, so in case of cloud downtime or failure, the cluster operation will not be disrupted. Furthermore, in this situation, we can dynamically deploy new cluster nodes in a different cloud to avoid the degradation of the cluster performance.
 - Infrastructure cost reduction, since different cloud providers can follow different pricing strategies, and even variable pricing models (based on the level of demand of a particular resource type, daytime versus night-time, weekdays versus weekends, spot prices, and so forth), the different cluster nodes can change dynamically their locations, from one cloud provider to another one, in order to reduce the overall infrastructure cost [3].
- The rest of this paper explains about the Background, Related work, MTC application through programming model and Conclusion

2. Background

Many-task computing aims to bridge the gap between two computing paradigms, high throughput computing and high performance computing. Many task computing differs from high throughput computing in the emphasis of using large number of computing resources over short periods of time to accomplish many computational tasks (i.e. including both dependent and independent tasks), where primary metrics are measured in seconds (e.g. FLOPS, tasks/sec, MB/s I/O rates), as opposed to operations (e.g. jobs) per month. Many task computing denotes high-performance computations comprising multiple distinct activities, coupled via file system operations [4]. Tasks may be small or large, uniprocessor or multiprocessor, compute intensive or data-intensive. The set of tasks may be static or dynamic, homogeneous or heterogeneous, loosely coupled or tightly coupled. The aggregate number of tasks, quantity of computing, and volumes of data may be extremely large. Many task computing includes loosely coupled applications that are generally communication-intensive but not naturally expressed using standard message passing interface commonly found in high performance computing, drawing attention to the many computations that are heterogeneous but not "happily" parallel.

3. RELATED WORK

We have found many applications that are a better fit for MTC than HTC or HPC. Their characteristics include having a large number of small parallel jobs, a common pattern observed in many scientific applications [5]. They also use files (instead of messages, as in MPI) for intra-processor communication, which tends to make these applications data intensive. While we can push hundreds or even thousands of such small jobs via GRAM to a traditional local resource manager (e.g. Condor [6]), the achieved utilization of a modest to large resource set will be poor due to high queuing and dispatching overheads, which ultimately results in low job throughput. A common technique to amortize the costs of the local resource management is to “cluster” multiple jobs into a single larger job. Although this lowers the per job overhead, it is best suited when the set of jobs to be executed are homogenous in execution times, or accurate execution time information is available prior to job execution; with heterogeneous job execution times, it is hard to maintain good load balancing of the underlying resource, causing low resource utilization. We claim that “clustering” jobs is not enough, and that the middleware that manages jobs must be streamlined and made as light-weight as possible to allow applications with heterogenous execution times to execute without “clustering” with high efficiency.

In addition to streamlined task dispatching, scalable data management techniques are also required in order to support MTC applications [7]. MTC applications are often data and/or meta-data intensive, as each job requires at least one input file and one output file, and can sometimes involve many files per job. These data management techniques need to make good utilization of the full network bandwidth of large scale systems, which is a function of the number of nodes and networking technology employed, as opposed to the relatively small number of storage servers that are behind a parallel file system or GridFTP server. We have identified various applications (as detailed below) from many disciplines that demonstrate characteristics of MTC applications. These applications cover a wide range of domains, from astronomy, physics, astrophysics, pharmaceuticals, bioinformatics, biometrics, neuroscience, medical imaging, chemistry, climate modeling, economics, and data analytics. They often involve many tasks, ranging from tens of thousands to billions of tasks, and have a large variance of task execution times ranging from hundreds of milliseconds to hours. Furthermore, each task is involved in multiple reads and writes to and from files, which can range in size from kilobytes to gigabytes. These characteristics made traditional resource management techniques found in HTC inefficient; also, although some of these applications could be coded as HPC applications, due to the wide variance of the arrival rate of tasks from many users, an HPC implementation would also yield poor utilization. Furthermore, the data intensive nature of these applications can quickly saturate parallel file systems at even modest computing scales.

4. . Mtc Application Through Programming Model

MTC applications through programming models and tools give the feasible and efficient solution for MTC application.

4.1 MTC through Programming model (MapReduce)

MapReduce [8] is one of the most popular programming models designed to support the development of compute and data-intensive applications which have high storage and processing demands which were initially met by large scale computer systems. MapReduce was initially created by Google for simplifying the development of large scale web search applications in data centers and has been proposed to form the basis of a ‘Data center computer’. MapReduce model is suitable for web search services, scientific research projects referred to as e-Science and data mining application amongst others.

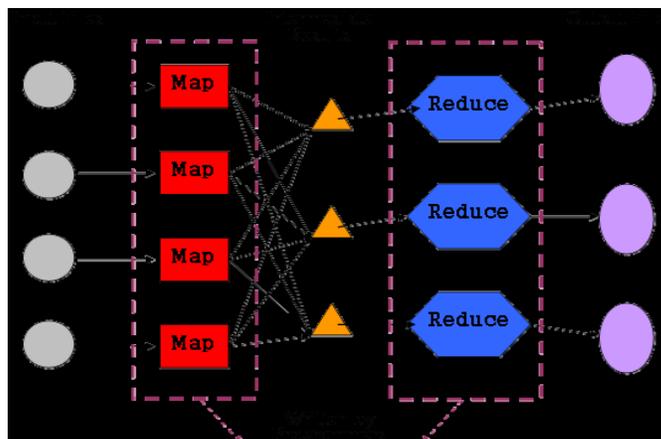


Fig 1: MTC Applications In Map Reduce Model

MTC application is executed in a parallel manner through two phases in Map reduce model . In the first phase ,MTC application are submitted to map functions and those operations are known as map operations. all map operations can be executed independently with each other. In the second phase, each reduce operation may depend on the outputs generated by any number of map operations. However, similar to map operations, all reduce operations can be executed independently.

From the perspective of dataflow, MapReduce execution for MTC applications consists of m independent map tasks and r independent reduce tasks, each of which may be dependent on m map tasks. Generally the intermediate results are partitioned into r pieces for r reduce tasks[9].

The MapReduce runtime system schedules map and reduce tasks to distributed resources. It manages many technical problems: parallelization, concurrency control, network communication, and fault tolerance. Furthermore, it performs several optimizations to de-crease overhead involved in scheduling, network communication and intermediate group-ing of results.

Thus the MTC application runs very efficiently through map reduce model.

5. Conclusion

In this paper we have shown how MTC applications in multi cloud will be more efficient and feasible than in single cloud and also the applications with support of Programming model give the feasible and efficient solution. It takes even very less time to compute them.

6. References

Journals:

- [1] P. Mell, 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] I. Raicu, I. Foster, Y. Zhao, Many-Task Computing for Grids and Supercomputers, Workshop on Many-Task Computing on Grids and Supercomputers, 2008, pp. 1–11
- [3] E. Walker, The Real Cost of a CPU Hour. Computer 42(4): 35–41, 2009.
- [4] J. Ousterhout, “Scripting: Higher Level Programming for the 21st Century”, IEEE Computer, March 1998
- [5] Y. Zhao, M. Hategan, B. Clifford, I. Foster, G. von Laszewski, I. Raicu, T. Stef-Praun, M. Wilde. “Swift: Fast, Reliable, Loosely Coupled Parallel Computation”, IEEE Workshop on Scientific Workflows 2007
- [6] J. Frey, T. Tannenbaum, I. Foster, M. Frey, S. Tuecke. “Condor-G: A Computation Management Agent for Multi-Institutional Grids”, Cluster Computing, 2002
- [7] E.V. Hensbergen, Ron Minnich. “System Support for Many Task Computing”, IEEE Workshop on Many-Task Computing on Grids and Supercomputers (MTAGS08) 2008
- [8] J. Dean, S. Ghemawat. “MapReduce: Simplified data processing on large clusters.” In OSDI, 2004
- [9] M. Cole, Parallel Programming with List Homomorphisms, Parallel Processing Letters 5 (1995) 191–203.



Vanitha Reddy Challa

received the Bachelor's degree from Aurora's Scientific and Technological Institute, affiliated to JNTU Hyderabad University in 2009. I'm currently pursuing my Masters from Anurag Group of Institutions affiliated to JNTU Hyderabad University in Computer Science & Engineering department



Sudheer Kumar Battula

received the Bachelor's degree from Aurora's Scientific and Technological Institute affiliated to JNTU Hyderabad University in 2009. I'm currently pursuing my Masters from the Anurag Group of Institutions affiliated to JNTU Hyderabad University in Computer Science & Engineering department.

AN OVERVIEW OF SMART ANTENNAS AND ITS TECHINQUES BEAMFORMING AND DIVERSITY

SURAYA MUBEEN
KL UNIVERSITY
VIJAYAWADA

DR.A.M.PRASAD
JNTU KAKINADA

DR.A.JHANSI RANI
VRSEC VIJAYAWADA

ABSTRACT:

Smart antennas are systems attract lot attentions now and believably more in the future, as it can increase the capacity of mobile communication systems dramatically. Design of smart antenna systems combines the technologies of antenna design, signal processing, and hardware implementation. : A smart antenna is therefore a phased or adaptive array that adjusts to the environment. That is, for the adaptive array, the beam pattern changes as the desired user and the interference move; and for the phased array the beam is steered or different beams are selected as the desired user moves. The early smart antenna systems were designed for use in military applications to suppress interfering or jamming signals from the enemy. The proposed research work gives us an overall view of basic smart antennas and its techniques.

KEYWORDS: Antenna, Miso, Diversity, Beamforming

1. Introduction: A single antenna element is mostly omni-directional (a). This means it receives and sends in and from all directions around it. If a base station uses an omni-directional antenna and a user communicates with this station, every signal that is send back and forth between the two devices is at the same time a source of interference for any other communication taking place within the same cell.

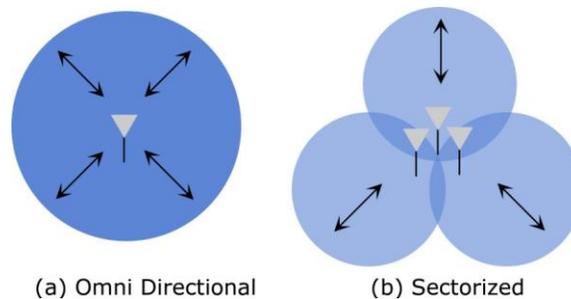


Figure 1: Non-smart antenna systems

If traffic increases, because devices require higher bandwidth or larger numbers of devices communicate within the same cell, the sources of interference increase equally. This shows the necessity for smart antenna techniques. A simple way of utilizing the availability of multiple antennas is to split the cell into different sectors by using sectorized antennas (b). These kinds of antennas are directional antennas. When positioned correctly on the base station they can be used to subdivide a cell into smaller regions, while the base station still receives signals from all cells. With the use of these antennas (b) the amount of interference is limited to smaller regions instead of the whole cell. Smart antenna techniques however offer a much larger increase in the signal to noise ratio. Already with less complex scenarios than MIMO, such as SIMO or MISO, an important concept of smart antennas can be realized and has been realized for a long time. This concept is called beam forming.

Beam forming is a signal processing technique that is used to steer a signal from a set of non-directional antennas into a certain direction. By combining the signals of these antennas, a new, simulated antenna is created. This simulated antenna can then be electrically pointed in a certain direction, without physically moving the “real” antennas [11]. Beam forming can be used in all applications that rely on receiving signals from a certain direction ,while ignoring other signals from the surroundings. Certainly beam forming has its origins in radio technologies and is therefore not a technology that is solely being used in networks. But the principle behind it is extremely able to be used in WMNs, a lot of problems with wireless networks result from environmental or external conditions that interfere with or completely destroy communication paths. Being able to use multiple antennas to “steer” optimized simulated antennas in space to improve sending and reception quality is a powerful method to overcome and adapt to many of these interferences. The range of applications in Wireless Mesh Networks, where nodes have to perform many different tasks simultaneously and adapt to constantly changing surroundings, is easily imaginable.

2. Antenna Radiation Patterns: Every antenna that transmits a signal generates an asymmetrical amount of electromagnetic waves. This amount is always stronger in some directions than in others. This balance and ratio of “field strength” vs.

“direction” is called “radiation pattern”. Each part of an antenna radiates these electromagnetic waves with a different amplitude and phase, therefore each of these waves travels a different way. Depending on the construction of the antenna, in some directions there are enough waves traveling a similar distance at similar amplitude, so they can add up and create a “gain”, i.e. a stronger signal or better reception capability. In other directions these waves can absorb each other and create a “loss”, i.e. a weak signal or no reception capability [11]. This principle is used to create directional antennas. Simply by increasing the size of an antenna the gains and losses can be controlled more effectively, resulting in a narrower, more directed beam.

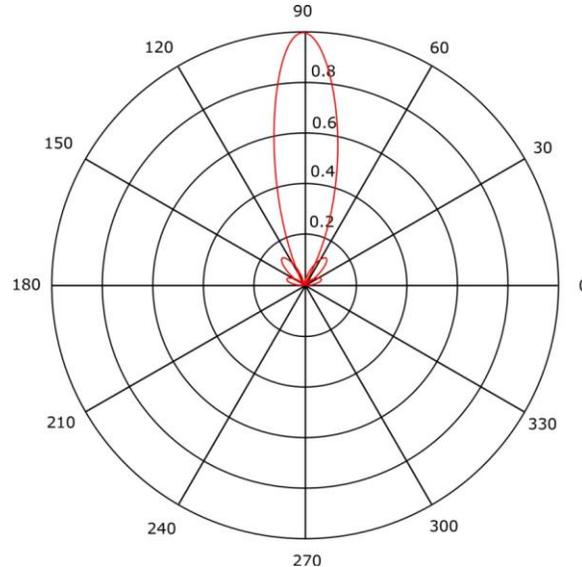


Figure 2: Three Wavelength Circular Aperture - Field Strength vs. Direction .

In an efficient design of a simple directional antenna, linear arrays of small omnidirectional radiating elements are used. These array elements each receive exactly the same signal and the same phase. This is done by feeding them with signals from the same transmitter, over cables that are identical in manufacturing and length. In order to further narrow the main beam and to reduce the size of the side lobes the number of radiating elements has to be increased and the distance between two neighboring elements has to be decreased to $1/2$ of the wavelength. If this distance becomes larger, the side lobes increase, and at the same time the beam becomes narrower, because the overall length of the antenna increases as well. If the gaps are then filled with additional radiating elements in order to reduce the distance to $1/2$ of the wavelength again, the beam dimensions persist, while the side lobes decrease again. The following Figure 3 shows an antenna array of 7 radiating elements, each placed with a distance of $1/2$ of the wavelength to its neighboring elements and each connected to the transmitter with identical cables.

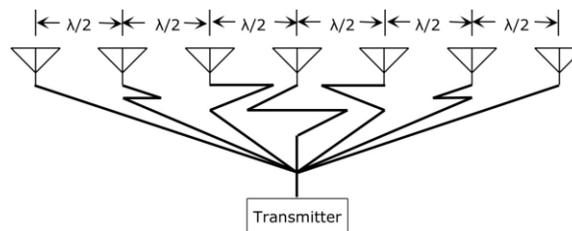


Figure 3: 7-Element 3λ Linear Array

3. Steered Beams and Beam forming:

If the phases of the individual elements are shifted, the beam points into a different direction, without the antenna having moved or rotated in any way. A simple way to do this would be to alter the length of the cables, connecting the elements to the transmitter. This would result in a phase shift, thus resulting in a beam that is steered to a different direction. Certainly this can be done electronically as well. By programming the transmitter to send out the signal over each of its cables with different delays, an interference in the construction of an antenna becomes unnecessary, and furthermore the steering of the beam can be completely controlled and changed extremely quickly.

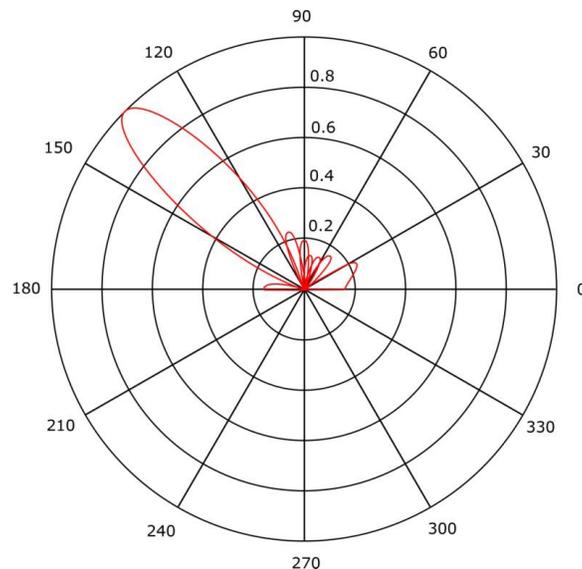


Figure 4: 3.5 Wavelength 8 Element Linear Array, Progressive 0.7π Phase Shift

In this way, simple beam forming is done. Additionally to the phase of each antenna element, also the amplitudes are controlled. By this not only the direction of the beam, but also the sizes of the side lobes and the direction of the nulls can be controlled. With a sufficient amount of array elements, a large number of beams and nulls can be created and steered simultaneously, thus making large antennas extremely efficient. But just by being able to create and steer beams, a network with mobile, moving devices is not yet effectively designed. Beam forming techniques are necessary to program the stations and devices which dispose of multiple antennas. In the scenario of a MISO system, the base station, which is equipped with multiple antennas, could use a beam forming technique to always point the beams towards the direction of arriving signals and point the nulls towards the directions of interfering signals.

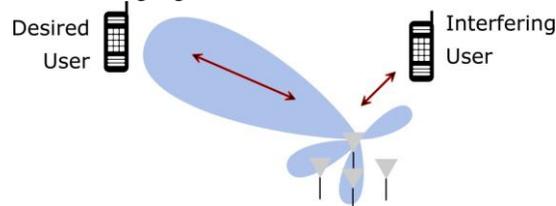


Figure 5: Smart antenna techniques – Beamforming

Again, this is easily done in a static network. In a network with mobile devices this becomes more complex. Two of the traditional techniques that are used in MISO scenarios are the “adaptive array” and the “switched-beam” technique [12]. “The adaptive array scheme tracks each user in a given cell with an individual adaptive beam pattern, while the switched-beam scheme selects one beam pattern for each user out of a number of preset fixed beam patterns, depending on the location of the user.” [9]

To achieve an adaptive beam forming a sequence of steps is necessary.

1. Each element of the adaptive array has to independently weight the signal it receives.
2. The factor for each of these signals has to be continuously controlled.
3. All signals are then combined and output.
4. A decision about the adaptation of the beam is made, based on this output and send back to the weight control [13].

In comparison to the switched-beam scheme where preset patterns are chosen based on the nature of the received signal, this scheme is considerably more complex. It generally gives better results, but is more expensive to implement and relies on the accuracy of the estimation techniques and the speed of the beam forming algorithms [9]. If applied to the scenario of a Wireless Mesh Network the advantages of adaptive array beam forming over switched beams become more obvious. In a switched beam system a mobile station moves throughout different sectors. The beams do not adapt, they just “switch” to and from one another once a mobile station leaves the coverage of one beam and enters another. In an adaptive array system, not only the beams are directly pointed at the mobile stations, but also the nulls are directly pointed towards the interfering signals. In a Wireless Mesh Network this reduces many problems that may occur due to unexpected interferences. Still a major problem that arises due to overly large amounts of clients in a sector is not satisfyingly solved with adaptive array beam

forming only. A third important technique that focuses exactly on this problem of large numbers of clients in multiuser environments is called “opportunistic beam forming”, also referred to as “random beam forming”. For the duration of each transmission a random beam is formed. During a communication training sequences are transmitted back and forth. The mobile station estimates a signal-to-interference-plus-noise-ratio, on which the base station bases a schedule for the specific user traffic. This way time slots are created to make sure the mobile stations can transfer data when the most effective beams for their position are generated. While offering great advantages in networks with a large amount of mobile stations, in networks with few mobile stations, this technique is not as effective as the previously mentioned more conventional techniques.

4.DIVERSITY: Another technique that is used to improve transmission quality and combat the fading of individual channels is diversity. Diversity is used to “exploit the multiple fading paths between the transmitter and the receiver.” [9] This is commonly used in simple SISO scenarios, but with the installation of multiple antennas this diversity can be further exploited on a spatial level.

The basic concept of diversity is to transmit the same signal via several independent diversity branches to receive independent signals that can be compared by the receiver. Since there is a high probability that not all the signals fade simultaneously and equally the receiver has several differently faded signals at its disposal and can then reconstruct the original signal much better as if only one signal had been transmitted. Possible diversity dimensions can be time-, frequency, or space-based. Time and frequency diversity have the disadvantage of bringing about a penalty in data rate. Spatial diversity on the other hand can be very effectively used with multiple antennas.

In general spatial diversity can be subdivided in the fields of transmit diversity and receive diversity, which indicates that this technique can take place on sender and receiver side. At the receiver side (e.g. in a MISO scenario) the signals could be received over different antenna elements and then be combined following different schemes. Three possible schemes are “selection combining”, “maximal ratio combining”, and “equal gain combining”, all of which are long established techniques [16]. In selection combining only one of the received signals is chosen. Selected is the channel with the highest signal-to-noise-ratio. Maximal ratio combining weighs all the signal based on their signal-to-noise-ratio and combines them accordingly. Equal gain combining simply sums up all received signals. If implemented on the transmitter side (e.g. in a SIMO scenario), spatial diversity can also be utilized by sending the same data over different antennas.

5.CONCLUSION:

This paper offers an overview of several aspects of Multiple Antenna Techniques and possible application scenarios in Wireless Mesh Networks. It explains the basics of these types of networks and specializes on Smart antenna techniques with a focus on digital beam forming. The functionality of simple antennas, as well as antenna arrays and their radiation patterns are explained and put into context with different explanations on how to utilize their characteristics to establish different types of connections and transmission channels. Based on the actuality of the cited works it appears that Multiple Antenna Techniques are not in the main focus of today’s Wireless Mesh Networking research. The high costs of modern MIMO systems could be a reason for that. Still smart antenna techniques are very capable of making Wireless Mesh Networks more efficient in terms of increased capacity and throughput, better routing performance, and increased energy efficiency, and are a foundation that much of the research in higher layer protocols.

VI.REFERENCES:

- [1]. Roos,David.<http://communication.howstuffworks.com/how-wireless-mesh-networks-work.htm>.
- [2]. Akyildiz, Ian F. and Wang, Xudong. A Survey on Wireless Mesh Networks. *IEEE Communications Magazine*. September 2005, Vol. 43, pp. 23-30.
- [3]. Akyildiz, Ian F. and Wang, Xudong. *Wireless Mesh Networks*. s.l. : John Wiley and Sons, 2009. pp. 1, 45-46.
- [4]. Krag, Tomas and Buettrich, Sebastian. O’Reilly Wireless DevCenter. [Online] January 22, 2004. <http://www.oreillynet.com/lpt/a/4535>.
- [5]. Gungor, V. C., et al. Challenges and Issues in Designing Architectures and Protocols for Wireless Mesh Networks. [book auth.] Ekram Hossain and Kin K. Leung. *Wireless Mesh Networks*. s.l. : Springer, 2008, pp. 1-12.
- [6]. Gkelias, A. and Leung, Kin K. Multiple Antenna Techniques for Wireless Mesh Networks. [book auth.] Ekram Hossain and Kin K. Leung. *Wireless Mesh Networks*. s.l. : Springer, 2008, pp. 277-291.
- [7]. Miniwatts Marketing Group. World Internet Usage Statistics. [Online] March 2009..<http://www.internetworldstats.com/stats.htm>.
- [8]. Hu, Honglin, Luo, Jijun and Zhang, Xiaodong. Multiple Antenna Techniques for Wireless Mesh Networks. [book auth.] Yan Zhang, Jijun Luo and Honglin Hu. *Wireless Mesh Networking. Architectures, Protocols and Standards*. s.l. : Auerbach Publications, 2007, pp. 361-384. 20
- [9]. Haynes, Toby. Spectrum Signal Processing - A Primer on Digital Beamforming. [Online] March 1998. http://spectrumsignal.com/publications/beamform_primer.pdf.

- [10]. Ho, Ming-Ju, Stüber, Gordon L. and Austin, Mark D. Performance of Switched Beam Smart Antennas for Cellular Radio Systems. *IEEE Transactions on Vehicular Technology*. 1996, Vol. 47, p. 545.
- [11]. Ohira, Takashi. Adaptive array antenna beamforming architectures as viewed by a microwave circuit designer. *Microwave Conference, 2000 Asia-Pacific*. 2000, pp. 828-833.
- [12]. Hu, Honglin, Luo, Jijun and Chen, Hsiao-Hwa. Radio Resource Management for Cooperative Wireless Communication Systems with Organized Beam-Hopping Techniques. *IEEE Wireless Communications*. 2008, Vol. 15, pp. 100-109.
- [13]. Hu, Honglin, Zhu, Jinkang and Yang, Lie-Liang. Performance evaluation of distributed-antenna communications systems using beam-hopping. *Wireless Communications and Mobile Computing*. 2005, Vol. 5, pp. 45-56.
- [14]. Brennan, D. G. Linear diversity combining techniques. *Proc. IRE*. 1959, Vol. 47, pp. 1075-1102.
- [15]. Zheng, L. and Tse, D. N. C. Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels. *IEEE Transactions on Information Theory*. May 2003, Vol. 49, pp. 1073-1096.

ACKNOWLEDGEMENT

This work was supported by our research guides DR.A.M.Prasad and DR.A.Jhansi Rani and is contributed of my own work I thank them for my development in doctoral work and financially supported by KL university. Author profile:



1.SURAYA MUBEEN working in K.L.UNIVERSITY as Assistant Professor and currently working in field of antennas as (PH.D)research scholar in JNTU Kakinada. Published 6 journals and 5 national conferences and 3 international conferences. My field of interests are Antennas, microwaves and radar systems

CO-AUTHORS



2.DR.A.M.PRASAD Professor in E.C.E is working in jntukakinada for the past 10 years.He has won the best teacher award in the year 2008 and now he is the controller of Examination . Dr. A.Mallikarjuna prasad did his B.Tech in ECE from Nagarjuna University, during 1984-88. He did his M.Tech in Electronics&Instrumentation from Andhra University in 1992 and completed his Ph.D in 2009 from JNTU in the field of Antennas . He has joined JNT University service as Associate Professor of ECE in June 2003. He got promoted as Professor in ECE during Nov 2011.

3.DR.A.JHANSI RANI Professor in E.C.E at V.R.siddharatacollegeofengineering and guided many students for research and more than 14 years of experience in teaching and attended many seminars and workshops.

“A Comparative Analysis of Fuzzy C-Means Clustering and K Means Clustering Algorithms”

Mrs. Bharati R.Jipkate and Dr.Mrs.V.V.Gohokar

SSGMCE, Shegaon, Maharashtra-443101 (India)

Abstract— Segmentation of an image entails the division or separation of the image into regions of similar attribute. The most basic attribute for segmentation of an image is its luminance amplitude for a monochrome image and color components for a color image. Clustering is one of the methods used for segmentation. The objective of this paper is to compare the performance of various segmentation techniques for color images. K-means clustering and Fuzzy C-Means clustering techniques are compared for their performance in segmentation of color images.

Keywords— K-Means clustering, Fuzzy C- Means Clustering.

I. Introduction

Segmentation is the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful easier to analyze. Clustering is one of the methods used for segmentation. Clustering can be considered the most important *unsupervised learning* problem. it deals with finding a *structure* in a collection of unlabeled data. A loose definition of clustering could be “the process of organizing objects into groups whose members are similar in some way”. A *cluster* is therefore a collection of objects which are “similar” between them and are “dissimilar” to the objects belonging to other clusters.

Clustering algorithms may be classified as listed below:

- Exclusive Clustering
- Overlapping Clustering

In the first case data are grouped in an exclusive way, so that if a certain datum belongs to a definite cluster then it could not be included in another cluster. On the contrary the second type, the overlapping clustering, uses fuzzy sets to cluster data, so that each point may belong to two or more clusters with different degrees of membership. K-means is an *exclusive clustering* algorithm, Fuzzy C-means is an *overlapping clustering* algorithm.

The specific criterion to be used for segmentation depends on the application. Pixels may belong together on the

basis of the same color and/or the same texture and/or distance.

For segmentation Input image is transformed into feature space. The clustering task separates the data into number of partitions, which are volumes in the n-dimensional feature space. These partitions define a hard limit between the different groups and depend on the functions used to model the data distribution Clustering is the search for distinct groups in the feature space [7].

It is expected that these groups have different structures and that can be clearly differentiated.'

II K-means clustering

It is an iterative technique that is used to partition an image into *K* clusters.K-means (MacQueen, 1967) is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume *k* clusters) fixed a priori. The main idea is to define *k* centroids,[7] one for each cluster. These centroids should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate *k* new centroids as barycenters of the clusters resulting from the previous step. After we have these *k* new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop we may notice that the *k* centroids change their location step by step until no more changes are done. In other words centroids do not move anymore. K-Means clustering generates a specific number of disjoint, flat (non-hierarchical) clusters. It is well suited to generating global clusters. The K-Means method is numerical, unsupervised, non-deterministic and iterative.

This algorithm aims at minimizing an *objective function*, e.g. a squared error function. The objective function

$$\sum_{j=1}^k \sum_{i=1}^n |x_i^{(j)} - c_j|^2$$

where $|x_i^{(j)} - c_j|$ is a chosen distance measure between a $x_i^{(j)}$ data point and the cluster centre c_j , is an indicator of the distance of the n data points from their respective cluster centers [1-2].

III Fuzzy c-means clustering

In hard clustering, data is divided into distinct clusters, where each data element belongs to exactly one cluster. In **fuzzy clustering** (also referred to as **soft clustering**), data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels, and then using them to assign data elements to one or more clusters.

One of the most widely used fuzzy clustering algorithms is the Fuzzy C-Means (FCM) Algorithm (Bezdek 1981). The FCM algorithm attempts to partition a finite collection of n elements $X = \{x_1, \dots, x_n\}$ into a collection of c fuzzy clusters with respect to some given criterion. Given a finite set of data, the algorithm returns a list of c cluster centres $C = \{c_1, \dots, c_c\}$ and a partition matrix

$$J=U = u_{i,j} \in [0,1], i = 1, \dots, n, j = 1, \dots, c$$

where each element u_{ij} tells the degree to which element x_i belongs to cluster c_j . Like the k-means algorithm, the FCM aims to minimize an objective function. The standard function is:

$$u_k(x) = \frac{1}{\sum_j \left(\frac{d(\text{center}_k, x)}{d(\text{center}_j, x)} \right)^{2/(m-1)}}$$

which differs from the k-means objective function by the addition of the membership values u_{ij} and the fuzzifier m . The fuzzifier m determines the level of cluster fuzziness. A large m results in smaller memberships u_{ij} and hence, fuzzier clusters. In the limit $m = 1$, the memberships u_{ij} converge to 0 or 1, which implies a crisp partitioning. In the absence of experimentation or domain knowledge, m is commonly set to 2.

In **fuzzy clustering**, each point has a degree of belonging to clusters, as in **fuzzy logic**, rather than belonging completely to just one cluster. Thus, points on the edge of a cluster, may be *in the cluster* to a lesser degree than points in the center of cluster.

Any point x has a set of coefficients giving the degree of being in the k th cluster $w_k(x)$. With fuzzy c -means, the

centroid of a cluster is the mean of all points, weighted by their degree of belonging to the cluster:

$$C_k = \frac{\sum_x w_k(x)x}{\sum_x w_k(x)}$$

The degree of belonging, $w_k(x)$, is related inversely to the distance from x to the cluster center as calculated on the previous pass. It also depends on a parameter m that controls how much weight is given to the closest center [8].

IV The algorithm of k-means clustering

Step1. Choose the number K of clusters either manually, randomly or based on some heuristic.

Step2. Generate K clusters and determine the cluster's center.

Step3. Assign each pixel in the image to the cluster that minimizes the variance between the pixel and the cluster center

Step4. Re-compute cluster centers by averaging all of the pixels in the cluster.

Step5. Repeat steps 3 and 4 until some convergence criterion is met

Figure 1 shows segmentation in 2 and 3 and 5 clusters using this method

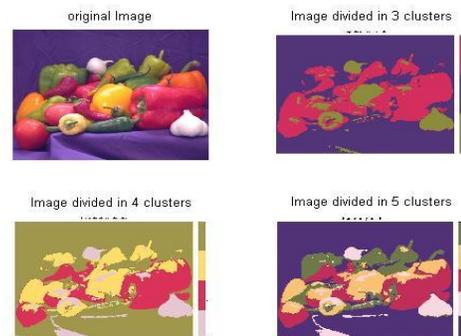


Fig:1 Kmeans Segmentations

V The algorithm of fuzzy c-means clustering

Step1. Choose a number of clusters in a given image.

Step2. Assign randomly to each point coefficients for being in a cluster.

Step3. Repeat until convergence criterion is met.

Step4. Compute the center of each cluster.

Step5. For each point, compute its coefficients of being in the cluster [4-5].

Fig2 shows the image segmented by c means algorithm



Fig :2. C means

3.Result and Conclusion

The paper compares k-means and fuzzy c-means clustering image segmentation algorithms .The algorithms are developed in MATLAB for analysis and comparison. K-means clustering produces fairly higher accuracy and requires less computation. C means clusyering produces close results to K-means clustering, yet it requires more computation time than K-means because of the fuzzy measures calculations involved in the algorithm.

References

- [1] ^a ^b MacQueen, J. B. (1967). "Some Methods for classification and Analysis of Multivariate Observations". **1**. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. pp. 281–297. MR0214227. Zbl 0214.46201. <http://projecteuclid.org/euclid.bsm/1200512992>. Retrieved 2009-04-07..
- [2] Dr.Teknomo, Kardi. K-Means Clustering Tutorials. <http://people.revoledu.com/kardi/tutorial/kMean/>July 2007.
- [3] S. Adaekalavan,C.Chandrasekar'A Comparative Analysis of Fuzzy C-Means Clustering and Harmonic K Means Clustering Algorithms'European Journal of Scientific Research ISSN 1450-216X Vol.65 No.1 (2011), pp. 45-49© EuroJournals Publishing, Inc. 201 <http://www.europeanjournalofscientificresearch.com>
- [4] J. C. Dunn (1973): "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", *Journal of Cybernetics* 3: 32-57

- [5] J. C. Bezdek (1981): "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York
- [6] Tariq Rashid: "Clustering" http://www.cs.bris.ac.uk/home/tr1690/documentation/fuzzy_clustering_initial_report/node11.html
- [7] A Tutorial on Clustering Algorithmshome.dei.polimi.it/matteucc/Clustering/tutorial_html/
- [8] [Fuzzy clustering - Wikipedia, the free encyclopedia](http://en.wikipedia.org/wiki/Fuzzy_clustering) en.wikipedia.org/wiki/Fuzzy_clustering -

Performance Analysis of Timing Attack on Elliptic Curve Cryptosystem

Mr. Praful V. Barekar

Master of Technology Scholar
Department of CSE
G. H. Rasoni College Of Engineering,
Nagpur

Prof. K. N. Hande

Assistant Professor
Department of CSE
G. H. Rasoni College Of Engineering,
Nagpur

1. Abstract

Cryptosystems often take slightly different amounts of running time depending on the input and the used key. This timing information, extracted from the decryption process, can be used to derive information about the secret key. This new class of attacks on implementations of cryptosystems is named Timing Attacks. Timing attacks attempt to exploit the variations in computational time for private key operations to guess the private key. This type of attack is primitive in the sense that no specialized equipment is needed. An attacker can break a key by simply measuring the computational time required by the user inputs and recording those user inputs. This paper is aimed to analyse the performance of Timing Attack on Elliptic Curve Cryptosystem. The main advantage of Elliptic Curve Cryptography is smaller key size, it is mostly used for public key infrastructure

Keywords: Cryptosystem, Timing Attack, Running Time, Elliptic Curve Cryptography, Public key Infrastructure.

2. Introduction

Timing Attacks were first introduced in a paper by Kocher [4]. Kocher describes the general idea of Timing Attack and shortly reports on some results for the RSAREF implementation of the RSA cryptosystem. He also claims that the same idea can be used for discrete logarithm based cryptosystems like Diffie-Hellman, DSS, and other systems. Later Kocher's timing attack on RSA was modified and practically examined, e.g. by Dhem et.al. on a RSA smart card implementation [6], and by Brumley and Boneh on the RSA implementation of the OpenSSL library [1]. We do however not know about any successful practical results for timing attacks when applied to elliptic curve cryptosystems.

Elliptic curve (EC) cryptosystems have gained large support in recent years after several standard documents on public key cryptography included EC cryptosystems and EC signature schemes [7]. Their main advantage in comparison to RSA is significantly smaller key sizes for similar security levels. The Wireless Transport Layer Security (WTLS) specification for securing wireless applications therefore also explicitly supports elliptic curves cryptography for wireless applications [8]. In this paper, we report on experimental results for timing attacks when applied to elliptic curve cryptosystems. We use a software implementation of the basic scalar point multiplication algorithm for points defined over finite prime fields $GF(p)$ as described in [7]. We start with a short explanation of the necessary facts about elliptic curves, timing attack mechanism, Finally we describe practical results from our simulation of timing attacks.

3. Basics of Elliptic Curve

This section explains some necessary background information on elliptic curves (for a more detailed description, see [7]). Let $GF(p)$ be the finite prime field with p elements, where p is prime. We define the group of points $E(GF(p))$ on an **elliptic curve** $(a, b) \in GF(p)^2$ as the set of solutions $(x, y) \in GF(p)^2$ to the equation

$$y^2 \equiv x^3 + ax + b \pmod{p} \quad (1)$$

Together with a point at infinity $O = (\infty, \infty)$. These points form an abelian (additive) group where the group operation is defined by the following formulas:

- O is the zero element.
- The negative point of (x, y) is the point $(x, -y)$.
- For two non-zero points $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$ with $P_1 \neq P_2$, we determine the sum $P_3 = (x_3, y_3) = P_1 + P_2$ as

$$x_3 \equiv -x_1 - x_2 - l^2 \pmod{p} \quad (2)$$

$$y_3 \equiv -y_1 + l(x_1 - x_3) \pmod{p} \quad (3)$$

Where $l \equiv (3x_1^2 + a) / (2y_1) \pmod{p}$ if $P_1 = P_2$, and $l \equiv (y_2 - y_1) / (x_2 - x_1) \pmod{p}$ otherwise. For an integer $k > 0$ and a point P , we define scalar multiplication $k \cdot P$ for the point P as adding P ($k - 1$) times to itself. There exist various algorithms for computing scalar multiplication (see [3]). In this paper, we consider only the following basic binary left-to-right fast multiplication algorithm:

Algorithm 1: (Scalar Point Multiplication)

Input: Point P, integer $k > 0$ with binary representation $k = (1 k_{w-2} \dots k_0)_2$.

Output: Point $k \cdot P$

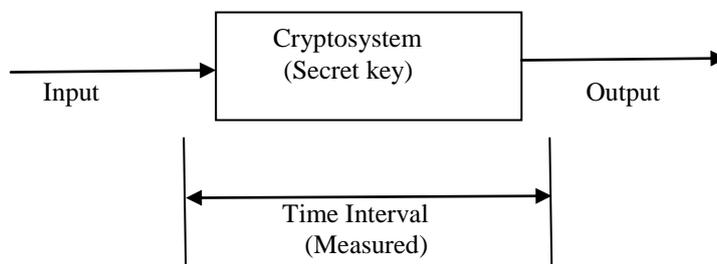
1. Let $Q = P$
2. for $j = w - 2$ downto 0 do
 - $Q = 2 \cdot Q$
 - if $k_j = 1$ then $Q = Q + P$
- done
3. return Q

Definition 1: The Elliptic Curve Discrete Logarithm Problem (ECDLP) is defined as the problem computing the integer k only with knowledge of the elliptic curve E and the two points P and $Q = K \cdot P \in E(\text{GF}(p))$ without prior knowledge of k . ECDLP is generally considered to be a very difficult problem if the field characteristic p is sufficiently large (at least 160 bit), and the order of the group of points on the elliptic curve is prime (or "almost prime"). Up to now, there exists no algorithm that can solve ECDLP for such curves in reasonable time. Therefore point multiplication was chosen as the "trapdoor one-way function" used in elliptic curve public key cryptosystems (ECC) and elliptic curve signature schemes (ECDSA). For more details on applications of elliptic curves for public key cryptography, we refer to [7]. Timing attack use the following scenario: Assume that the unknown secret key k is stored securely and cannot be accessed directly. We can compute $k \cdot P_i$ for arbitrarily many random points P_i . Moreover we are able to determine the running time T_i of each of these computations (in our simulation, we uses Algorithm 1 for fast multiplication). We might also be able to determine running times for a single point addition or point doubling for arbitrarily chosen points.

4. Timing Attack

Recently, a new class of cryptanalysis aimed at a cryptosystem’s implementation-specific weaknesses has attracted great interest. This kind of cryptanalysis exploits the leak of information such as timing, power consumption, and electromagnetic radiation from system operations to facilitate attacks on the cryptosystem. Since the information used by the attack is not in the “main channel”, the input or output, we call these types of attacks “side-channel” attacks. In this paper, we will focus on timing attacks.

Let’s think the cryptosystem as a black box with input and output which constitute the “main channel” of the system. We can measure the time it takes for the system to give an output after given an input. The time required for different inputs may vary, forming a timing distribution. If this timing distribution is related to the secret (key bits) in the system, we may have a way to reveal the secret key.



5. The Attacker’s Task

The attacker has the ability to observe a sequence of elliptic curve operations, thus, the attacker’s aim is to calculate and exploit the probabilities of certain sequences of bits given an observed sequence of elliptic curve operations. Using the information of such conditional probabilities, the key-space that has to be searched to find the correct ephemeral key, can be significantly reduced. This is because certain combinations of patterns in the power trace and certain combination of digits are less likely than the others (or even not possible at all). The attacker’s task can be stated in a more formal way. Let X be a random variable that denotes a sequence of elliptic-curve operations and $|X|$ the length of X (i.e. the number of elliptic-curve operations in this sequence). For example, $X = \text{“DDD”}$ (i.e. the realization of the random variable X consists of three consecutive elliptic-curve point-double operations) thus $|X| = 3$, or $X = \text{“DAD”}$ (i.e. the realization of the random variable X consists of an elliptic-curve point-double operation, an elliptic-curve point-addition operation and an elliptic-curve point-double operation) thus $|X| = 3$. Let Y be a random variable that denotes a sequence of digits in the digit representation of k and $|Y|$ the length of Y (i.e. the number of digits). For example $Y = \text{“000”}$ (i.e. the realization of the random variable Y consists of three consecutive zeros) thus $|Y| = 3$, or $Y = \text{“01”}$ (i.e. the realization of the random variable Y consists of a zero and a one digit) thus $|Y| = 2$. Then the attacker’s goal is to calculate and exploit the conditional probability.

6. Mathematical Model

Let us denote a set of inputs (plaintexts) to the system by $S_M = \{M_1, M_2, M_3, \dots, M_n\}$ All the possible keys compose the key set denoted by $S_K = \{K_1, K_2, K_3, \dots, K_d\}$ where d is the number of possible keys. If the cryptosystem

implementation we want to attack is vulnerable to timing attacks, the timing distribution of the input will be dependent on the key used in the system. Thus for key K_i , we will have a timing distribution denoted by $P_i(t) = F(S_M, K_i)$ which is different from that of other keys.

For the system we want to attack, we measure the timing information for a set of input values from the set S_M , and form a timing distribution $P(t)$. The attack to the system will be reduced to a usual detection problem which tries to detect K_i knowing $P_i(t)$ and $P(t)$. We can apply, at least in theory, regular detection solutions to solve the problem. For example, the detection problem has a general form of the solution: if $T(P(t), K_i) > \text{Threshold}(K_i, S_M)$, K_i is detected. As long as we find the proper transform function $T()$ and the threshold functions, we break the system.

7. Timing Attack applied to Elliptic Curves

Timing attack is based on the following idea. Denote by k_j the j -th bit of the secret key k . Then we get the following equation for the total running time in Algorithm 1 with input point P_i :

$$T_i = e_i + \sum_{j=0}^{w-1} (D_{i,j} + k_j A_{i,j}). \tag{4}$$

In this formula, $D_{i,j}$ denotes the time needed for a point doubling operation for bit j , and $A_{i,j}$ denotes the time for an addition operation for bit j , e_i is some "noise" (run time for looping, if-operation, and other external influences). Note that both the doubling time $D_{i,j}$ and the addition time $A_{i,j}$ depend on the chosen random point P_i (index i) and the iteration index j . Let $0 < r < w$, and assume that we already know the "upper" bits k_{w-1}, \dots, k_{r+1} of the binary representation of the secret key k . Kocher's fundamental idea is the fact that we can determine the values of $D_{i,r}$ and $A_{i,r}$ with this information. For every sample point P_i we use the known bits of k to determine the value of the point Q at the beginning of iteration $j = r$ in Algorithm 1. Once we know these points we can determine both $D_{i,r}$ and $A_{i,r}$ using the decryption device (we assume that we can determine the run time for a single point operation). The next step of timing attack is the usage of $D_{i,r}$ and $A_{i,r}$ to determine a (probable) value for the next bit k_r using statistical methods. Similar to the above method, we can assume that we also know all intermediate timing values $D_{i,j}$ and $A_{i,j}$ for all $r \leq j \leq w-1$ and all samples indices i . Therefore we can compute T_i minus the "upper part" of the sum in (4).

8. Experimental Results

Table 1 describes some practical information applied to random scalars of different sizes. We show the length of prime number, the complete running time, and the average number of iterations for one bit of the scalar. It should be noted that all the point computations and field inversions were repeated several times, until the variance of the single timings for these operations was sufficiently small (and hopefully the timing error sufficiently small). Seeing the number of iterations (i.e. backtracks) especially for large scalars, we dare to conclude that several parameters of our implementation of the timing attack algorithm (e.g. the sample number, the definition of "sufficiently different") are not yet chosen optimally. Further examinations and determination of optimal parameters have to be done in future.

Length of Prime Number	Average number of iterations per bit	Total Running Time
$2^{24} - 1$	3.54	11 min 13 sec
2^{24}	0.71	1 min 20 sec
$2^{32} - 1$	1.16	18 min 37 sec
2^{32}	0.84	7 min 1 sec
$2^{48} - 1$	1.10	27 min 15 sec
2^{48}	0.94	12 min 8 sec
$2^{96} - 1$	2.33	53 min 32 sec
2^{96}	1.28	22 min 7 sec
$2^{128} - 1$	4.06	58 min 8 sec
2^{128}	1.37	43 min 32 sec
$2^{159} - 1$	7.38	59 min 26 sec
2^{159}	1.88	53 min 15 sec

9. Conclusion

The running time of the attack can be several hours; it always succeeded to determine the secret scalar. We are optimistic that further experiments can greatly improve the still large running time by searching for more optimized parameters. Therefore timing attacks should be considered as a serious threat for EC security system implementations in mobile applications and hence anticipated. Anticipating timing attacks is quite simple

10. References

- [1] Ekambaram Kesavulu Reddy, "Elliptic Curve Cryptosystems and Side-channel Attacks", Published in international Journals of Network Security, 2009.
- [2] Dhem J.-F., Koeune F., Leroux P.-A., Mestre P., Quisquater J.-J., and Willem J.- L.s, "A practical implementation of the timing attack", Proc. of the Third Working Conference on Smart Card Research and Advanced Applications (CARDIS 1998), LNCS 1820, 1998.
- [3] Marc Joye, "Elliptic curves and side channel analysis", published in ST Journal of System Research, 2003.
- [4] Boneh D. and Brumley D. Remote timing attacks are practical, To appear in the 12th Usenix Security Symposium, 2003.
- [5] Lu Z., Mah M., Neve M., and Peeters E., "Timing Attacks on Elliptic Curve Cryptosystems", Project Presentation for course "CS588: Cryptology Principles and Applications, Fall 2001", University of Virginia, Department of Computer Science, available at
- [6] Kocher P., Timing Attacks on Implementations of Diffie-Hellman, RSA, DSS, and the Systems, In N. Koblitz, editor, Advances in Cryptology - CRYPTO'96, Lecture Notes in Computer Science vol. 1109, pp. 104 - 113, 1996.
- [7] Wireless Application Protocol Forum, Wireless Transportation Layer Security, Version WAP-261-WTLS-20010406-a,21. April 2001,
- [8] An Implementation Tutorial on "Elliptic Curve Cryptography", By Anoop MS.

Hand Held Emergency Wireless Telemedicine System

Suganthi.J. # , N.V.Umareddy # and Sridharan.B *

Student AMC Engineering college, VTU Karnataka, India # professor AMCEC, VTU Karnataka, India * group lead, DRAX,C-DOT, Bangalore, India

Abstract—with the rapid development of computer science and communication technologies, doctors will employ electronic communication to facilitate patient care more and more. We have developed a portable telemedicine system which is much more flexible, robust and easier to use. It helps eliminate distance barriers and can improve access to medical services that would often not be consistently available in distant rural communities. It is also used to save lives in critical care and emergency situations. There are many healthcare technologies which have been implemented around the globe. Amongst these technologies, very few are used for an emergency case. In this paper we are going to introduce Portable Emergency System which is based on Locate-Diagnose-Move technique. This system is the collaboration of GSM/GPRS, GPS, sensors (wearable device) and P2P technology. This system is will be useful for all people especially for the emergency movement and diagnoses the roaming heart (Cardiac) patients, diabetes patients, elder peoples as well as accidental victim. The main aim of this system is to provide urgent provisional medication and movement of patient to the hospital which can save lives of many before the contact of expert doctors.

Keywords—India; TeleHealth; Telemedicine; GSM, GPRS; GPS;P2P (Peer to Peer); , mobile, wireless, Cardiac

I. INTRODUCTION

Recent developments in telecommunications and information technology enhanced availability of telemedicine system in health care fields. Telemedicine is defined as the delivery and sharing of medical information of patient over a distance using telecommunication. Existing telemedicine systems only exchange the information in a limited location. Using the portable device, the telemedicine service is feasible on the move.

Telemedicine provides the facility for patients to receive medical treatment from their own desired place. It helps to save the time and money for such people who cannot afford the travel along with the factor of cost. Telemedicine helps in providing significant services to those areas where medical facilities are not available commonly. Apart from benefits of telemedicine, there are such obstacles that exist in this area. Those barriers are legal issues of physician, patient confidentiality and so on. Currently telemedicine systems involve an integration of networking technologies with healthcare processes. The interoperability problem in telemedicine is clear in patient monitoring, diagnostic, decision support and communication systems needed at the point of care. Different telemedicine systems are made by

different vendors and they apply different technology, standards and information [1]

II. EXISTING TELEMEDICINE SYSTEM

Early forms of telemedicine achieved with telephone and radio have been supplemented with video telephony, advanced diagnostic methods supported by distributed client/server applications, and additionally with telemedical devices to support in-home care.

Telemedicine can be broken into three main categories: store-and-forward, remote monitoring and (real-time) interactive services. Store-and-forward telemedicine involves acquiring medical data (like medical images, bio signals etc.) and then transmitting this data to a doctor or medical specialist at a convenient time for assessment offline. It does not require the presence of both parties at the same time.

Dermatology (tele dermatology), radiology, and pathology are common specialties that are conducive to asynchronous telemedicine. A properly structured medical record preferably in electronic form should be a component of this transfer. A key difference between traditional in-person patient meetings and telemedicine encounters is the omission of an actual physical examination and history. The 'store-and-forward' process requires the clinician to rely on history report and audio/video information in lieu of a physical examination.

Remote monitoring, also known as self-monitoring or testing, enables medical professionals to monitor a patient remotely using various technological devices. This method is primarily used for managing chronic diseases or specific conditions, such as heart disease, diabetes mellitus, or asthma. These services can provide comparable health outcomes to traditional in-person patient encounters, supply greater satisfaction to patients, and may be cost-effective.

Interactive telemedicine services provide real-time interactions between patient and provider, to include phone conversations, online communication and home visits many activities such as history review, physical examination, psychiatric evaluations and ophthalmology assessments can be conducted comparably to those done in traditional face-to-face visits. In addition, "clinician-interactive" telemedicine services may be less costly than in-person clinical visit[2] Telemedicine provides the facility for patients to receive medical treatment from their own desired place. It helps to save the time and money[3]

Tele monitoring is a medical practice that involves remotely monitoring patients who are not at the same location as the health care provider. In general, a patient will have a number of monitoring devices at home, and the results of these devices will be transmitted via telephone to the health care

provider. Tele monitoring is a convenient way for patients to avoid travel and to perform some of the more basic work of healthcare for themselves.[4]

In addition to objective technological monitoring, most tele monitoring programs include subjective questioning regarding the patient's health and comfort. This questioning can take place automatically over the phone, or tele monitoring software can help keep the patient in touch with the health care provider. The provider can then make decisions about the patient's treatment based on a combination of subjective and objective information similar to what would be revealed during an on-site appointment.[5]

Some of the more common things that tele monitoring devices keep track of include blood pressure, heart rate, weight, blood glucose, and haemoglobin. Tele monitoring is capable of providing information about any vital signs, as long as the patient has the necessary monitoring equipment at his or her location. Depending on the severity of the patient's condition, the provider may check these statistics on a daily or weekly basis to determine the best course of treatment.

III. PROPOSED SYSTEM

Heart failure is generally regarded as the inability of the heart to provide adequate blood flow to the body. This progressive disorder affects over 5 million people in the United States and around 15 million worldwide. There is a 20% lifetime risk of developing heart failure for both men and women. Mortality rates of 30%–40% during the first year after diagnosis have been published, and after 5 years, this percentage increases to 60%–70%. Out of these, more than 50% dies without the in-time treatment or emergency movement.

If we talk about road accidents in metros, average 30% to 40% victims die on road because of lack in communication and tracking systems. 10% die because of first-aid treatment. To avoid this type of death, a strong emergency system is required. [1]

In the case of emergency situation normal telemedicine system will not workout. Because in normal condition the doctors and patient will be present at both the ends. But in emergency situation we cannot expect the doctor at the other end. For that we have to add GSM module to the send SMS the patient information to the doctor and the nearby hospital. So they will send ambulance and doctor immediately to the patient place. In this way we can save the patient easily

• THE GSM AND GPRS SYSTEMS

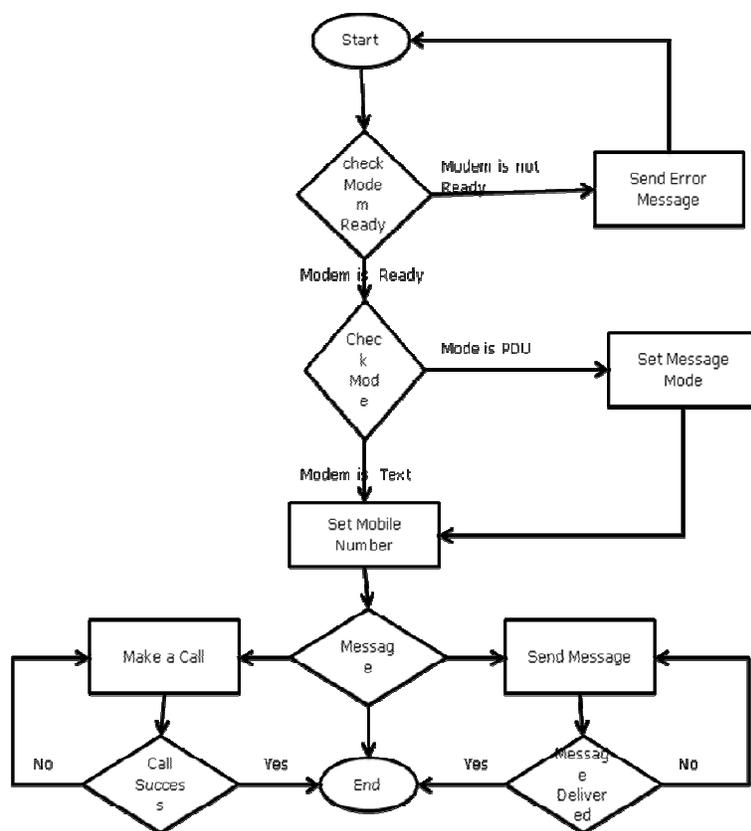
The main wireless technologies that are being used in wireless telemedicine systems are the GSM, GPRS, satellite, Wireless Local Area Network (WLAN) and Bluetooth. The GSM is considered as the second generation (2G) of the mobile communication networks.

**TABLE I
GSM AND GPRS FREQUENCY BANDS AND DATA TRANSFER RATES**

Type	Frequency band	Data transfer rates
GSM	900/1800/1900 MHz	9.6 – 43.3 kbps
GPRS	900/1800/1900 MHz	171.2 kbps

When GSM is in the standard mode of operation, it provides data transfer speeds of up to 9.6 Kbps (see Table I). Throughout the years a new technique was introduced in the GSM standard called HSCSD (High Speed Circuit Switched Data). This technology makes it possible to use several time slots simultaneously when sending or receiving data, so that the user can increase the data transmission up to 43.3 Kbps [3].

The theoretical maximum downlink data rate for GPRS is 171.2 kbps



Flow chart of GSM Operation

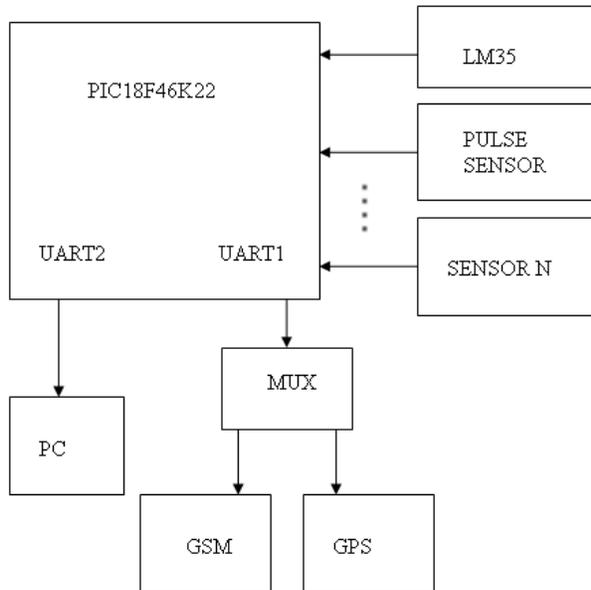


Figure 1. Block diagram

IV WORKING OF THE SYSTEM

These technologies permit communications between patient and medical staff with both convenience and fidelity, as well as the transmission of medical, imaging and health informatics data from one site to another. This device works as telephone using GSM/GPRS, for locating patient GPS etc. as well as to measure few biological parameters like pulse rate (PR), blood pressure (BP), body temperature etc.[1]

The heart of the circuit is PIC18F46K22 microcontroller. This family offers the advantages of all PIC18 microcontrollers – namely, high computational performance at an economical price – with the addition of high-endurance, Flash program memory. On top of these features, the PIC18 (L) F2X/4XK22 family introduces design enhancements that make these microcontrollers a logical choice for much high performance, power sensitive applications.

All of the devices in the PIC18(L)F2X/4XK22 family offer ten different oscillator options, allowing users a wide range of choices in developing application hardware.

The entire sensor's output is given to the microcontroller. The wearable system consists of an PIC processor. Two links run at RS-232 levels, and are for dedicated RS-232 devices. Any commercially available device that runs RS-232 (such as a GPS, GSM) can be connected.

If the doctor at remote distance wants to communicate with the patient, he or she can do so through a microphone-loudspeaker that's incorporated into the device. If necessary, an ambulance can be quickly dispatched to the patient's location. The patient is easily found with the built-in GPS receiver

- Red Alert Button, which will alert the heart patient that something is going wrong. That might be abnormal pulse rate or BP or body temperature or combination of any parameters. This alert will be for

specific time of 30 sec., after that automatic call goes to Monitoring Center.

- Speaker, which is used for telephonic conversation.
- Call Button, which is useful for talking with Monitoring Center or with the relatives in emergency when the red alarm starts.
- Display, which is useful to display normal time, phone numbers as well as other bio-parameters like BP, pulse rate, body temperature etc. time-to-time with the help of motion keys.

V EXPECTATION

The Telemedicine system makes health care system more affordable and easily accessible to everyone. It can be life saver in cases of emergency situation. The telemedicine system approach is a way to enable extremely powerful strategies for preventing diseases and improving health and the overall quality of life for elderly and chronically ill persons living alone into their homes. When taken one by one, the architecture of each system component. With the hope that Telemedicine can address rural and urban healthcare needs, The system offers a realistic, face-to-face experience remotely while concurrently offering high resolution video images, as well as high-fidelity telephonic audio. In the case of emergency situation normal telemedicine system will not work out. Because in normal condition the doctors and patient will be present at both the ends. But in emergency situation we cannot expect the doctor at the other end. For that we have to add GSM module to the send SMS the patient information to the doctor and the nearby hospital. So they will send ambulance and doctor immediately to the patient place. In this way we can save the patient easily.

Telemedicine is a generic term covering the application of a variety of proven electronic and communication techniques in providing healthcare. The techniques have already been applied in the context of teletriage, telediagnosis, telefollow up and telemonitoring.

Telemedicine has a variety of applications in patient care, education, research, administration and public health. It has the potential to deliver several benefits to patients, clinicians and the health service as a whole.

VI. CONCLUSION AND FUTUREWORK

The small portable device will play an important role in the TeleHealth system by providing Locate-Diagnose-Move services to the roaming heart patients and accidental victims to save their lives. If we utilize the available resources in Indian metros (initially) like GSM/GPRS network. GPS services (ISRO) etc. Using this system, we can save up to 50% lives that are losing it without primary treatment in emergency.

In future, we will come with P2P- system. Further we would like to propose Wearable Tele-Bio Watch with additional facilities to reduce the rate of mortality without treatment. We

are supposed to add facilities like speed alarm, which will alert you if you are crosses the speed limit to avoid accidents. Another facility of display your location, which will help you to locate yourself in emergency. In future this device can replace your mobile phone completely.

REFERENCES

- [1]. TeleHealth: Healthcare Technologies and TeleHealth Emergency (THE) System
Minesh Ade, Nikolaos Doulamis, Shyam S. Wagle, M. Ghazanfar Ullah
978-1-4577-0787-2/11/\$26.00 ©2011 IEEE
- [2]. Portable emergency telemedicine system over wireless broadband and 3G networks
SungHye Hong, SangYong Kim, Jungchae Kim, DongKyu Lim, SeokMyung Jung, DongKeun Kim, Sun K. Yoo September 2-6, 2009
- [3]. The Role of Service Oriented Architecture in Telemedicine Healthcare System
Asadullah Shaikh, Muniba Memon, Nasrullah Memon, Muhammad Misbahuddin
978-0-7695-3575-3/09 \$25.00 © 2009 IEEE DOI 10.1109/CISIS.2009.181
- [4.] Requirement Study of Telemedicine System and product, 2010 3rd International Conference on Biomedical Engineering and Informatics (BMEI 2010)
- [5]. Transforming Telemedicine for Rural and Urban Communities
Telemedicine 2.0 - Any Doctor, Any Place, Any Time, 978-1-4244-6376-3/10/\$26.00 ©2010 IEEE
- [6]. W. M. Omar and A. Taleb-Bendiab, "Service oriented architecture for e-health support services based on grid computing," Proceedings of the IEEE International Conference on Services Oriented Computing, Chicago, IL, September 2006
- [7]. The Study of Auxiliary Telemedicine system, 2010 International Conference on Advances in Energy Engineering
- [8]. Adaptive bandwidth reservation and scheduling for Efficient wireless telemedicine traffic transmission 0018-9545/\$26.00 © 2011 IEEE
- [9]. Fair and Efficient Scheduling for Telemedicine Traffic Transmission over Wireless Cellular Networks 978-1-4244-2517-4/09/\$20.00 ©2009 IEEE
- [10]. A Multimedia Telemedicine System 978-1-4244-5316-0/10/\$26.00 ©2010 IEEE
- [11]. HCPP: Cryptography Based Secure EHR System for Patient Privacy and Emergency Healthcare 1063-6927/11 \$26.00 © 2011 IEEE DOI 10.1109/ICDCS.2011.83
- [12]. Systems of systems applications for telemedicine 9th RoEduNet IEEE International Conference 2010
- [13]. Wireless Networked Chinese Telemedicine System: Method and Apparatus for Remote Pulse Information Retrieval and Diagnosis 0-7695-3113-X/08 \$25.00 © 2008 IEEE DOI 10.1109/PERCOM.2008.45
- [14]. Advanced Telemedicine Services through Context-aware Medical Networks Charalampos Doukas, Ilias Maglogiannis, and George Krementz's

Secure and Reliable Data Transmission in Wireless Sensor Network: A Survey

Rudranath Mitra¹, Tauseef Khan²

^{1,2}Department of Information Technology
Heritage Institute of Technology
Kolkata – 700107, West Bengal, INDIA.

Abstract

Wireless sensor network increases its application in industrial field as well as in consumer application very rapidly. Its growth increases day by day. Sensor node normally senses the physical event from the environment such as temperature, sound, vibration, pressure etc. Sensor nodes are connected with each other through wireless medium such as infrared or radio waves it depends on applications. Each node has its internal memory to store the information regarding the event packets. Basically this whole sensor network called sensor net is working in a distributive manner, sensor nodes are deployed in a huge area and use to send data packet in broadcast manner. This data packet finally reaches to the base station or called sink and vice versa. Nodes are deployed over a huge region in an ad-hoc based manner and use to sense the physical events. If any region cannot be sensed by any nodes then that region is called blind area. If blind area is too large then data retrieval is become unreliable. Nodes normally works in a collaborative manner to perform a specific task by transferring data packet to its neighbor nodes and so on until it reached to the base station. Every node has its own transmission range and within this transmission range node can transmit data packet. The event packet which sensor node transmit may be secret or confidential for the application ; so the data transmission must be secured to maintain the confidentiality of data packets.

Key word – One way hash chain (OHC), Request for missing packet (RMP), Message authentication code (MAC), Base Station (BS).

1. Introduction

Wireless sensor network is an emerging field where lots of research work has been done involving hardware and system design, networking, security factor and distributed algorithm[1,2,3]. Sensor nodes normally sense the data packet and transfer it to the base station via some intermediate nodes. The sensor nodes are low cost, low power and short transmission range [4]. Nodes use to send data packet locally to its single hop neighbor nodes and so on and finally it reaches to its base station. Initially nodes are deployed flying from aircrafts or randomly and some time node changes its initial position (the time of deployment) and moves across the region based on the requirement; so this type of nodes is called mobile nodes. So there are two types of data transmission in wireless sensor network, these are – direct transmission and multi-hop data transmission. In direct transmission data are send directly to the sink where as multi-hop transmission data send via no of intermediate nodes lies between source node and base station. In sensor network the flow of data is very important aspect because each data packet contains the event which may be very important for some application. So data transmission must be secured. But sensor node has limited energy and limited memory capacity so maintaining security is difficult for them. [5] It should be made sure that, the reports from the ‘sensors in action’ are authentic and reach the base station (BS) without any fabrication or modification. The task of securing wireless sensor networks is however, complicated because sensors are highly anonymous devices with a limited energy and memory capacity, and initially they have no knowledge of their locations in the deployment environment. To make the data transmission secure some basic aspects of security has to be maintained during transmission. Here we discussed how the authentication and confidentiality maintained during data transmission because without this two parameter data transmission cannot be reliable; also we discussed how the missing packets can be detected during transmission by some efficient methods. In this paper we will discuss various ways to make the data transmission secure and efficient; also we will discuss some mechanism and protocol used in secure data transmission.

2. Mechanism for secure data transmission

Basically to make the data transmission secure first we have to maintain two basic fields these are – Authentication and confidentiality. Authentication means it has to make sure that data packet comes from the intended sender or packet received by the intended receiver those which are involved in the transmission process. Confidentiality refers preventing the data packets from any unauthorized access.

So here we discuss a scheme which makes the data transmission secure from base station to sender node[5]. One-way hash chain (OHC) and shared secret key; OHC is mainly used for authentication and secret key used for confidentiality which are pre stored at the time of deployment of nodes initially. OHC is basically series of hash function and for each data packet there is identical

hash number is generate which spread through all over the network to make data transmission authenticate. Zhu et al. [6] Proposed the interleaved hop-by-hop authentication scheme that detects false reports through interleaved authentication. Lee and Cho [6, 7] proposed an enhanced interleaved authentication scheme called the key inheritance-based filtering that prevents forwarding of false reports.

2.1 One-way hash chain (OHC) based security scheme

One way hash chain is basically used for authentication. It consists of sequence of number generated by hash function F.[8] At first base station generate a random number and then applied F on it to make the other sequence number. The first phase is used for initializing the one-way hash chain number in the network. We create a secure path from the base station to the source nodes (any sensor in the network). Along the path, the OHC plays the major role to provide authenticity of the reported data.

Assume initially all the nodes and base station are in same transmission range and each node locally broadcast its data packet to its immediate neighbor nodes. The data transmission is bi- directional that means data can be send from both side. All the sensors and the base station have shared secret keys that are pre-stored before deployment of the network. So, when the sensors are deployed in the target area randomly, each sensor contains a shared secret key with the base station. To provide data transmission authenticity, all the intermediate nodes between any source and the base station must be initialized with the basic one-way hash chain number. Let us suppose the initial hash chain number is Hs0 which is initialized by base station and it broadcast this number to all the nodes in the network.

At first base station send a control message to its one hop neighbor node which is called base station control message denoted by (bcm) . This control packet contains the initial hash number (Hs0), message authentication code (MAC) which is generated by key K_i , where k_i is key generated at the time slot t_i . The format of the control message is:

$$bcm: B|HS0|MACK_i(B|HS0)$$

Here B is the id of base station.

Now this control message is send to the immediate neighbor nodes. After receiving this control message the nodes set the base station as its forwarder and add its own id and rest of the message is remain same. This control message is then again send to its one hop neighbor node denoted by (ncm). The format of this message is

$$ncm: sid|fid|B|HS0|MACK_i(B|HS0)$$

Here sid is the node own id and fid is forwarder id which is here base station id for its immediate neighbor nodes.

Now this message is again locally broadcast by the sender nodes and received by its immediate neighbor nodes and so on. This process is continues until all the nodes in the network gets the message. If any node already got the control message then it rejects the later control message.

Now, any node that has not received the message earlier (i.e., two hops away from the base station) receives it and stores the initial OHC number, Hs0. It then sets the id of the sender node as its forwarder node and again locally broadcasts the control message with its own id as *sid*. In case, it has got the control packet from two or more sender nodes, it picks up the message which it receives first and discards all other messages. However, this node stores the ids of the other senders. This knowledge is necessary to repair a broken path, which we will discuss later in this paper.

If any node set as a forwarder of any downstream nodes then eventually it knows by the bi-directional data transmission. If any downstream node already set its forwarder then it rejects other upstream immediate neighbor nodes forwarder request.

In fig (1) it shows that in this sensor network node 1, 2, 3 are the immediate one hop neighbor nodes of base station. Here total three paths which are BS-1, BS-2-4-7, and BS-3-5-6.

At first control packet send by the base station and received by the nodes 1,2,3; then data packet locally broadcast from each of the data packet. In this fig (1) node 4 set node 2 as its forwarder. Though node 4 is within the range of node 2 as well as node 3. So both nodes can send data packet by local broadcasting. But node 2 already set as the forwarder of node 4 , so node 4 rejects the request of node 3 which wants to become the forwarder. This process is maintained in the whole network.

Now node 1 and node 2 are the one-hop neighbors of the base station and they both get the control message from the same source (which is in this case the base station itself). So, when the local broadcasts of node 1 or node 2 reach each other, as previously stated they simply ignore the messages. This process continues until all the nodes in the network are authenticated using the OHC mechanism.

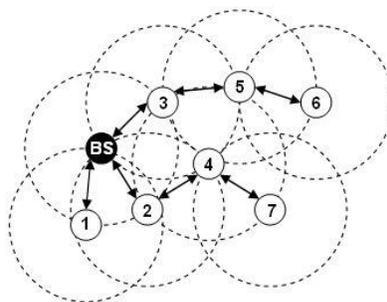


Fig (1)

So here each node locally broadcast the control message to its neighbor node which contain the hash number and authentication code and this process is continues until all the nodes are authenticated.

2.2 Secure data transmission

After initializing the hash number and getting the hash number by all the nodes which is sent by the base station now all the nodes ready to start data transfer securely. Here each source nodes maintains a unique hash chain number $H_s: \langle H_{sn}.. H_{sn-1}.. H_{s1}.. H_{s0} \rangle$.

it encrypts the packet with its shared secret key with the sink (or base station), includes its own id and an OHC sequence number from HS in the packet. It attaches HS_1 for the first packet, HS_2 for the second packet, and so on. To validate an OHC number, each intermediate node n_1, \dots, n_m maintains a verifier IS for each source node, ns . Initially, IS for a particular source node is set to HS_0 . When source node sends the i th packet, it includes HS_i with the packet. When any intermediate node receives this packet, it verifies, if $IS = F(HS_i)$. If so, intermediate node validates the packet, it forwards it to the next intermediate node, and sets IS to HS_i . In general, intermediate node can choose to apply the verification test iteratively up to a fixed number w times, checking at each step whether, $IS = F(F\dots(F(HS_i)))$. If the packet is not validated after the verification process has been performed w times, intermediate nodes simply drops the packet.

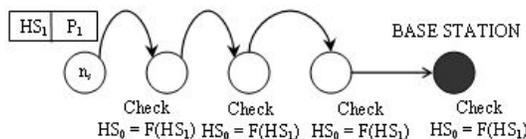


Fig (2.a)

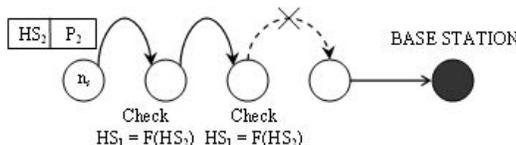


Fig (2.b)

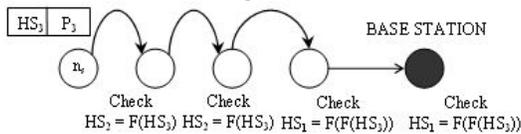


Fig (2.c)

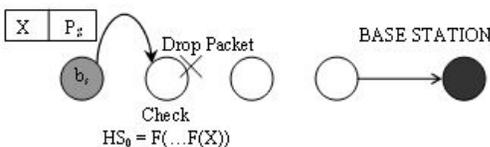


Fig (2.d)

Here it shows in fig (2) how the data are transmitted securely from source node to base station securely.

At first in fig (2.a) source node send data packet 1 along with its corresponding hash number Hs1 to the base station via some intermediate nodes. Here each intermediate node has verifier (Is) as described earlier and initially verifier set as Hs0 (Is=Hs0). At each intermediate node data packet checks for validation that is $Is=F(Hs1)$. If true then validation done and data packet forward to next node and again checks for validation and so on. Finally it reaches to base station. If in between any intermediate node then validation fails the data packet discarded.

Now in fig (2.b) the data packet couldn't reach to the base station because of some reason but still it doesn't hamper the data packet validation using verifier (Is). Here in this fig second data packet couldn't reach to the base station.

In fig (2.c) here node send the data packet 3 along with its corresponding hash number (Hs3). Now in previous case packet 2 with its hash number is been discarded in some intermediate node and it couldn't reach to the base station. So here data packet 3 will be verified by the verifier (Is) which is set to Hs2 up to the intermediate node which already received the previous data packet along with its hash number(Hs2) and from the intermediate node where Hs2 was discarded (in fig 2) the verifier set to Hs1 and checks the new data packet 3. So here basically two validation is done during the transmission and finally at base station two times validation is done for the received data packet. Finally in fig (2.d) source node send false packet which detects by its immediate neighbor node during validation then it immediate discarded.

So by this process data packet securely transmitted to the base station or vice versa. Source node encrypts the data packet by its shared secret key and after receiving the packet base station decrypts its by its secret key which was pre stored as stated earlier.

2.3 Optional key refreshment Mechanism for data freshness

To make the data freshness and increase the level of security another mechanism is used which is called optional key refreshment. Here base station broadcast a session key (Ks) to all nodes.

The format of this message is

$$B|Ks| \text{ MACKs}(B|Ks))$$

Initially this session key is encrypted by its secret key and transfer to its immediate neighbor nodes, now neighbor nodes after receiving the session key it performs X-OR operation with its own old key and form a new derived key.

This newly derived key is again used for sub sequent data packets.

So by this process every time key is getting updated and it increases the level of security every time. Because here same secret key is not used for secure data transmission, every time a newly derived key is used.

3. Mechanism for repairing the broken path in sensor network

If any intermediate node is crashed or fails to transmit then the whole path along that node is broken, all the downstream nodes get detached from the base station. So here we will use some mechanism to repair the broken path and create another alternative path to maintain the transmission of data packets.

In fig (1) if node 2 is crashed or fails to transmit data then all the downstream nodes here 4 and 7 gets detached from base station, so no data will flow along that path. This failure could first be detected by the one-hop neighbors of node 2 in the network i.e., node 4 and node 1.

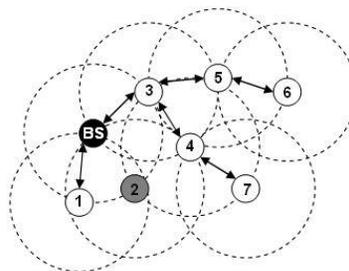


Fig (3)

Here in fig (3) it shows that node 2 is crashed or fails to send data so the entire path along that node gets down. Now to repair this broken path some technique has to be followed. Node 4 was within the range of node 2 and node 4 initially. But node 4 set

node 2 as its forwarder. Node 3 also locally broadcast the data packet to node 4 and initially wanted to become the forwarder of node 4. But node 4 rejects its request because it already set node 2 as its forwarder initially. But before rejecting the request node 4 initially in fig (1) stored the id of node 3 and also used to store other nodes id who wanted to become the forwarder of node 4. This phenomena is used to repair the broken path .Now in fig (3) node 4 has no forwarder because node 2 is crashed. So node 4 now send a message to node 3 and set it as its forwarder because initially node 4 stored the id of node 3. So in fig (3) it shows that new path has been created from Bs-3-4-7.

If node 4 is required to send any packet as a source node, it could simply send it using the OHC number in the sequence, HSk+1 which is next to its last used OHC number, that is Hsk. For node 3, node 4 is a new source, so it could save its HS value in I4. The subsequent transmissions from node 4 are verified by node 3 based on this initial knowledge.

4. Reliable data transfer with no data loss

Normally data transfer has to be reliable and make sure that no data is lost during the transmission. So there are two basic methods by which data is transfer from one node to another node reliably. This are- 1. Acknowledgement based method, 2. Non acknowledgement based method.

4.1 Acknowledgement based method vs. non acknowledgement based method [9]

For both ack based and non ack based method there are three kinds of message used.

1. Transmission of data packet pi .
2. Acknowledgement of packet pi for ack based method and RMP that is request for missing Packet in non ack based method
3. Re-transmission of packet pi .

4.2 Acknowledgement based method

A node waits for an acknowledgement from receiver after sending the packet to receiver. It waits for a time period (ta). If no acknowledgment receives within this time period then sender again re-transmits the data packet and this process continues until sender gets the acknowledgement for that data packet from receiver.

4.3 Non acknowledgement based method

In Non-ack based method sender send the first data packet to the receiver. After that it send the next data packet without waiting for acknowledgement of the previous sent data packet because here no acknowledgement is used. Now after receiving the next data packet the receiver can realize that previous data packet has not been received by the receiver; so it send a RMP request(request for missing packet) for that missing data packet. After receiving this RMP sender only send that requested data packet. So this is the main working mechanism for non-ack based method.

5. Reduction of delivery time by reducing the number of message transfer

5.1 Acknowledgement based method:

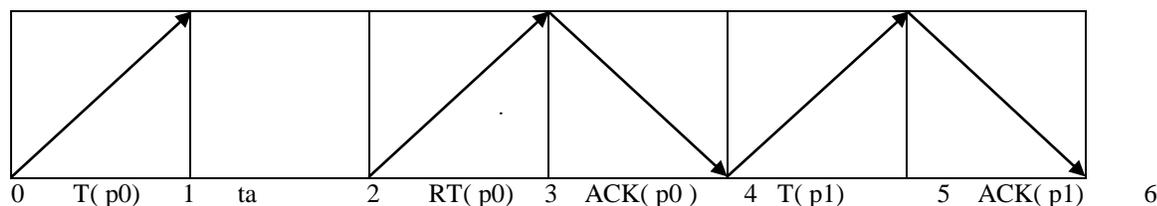


Fig (4.a)

In fig(4.a) first data packet 1 is send by the sender. Then it waits for time period (ta). After waiting for time period when sender doesn't any ack then it re-transmits the data packet. In next time slot it got then ack from receiver then sends the next data packet. Here in acknowledgement based method total no of message is five and total time slot required is 6.

5.2 Non-acknowledgement based method

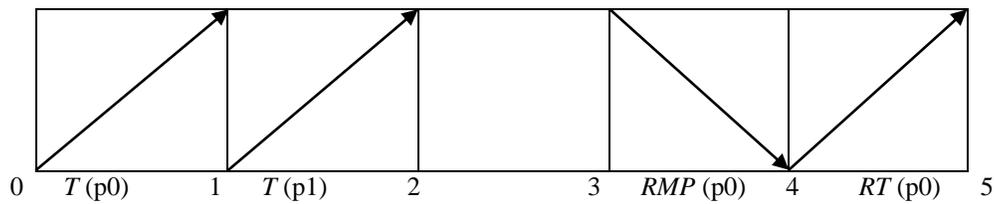


Fig (4.b)

Here it shown that in non acknowledgement based method number of message is less than Ack based method so total delivery time is less in case of non Ack based method. Here total time slot required is 5.

So it is proved that non Ack based method takes less time to deliver packet as compare to Ack based method.

6. Comparative study:

In this paper we have studied many different methods of data transmission which are used for various aspects of data transmission. These mechanisms are used for reliability, security, detection of missing packets etc. Some mechanism is used for authentication; some maintain confidentiality to secure data during transmission etc. In this paper we basically focused on OHC and pre-stored secret key mechanism for ensuring security also optional key refreshment mechanism increase the security level by changing the session key every time. Also there are methods which are used for detecting the missing packets during transmission.

So here we are trying to compare those mechanisms by some standard parameter. So the comparison table as follows:

Mechanism	Authentication	Confidentiality	No of message deliver	Delivery time	Reliability
OHC	High	—	—	Varies	High
Pre-stored secret key	—	high	—	—	High
Optional key refreshment	High	High	—	Varies	High
Ack-based method	—	—	More	More	High
Non-ack based method	—	—	Less	Less	Low

7. Conclusion

In this paper we discussed how the data transmission can be secured and what are the main mechanisms is used to make the data transmission secure also how the broken path can be repair by re-initialization of OHC. We discussed the reliability factor in terms of data packet loss detection by these two methods. There are lots of scopes for further work on this topic also we can enhance the level of security by some new approach and reduce the delivery time by reducing more number of message transfers. Here we basically discussed how efficiently data transmission can be secured by some mechanisms. Here the optional key refreshment mechanism basically increase the level of security and ensure data freshness.

So in future there are lots of scopes to work on this field and enhance security mechanism by different methods also we can increase the reliability by reducing the no of missing packets. In future we can also work on energy efficient data transmission so the life time of each sensor node can be increased.

Acknowledgement

We express our sincere gratitude to the senior members of Faculty, Department of Information Technology, Heritage Institute of Technology, Kolkata, for extending their valuable time and guidance for the completion of this paper.

References

1. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “Wireless Sensor Networks: A Survey”, *Computer Networks*, 38, pp. 393-422, 2002.
2. S. Dai, X. Jing, and L. Li, “Research and analysis on routing protocols for wireless sensor networks”, *Proc. International Conference on Communications, Circuits and Systems*, Volume 1, pp. 407-411, 27-30 May, 2005.
3. D. E. Culler, and W. Hong, “Wireless Sensor Networks”, *Communication of the ACM*, Vol. 47, No. 6, pp. 30-33, June 2004
4. Yu Wang, Hongyi Wu, and Nian-Feng Tzeng, “Energy-efficient Data Transmission in Wireless Sensor Networks”, in Center for Advanced Computer Studies University of Louisiana at Lafayette P.O. Box 44330, Lafayette, LA 70504. Email: {yxw1516, wu, [tzeng](mailto:tzeng@cacs.louisiana.edu)}@cacs.louisiana.edu.
5. Al-Sakib Khan Pathan and Choong Seon Hong, “An Efficient Scheme for Secure Data Transmission in Wireless Sensor Network.”, in Department of Computer Engineering, Kyung Hee University spathan@networking.khu.ac.kr and cshong@khu.ac.kr.
6. S. Zhu, S. Setia, S. Jajodia, P. Ning, “An Interleaved Hop-by-Hop Authentication Scheme for Filtering of Injected False Data in Sensor Networks”, In *Proc. of S&P*, pp. 259-271, 2004.
7. H. Y. Lee, and T. H. Cho, "Key Inheritance-Based False Data Filtering Scheme in Wireless Sensor Networks", *LNCS 4317*, Spr.-Ver., pp. 116-127, 2006.
8. L. Lamport, “Constructing digital signatures from oneway function.” In technical report SRI-CSL-98, SRI International, October 1979.
9. Al-Sakib Khan Pathan and Choong Seon Hong, “Reliable and Efficient Data Transfer in Wireless Sensor Networks via Out-of-Sequence Forwarding and Delayed Request for Missing Packets”, in *Computer Science Dept, Louisiana State University Baton Rouge, LA 70803, USA* {ddatta1@lsu.edu, kundu@csc.lsu.edu}.

Detection and Classification of Epileptic Seizures using Wavelet feature extraction and Adaptive Neuro-Fuzzy Inference System

Dr. D. Najumnessa* Dr. T. R. Rangaswamy

B S Abdur Rahman University, Chennai, India

*Corresponding author

Abstract - Epilepsy, a neurological disorder in which patients suffer from recurring seizures, affects approximately 1% of the world population. In this work, an attempt has been made to enhance the diagnostic importance of EEG using Adaptive neuro fuzzy inference system (ANFIS) and Wavelet transform coefficients. For this study, EEG for 20 normal and 30 seizure subjects under standard recording procedure is used. A method based on wavelet transform and ANFIS is used to detect the epileptic seizures. Further, BPN algorithm is used to study and compare the datasets. Average specificity of 99% and sensitivity of 97% are obtained. Results show that the ANFIS is able to detect seizure. It appears that this method of detection makes it possible as a real-time detector, which will improve the clinical service of Electroencephalographic recording.

Keywords - ANFIS, ANN, BPN, Discrete Wavelet Transform, Epileptic seizure.

1. INTRODUCTION

Surveys report that about 1% of the people in the world suffer from epilepsy and about 30% of epileptics are not helped by medication [1]. It is characterized by recurrent, paroxysmal (short-term) electrical discharges of the cerebral cortex that result in intermittent disturbances of brain function as reported by Neidermeyer [2]. Electro-encephalography (EEG) is an inexpensive and an important clinical tool for the evaluation and treatment of neurophysiologic disorders [3].

The only reliable and quantitative method of assessment of epileptiform discharges occurring in the EEG between seizures is an important component in the diagnosis of epilepsy is by digital recording with subsequent analysis [4, 5]. This paper will concentrate in building a computer aided diagnosis system to classify the normal signals and epileptic seizure signals by analyzing the EEG. This system consists of two stages. The first stage is the feature extraction from EEG signals and the second stage is the classification of these features, as shown in figure 1.

Significant diagnostic information can be obtained from the frequency distribution of epileptic EEG. However, selecting the signal processing technique is very important. Many efforts have been reported in literature on the classification of epileptic seizures using frequency analysis [6-10]. One popular approach as shown by Sharma et al.[6] is to use the Fast Fourier Transform (FFT) for the first stage and using FFT calculate the features from the FFT such as relative power in various frequency bands. But Fast Fourier Transform (FFT) suffers from large noise and sensitivity. The second stage then uses an Artificial Neural Network (ANN) to generate a single number that indicates the degree to which the event is a seizure. A few workers like Pritchard et al [7] , Anderer et al [8] have compared ANNs to some of the statistical methods on the same EEG data when looking for changes in the background EEG and have generally found that the ANNs perform better; that is they more closely model human performance in observing differences between normal and abnormal subjects.

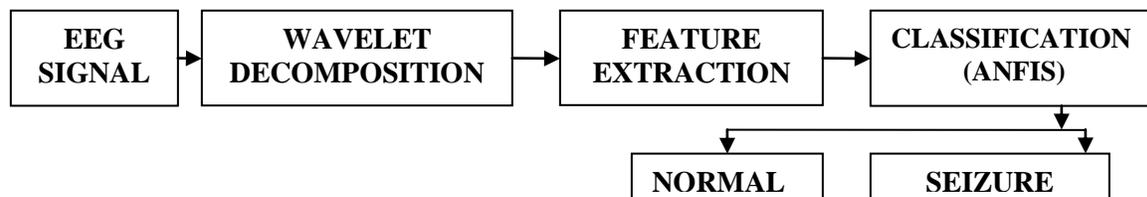


Fig.1. Schematic of the Classification method

Comparing techniques of seizure detection developed by diverse groups is filled with problems because of no objective definition of the appearance of a seizure in the EEG. It has been shown by Gotman et al [9], that there are significant differences among EEG readers regarding which individual events to mark, even when there is overall agreement on the

elucidation of an EEG record. A powerful method to perform time-scale analysis of signals is the wavelet transforms (WT). This technique provides a consistent framework for different techniques that have been developed for various applications [4, 10 -12]. It should also be emphasized that the WT is appropriate for analysis of non-stationary signals, and this

represents a major advantage over spectral analysis. Hence the WT is suitable for locating transient events. Such transient events as spikes can occur during epileptic seizures. Wavelet is an effective time–frequency analysis tool for analyzing transient signals. Its properties and feature extraction can be used to analyze various transient events in biological signals as reported by Adeli et al. [4] gave an overview of the DWT developed to be renowned with and quantifying spikes, sharp waves and spike-waves. They used wavelet transform to analyze and characterize epileptiform discharges in the form of 3-Hz spike and wave complex in patients with absence seizure. Through wavelet decomposition of the EEG records, transient features are accurately captured and localized in both time and frequency context. The capability of this mathematical microscope to analyze different scales of neural rhythms is a powerful tool for investigating small-scale oscillations of the brain signals. An understanding of the dynamics of the human brain through EEG analysis can be obtained through further analysis of such EEG records. Recently, a work on time-frequency analysis of EEG signals for detecting seizures using WT has been reported [13 - 17].

Neural networks are routinely employed in signal classification algorithms including automatic seizure detection. Gotman, Ives and P Gloor., [18] used template matching for the detection of epileptic seizures. Cheng-wen et al [19] employed an Artificial Neural Network (ANN) for identifying seizures using radial basis function. [20 – 22] used wavelet coefficients as an input to a feed forward neural network for the detection of epileptogenic transient waveforms. As EEG signals are non-stationary, Proakis and Manolakis,[23] apply the conventional method of visual analysis but is not highly successful in diagnostic classification. Temporal patterns are processed by the approach of using recurrent neural networks (RNNs) which have memory to encode past history [24,25] extracted features and used back propagation algorithm to classify normal and abnormal EEG.

Fuzzy set theory plays a vital role in dealing with uncertainty when making decisions in biomedical applications. Introduced by Zadeh [26], fuzzy logic and fuzzy set theory are employed to describe human thinking and reasoning in a mathematical framework. Fuzzy-rule based modeling is a qualitative modeling scheme where the system behavior is described using a natural language [27]. Fuzzy sets have attracted the growing attention and interest in modern information technology, production technique, decision making, pattern recognition, diagnostics, data analysis, etc. [28, 29]. These intelligent computational methods offer real advantages over conventional modeling, especially when the underlying physical relationships are not fully understood.

In recent years, the integration of neural networks and fuzzy logic has given birth to new research into neuro-fuzzy systems. Neuro-fuzzy systems have the potential to capture the benefits of both these fields in a single framework. Neuro-fuzzy systems eliminate the basic problem in fuzzy system

design (obtaining a set of fuzzy if–then rules) by effectively using the learning capability of an ANN for automatic fuzzy if–then rule generation and parameter optimization. As a result, those systems can utilize linguistic information from the human expert as well as measured data during modeling. Such applications have been developed for signal processing, automatic control, information retrieval, database management, computer vision and data classification [30–34].

A specific approach in neuro-fuzzy development is the adaptive neuro-fuzzy inference system (ANFIS), which has shown significant results in modeling nonlinear functions. In ANFIS, the membership function parameters are extracted from a data set that describes the system behavior. The ANFIS learns features in the data set and adjusts the system parameters according to a given error criterion [34]. Successful implementations of ANFIS in biomedical engineering have been reported, for classification [30–33, 35–39], for modeling and controlling real systems [40] and data analysis [41].

In this study, a simple approach based on ANFIS is implemented to classify the EEG signal to one of the categories: epileptic or normal. The aim of this study is to develop a simple algorithm for the detection of epileptic seizure which could also be applied to real-time. This paper aims to extract features using WT and to detect and classify seizure using ANFIS and neural network techniques. The choice of the network is based on the fact that it is the most popular type of ANNs and the most powerful networks commonly used in solving classification / discrimination problems. The accuracy of the classifiers will be assessed and compared. Finally, some conclusions is drawn concerning the superiority of the ANFIS over the ANNs and impacts of features on the detection of epileptic seizure through analysis of the ANFIS.

2. MATERIALS AND METHOD

2.1. Subjects and data recording

Subjects within the age group from 21 to 40 were selected for this study. The EEG is collected using Nihon Kohden digital EEG system comprising of a data acquisition system, signal processor and a personal computer from Sri Ramachandra Medical University and Research Institute, Chennai. The 10 second scalp EEG data used in this study is sampled at a rate of 500 Hz after filtering between 1 and 70 Hz. A bipolar electrode montage of 16 channels is used in the analysis. The EEGs were recorded with Ag/AgCl electrodes placed at the F₄, C₄, P₄, O₂, F₃, C₃, P₃, O₁, Fp₂, F₈, T₄, T₆, Fp₁, F₇, T₃, & T₅, loci of the 10–20 International System. Impedance is kept below 5 k Ω to avoid polarization effects. All data are stored for off-line processing. All EEGs with artifact, electrode movement and bursts of alpha waves is discarded. In order to assess the performance of the classifier, we selected 900 EEG segments containing spike and wave complex, artifacts, and background normal EEG.

2.2. Visual inspection and validation

Two EEG technologists with experience in the clinical analysis of EEG signals separately inspected every recording included in this study to score ictal and normal signals. The two experts jointly revised the signals to solve disagreements and set up the training set for the program for the epileptic seizure detection. The agreement between the two experts was evaluated for the testing set. A further step was then performed with the aim of checking the disagreements and setting up a 'gold standard' reference set as done by Subasi, [14]. When revising this unified event set, the human experts, by mutual consent, marked each state as ictal or normal. They also examined each recording completely for epileptic seizures. This validated set provided the reference evaluation to estimate the sensitivity and specificity of computer scorings.

2.3. Wavelet analysis and feature extraction

2.3.1. Denoising the EEG signals

One of the major difficulties in analysis of EEG signal is the presence of artifacts in the signals. This nature of disturbance is a serious obstructing factor that prohibits further processing to identify useful diagnostic features. Hence denoising of the signals is a must for effective utilization for diagnosis. Since the frequency bands of these noises may overlap with the seizure signal, conventional method of using filters is not suitable for removal of noise. In this work, DWT based denoising technique, namely wavelet shrinkage denoising, as reported by Kandaswamy et al [42] is used.

2.3.2. Multiresolution decomposition of EEG signals

The Discrete Wavelet Transform (DWT) is a versatile signal processing tool that finds many engineering and scientific applications.. DWT analyzes the signal at different frequency bands, with different resolutions by decomposing the signal into a coarse approximation and detail information [10 - 13]. DWT employs two sets of functions called scaling functions and wavelet functions, which are associated with low-pass and high-pass filters, respectively. The decomposition of the signal into the different frequency bands is simply obtained by successive high-pass and low-pass filtering of the time domain signal. Selection of wavelet that is suitable and the number of levels of decomposition is very vital in analysis of signals using DWT. We visually inspect the data first, and if the data are discontinuous, Haar or other sharp wavelet functions are applied [44] or else a smoother wavelet can be employed. Usually, tests are performed with different types of wavelets and the one which gives maximum efficiency is selected for the particular application. In this study db4 is chosen [4,5]. The levels are chosen such that those parts of the signal that correlate well with the frequencies required for classification of the signal are retained in the wavelet coefficients. Since the EEG signals do not have any useful frequency components above 30 Hz, the number of levels was chosen to be 8. Thus the signal is decomposed into the details D1–D8 and one final

approximation, A8 [5, 42]. The ranges of various frequency bands are shown in Table 1.

Table 1: Frequencies corresponding to different levels of decomposition for Daubechies 4 filter wavelet with a sampling frequency of 500 Hz

Decomposed signal	Frequency range (Hz)
D1	125 – 250
D2	62.5 – 125
D3	31.25 – 62.5
D4	15.625 - 31.25
D5	7.8125 - 15.625
D6	3.9063 - 7.8125
D7	1.9531 - 3.9063
D8	0.9766 - 1.9531
A8	0 - 0.9766

2.3.3. Feature extraction

The extracted wavelet coefficients provide a compact representation that shows the energy distribution of the signal in time and frequency. It is anticipated that the coefficients of the seizure frequency spectrum ranges from 0.5 to 30 Hz. So the coefficients corresponding to the frequency bands, D1- D3 were discarded, thus reducing the number of feature vectors representing the signal. In order to further lessen the dimensionality of the extracted feature vectors, Gotman [9] features are used from the wavelet coefficients. These feature vectors, calculated for the frequency bands A8 and D4–D8, is used for classification of the EEG signals.

2.4. Adaptive neuro-fuzzy inference system (ANFIS)

The ANFIS first introduced by Jang in 1993 [34]. It is a model that maps inputs through input membership functions (MFs) and associated parameters, and then through output MFs to outputs. The initial membership functions and rules for the fuzzy inference system can be designed by employing human expertise about the target system to be modeled. ANFIS can then purify the fuzzy if–then rules and membership functions to describe the input–output behavior of a complex system. Jang showed that even if human expertise is not available it is possible to intuitively set up practical membership functions and employs the neural training process to generate a set of fuzzy if–then rules that approximate a desired data set [30–33,35].

Five layers are used to create this inference system. Each layer involves several nodes described by node function. The output signals from nodes in the previous layers will be accepted as the input signals in the present layer. After manipulation by the node function in the present layer will be served as input signals for the next layer. Here square nodes, named adaptive nodes, are adopted to represent that the parameter sets in these nodes are adjustable. Whereas, circle nodes, named fixed nodes, are adopted to represent that the

parameter sets are fixed in the system. For simplicity to explain the procedure of the ANFIS, we consider two inputs x , y and one output f in the fuzzy inference system. And one degree of Sugeno's function [34] is adopted to depict the fuzzy rule. Hence, the rule base will contain two fuzzy if-then rules as shown in equations (1) and (2):

Rule 1: if x is $A1$ and y is $B1$ then $f = p1x + q1y + r1$. (1)

Rule 2: if x is $A2$ and y is $B2$ then $f = p2x + q2y + r2$. (2)

For detailed literature refer [34].

3. Results and discussion

In this study, the potential of neuro-fuzzy computing techniques for epileptic seizure detection is investigated by using an ANFIS -Takagi Sugeno Model. Since an ANFIS model is essentially a fuzzy model with learning capabilities; we consider the various methods that are involved in the design of a fuzzy model. To construct a fuzzy model we need to perform knowledge acquisition, which takes a human operators knowledge about how to control the system and generates a set of fuzzy if-then rules as the backbone for a fuzzy model that behaves like the original human operator [34].

The design method for the proposed neurofuzzy model makes use of the numerical information method to create a fuzzy inference system that learns the input-output data pairs presented to it. The inputs to the ANFIS network are the extracted features, obtained from the WT at various levels. In this study, training and test sets were formed by 900 feature vectors. The 700 feature vectors (from 20 epileptic and 15 normal subjects) were used for training and the 200 feature vectors (from 10 epileptic and 5 normal subjects) were used for testing. The type distribution of the samples in the training and validation data set is summarized in Table 2. In order to improve the generalization capability of the ANFIS, the training and the test sets were formed by data obtained from different subjects. The training data set was used to train the ANFIS, whereas the testing data set was used to verify the accuracy and the effectiveness of the trained ANFIS model for the detection of epileptic seizure in patients with generalized tonic-clonic (GTC) seizures.

Table 2 Type distribution of the samples in the training and test data sets

Type	Training set	Test set	Total
Epileptic	450	120	570
Normal	250	80	330

3.1 Design of ANFIS model

The input output data set is features extracted from the wavelet coefficients. The first step involved in the training of an ANFIS is the creation of the initial Sugeno type FIS based on the input output data set. This was accomplished by using the *genfis1* function in MATLAB. This function partitions the fuzzy space into various regions based on the number of

membership functions given for each input. Three membership functions (Low, Medium, and High) were used for each input variable.

The three membership functions correspond to low, medium, high respectively. Here the fuzzy space which is a plane with input features as the rectangular co-ordinates is divided into 243 regions; each corresponding to a rule and an output membership function which is linear in nature. The Sugeno FIS thus created was trained by invoking the *ANFIS editor* GUI.

Input membership function:

$$\mu_{Ai}(x) = \frac{1}{1 + \{((x - c_i) / a_i)^2\} b_i} \quad (3)$$

Output membership function: $Z = px + qy + r$ (4)

Where a , b , c are the antecedent parameters and p , q , r are the consequent parameters. Training proceeds in two phases; in first phase the premise parameters a , b , c of the model is fixed and the consequent parameter is estimated by the method of Recursive Least Squares (RLSE). In the second phase the consequent parameters is fixed and the premise parameters is tuned by the back propagation algorithm. This is done until the specified numbers of epochs are reached or the error between the input and the target patterns is lesser than the tolerance value specified [34]. At the end of training the parameters i.e. a , b , c & p , q , r obtained, represents the optimum parameters of the ANFIS network. Figures 2 - 7 show the changes in the shape of the membership functions before and after training for the input features. Figures 8-9 show the rule outputs before training and after training.

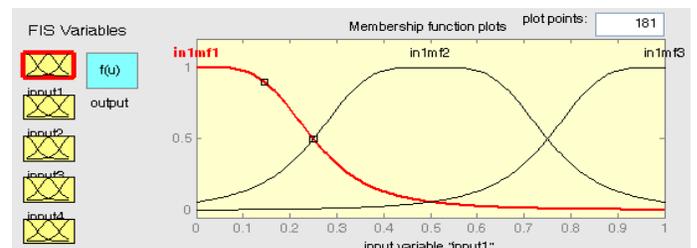


Fig 2. Membership functions before training

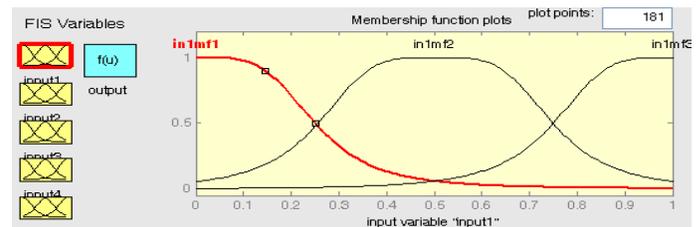


Fig 3. Membership input1 functions after training

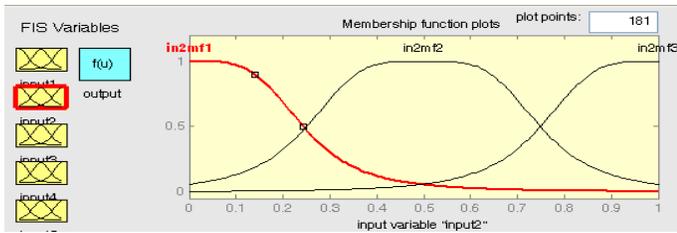


Fig 4. Membership input2 functions after training

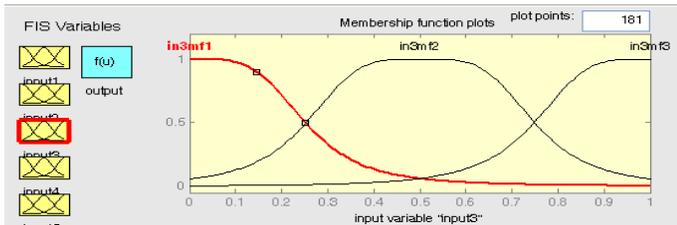


Fig 5. Membership input3 functions after training

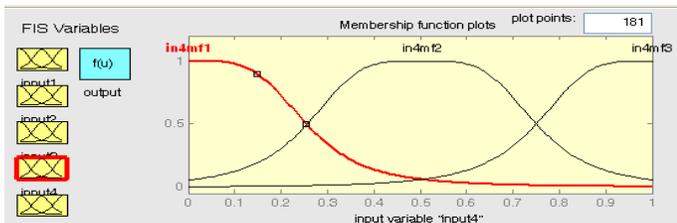


Fig 6. Membership input4 functions after training

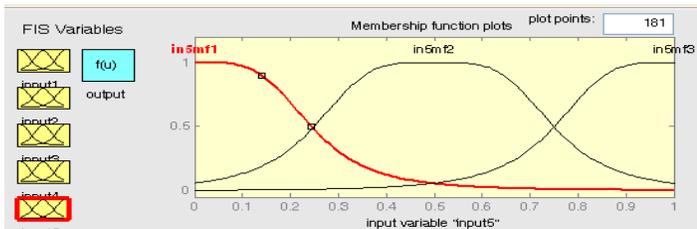


Fig 7. Membership input5 functions after training

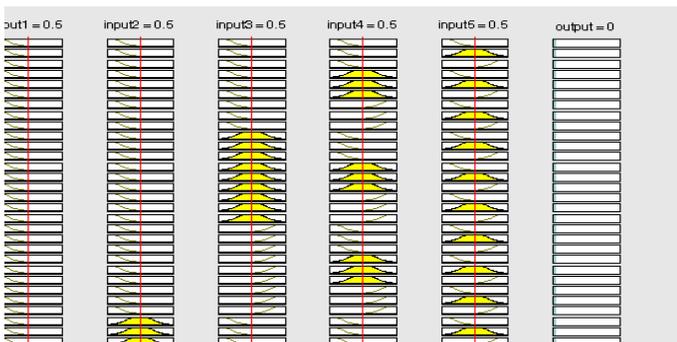


Fig 9. Rule outputs before training

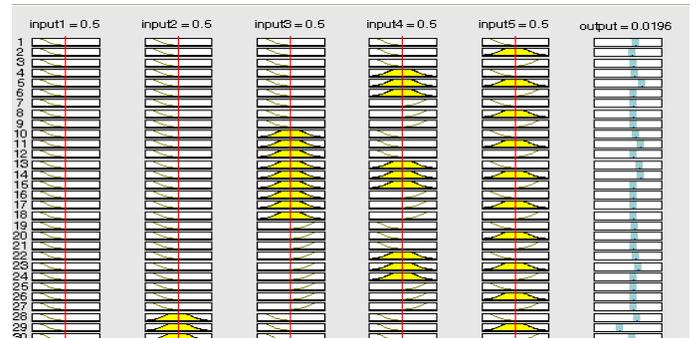


Fig 10. Rule outputs after training

3.2 Performance analysis

The test performance of the ANFIS model was determined by the computation of the statistical parameters such as sensitivity, specificity and total classification accuracy. The sensitivity, specificity and total classification accuracy are defined as follows:

Sensitivity: The ratio of the number of true positives to the sum of true positives and false negatives.

Specificity: The ratio of the number of true negatives to the sum of true negatives and false positives

Accuracy: It is the number of patterns detected by the total number of patterns.

Out of 120 feature vectors of seizure 117 patterns were detected (TP) and 3 were not detected (FN). Out of 80 non-seizures feature vectors 78 were detected (TN). The performance of the classifier for normal and seizure conditions are shown in Table 3 and success rate in Table 4.

Table 3. The performance of the classifier for normal and seizure conditions

Type of seizure	Total feature	Correctly detected	Incorrectly detected
Normal	80	78	2
Tonic-Clonic	120	117	3

Table 4. The accuracy of the network

Test outcome	Total feature	Correctly detected	Accuracy (%)
Specificity	80	79	98.75
Sensitivity	120	116	96.7
Accuracy	200	195	98

Table 5 Comparison of BPN and ANFIS models for EEG signal classification

Classifier type	Correctly classified (%)	Specificity (%)	Sensitivity (%)
BPN	95	95.7	94.3
ANFIS	98	98.5	96.7

As it is seen from Table 5, the ANFIS model classified normal subjects and epileptic subjects with the accuracy of 98.75.7% and 96.7%, respectively. The correct classification rates of the neural network (BPN) were 95.7% for normal subjects and 94.3% for epileptic patients. The total classification accuracy of the stand-alone neural network was 95%. Thus, the accuracy rates of the ANFIS model presented for this application were found to be higher (98%) than that of the stand-alone neural network model. These results indicate that the proposed ANFIS model has some prospective in epileptic seizure detection

4. CONCLUSION

Diagnosing epilepsy is a difficult task requiring observation of the patient, an EEG, and gathering of additional clinical information. In this work, a simple and new technique of an automatic diagnostic classification system is designed and executed. Conventional method of classification of EEG signals using mutually exclusive time and frequency domain representations does not give efficient results. The epileptic seizure signals were decomposed into time-frequency representations using wavelet transform and statistical features were calculated to describe their distribution. An ANFIS-based system was implemented for the classification of epileptic seizure using the statistical features as inputs. The proposed technique involved training the ANFIS classifier to detect epileptic seizure in EEG when the statistical features extracted from the wavelet sub-bands of EEG signals were used as inputs. The presented ANFIS classifier combined the neural network adaptive capabilities and the fuzzy logic qualitative approach. Some conclusions relating to the impacts of features on the epileptic seizure detection were obtained through analysis of the ANFIS. The classification results and statistical measures were used for evaluating the ANFIS. The classification of normal subjects and epileptic patients were done with the accuracy of 99% and 97%, respectively.

Acknowledgements

We would like to thank the EEG experts and the team of technicians for their technical support especially for the data acquisition and handling, from the department of Neurology, Sri Ramachandra University, Chennai.

5. REFERENCES

[1] NINDS. 'Seizure and epilepsy: hope through research', National Institute of Neurological Disorders & Stroke, from: http://www.ninds.nih.gov/health_and_medical/pubs/seizures_and_epilepsy_htr.html.

[2] Niedermeyer E. Historical aspects: Niedermeyer E, Silva F, editors. *Electroencephalography, basic principles, clinical applications, and related fields*. Baltimore, Maryland: Williams & Wilkins, 1987.

[3] John G. Webster, *Medical instrumentations: Application and design book*, third edition, pp. 173-175, John Wiley & Sons, Inc., 1998.

[4] Hojjat Adeli, Ziqin Zhou, and Nahid Dadmehr, "Analysis of EEG records in an epileptic patient using wavelet transform", *Journal of Neuroscience Methods*, Vol. 123, pp.69-87, 2003.

[5] Subasi A., Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Systems with Applications*, 28, pp 701-711, 2005.

[6] Sharma A, Wilson, S.B. and Roy.R, EEG classification for estimating anesthetic depth during halothane anesthesia. *Int. Proc.14th Annual International Conference. IEEE, Newyork.*: pp 2409 -2410, 1992.

[7] Pritchard W.S., Duke, D.W., Coburn, K.L., MooreN.C. TuckerK.A., JoanM.W. and Hosteler, W.E EEG-based, neural-net predictive classification of alzheimers disease versus control subjects is augmented by non-linear EEG measures. *Electroenceph.clin. Neurophysiol.*, 91:pp 118 -130, 1994.

[8] Anderer,P, Saletu, B, KloppeI.B, Semlitsch. H.V. and Werner H, Discrimination between demented patients and normals based on topographic EEG slow wave activity: comparison between statistics discriminant analysis and artificial neural network classifiers. *Electroencephalograph. Clin. Neurophysiol.*, 91: pp 108 – 117,1994.

[9] Gotman,J., Gloor.,P,and Schaul, N. Comparison of traditional reading of the EEG and automatic recognition of interictal epileptic activity. *Electroencephography clinical. Neurophysiol.*, 44: pp 48 – 60, 1978.

[10] Quian Quiroga R and Schürmann M., Functions and sources of event-related EEG alpha oscillations studied with the Wavelet Transform. *Clin. Neurophysiol.*, 110: pp 643-655, 1999.

[11] Abdulhamit Subasi "Epileptic seizure detection using dynamic wavelet network", *Expert Systems with Applications* 29, 343-355, 2005.

[12] Quian Quiroga R, Rosso OA and Basar E. Wavelet-entropy: a measure of order in evoked potentials. *EEG Suppl.* 49: pp 298-302; 1999a.

[13] Abdulhamit Subasi ,EEG signal classification using wavelet feature extraction and a mixture of expert model *Expert Systems with Applications*, Volume 32, Issue 4, pp 1084-1093, 2007.

[14] Abdulhamit Subasi ,Application of adaptive neuro-fuzzy inference system for epileptic seizure detection using wavelet feature extraction *Computers in Biology and Medicine*, Volume37, Issue2, pp227-244,2007.

[15] Abdulhamit Subasi ,Automatic detection of epileptic seizure using dynamic fuzzy neural networks, *Expert Systems with Applications*, Volume 31, Issue 2, pp 320-328, 2006.

- [16] Abdulhamit Subasi, Ergun Erçelebi, Ahmet Alkan , Comparison of subspace-based methods with AR parametric methods in epileptic seizure detection, *Computers in Biology and Medicine*, Volume 36, Issue 2, pp 195-208, 2006.
- [17] Hasan Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy *Expert Systems with Applications* 36 , pp 2027–2036, 2009.
- [18] Gotman,J., Gloor,P,and Schaul, N.Comparison of traditional reading of the EEG and automatic recognition of interictal epileptic activity. *Electroencephography clinical. Neurophysiol.*, 44:pp 48 – 60, 1978.
- [19] Cheng-wen ko, Yue-Der Lin, Hsiao-Wen Chung, Gwo-Jen Jan “An EEG spike detection Algorithm using artificial neural network with multi-channel correlation” *Engineering in medicine and biology society*, Vol 4 , No. 4, pp 2070-2073, 1998.
- [20] Baikun WAN, Bikash Dhakal, Hongzhi QI, Xin zhu, Multi-method synthesizing to detect and classify epileptic waves in EEG, *Proceedings of the Fourth International Conference on Computer and Information Technology*, vol 10, pp 922 -926, 2004.
- [21] Hyun S Park, Yong H Lee, Doo S Lee, Sum I Kim, Detection of Epileptiform activity using wavelet and neural network, *Proceedings of the 19th Annual International Conference of the IEEE / EMBS Vol 3*, pp 1194 – 1197, 1997.
- [22] Tulga Kalayer, Ozcan Ozdamar, Nurgun Erdol, The use of wavelet transform as a preprocessor for the neural network detection of EEG spikes, *IEEE engineering in medicine and biology* , pp 1 – 3, 1994.
- [23] Proakis, J., and Manolakis, D., *Digital Signal Processing Principles, Algorithms and Applications*, Prentice Hall, New Jersey, 1996.
- [24] Guler, N.F., Ubeyli, E.D. and Gulaer, I. ‘Recurrent neural networks employing Lyapunov exponents for EEG signals classification’, *J. Expert Systems with Applications*, Vol. 29, pp.506–514, 2005.
- [25] Yaunanghi, N., Najumnissa, D. and Shenbagadevi, S. (2006) ‘Feature extraction and detection of epileptic seizure’, *Proceedings of the National Conference on Signals, Systems and Communication*, NCSSC-2006, Chennai, India, pp.8–12, 2006.
- [26] L.A. Zadeh, Fuzzy sets, *Inf. Control* 8 (3) ,338–353, 1965.
- [27] M. Sugeno, T. Yasukawa, A fuzzy-logic based approach to qualitative modeling, *IEEE Trans. Fuzzy Syst.* 1 (1) pp7–31, 1993.
- [28] C.A. Pena-Reyes, M. Siper, A fuzzy-genetic approach to breast cancer diagnosis, *Artif. Intell. Med.* 17 131–155, 1999.
- [29] D. Nauck, R. Kruse, Obtaining interpretable fuzzy classification rules from medical data, *Artif. Intell. Med.* 16 149 –169, 1999.
- [30] I. Guler, E.D. Ubeyli, Automatic detection of ophthalmic artery stenosis using the adaptive neuro-fuzzy inference system, *Eng. Appl. Artif. Intell.* 18, 413–422, 2005.
- [31] E.D. Ubeyli, I. Guler, Automatic detection of erythematous-squamous diseases using adaptive neuro-fuzzy inference systems, *Comput. Biol. Med.* 35,421–433, 2005.
- [32] I. Guler, E.D. Ubeyli, Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction, *Expert Syst. Appl.* 27 ,323–330,2004.
- [33] E.D. Ubeyli, I. Guler, Adaptive neuro-fuzzy inference systems for analysis of internal carotid arterial Doppler signals, *Comput. Biol. Med.* 35 (2005) 687–702.
- [34] J.S.R. Jang, ANFIS: adaptive network based fuzzy inference system, *IEEE Trans. Syst., Man Cybern.* 23 (3), 665–683,1993.
- [35] J.S.R. Jang, Self-learning fuzzy controllers based on temporal backpropagation, *IEEE Trans. Neural Networks* 3 (5) 714–723,1992.
- [36] J. Usher, D. Campbell, J. Vohra, J. Cameron, A fuzzy logic-controlled classifier for use in implantable cardioverter defibrillators, *Clin. Electrophysiol.* 22, 183–186, 1999.
- [37] S.Y. Belal, A.F.G. Taktak, A.J. Nevill, S.A. Spencer, D. Roden, S. Bevan, Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptive-network-based fuzzy inference system, *Artif. Intell. Med.* 24, 149–165, 2002.
- [38] E. Besdok, A new method for impulsive noise suppression from highly distorted images by using ANFIS, *Eng. Appl. Artif. Intell.* 17 , 519–527, 2004.
- [39] X. Lou, K.A. Loparo, Bearing fault diagnosis based on wavelet transform and fuzzy inference, *Mech. Syst. Signal Process.* 18 , 1077–1095, 2004.
- [40] J. Vieira, F.M. Dias, A. Mota, Artificial neural networks and neuro-fuzzy systems for modelling and controlling real systems: a comparative study, *Eng. Appl. Artif. Intell.* 17 ,265–273, 2004.
- [41] I. Virant-Klun, J. Virant, Fuzzy logic alternative for analysis in the biomedical sciences, *Comput. Biomed. Res.* 32 ,305–321, 1999.
- [42] A. Kandaswamy, C.S. Kumar, R.P. Ramanathan, S. Jayaraman, N. Malmurugan, Neural classification of lung sounds using wavelet coefficients, *Comput. Biol. Med.* 34 (6) ,523–537, 2006.
- [43] D.L. Donoho, I.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, *J. Am. Stat. Assoc.* 90 (432) 1200–1224, 1995.
- [44] M. Akay, Wavelet applications in medicine, *IEEE Spect.* 34 (5) , 50–56.

Customer and User Requirements Modeling Enhanced Software Development

Tawfik Saeed Zeki
IAU University

Abstract

This paper explores various software models for business globalization and the nature of customer requirements. It discusses building a new model which the customer need involve System Development Models as well as the system user requirement before building customer applications.

Introduction

There are many device industry is globalizing. As part of that trend, companies are developing products for sale worldwide rather than developing devices just for a particular market. This shift has placed new demands on design teams. The purpose of this roundtable was to identify these demands and to share opinions on how to best address them. The participants in the roundtable have backgrounds in human factors and user-interface design. We discussion focuses specifically on the best or most practical way to develop customer and requirements simultaneously. I agree that many companies are designing an increasing proportion of products for customer. Approximately 50% of sales are international [1]. It's a mind shift for a company that started off focused primarily on the U.S. market, and so we are reminded every day that we're working on international products. Consistency in user interfaces lets users make less error and they feel more secure in what they do. In theory all clients should conform to the requirements, in practice mileage depends on what the client technology gives [2].

In the innovation process, we have identified three distinct types of information that must be captured from customers to measure performance. They are the "jobs" customers are trying to get done when using a product or service; the "outcomes" they are trying to achieve when performing these jobs in a variety of circumstances and the "constraints" that stand in the way of them adopting or using a new product or service. Jobs, outcomes and constraints are critical inputs to the innovation process because they represent the three primary avenues in which managers are able to achieve growth: (1) by helping customers perform ancillary jobs or new jobs that could not be performed before, (2) by helping customers better perform a specific job, or (3) by helping customers overcome the roadblocks that stand in the way of them performing a specific job altogether.

The basis for the application of different user requirements methods is a simple process as shown in Figure 1 below encompassing 4 elements:

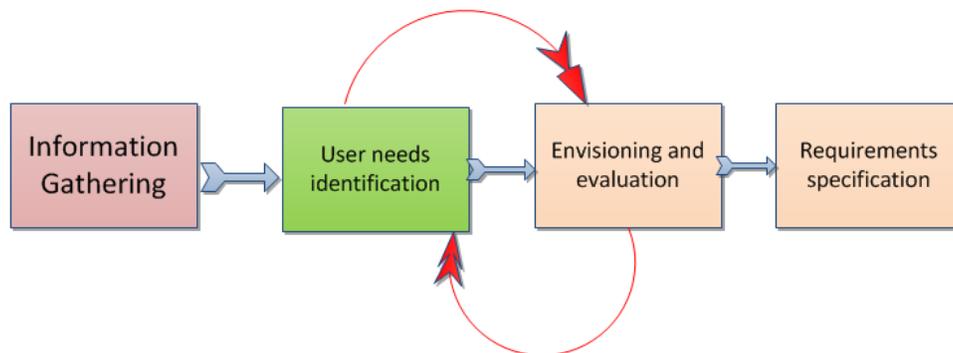


Figure 1.1: General process for user requirements analysis

The four stages, and methods used to support the stages, are described in the next sections, followed by a summary table highlighting the advantages and disadvantages of each technique.

Effectiveness of user requirements analysis

The effectiveness of user requirements analysis in the beginning of a development project depends to a large extent on the type of project.

Collecting user requirements for consumer products requires much effort, and the risk to fail is still very high. As long as consumers have no idea of the innovative product or service, it will be very difficult for them to state their needs. Creativity of designers is required for the transfer of user requirements into innovative consumer products.

For the development of professional applications precise user requirements analysis and specification is essential. Professionals often are available who perform the tasks under investigation.

Task analysis is mandatory for the development of safety critical applications. A characteristic of safety critical work domains is that tasks and procedures are precisely defined before new support tools are built. This is a good precondition for the specification of functional and non-functional requirements.

Developing requirements for customer applications

Collecting user requirements for consumer products requires much effort, and the risk to fail is still very high. As long as consumers have no idea of the innovative product or service, it will be very difficult for them to state their needs. Creativity of designers is required for the transfer of user requirements into innovative consumer products.

It is a much more complex process to develop customer requirements. You have many more players who desire input. And once you've collected and interpreted the input, you have to go back and verify that you got it right.

If the customer requirements are not fulfilled, the customer will be extremely dissatisfied. On the other hand, as the customer takes these requirements for granted, their fulfillment will not increase his satisfaction. The must-be requirements are basic criteria of a product. Fulfilling the must-be requirements will only lead to a state of "not dissatisfied". The customer regards the must be requirements as prerequisites, he takes them for granted and therefore does not explicitly demand them.

They also get feedback from the people who train customers to use products, as well as service products. So, while their information may be biased, it still has considerable value. You just have to further validate the information by means of follow-up analysis. Lacking input from managers, the task of developing customer requirements would take much more time and effort, which is often unacceptable because of a tight development schedule. But with new technology used by all people it could be help to almost too tight the schedule.

Conducting Web-based surveys

They have a lot of experience hosting surveys, compiling the data, and getting the results back in short order. But, you still have to recruit the right people to complete the survey. Current customers are much more likely to participate in the survey than noncustomers. If our study concerns cardiology, I simply pull from our database of cardiology customers. In my company, we've had success sending software user-interface prototypes to prospective users via e-mail. Recipients can run the prototypes in a standard browser. Using Web conferencing tools, we can direct a pretty effective walkthrough or even a full-blown usability test without leaving the office.

Surveys can help determine the needs of users, current work practices and attitudes to new system ideas. Surveys are normally composed of a mix of 'closed' questions with fixed responses and 'open' questions, where the respondents are free to answer as they wish. This method is useful for obtaining quantitative as well as some qualitative data from a large number of users about the problems of existing tasks or the Questionnaires are useful for collecting information especially we collect in the World Wide Web customers. Sometimes the questionnaires are not promptly replied and several follow-ups/personal interviews may be required to get questionnaires back from respondent's current system.

Even the survey takes more time and that will affected to the time scheduling of the project need more time, but give's a successful project lives more longer than a normal design. And to be successful at requirements gathering and to give the project an increased likelihood of success follows these rules:

1. Don't assume you know what the customer wants, ask.
2. Involve the users from the start.
3. Define and agree the scope of the project.
4. Ensure requirements are specific, realistic and measurable.
5. Gain clarity if there is any doubt.
6. Create a clear, concise and thorough requirements document and share it with the customer.
7. Confirm your understanding of the requirements with the customer (play them back).
8. Avoid talking technology or solutions until the requirements are fully understood.
9. Get the requirements agreed with the stakeholders before the project starts.
10. Create a prototype if necessary to confirm or refine the customers' requirements.

Model of User and Customer Satisfaction

User requirements analysis is an error prone part of the development process and errors not detected at this stage may lead to expensive system failures later. For this reason, user requirements should be verified as soon as design solutions and prototypes are available.

SDLC is not limited to technical activity but actually begins with customer needs and evolves through processes and user requirements to develop a solution or support process.

The primary objective of implementing a standardized SDLC policy is to provide coordinated excellent service, at reduced costs, to support the activity of customers.

A simplified and common framework for implementing SDLC will improve communications and promote coordination across projects Figure 1.2

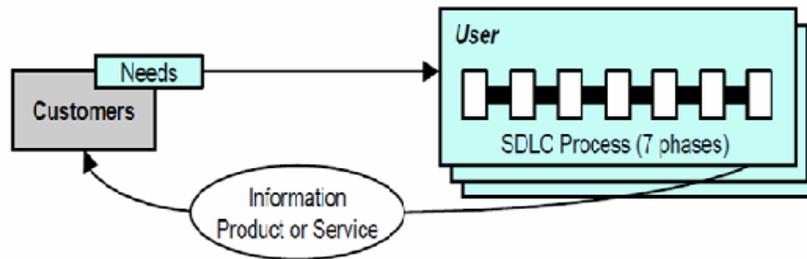


Figure 1.2 Customer need

As illustrated in Figure 1.3, the traditional SDLC is divided into five phases: planning, analysis, detailed systems design, implementation, and maintenance. The SDLC is an iterative rather than a sequential process [4].

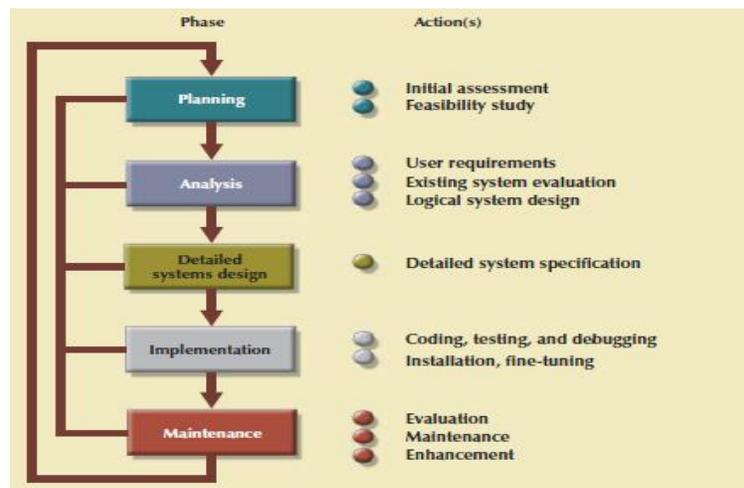


Figure 1.3 The system development life cycle (SDLC)

During this phase Implementation we use state of the art bug tracking and customer feedback systems to make sure that every issue is properly addressed. Quality assurance department is also responsible for producing statistical data which later is used to improve the development process.

Software development model is carefully selected based on the user requirements as well as the customer requirements. Most customers are not technologists, engineers or scientists and do not know the best solutions as a result, giving customers the solutions they request often leads to customer disappointment.

A specification is an input in which the customer states the desired size, weight, color, shape, look, feel or other product and service performance characteristics – in an attempt to shape the solution.

Customer need is a universal form of input and is typically stated as a high-level descriptor of quality. It is not uncommon for a customer to say that they want a product or service to be “reliable”, “effective”, “robust”, “stable”, “resilient”, “consistent”, “powerful”, “resistant”, “serviceable” or “dependable”.

During the designing the system usually the software developer depend on the user requirement not to the customer requirements which can be the new software system developer not satisfy the customer needs. Then the final system in reality cannot proceed as successful system.

The new basis for the application of different user requirements methods and different customers is a process as shown in Figure 1.4 below encompassing 4 elements:

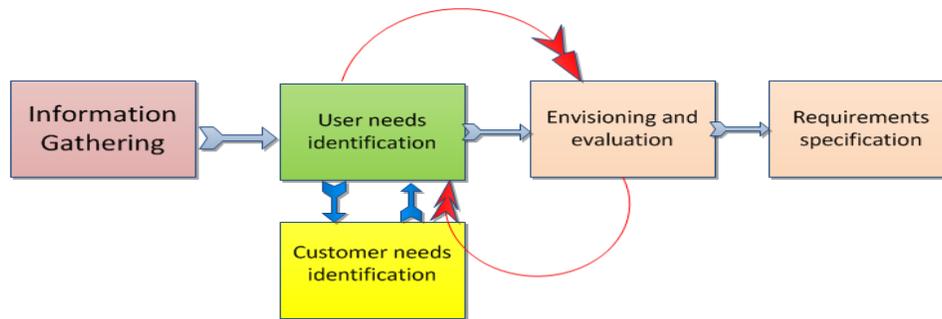


Figure 1.4: General process for user requirements analysis

My model Figure 1.5 suggests that the entire requirement should satisfy both user and customer needs. And the evolution the system before the release it comes from the both (User & Customer) in stage of Acceptance & Installation. If the user and customer acceptance the system release otherwise should go back to user and customer for requirements change that make the SDLS is more “reliable”, “effective”, “robust”, “stable”, “resilient”, “consistent”, “powerful”, “resistant”, “serviceable” or “dependable”.

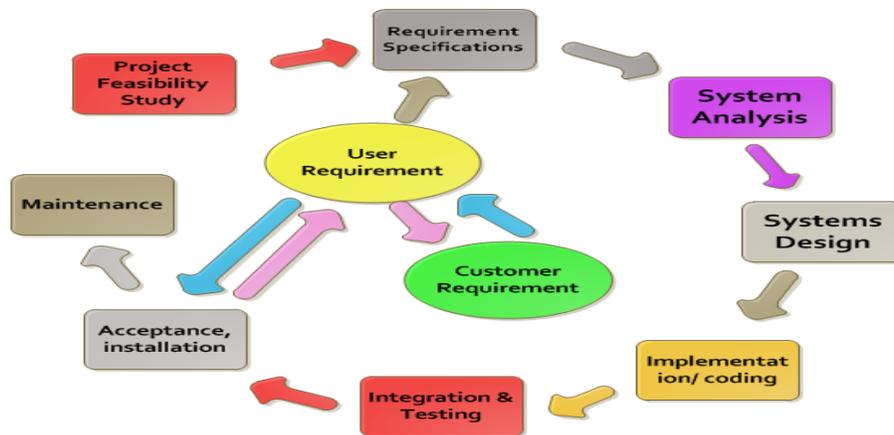


Figure 1.5 Circles - Model

Recommendations and Conclusion

The study found it imperative to recommend the following:

These recommendations are based on the finding of this research and this research and it is hoped that they would go a long way in making the use of the development the software system user and customer much easier. The analysis of user’s requirements and customer requirements studies would help not only the clients and customer, etc., but also the software developers/designers; creators, software creators, programmers and other information professionals in determining what information and delivery package are required by users and customer. The recommendations are:

1. The study found that software system development need user requirements as well as the customer requirements.
2. The findings that system developing depend on user and customer at the same time is more secure for a system to live more longer as well as been a great robust.
3. With most of the Internet users comfortable/satisfied using the Internet, it goes to show that it is the medium of the moment

References

1. <http://www.exportsolutions.com/ExportTipsDetails.aspx?id=144&title=Carrefour%20at%20a%20Crossroads>
2. <http://wiki.list.org/display/DEV/global+requirements>
3. <http://searchsoftwarequality.techtarget.com/definition/systems-development-life-cycle>
4. Carlos Coronel, Steven Morris, and Peter Rob, Database Systems: Design, Implementation, and Management, Ninth Edition
5. Anthony W. Ulwick, The Strategic Role of Customer Requirements in Innovation
6. Beyer, H. & Holtzblatt, K. (1998), *Contextual design: defining customer-centered systems*, Morgan Kaufmann Publishers.
7. Martin Maguire, User requirements analysis *A review of supporting methods*

Enhanced Clusterhead Selection Algorithm Using LEACH Protocol for Wireless Sensor Networks

Rudranath Mitra¹, Anurupa Biswas²

^{1,2}Department of Information Technology
Heritage Institute of Technology
Anandapur, Kolkata – 700107.INDIA.

Abstract

WSN is now a days a vast field for research. Its growth increases day-by-day. Routing protocol with energy efficiency has been a challenging issue in the design of wireless sensor networks. Efficiency and security are two topics in the design of routing protocol for WSNs. Heinzelman, ET. Al. introduced a hierarchical clustering algorithm for sensor networks; called Low Energy Adaptive Clustering Hierarchy (LEACH).The new proposed scheme describes two new ways to select Cluster head. Analysis shows that the improved or enhanced LEACH protocol balances the energy expense, saves the node energy and hence prolongs the lifetime of the sensor network.

Key Words: Clusterhead, , Energy conservation, Energy level, Optimum distance.

Introduction

LEACH is a cluster-based protocol. LEACH is one of the first hierarchical routing approaches for sensors networks. LEACH randomly selects a few sensor nodes as clusterheads (CHs) and rotate this role to evenly distribute the energy load among the sensors in the network. In LEACH, the clusterhead (CH) nodes compress data arriving from nodes that belong to the respective cluster, and send an aggregated packet to the base station in order to reduce the amount of information that must be transmitted to the base station(negotiation). WSN is considered to be a dynamic clustering method.

Operation :

It has 2 phases :

1. Set up State Phase
 2. Steady State Phase
- In the setup phase, the clusters are organized and CHs are selected.
 - In the steady state phase, the actual data transfer to the base station takes place.
 - The duration of the steady state phase is longer than the duration of the setup phase.
 - During the setup phase, a predetermined fraction of nodes, p , elect themselves as CHs.
 - A sensor node chooses a random number, r , between 0 and 1. Let a threshold value be $T(n)$. If this random number is less than a threshold value, $T(n)$, the node becomes a cluster-head for the current round. The threshold value is calculated based on an equation that incorporates the desired percentage to become a cluster-head, the current round, and the set of nodes that have not been selected as a cluster-head in the last $(1/P)$ rounds, denoted by G .

It is given by [2]:

$$T(n) = \frac{p}{1 - p(r \bmod (1/p))} \quad \text{if } n \in G$$

where G is the set of nodes that are involved in the CH election.

Each elected CH broadcast an advertisement message to the rest of the nodes in the network that they are the new cluster-heads. All the non-cluster head nodes, after receiving this advertisement, decide on the cluster to which they want to belong to. This decision is based on the signal strength of the advertisement. The non cluster-head nodes inform the appropriate cluster-heads that they will be a member of the cluster. After receiving all the messages from the nodes that would like to be included in the cluster and based on the number of nodes in the cluster, the cluster-head node creates a TDMA schedule and assigns each node a time slot when it can transmit. This schedule is broadcast to all the nodes in the cluster.

During the steady state phase, the sensor nodes can begin sensing and transmitting data to the cluster-heads. The cluster-head node, after receiving all the data, aggregates it before sending it to the base-station. After a certain time, the network goes back into the setup phase again and enters another round of selecting new CH.

Related Works

The idea proposed in LEACH has been an inspiration for many hierarchical routing protocols, although some protocols have been independently developed [1]. Taxonomy of the different architectural attributes of sensor networks is developed [3]. Further improvements on LEACH protocol for wireless sensor networks has been developed where both security & efficiency features have been dealt with[4]. Here the sensing area has been divided into a number of equilateral areas, called as clusters. Each cluster consists of six equilateral triangles called cells. The protocol consists of a number of rounds but after forming the clusters they do not change in each round. Both each equilateral triangle & each equilateral hexagon has same number of nodes. In each cell one cell head is selected & one CH is selected is chosen from six cell heads. The data are sent to the base station by using the multi-hop manner through a secure path consisting of cluster heads. The analysis shows that the improved protocol saves nodes energy, prolongs WSN lifetime, balances energy expenses and enhances security for WSNs.

To allow a single-tier network to cope with additional load & to be able to cover a large area of interest without degrading the service, networking clustering has been pursued in some routing approaches [7]. The hierarchical routing protocols involve nodes in multi-hop communication within a particular cluster to efficiently maintain the energy consumption of sensor nodes as well as perform data aggregation fusion to decrease the number of transmitted messages in the sink. Cluster formation is typically based on the energy reserve of sensors and sensor's proximity to the cluster head [5][6]. LEACH is one of the first hierarchical routing approaches for sensors networks [1].

The aim of LEACH protocol is to minimize energy consumption or in other words, to maximize the network lifetime. To make this happen several ideas are proposed for CH selection but they were based on mainly the node's (to be selected as CH) energy level. The node having greater energy level will be selected as CH most of the times. But here in the new proposed scheme not only the node's energy level is considered but also it's location or position both within the CH & from outside the cluster(neighbour clusters) are considered.

Proposed Scheme

We know that there may a number of nodes in a cluster & there is always a CH. Suppose for example, if the CH lies at a distant position from the majority of nodes. So to communicate between CH & sensor nodes, since the distance between them is high, energy consumption for the communication is also high. That means, the higher the distance between CH & sensor nodes the greater the energy consumption.

Here a new idea to select the CH is given below :

1. Select the CH in the dense node zone.

To illustrate this say for example, you are announcing something. If the persons, for whom your announcement is, are very far from you, you have to shout more to make them listen to it but if those persons are near to you, you won't have to shout that much.

That means, if nodes are near to the CH, energy consumption is less.

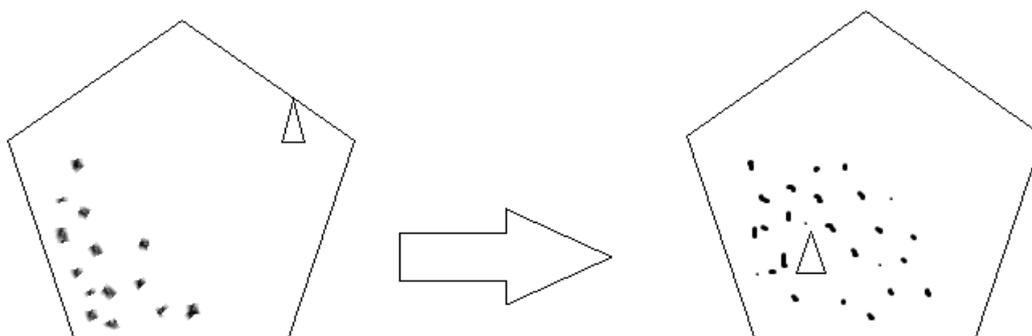


Fig. 1

2. Suppose a cluster is surrounded by 6 clusters. So 6 CH can communicate with the central CH. This central CH should be at an optimum distance from those CH. That means the distance between them should be balanced or on average.

Say, C0,C1,C2,C3,C4,C5,C6 are the CH of cluster 0(central cluster),cluster 1, cluster 2, cluster 3, cluster 4, cluster 5, cluster 6 respectively. There should not be a huge difference among distances between C0-C1,C0-C2,C0-C3,C0-C4,C0-C5,C0-C6.

Hence, energy consumption will be in control.

- New Proposed Mathematical formula:

$$\frac{P}{1 - P \times (r \bmod P^{-1})} \frac{E_{n_current}}{E_{n_max}} \frac{D_{avg}}{\sum D_{inter_node}} \frac{D_{ch_avg}}{D_{centre_avg}}$$

Where:

$E_{n_current}$ is the current amount of energy

E_{n_max} is the initial amount of energy

D_{avg} is the average distance from all other nodes in the cluster

E_{inter_node} is the distance between any two nodes in the cluster

D_{ch_avg} is the average distance from the node to the neighbouring cluster heads

D_{centre_avg} is the average distance of the neighbouring cluster heads from the centre of the cluster

Illustration:

Here with the original formula 2 factors are multiplied.

- I. [Average distance from other nodes in same cluster/ \sum inter-node distance]

This factor checks whether the node, to be selected as CH, belongs to a density popular area as well as the distance from the node to the other nodes within the cluster is on average.

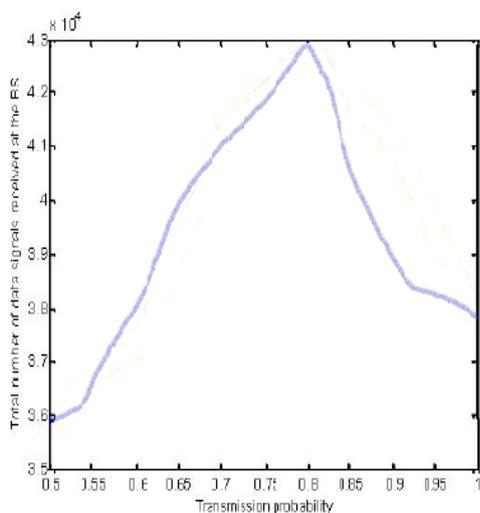
- II. [Average distance from neighbouring clusterheads / average distance from the centre of the cluster to the neighbouring clusterheads]

This factor measures the distance from the node, to be selected as CH, to its neighbouring CHs as well as the distance from the centre point of the cluster to the neighbouring CHs.

If we follow the new proposed formula, then the clusterhead will be selected in such a way that will be very energy efficient. That means if we select the clusterhead by following the new proposed formula instead of the original formula of LEACH then the fact energy consumption to select the clusterhead will be emphasised this is a very important factor.

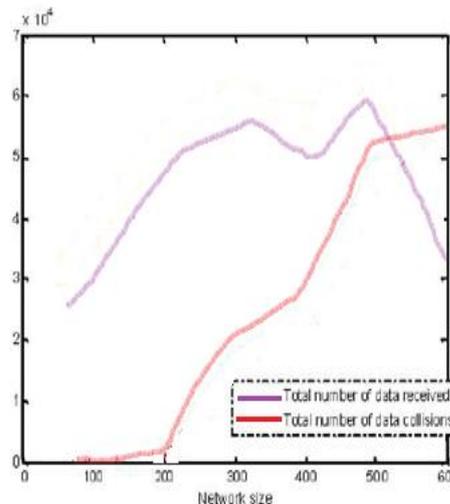
Results :

Fig. 2



Total amount of data received at the BS at the end of the network lifetime over transmission probability of nodes sending data. Network size is 100 nodes.

Fig. 3



Total number of data signals received at the BS and the total number of data collisions over network size.

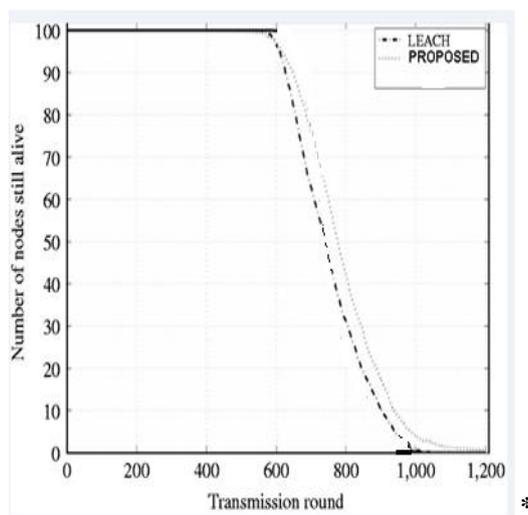


Fig. 4

Comparative Study:

According to **Stochastic Threshold Algorithm** :

- Cluster-heads can be chosen *stochastically* (randomly based) on this algorithm:

$$T(n) = \frac{P}{1 - P \times (r \bmod P^{-1})} \quad \forall n \in G$$

$$T(n) = 0 \quad \forall n \in G$$

Where n is a random number between 0 and 1
 P is the cluster-head probability and
 G is the set of nodes that weren't cluster-heads the previous rounds

- If $n < T(n)$, then that node becomes a cluster-head
- The algorithm is designed so that each node becomes a cluster-head at least once

According to **Deterministic Threshold Algorithm**:

- A modified version of this protocol is known as LEACH-C (or LEACH Centralized)
- This version has a *deterministic* threshold algorithm, which takes into account the amount of energy in the node...

$$T(n)_{new} = \frac{P}{1 - P \times (r \bmod P^{-1})} \frac{E_{n_current}}{E_{n_max}}$$

Where $E_{n_current}$ is the current amount of energy and
 E_{n_max} is the initial amount of energy

- ...and/or whether or not the node was recently a cluster-head

$$T(n)_{new} = \frac{P}{1 - P \times (r \bmod P^{-1})} \left[\frac{E_{n_current}}{E_{n_max}} + (r_s \div P^{-1}) \left(1 - \frac{E_{n_current}}{E_{n_max}} \right) \right]$$

Since the goal is to maximize the lifetime of the network or to minimize the energy consumption, according to the new proposed formula the lifetime of the network will be greater than the above mentioned formulae.

Comparison Between SPIN, LEACH, Directed Diffusion and the Enhanced LEACH :

	SPIN	LEACH	Directed Diffusion	Enhanced LEACH
Optimal Route	No	No	Yes	No
Network Lifetime	Good	Very Good	Good	Better than LEACH
Resource Awareness	Yes	Yes	Yes	Yes
Use of Meta-data	Yes	No	Yes	No

Advantages

- Optimises the distance between head nodes and other nodes reducing energy consumption and chances of loss of signal strength.
- Optimises the distance between the inter cluster head nodes, thus optimising the communication between head nodes and central server.
- The new approach selects the optimised node as head node which has the minimal cost in terms of energy while communicating with other nodes, thus increasing the lifetime of the network

Conclusion

The LEACH routing protocol was designed to increase network lifetime. Since nodes are active until energy becomes finished in nodes battery. LEACH protocol takes in concern only the current energy level and calculates how much possibility a node has, to be selected as clusterhead. The Enhanced clusterhead selection algorithm using LEACH not only concerns about the current energy level but also takes into account the position or location of that node to be selected as clusterhead. If the node, to be selected as CH is near rest nodes within or outside (neighbour clusters) a cluster, then it will consume less energy for communication. So the new proposed scheme calculates the distance of the node from other nodes within the cluster as well as the distance of the node from the neighbour clusterheads and the distance from the centre position of the cluster to the neighbour clusterheads. These factors are there in the new proposed scheme. Hence the proposed one is improved compared to the previous LEACH algorithm in terms of energy conservation.

If we analyse the new mathematical formula for increasing network lifetime, we will find enhanced results with the new. The new proposed scheme obviously has future scope for betterment of increasing network lifetime. There will be more advancement in placing CH over the cluster to minimize energy conservation. The two new factors need further studies and practical implementation to understand their exact importance and efficiency. After further study and research they can be raised to the power n and m respectively where n and m can be any rational number.

Acknowledgements

We express our sincere gratitude to Prof.(Dr.) Tapan Chakraborty, Head, Department of Information Technology, Heritage Institute of Technology, for promoting research activities. And also special thanks to Prof.(Dr.) Somenath Mitra, Shri Souren Pathak and Shri Indranath Mitra for constant support and encouragement.

References

- 1) W. Heinzelman, A. Chandrakasan and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Multi-sensor Networks," Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS '00), January 2000.
- 2) Jamal N. Al-Karaki Ahmed E. Kamal – "Routing Techniques in Wireless Sensor Networks: A Survey" Dept. of Electrical and Computer Engineering Iowa State University, Ames, Iowa 50011.
- 3) S. Tilak et al., "A Taxonomy of Wireless Microsensor Network Models," in *ACM Mobile Computing and Communications Review (MC2R)*, June 2002.
- 4) ZHANG Yu-quan, WEI Lei – "IMPROVING THE LEACH PROTOCOL FOR WIRELESS SENSOR NETWORKS"; School of Information and Electronics, Beijing Institute of Technology, Beijing 100081.
- 5) A. Buczak and V. Jamalabad, "Self-organization of a Heterogeneous Sensor Network by Genetic Algorithms," *Intelligent Engineering Systems Through Artificial Neural Networks*, C.H. Dagli, et.(eds.), Vol. 8, pp. 259-264, ASME Press, New York, 1998.
- 6) C.R. Lin and M. Gerla, "Adaptive Clustering for Mobile Wireless Networks," *IEEE Journal on Selected areas in Communications*, Vol. 15, No. 7, September 1997.
- 7) Kemal Akkaya and Mohamed Younis – "A Survey on Routing Protocols for Wireless Sensor Networks"; Department of Computer Science and Electrical Engineering University of Maryland, Baltimore County, Baltimore, MD 21250.

Implementing VGA Application on FPGA using an Innovative Algorithm with the help of NIOS-II

Ashish B. Pasaya¹

¹E & C Engg. Department, Sardar Vallabhbhai Patel institute of technology, Vasad, Gujarat, India.

Kiritkumar R. Bhatt²

²Associate Professor, E & C Engg. Dept.,
Sardar Vallabhbhai Patel Institute of Technology, Vasad-388306.

Abstract

Basically, here we have used VGA for implementing basic graphics applications that can be either used in a single user game or either in advertisements that deals with real-time application. Further expanding the logic with the coding part even double user game could be developed. So, we thought of using VGA as a standard for this implementation as it is the basic graphics array and compatible with other graphical arrays. Here, we used HDL language on Quartus-II software for interfacing the required peripheral to the NIOS-II Soft-core processor through FPGA Cyclone-II Processor. Where, we made use of the Innovative Algorithm for implementing the application of VGA with the help of C language on NIOS-II Soft-core processor that will contain the logic part. Finally, the results that we obtained for VGA application implementation.

Keywords: VGA, FPGA, HDL, Quartus-II, NIOS-II, DE2 Education Board.

I. Introduction

In this project the most required things will be VGA, DE2 Education Board on which the FPGA and NIOS-II soft-core processor is embedded along with other components. Here, VGA stands for "Video Graphics Array" (en.wikipedia.org/wiki/Video_Graphics_Array). It is the standard monitor or display interface used in most PCs. Therefore, if a monitor is VGA-compatible, it should work with most new computers. The VGA standard was originally developed by IBM in 1987 and allowed for a display resolution of 640x480 pixels. The VGA supports both All Points Addressable graphics modes, and alphanumeric text modes. There are two kinds of VGA interface signals to display. One is data signal, and the other one is control signal. There are three data signals red, green and blue and two control signals horizontal synchronization and vertical synchronization signals [1]. There are different frequencies of the horizontal synchronization signal and vertical synchronization signal for the changeable output resolution [1]. Thus, if someone wants to implement any application on any higher graphics arrays, the try could be given on the VGA first. So, we did the same thing by implementing the application on the VGA.

Now, if we talk of FPGA then we used Cyclone-II FPGA of family EP2C35F672C6. For this we used Quartus-II software in which you can specify your design by three ways i.e. schematic entry, Verilog HDL and VHDL. In our case we have used the Verilog HDL to specify the design. It is possible to make use of even both the HDL languages together. Also, it is possible to implement the required application by just using the HDL languages but then this will make your code very lengthy and will also increase the complexity. Due to which it will reduce the easiness to understand the code and it will also put burden of the memory associated with the FPGA. Now, we know that the memories associated with the FPGA will have sufficient memory space for small information's like limited number of characters and some image requiring very small memory to be displayed on screen, but would be insufficient to display an animated characters or images. For this kind of situation we will require some additional processor that can reduce the length of code and other information while developing some real-time applications. So, we do have a processor named NIOS-II soft-core processor which allows the input in both the C and C++ languages. This processor is embedded on the DE2 Board itself.

Thus, what we need to do is just to interface the peripheral required for our application with the NIOS-II soft-core processor. For that here we use the HDL language to interface the NIOS-II with required peripheral through FPGA. We had used C as an input language on the NIOS-II soft-core processor for implementing the innovative algorithm that we developed for our required application. Here, our required application is a single user game. Where there is a mouse as an input for single user and VGA screen as an output. On screen we will have one paddle which in our case is a rectangular block and one small ball like structure which in our case is small square block. Paddle will move according to the movement of the mouse that is connected to the USB blaster of the DE2 Board and the ball will do the movement depending on whether there has been a click on a mouse or not.

This paper is organized as follows: Section 2 describes the innovative algorithm used for implementing VGA application. Section 3 shows the VGA application implementation results. Section 4 concludes the paper.

II. Innovative Algorithm for VGA application

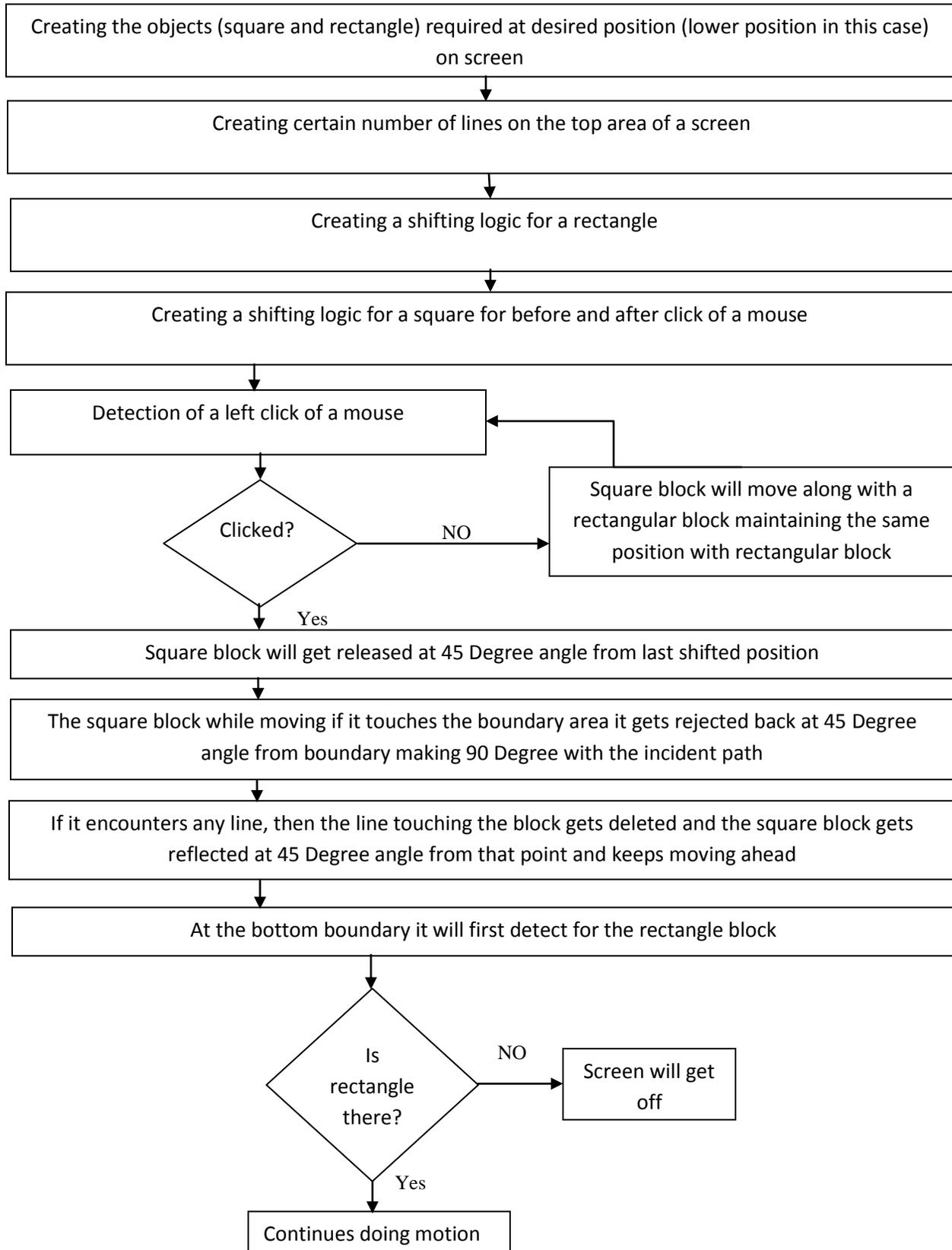


Figure 1. Innovative Algorithm for VGA Application

Here, first of all we will create the objects that are required in our project to be displayed on the screen i.e. a small square block, a rectangular block and few lines. Where we will have these lines at the top most part of the screen and these rectangle block will do movement in horizontal plane on the bottom part of the screen and will remain at the fixed at vertical position of the screen, while the small square in this project will have both the horizontal movement as well as other random movement depending on the situation.

Now, we will create a shifting logic in which the rectangle which shift in horizontal direction on screen being fixed in vertical direction. This shift will be according the movement of the mouse, where the movement of the mouse decides whether to get shifted in on left or right. That is in wherever direction the mouse moves the rectangle block on the screen also moves, just because the rectangle block is in synchronization with the position of the mouse. In other sense we can say that the rectangular block represents the cursor of the mouse but not exactly. The movement of the rectangle block in the screen would be only within the boundaries set by the programmer on the screen. For example, you can take the boundary as 620×460.

Now creating a shifting logic for a small square block. This task will be more complex than that of creating a sifting logic for a rectangular block. Here there will be two kind of motion of a square block depending on the click of the mouse. Here the value will be assigned to one variable which will detect whether the click has occurred or not. Like its value will be set if the click has already occurred and unset if not occurred. This part will be the detection part of whether the click has occurred or not. Now during detection if no click has occurred then it will just keep shifting in horizontal direction along with the rectangular block, where the position of a square block will remain constant with respect to the position of the rectangular block. These detection will occur continuously. If the click occurs during detection then the square block will now get released at 45 degree angle from the last moved position and will start to get incident on the boundary wall set by the programmer. Now this angle of incident and angle of reflection at the wall will make 90 degrees with each other. After getting reflected the square block will keep moving ahead and will touch the bottommost line at the top of the screen. When the square block touches the line it gets reflected from the like at 45 degrees again making 90 degree angle with the incident path to the line and the line will get deleted/erased from that position. The square block will keep moving ahead and now will touch the another boundary wall and get reflected back at 45 degree angle making 90 degree with the incident path.

Now it keeps doing motion ahead in the bottom direction where the rectangle block is already doing motion as per the movement of the mouse in horizontal direction. The rectangular block must catch the square block on its top otherwise if it misses the square block then the screen will get off. There will be detection by square block logic for the presence of the rectangular block on the bottom surface of the screen. If the rectangular block catches the square block on the top then the square block will get reflected from the top surface of the rectangular block at 45 degree angle making 90 degree angle with the incident path and keeps moving ahead towards wall and will follow the same scenario as ahead it did and will erase each lines when it gets touched to line and will then incident to the wall after each line gets erased by the touch. Here when the square block is shifting then the one pixel from back position will get and one pixel in front of the rectangular block will be displayed which shows that the rectangular block is doing motion if speed is increased for the shift.

The rectangular block will move according to positive X and negative X direction. While the square block move according to both the positive and negative direction of X and Y. so, at whatever wall it gets incident to the wall or any other object it will get reflected in its negative direction.

Here the square block will become independent of the movement of the mouse when the click occurs, so there will be no dependence of the square block on the rectangular block and the mouse as well. The speed of the movement of the rectangular block and square block can be set between the minimum 50 MHz to maximum 400 MHz in our project.

III. Results of VGA application implementation



FIGURE 2. Motion of a square and a rectangular block



FIGURE 3. Release of a square block at 45 degree angle

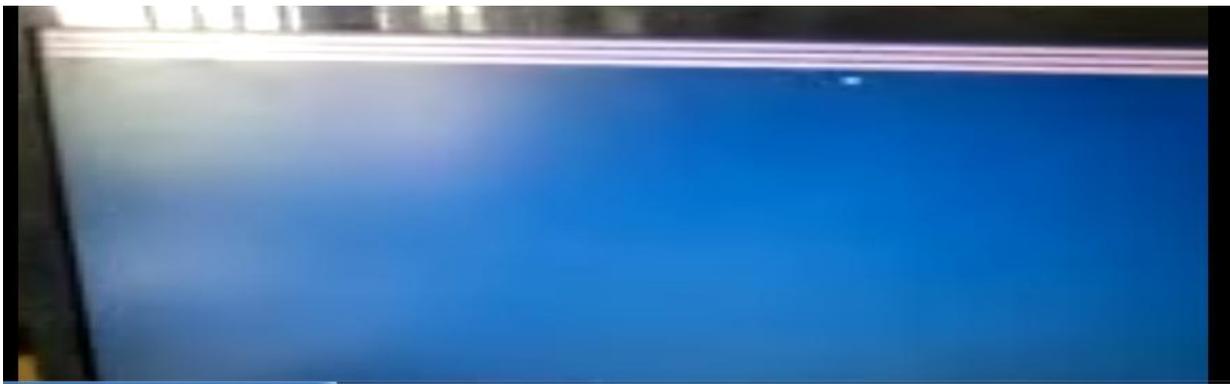


FIGURE 4. Motion of a square block towards the lines after getting reflected from the side wall of a screen



FIGURE 5. Erasing of a line with which the square block got incident

IV. Conclusion

In this paper we presented an innovative algorithm for implementing basic application of VGA for a single user game and the output that we attained. It further can be extended into an advertisement with real time application or a double user game with respective addition of logic through NIOS-II processor. We interfaced NIOS-II with required peripheral through FPGA using HDL languages and used C as a language for implementing the innovative algorithm. So here NIOS-II adds more flexibility along their use with FPGA and reduces the complexity in coding part especially, which makes it suitable to meet both our hardware and software real-time requirements.

Acknowledgement

Words are often less to reveal one's deep regards. An understanding of work like this is never the outcome of the efforts of a single person. First off all i would like to thank the Supreme Power, one who has always guided me to work on the right path of the life. Without his grace this would never come to be today's reality. This work not have been possible without the encouragement and able guidance of my supervisor Prof. K.R.Bhatt, his enthusiasm and optimism made this experience both rewarding and enjoyable. Also my sincere thanks to Mihir Sir and Nirav Sir (smartech solutions), Vadodara for providing me tools and guidance required in my work.

References

- [1] Guohui Wang and Yong Guan, "Designing of VGA Character String Display Module Base on FPGA", IEEE 2009.
- [2] Cyclone II Device Handbook (PDF) – Altera.
- [3] DE2 Development and Education Board User Manual.
- [4] NIOS II Software Developer's Handbook (PDF) – Altera.
- [5] Fawcett, B., "FPGAs as Reconfigurable Processing Elements", Circuits and Device Magazine, March 1996, pp. 8-10.
- [6] Van-Huan Tran and Xuan-Tu Tran, "An Efficient Architecture Design for VGA Monitor Controller", IEEE 2011.
- [7] Babu T Chacko and Siddharth Shelly "Real-Time Video Filtering and Overlay Character Generation on FPGA", IEEE 2011.
- [8] A. D. Ioan, "New Techniques for Implementation of Hardware algorithms inside FPGA Circuits", Advance in Electrical and Computer Engineering, Vol. 10, No. 2, Suceava, Romania, 2010.

VIBRATION ANALYSIS OF GLASS FIBER REINFORCED COMPOSITES

R. K. Mishra

Department of Mechanical Engineering, Gautam Buddha University, Gautam Buddha nagar, Greater Noida, India

Abstract:

Glass fiber reinforced resol/vac-eha composites have been fabricated in laboratory to determine the dynamic behavior of glass fiber reinforced composites. Resol solution was blended with vinyl acetate-2-ethylhexyl acrylate (vac-eha) resin in an aqueous medium with varying volume fraction of glass fibers. The role of fiber/matrix interactions in glass fibers reinforced composites were investigated to predict the stiffness and damping properties. In order to study the static and dynamic response of Resol, Resol/VAC-EHA blend and glass fibers reinforced composites, a multiquadric radial basis function (MQRBF) method is developed. MQRBF is applied for spatial discretization and a Newmark implicit scheme is used for temporal discretization. The discretization of the differential equations generates a larger number of algebraic equations than the unknown coefficients. To overcome this ill conditioning, the multiple linear regression analysis, which is based on the least square error norm, is employed to obtain the coefficients. Simple supported and clamped boundary conditions are considered. Numerical results are compared with those obtained by other analytical methods.

Notations : a, b -Dimension of plates, h ---Thickness of plates, R ---Aspect ratio (a/b) , ν_{LT} --- Major Poisson's ratio, C_v^* , C_v ---Viscous damping, dimensionless viscous damping, D_{11} , D_{22} , D_{12} , D_{66} ---Flexural rigidity of plates, E_L , E_T ---Young's modulus in x^* , y^* direction respectively, G_{LT} ---Shear modulus relative to $x^* - y^*$ direction, q, Q ---Transverse load, dimensionless transverse load, t^* , t ---Time, dimensionless time, w^* ---Displacement in z^* direction, w ---Dimensionless displacement in z direction

Key words: *Glass fiber, Resol, vinyl acetate-2-ethylhexyl acrylate, multiquadric radial basis function glass fiber-reinforced composites.*

Introduction:

Fiber-reinforced composites based upon thermoplastic polymer matrices and thermoplastic-thermoset interpenetrating network (IPN) polymer matrices potentially offer several advantages compared to those based upon thermosetting resin matrices. The early 1980s thermosetting matrix composites were used in significant positions of the primary structure of military aircraft and secondary structure of commercial aircraft. Composites are used because of their weight savings compared with metals, and hence reduced fuel consumption, longer-range, higher performance or greater payload. It was recognized that thermoset composites were sensitive to damage, and in particular to low energy and low velocity impacts, which occur in routine service. Thermoset Phenolic resin is one of the earliest synthetic resins to possess excellent fire retardancy, low smoke density and toxic emission [1-3], though its application is limited because unmodified phenolic is a brittle material. A lot of research work has been conducted to toughen phenolic resin. The polyester-phenolic copolymer has been synthesized to improve the mechanical properties and heat resistance [4] the flexural strength improves when phenoxy resin is blended with resol-type phenolic resin [5]. Various types of fibers are used for reinforcement; these are glass fibers, carbon fibers and polyethylene fibers etc. but glass fiber-reinforced composites are used extensively in various applications. Glass fibers have high mechanical and bending properties. Mechanical properties of the fiber reinforced composites are different from mechanical properties of the fibers [6-7]. It depends upon various factors. These are fiber orientation, weight and volume fraction of the fibers and matrix, bonding between fibers and matrix. To predict the static and dynamic behavior of the composite is very important when it is used in structural applications. The static and dynamic behavior of the unidirectional banana fibers reinforced high-density polyethylene/ poly (C-caprolactone) composites have been presented by Misra et al [8] Normally finite element method is used to predict the behavior of composites. But in finite element method generation of mesh is tedious and time-consuming task. To avoid this problem meshless methods have been developed. In various meshless methods importance of the multiquadric radial basis function method (MQRBF) is increasing day by day. In 1971, it was proposed by Hardy [9] for the interpolation of geographically scattered data. Franke [10] has ranked MQRBF is the best interpolation method based on its accuracy, visual aspect, sensitivity to parameters, execution time and ease of implementation. Later on this method was used by Kansa [11] to solve partial differential equations. Ferreira applied the radial basis function for the analysis of composite beams [12] and plates [13]. Misra et al [14-15] used this method for the analysis of anisotropic plates and laminate. The properties of the resol improve by blending vinyl acetate-2-ethylhexyl acrylate copolymer (VAC-EHA) by IPN technique. This paper reports fabrication and evaluation of the mechanical properties of glass fibers reinforced Resol/VAC-EHA composites at different volume fraction of the fibers. After obtaining the properties experimentally, MQRBF method is employed for static and dynamic analysis of unidirectional glass fibers reinforced Resol/VAC-EHA composite plate.

Experimental Studies:

• **Materials**

- ❖ Resol was prepared by the method cited in the literature [16-17].
- ❖ The hardener for resol was para-toluene sulphonic acid (PTSA), 0982 H from Bakelite AG.
- ❖ VAC-EHA copolymer in emulsion form was obtained from macromoles, India.
- ❖ Hydroxy methoxy methyl melamine (HMMM) was prepared in the laboratory using standard procedure [18].
- ❖ Glass Fiber (433 BF-225) supplied by Indian Petrochemicals Corporation Ltd., Boroda, India.

Preparation of Composites:

The individual polymers were first separately diluted with distilled water to maintain a solid content of 50 % by weight for convenience, under well-stirred condition. Then a weighed amount of resol was taken in a three-necked round bottom flask. VAc-EHA copolymer was then accurately weighed into the flask and the contents were stirred to give a homogeneous mixture in the desired blend ratio of the components. Para-toluene sulfonic acid (PTSA) 7 % by weight based on resol was thoroughly mixed for 20 minutes [19]. Then hydroxy methoxy methyl melamine (HMMM) was added at 20 % by weight of VAc-EHA copolymer. The prepreg of glass fiber was prepared previously. For the use of treated fiber in making composites, the fibers were treated with chromic acid at room temperature for 15 min. When the formation of the bubbles ceased, the viscous mass was poured into a glass mold prepared by clipping together two glass plates separated by a Teflon gasket in between, the thickness of which controls the thickness of the sample sheet formed. It was then initially kept at room temperature for about 24 hours and then heated at 80⁰ C for 4 hours [20]. Thus, the samples were produced.

• **Tensile Properties:**

An Instron universal – testing machine (Model 3360) was used for measuring the mechanical properties like tensile strength and tensile modulus. ASTM D638 method was followed. A crosshead speed of 5 mm/min was maintained. All testing were conducted under ambient conditions in an environmentally controlled room. The data reported are averages of at least six measurements, and typical scattering range of the results was ±5 %.

Theoretical Analysis:

After evaluating the mechanical properties of the unidirectional glass fibers reinforced Resol/VAC-EHA composites at different volume fraction of the fibers, Multiquadric radial basis function method is applied for static and dynamic analyses of the unidirectional glass fibers reinforced Resol/VAC-EHA composite plate to assess its response to external loading such as uniformly distributed load.

• **The Multiquadric Radial Basis Function Method**

Consider a general differential equation

$$Aw = f(x, y) \quad \text{in } \Omega \quad (1)$$

$$Bw = g(x, y) \quad \text{on } \partial\Omega \quad (2)$$

where *A* is a linear differential operator and *B* is a linear boundary operator imposed on boundary conditions for the orthotropic plate. In the present research work, it is assumed that the glass fiber reinforced composite plate behaves as an orthotropic plate since the diameter of glass fiber is so small that its contribution to the strength is predominantly along its longitudinal direction only.

Let $\{P_i = (x_i, y_i)\}_{i=1}^N$ be *N* collocation points in domain Ω of which $\{(x_i, y_i)\}_{i=1}^{N_i}$ are interior points; $\{(x_i, y_i)\}_{i=N_i+1}^N$ are boundary points. In MQRBF method, the approximate solution for differential equation (1) and boundary conditions (2) can be expressed as:

$$w(x, y) = \sum_{j=1}^N w_j \phi_j(x, y) \quad (3)$$

and multiquadric radial basis as:

$$\phi_j = \sqrt{(x - x_j)^2 + (y - y_j)^2 + c^2} = \sqrt{r_j^2 + c^2} \quad (4)$$

where $\{w_j\}_{j=1}^N$ are the unknown coefficients to be determined, and $\phi_j(x_j, y_j)$ is a basis function. Other widely used radial basis functions are:

- $\phi(r) = (c^2 + r^2)^{1/2}$ Multiquadrics
- $\phi(r) = (c^2 + r^2)^{-1/2}$ Inverse multiquadrics
- $\phi(r) = r^3$ Cubic

$\varphi(r) = r^2 \log(r)$	Thin plate splines
$\varphi(r) = (1 - r)^m + p(r)$	Wendland functions
$\varphi(r) = e^{-(cr)^2}$	Gaussian

Here c is a shape parameter, a positive constant. Ling and Kansa [21] discussed about the shape parameter.

Results and Discussion:

Figure 1 shows the theoretical and experimental Young’s modulus of the glass fiber reinforced composites in longitudinal direction at different volume fraction of the fibers. Young’s modulus (E_L) appears to be linearly dependent on volume fraction of the fibers (V_f). As soon as volume fraction of the fiber increases theoretical and experimental Young’s modulus increases, but theoretical Young’s modulus increases at higher rate as compared to experimental Young’s modulus.

The interesting feature of this study is that the distance between corresponding theoretical and experimental curves increases with V_f , and the deviation is minimum at lower range of V_f . In the present work, the hand lay-up technique has been used with a V_f range of 0.089-0.357. At a higher V_f , fiber interaction takes place, and either they tend to bundle up among themselves or they touch each other physically, which is due to the fact that hand lay-up technique produces more or less random nature of fiber distribution in the matrix. Due to the above facts, proper and uniform penetration of the matrix does not take place throughout the fiber surfaces, leaving interstitial voids. Thus, the degree of fiber misalignment and void content increases with the increase in V_f . These facts are reflected in the experiment by the deviation of the experimental curve from the theoretical one. Longitudinal and transverse Young’s modulus of the glass fiber reinforced composites at different volume fraction of the fibers has been shown in figure 2. When volume fraction of the fibers is 8.8 % difference between both Young’s modulus is less but as soon as volume fraction of the fiber increases difference between both Young’s modulus also increases. Figure 3 shows the stress- strain behavior of the glass fiber reinforced composites and the Resol/VAC-EHA matrix. The failure strain of glass fiber is about 2 times than the matrix (ϵ_m). The stress- strain curve display a point of inflection (Knee point) around ϵ_m , which enables the curve to be approximated by two regions – one indicating the elastic (below the point corresponding to ϵ_m), other the plastic region (beyond the point corresponding to ϵ_m). It is established that the elastic region of the composite is dependent on ϵ_m . The increase in fiber content leads to an increase in the first crack stress (corresponding to ϵ_m), and the ultimate strain. It is observed from the stress-strain curves that the main influence of the fibers is in the post-cracking zone, where the contribution of the matrix is small or even negligible.

The experimental and theoretical tensile strength of the glass fiber reinforced composites has been compared and shown in figure 4 at different volume fraction of the fibers. From the graph it is established that the elastic region of the composite is dependent on matrix. The increase in fiber content leads to an increase plastic deformation. The experimental tensile strength is less than theoretical tensile strength. Figure 5 shows the elongation at break of the glass fiber reinforced composites at different volume fraction of the fibers. It is observed that elongation shows linear relationship upto 17.6 % volume fraction of the fibers but becomes constant from 17.6 % to 25.4 % volume fraction of fibers and after 25.4 % volume fraction of fibers it increases slowly. Table 1 shows the tensile modulus of the Resol and Resol/VAC-EHA blends. Tensile modulus decreases due to increase blending of VAC-EHA copolymer content in Resol. The mechanical properties of the glass fiber reinforced Resol/VAC-EHA composites are studied to observe the effect of glass fiber by increasing its volume fraction has been shown in Table 2. As volume fraction of fiber increases from 8.8 % upto 35.3 %, longitudinal modulus of elasticity increases 53.75 %, 96.40 % and 142.42% respectively.

Case Study: Static and Dynamic Response of the Composite plate

Figure 6 shows the geometry, coordinate system and loading in glass fiber reinforced composites. Neglecting the transverse shear and rotary inertia, equation of glass fiber reinforced composites is expressed in non-dimensional form as:

$$(w_{xxxx} + 2R^2\eta w_{xxyy} + R^4\psi w_{yyyy}) + w_{tt} + c_v w_t - Q(x, y, t) = 0 \tag{5}$$

where the subscript denotes the partial derivative with respect to the suffix following. The non-dimensional quantities are defined by

$$w = w^* / h, x = x^* / a, y = y^* / b, R = a / b, t = t^* \sqrt{D_{11} / (\rho a^4 h)}$$

$$Q = qa^4 / (D_{11}h), C_v = (C_v^* / M) \sqrt{(\rho a^4 h) / D_{11}}, M = \rho abh$$

$$\eta = (D_{12} + 2D_{66}) / D_{11}, \psi = D_{22} / D_{11} \tag{6}$$

Boundary conditions for the plate simply supported at all four edges are:

$$x = 0,1 \quad w = 0 \tag{7a}$$

$$x = 0,1 \quad w_{xx} = 0 \tag{7b}$$

$$y = 0,1 \quad w = 0 \quad (7c)$$

$$y = 0,1 \quad w_{xx} = 0 \quad (7d)$$

Boundary conditions for the clamped edge plate are:

$$x = 0,1 \quad w = 0 \quad (8a)$$

$$x = 0,1 \quad w_x = 0 \quad (8b)$$

$$y = 0,1 \quad w = 0 \quad (8c)$$

$$y = 0,1 \quad w_y = 0 \quad (8d)$$

The governing equation (5) is solved using multiquadric radial basis function and boundary conditions, equations (7) and (8), for simply supported and clamped edge plates respectively, and has been presented in Appendix A. In this multiquadric radial basis function method number of algebraic equations creates more than the number of unknown coefficients $\{w_j\}_{j=1}^N$. Hence

this method creates ill conditioning. To overcome this ill-conditioning, multiple linear regressions analysis (Appendix B) based on least-square error norms is employed. Deflection of the pure Resol, Resol/VAC-EHA blends and glass fiber reinforced composite plate at various aspect ratios of the present method for simply supported boundary condition has been compared with those obtained by Timoshenko and Woinowsky-Krieger [22] and Whitney [23] as shown in Tables [3-11]. There is good agreement between the present results and reference solution. The following features are observed after analyzing the Tables 3-11 for simple supported boundary conditions:

- As aspect ratio increases deflection decreases due to decrease in load.
- Resol/VAC-EHA (F20) blends Plate gives maximum deflection and 35.3 % glass fiber reinforced composite plate gives least deflection. It means strength of the Resol/VAC-EHA (F20) blends plate is least and 35.3 % glass fiber reinforced composite plate is maximum.

• As percentage of the VAC-EHA copolymer increases deflection also increases and strength decreases. Similar behavior is observed at clamped edge boundary conditions as shown in Table 12 and 13, but clamped edge plate deflect less as compared to simple supported plate. A computer program based on the finite difference method (FDM) proposed by Chadrashkara [24] (Chadrashkara, 2001) is also developed. The damped response of the simple supported composite plates obtained by the present method and finite difference method has been compared and shown in Figure 7 at uniformly distributed load $q = 10 \text{ (N/m}^2\text{)}$, sides $a = b = 300 \text{ (mm)}$ and thickness $h = 10 \text{ (mm)}$ at non-dimensional viscous damping factor $C_v = 1.25$. There is good agreement in the results. The following features are observed after analyzing Figures 8-14 for the simple supported boundary condition:

- Resol/VAC-EHA blends plate takes more time to stabilize as compared to pure Resol and 8.8 % glass fiber reinforced composite plate.
- As soon as percentage of the glass fiber increases from 8.8 % upto 35.3 % damping property of the composite plate increases.
- At same volume fraction, aspect ratio (b/a) of the glass fiber reinforced composite plate increases stability time also increases.

After incorporation of glass fibers in composite materials, tensile properties, tensile modulus, stiffness and damping properties increase. Damping means capacity of the materials to dissipate energy. Here glass fibers dissipate more energy as compared to Resol/VAC-EHA matrix.

Similar behavior is observed at clamped edge boundary conditions as shown in Figure 15 , but clamped edge plate takes more time to stabilize as compared to simple supported plate because clamped edge plate dissipate less energy.

Conclusion:

Glass fiber reinforced Resol/VAC-EHA composites plate has been prepared in laboratory to evaluate the mechanical properties. Damping properties decrease after blending VAC-EHA copolymer content with Resol. Tremendous changes occur in mechanical properties when glass fibers are reinforced with Resol/VAC-EHA blends. Resol, Resol/VAC-EHA blends and glass fibers reinforced composite plate show different mechanical response at static and dynamic loading. As soon as quantity of the glass fibers increase in composite plate damping properties increase. Due to incorporation of glass fibers in matrix, tensile properties, stiffness and damping properties increase. It has been found that MQRBF method is most suitable to predict the mechanical behavior of composite plate at static and dynamic loading.

Appendix

A. Multiquadric radial basis function method for governing differential equation

Substitution of multiquadric radial basis function in equation (5) gives

$$\left(\sum_{j=1}^N w_j \frac{\partial^4}{\partial x^4} \varphi_j + 2R^2 \eta \sum_{j=1}^N w_j \frac{\partial^4}{\partial x^2 \partial y^2} \varphi_j + \sum_{j=1}^n R^4 \psi w_j \frac{\partial^4}{\partial y^4} \varphi_j\right) + \sum_{j=1}^N w_j \frac{\partial^2}{\partial t^2} \varphi_j + C_v w_j \frac{\partial}{\partial t} \varphi_j - Q = 0 \quad (9)$$

Substitution of multiquadric radial basis function in boundary conditions

For Simple supported edge

$$x = 0, a \quad \sum_{j=1}^N w_j \varphi_j = 0 \quad (10a)$$

$$y = 0, b \quad \sum_{j=1}^N w_j \varphi_j = 0 \quad (10b)$$

$$x = 0, a \quad \sum_{j=1}^N w_j \frac{\partial^2}{\partial x^2} \varphi_j = 0 \quad (10c)$$

$$y = 0, b \quad \sum_{j=1}^N w_j \frac{\partial^2}{\partial y^2} \varphi_j = 0 \quad (10d)$$

For clamped edge

$$x = 0, a \quad \sum_{j=1}^N w_j \varphi_j = 0 \quad (11a)$$

$$y = 0, b \quad \sum_{j=1}^N w_j \varphi_j = 0 \quad (11b)$$

$$x = 0, a \quad \sum_{j=1}^N w_j \frac{\partial}{\partial x} \varphi_j = 0 \quad (11c)$$

$$y = 0, b \quad \sum_{j=1}^N w_j \frac{\partial}{\partial y} \varphi_j = 0 \quad (11d)$$

B. Multiple regression analysis

$Aa = p$; Where $A = (l*k)$ coefficient matrix, $a = (k*1)$ vector and $p = (l*1)$ load vector

Approximating the solution by introducing the error vector e , we get $p = Aa + e$

where $e = (l*1)$ vector

To minimize the error norm, let us define a function S as

$$S(a) = e^T e = (p - Aa)^T (p - Aa)$$

The least-square norm must satisfy

$$(\partial S / \partial a)_a = -2A^T p + 2A^T Aa = 0$$

This can be expressed as

$$a = (A^T A)^{-1} A^T P \text{ or } a = B.P$$

The matrix B is evaluated once and stored for subsequent usages.

References:

1. Sperling, L.H. (1981), "Interpenetrating Polymeric Networks and Related Material", Plenum Press, New York, USA.
2. Frisch, H.L., Frisch, K.C. and Klempner, D. (1981), "Advances in Interpenetrating Polymer Networks", Pure and Applied Chemistry, 53, 1557-1566.
3. Zaks, Y., Lo, J., Raucher, D. Pearce, E.M. (1982), "Some structure-property relationships in polymer flammability: Studies of phenolic-derived polymers", J. Appl. Polym. Sci., 27, 913-930.
4. Gardziella, A. and Miller, F.O. (1990), "Environmental importance of phenolic resins and six functions in applications," Kunststoffe, 4, 510-514.
5. Freilich, M.A., Meiers, J.C., Duncan, J.P. and Goldberg, A.J. (2000), "Fiber-reinforced Composites, Quintessence Publishing", Hong Kong, China, 9 – 21.
6. Herrera-Franco, P.J. and Valadez-Gonzalez, A. (2005), "A study of the mechanical properties of short natural-fiber reinforced composites", Composite Part B, 36, 597–608.
7. Ellakwa, A.E., Shorthall, A.C., Shehata, M.K., Marquis, P.M. (2001), "Influence of veneering composite composition on the efficacy of fiber-reinforced restorations (FRR)", Oper. Dent, 26 (5), 467– 475.

8. Misra, R.K., Sandeep, K., Kumar, S., Misra, A. (2008), “Dynamic Analysis of Banana Fibers reinforced high-density Polyethylene/ Poly (ε-caprolactone) Composites”, J. Mech. Mater. Struct, 3(1), 107-126.
9. Hardy, R.L. (1971), “Multiquadric equations of topography and other irregular surfaces”, Geophys. Res., 176, 1905–1915.
10. Franke, R. (1982), “Scattered data interpolation: tests of some methods”, Math. Comput, 48, 181–200.
11. Kansa, E.J. (1990), “Multiquadrics—a scattered data approximation scheme with applications to computational fluid dynamics: II. Solutions to parabolic, hyperbolic and elliptic partial differential equations”, Comput. Math. Appl., 19(8–9), 147–161.
12. Ferreira, A.J.M. (2003), “Thick composite beam analysis using a global meshless approximation based on radial basis functions”, Mech. Adv. Mater. Struct, 10, 271–284.
13. Ferreira, A.J.M. (2003), “A formulation of the multiquadric radial basis function method for the analysis of laminated composite plates”, Comput. Struct, 59, 385–392.
14. Misra, R.K., Sandeep, K., Misra, A. (2007), “Analysis of anisotropic plate using multiquadric radial basis functions”, Int. J. Eng. Anal. Bound. Element, 31, 28-34.
15. Misra, R.K., Sandeep, K., Misra, A. (2007), “Analysis of Cross-ply Laminate using Multiquadric Radial Basis Function”, Int. J. Comput. Meth. Eng. Sci. Mech., 8, 1–10.
16. Steiner, P. R. (1975), “Phenol-formaldehyde wood adhesive characterization by proton magnetic resonance spectroscopy”, J. Appl. Polym. Sci, 19(1), 215-225.
17. Rudin, A., Fyfe, C.A., Vines, S.M. (1983), “Gel permeation chromatographic analyses of resole phenolic resins”, J. Appl. Polym. Sci., 28(8), 2611-2622.
18. Melter, Y.L. (1981), “Water Soluble Polymers Development”, since 1978, Noyes Data Corporation, Park ridge, NJ.
19. Wolfrum, J., Ehrenstein, G.W. (1999), “Interdependence between the curing, structure and the mechanical properties of phenolic resins”, J. Appl. Polymer Science, 74(13), 3173-3185.
20. Datta, C., Basu, D., Banerjee, A. (2002), “Studies and Characterization of RESOL/Vac-EHA/HMMM IPN systems in aqueous medium”, J. Appl. Polym. Sci., 86(14), 3581-3588.
21. Ling, L., Kansa, E.J. (2004), “Preconditioning for radial basis functions with domain decomposition methods”, Math. Comput. Modell, 40, 1413-1427.
22. Timoshenko, S.P., Woinowsky-Krieger, S. (1989), “Theory of plates and shells”, II Edition, McGraw-Hill, Kogakusha Ltd, Tokyo.
23. Whitney, J.M. (1987), “Structural analysis of laminated anisotropic plates”, Technomic Publishing Company Inc, Lancaster, Pennsylvania, USA.
24. Chandrashekhara, K. (2001), “Theory of plates”, Universities Press (India) Limited: Hyderabad, India.

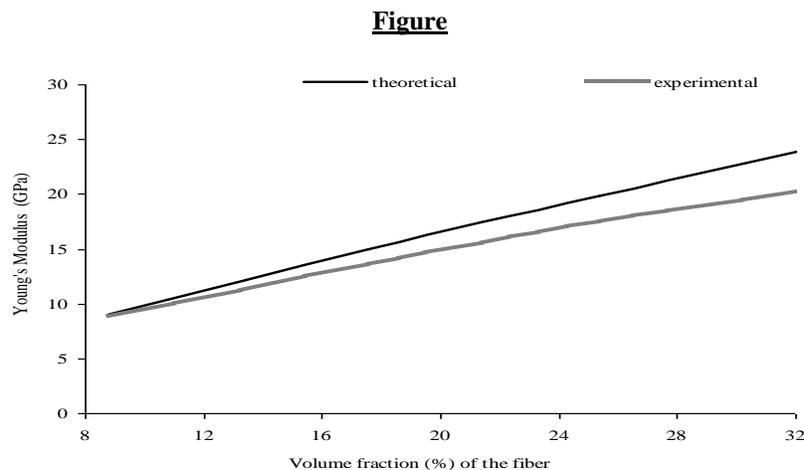


Figure 1: Theoretical and experimental Young’s modulus of the glass fiber reinforced composites in longitudinal direction at different volume fraction of the fibers.

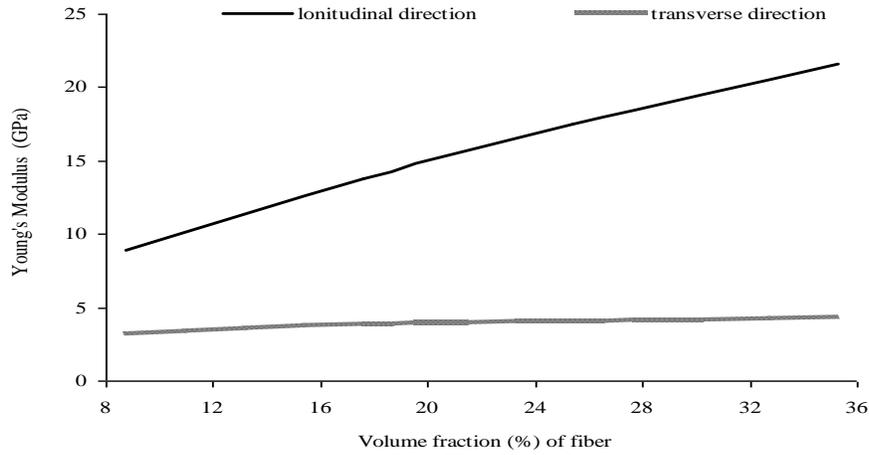


Figure 2: Young's modulus of the glass fiber reinforced composites in longitudinal and transverse direction at different volume fraction of the fibers.

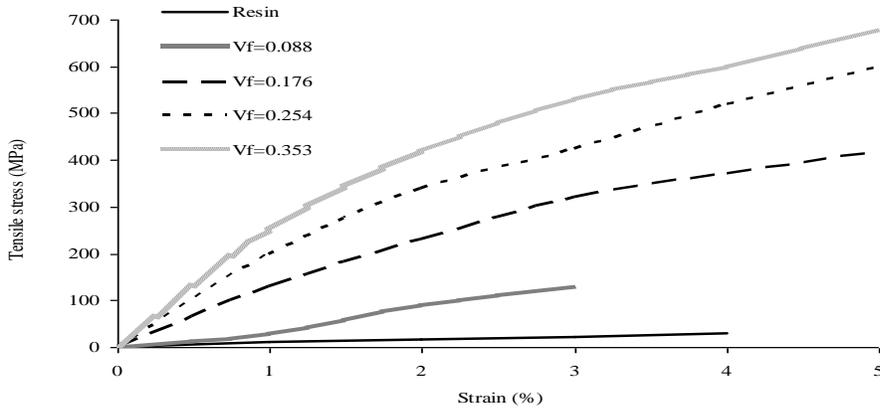


Figure 3: Stress- strain behavior of the glass fiber reinforced composites and the Resol/VAC-EHA matrix.

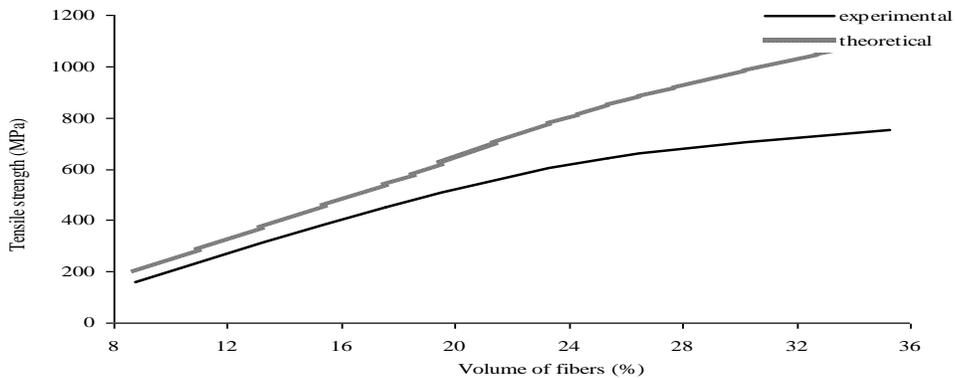


Figure 4: Experimental and theoretical tensile strength of the glass fiber reinforced composites at different volume fraction of the fibers.

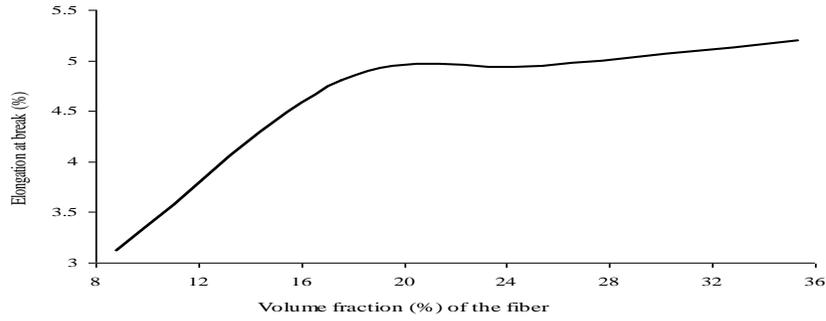
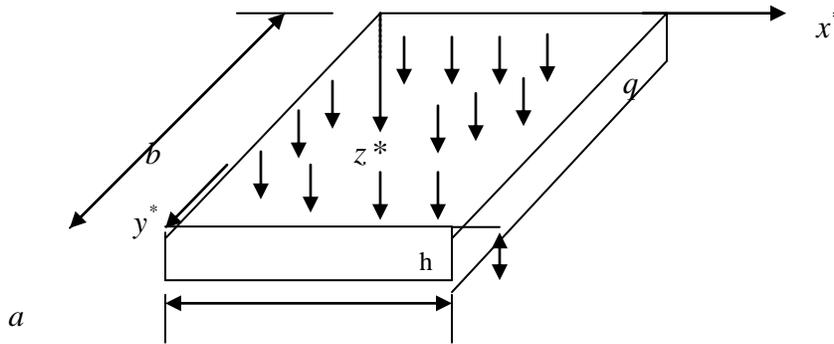


Figure 5: Elongation at break of the glass fiber reinforced composites at different volume fraction of the fibers.



Figures 6: Geometry of the glass fiber reinforced Resol/VAC-EHA blend Composite plate

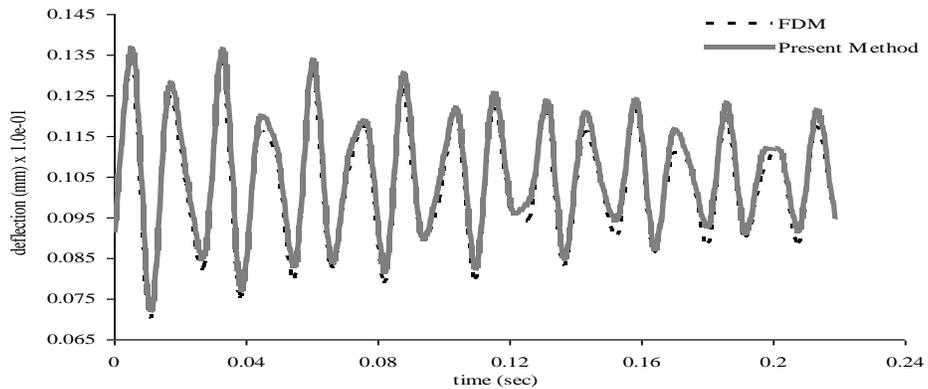


Figure 7: Damped response of a simple supported pure resol square isotropic plate at damping coefficient factor $C_v = 1.25$.

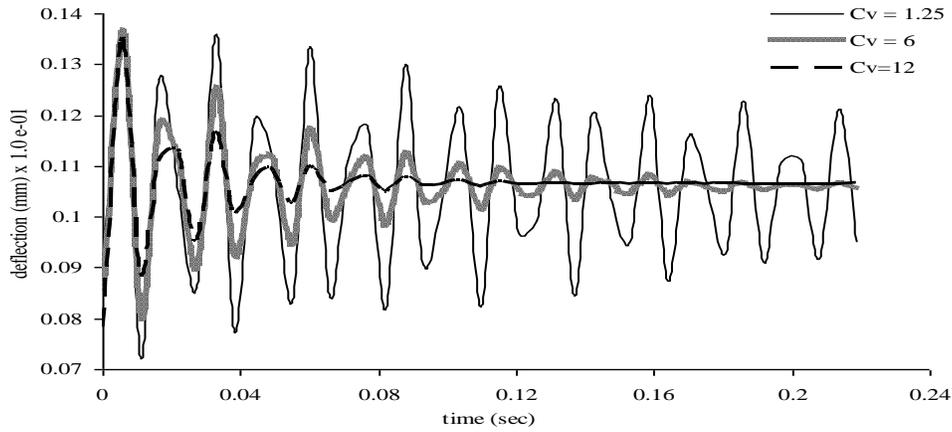


Figure 8: Damped response of a simple supported pure resol square isotropic plate at various damping coefficient factor.

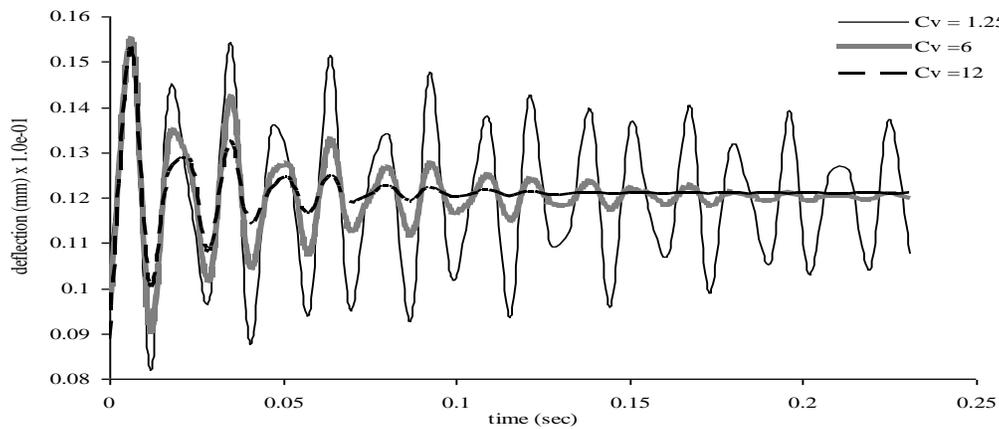


Figure 9: Damped response of a simple supported Resol/VAC-EHA (F20) blends plate at various damping coefficient factor.

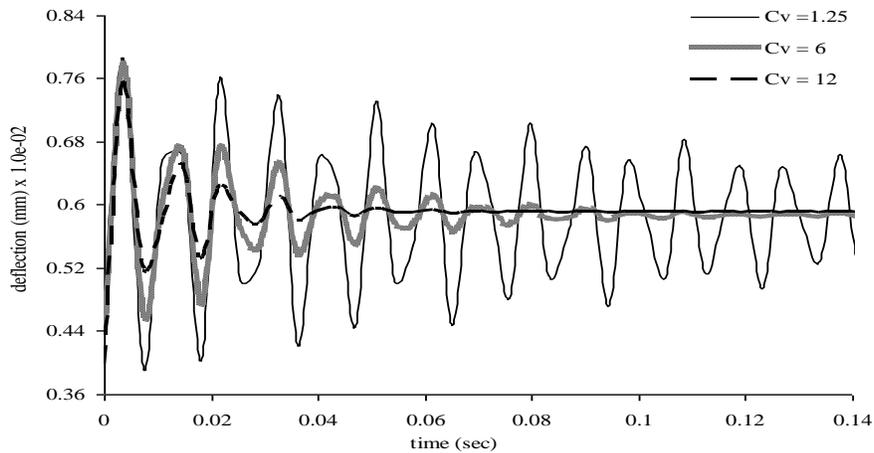


Figure 10: Damped response of a simple supported 8.8% glass fiber reinforced (F20GF8.8) Resol/VAC-EHA blends square composite plate.

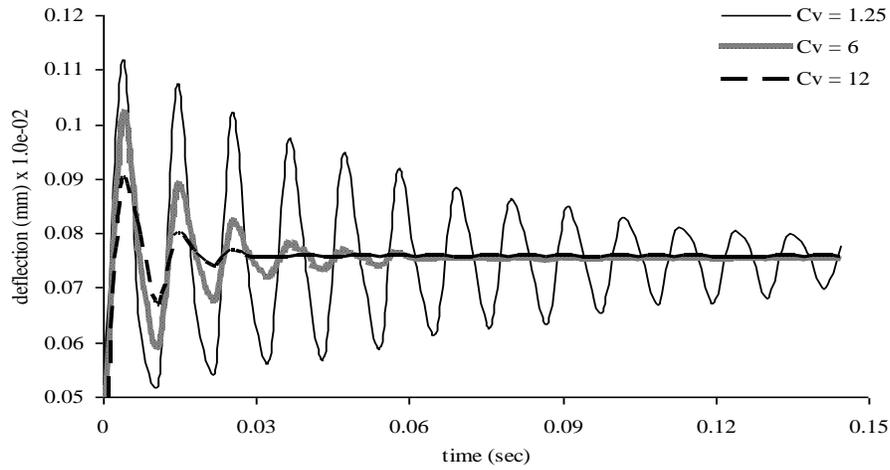


Figure 11: Damped response of a simple supported 8.8% glass fiber reinforced (F20GF8.8) Resol/VAC-EHA blends composite plate at aspect ratio $b/a = 2.0$

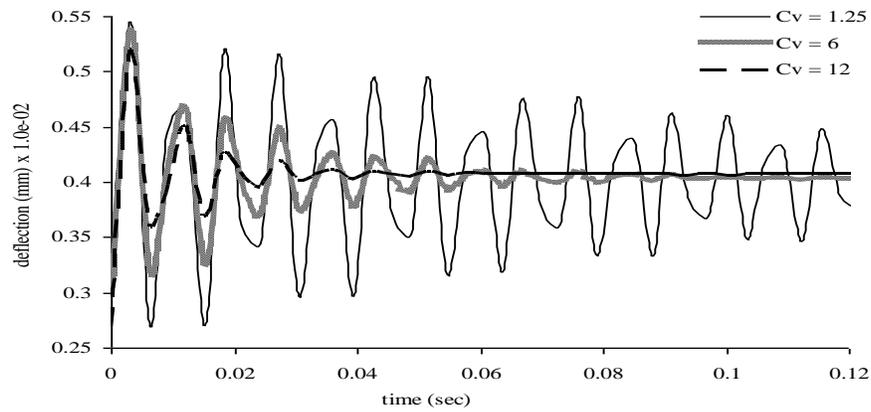


Figure 12: Damped response of a simple supported 17.6% glass fiber reinforced (F20GF17.6) Resol/VAC-EHA blends square composite plate.

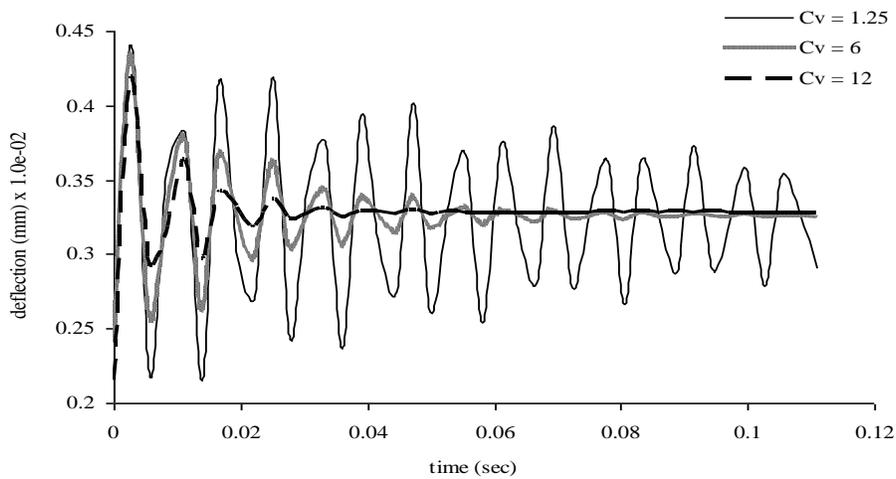


Figure 13: Damped response of a simple supported 25.4% glass fiber reinforced (F20GF25.4) Resol/VAC-EHA blends square composite plate.

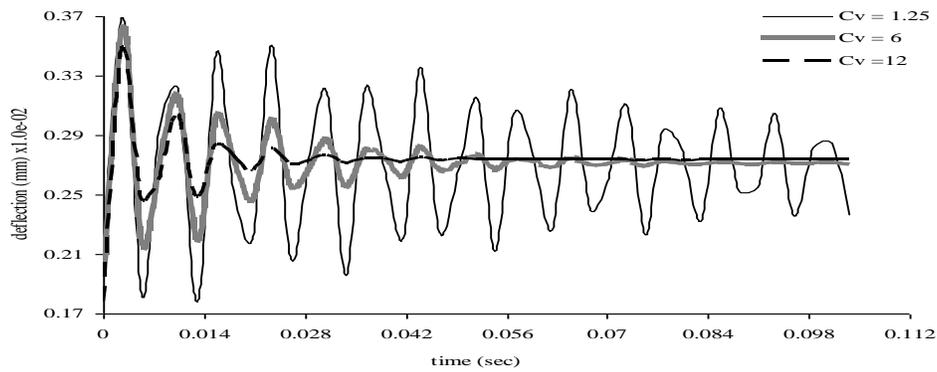


Figure 14: Damped response of a simple supported 35.3 % glass fiber reinforced (F20GF35.3) Resol/VAC-EHA blends square composite plate.

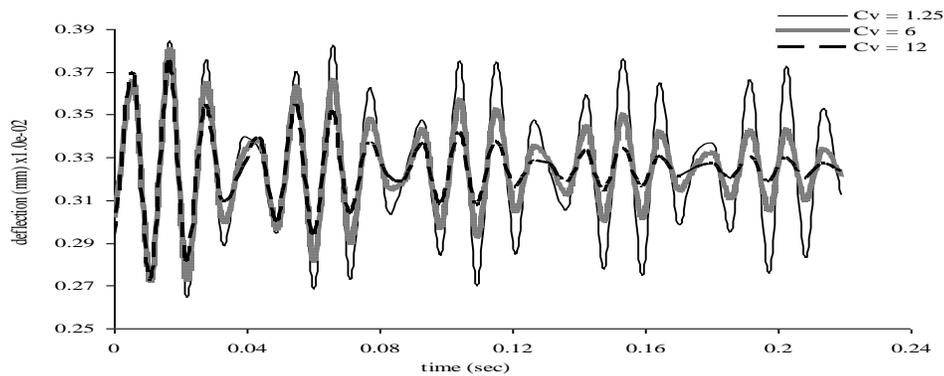


Figure 15: Damped response of a clamped supported pure resol square isotropic plate at various damping coefficient factor.

Table

Table No.1: Tensile modulus of the Resol and Resol/VAC-EHA blends

Code	Materials		Tensile Modulus (GPa)
	Resol % by weight	VAC-EHA % by weight	
Resol	100 %	Nil	3.46
F5	95%	5%	3.25
F10	90%	10%	3.21
F15	85%	15%	3.17
F20	80%	20%	3.05

Table No.2: Mechanical properties of the Glass fibers reinforced in Resol/VAC-EHA (F20) Composites

Code	Volume fraction of fiber in (%)	E_L (GPa)	E_T (GPa)	% Elongation at break
F20GF8.8	8.8	8.91	3.27	3.12
F20GF17.6	17.6	13.7	3.88	4.8
F20GF25.4	25.4	17.5	4.13	4.95
F20GF35.3	35.3	21.6	4.37	5.2

Table No.3: Deflection of the simple supported Resol Plate

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Timoshenko [22] w_{max} (mm)	MQRBF Method w_{max} (mm)
10	300	300	1.0	10	0.0011329	0.0011440
10	300	150	2.0	10	0.00017666	0.00017614
10	300	100	3.0	10	0.000042130	0.000042027

Table No.4: Deflection of the simple supported Resol/VAC-EHA (F5) blends Plate

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Timoshenko [22] w_{max} (mm)	MQRBF Method w_{max} (mm)
10	300	300	1.0	10	0.0010379	0.0010481
10	300	150	2.0	10	0.00016185	0.00016137
10	300	100	3.0	10	0.000038599	0.000038504

Table No.5: Deflection of the simple supported Resol/VAC-EHA (F10) blends Plate

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Timoshenko [22] w_{max} (mm)	MQRBF Method w_{max} (mm)
10	300	300	1.0	10	0.0011050	0.0011159
10	300	150	2.0	10	0.00017231	0.00017180
10	300	100	3.0	10	0.000041093	0.000040992

Table No.6: Deflection of the simple supported Resol/VAC-EHA (F15) blends Plate

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Whitney [23], $m=1, n=1, w_{max}$ (mm)	MQRBF Method w_{max} (mm)
10	300	300	1.0	10	3.9093e-004	3.945e-004
10	300	150	2.0	10	5.0269e-005	4.903e-005
10	300	100	3.0	10	1.1984e-005	1.168e-005

Table No.7: Deflection of the simple supported Resol/VAC-EHA (F20) blends Plate

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Timoshenko [22] w _{max} (mm)	MQRBF Method w _{max} (mm)
10	300	300	1.0	10	0.0011774	0.0011890
10	300	150	2.0	10	0.00018361	0.00018307
10	300	100	3.0	10	0.000043787	0.000043680

Table No.8: Deflection of the simple supported Glass fibers reinforced Resol/VAC-EHA Composite plate at 8.8% volume fraction of the fibers

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Timoshenko [22] w _{max} (mm)	MQRBF Method w _{max} (mm)
10	300	300	1.0	10	0.0011187	0.0011298
10	300	150	2.0	10	0.00017446	0.00017394
10	300	100	3.0	10	0.000041605	0.000041503

Table No.9: Deflection of the simple supported Glass fibers reinforced Resol/VAC-EHA Composite plate at 17.6 % volume fraction of the fibers

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Whitney [23]m= 1,n =1, w _{max} (mm)	MQRBF Method w _{max} (mm)
10	300	300	1.0	10	5.7270e-004	5.741e-004
10	300	150	2.0	10	7.5908e-005	7.420e-005
10	300	100	3.0	10	1.8203e-005	1.781e-005

Table No.10: Deflection of the simple supported Glass fibers reinforced Resol/VAC-EHA Composite plate at 25.4 % volume fraction of the fibers

q (N/m ²)	b (mm)	a (mm)	b/a	h (mm)	Whitney [23] m= 1,n =1, w _{max} (mm)	MQRBF Method w _{max} (mm)
10	300	300	1.0	10	2.6023e-004	2.648e-004
10	300	150	2.0	10	3.2448e-005	3.158e-005
10	300	100	3.0	10	7.6914e-006	7.472e-006

Table No.11: Deflection of the simple supported Glass fibers reinforced Resol/VAC-EHA Composite plate at 35.3 % volume fraction of the fibers

q (N/m ²)	b(m m)	a(m m)	b/a	h(mm)	Whitney [23]m= 1,n =1 w _{max} (mm)	MQRBF Method w _{max} (mm)
10	300	300	1.0	10	3.1475e-004	3.177e-004
10	300	150	2.0	10	3.9759e-005	3.872e-005
10	300	100	3.0	10	9.4471e-006	9.190e-006

Table No.12: Deflection of clamped edge Resol, Resol/VAC-EHA blends plate

q = 10 (N/m ²) a = b = 300 (mm) h = 10 (mm)	Sample Code	Whitney [23]	MQRBF Method w _{max} (mm)
	F20GF8.8	1.9491e-004	1.9820e-004
	F20GF17.6	1.3458e-004	1.3650e-004
	F20GF25.4	1.0910e-004	1.1119e-004
	F20GF35.3	9.0669e-005	9.0983e-005

Table No.13: Deflection of clamped edge glass fibers reinforced Resol/VAC-EHA Composite plate

q = 10 (N/m ²), a = b = 300 (mm) h = 10 (mm)	Sample Code	Analytical Method w _{max} [22] (mm)	MQRBF Method w _{max} (mm)
	Resol	3.2211e-004	3.2452e-004
	F5	3.4292e-004	3.4549e-004
	F10	3.4719e-004	3.4980e-004
	F15	3.5158e-004	3.5421e-004
	F20	3.6541e-004	3.6819e-004

Comparitive Study of Advanced Database Replication Strategies

A. Pramod Kumar¹ M.Tech, B.SATEESH² M.Tech

ACE Engineering College, JNTUH¹

CBIT, (OsmaniaUniversity)²

Abstract

In this paper we are comparing two advanced replication strategies namely materialized view replication and multi master replication are analyzed based on the Five partitioning algorithms namely range partitioning, hash partitioning, list partitioning, composite range-hash partitioning and composite range-list partitioning have been implemented for both the replication strategies. The performance of all these partitioning algorithms have been evaluated for each of the replication strategies with simulation results.

Keywords: Partitioning Algorithm , Replication, Distributed Database, Centralized Database.

1. Introduction

Database replication is nothing but copying of data from a Database in one server to a Database in another so that all users share the same level of information. The result is a Distributed Database in which users can access data relevant to their tasks without interfering with the work of others. Distributed Database Management Systems (DDBMS) ensures that changes, additions, and deletions performed on the data at any given location are automatically reflected in the data stored at all the other locations. Therefore, every user always sees data that is consistent with the data seen by all the other users. There are different strategies to replicate data among servers located in different places [1]. A proper replication method has to be determined before implementation of Database replication for huge installations

The centralized approach suffers from two major drawbacks:

- Performance problems due to high server load or high communication latency for remote clients.
- Availability problems caused by server downtime or lack of connectivity. Clients in portions of the network that are temporarily disconnected from the server cannot be serviced. The server load and server downtime problems can be addressed by replicating the Database servers to form a cluster of peer servers that coordinate updates.

To overcome above drawbacks Advanced replication strategy is implemented. Advanced replication is a fully integrated feature of the server; it is not a separate server. Replication uses Distributed Database technology to share data between multiple sites, but a replicated Database and a Distributed Database are not the same. In a Distributed Database, data is available at many locations, but a particular table resides at only one location. A centralized approach manages only one copy of the Database. This approach is simple since contradicting views between replicas are not possible[2].

Replication objects

A **replication object** is a Database object existing on multiple servers in a Distributed Database system. In a replication environment, any updates made to a replication object at one site are applied to the copies at all other sites.

Replication Groups

A replication group can exist at multiple **replication sites**. Replication environments support two basic types of sites: **master sites** and **materialized view sites**. One site can be both a master site for one replication group and a materialized view site for a different replication group. However, one site cannot be both the master site and the materialized view site for the same replication group.

2. Advanced Database Replication Strategies

The advanced Database replication strategies are materialized view replications and multimaster replication

2.1 Materialized view Replication

Materialized view replications are schema objects that can be used to summarize, precompute, replicate, and distribute data. E.g. to construct a data warehouse A materialized view replication provides indirect access to table data by storing the results of a query in a separate schema object. The existence of a materialized view replication is transparent to SQL, but when used for query rewrites will improve the performance of SQL execution. An **updatable materialized view replication** [3] lets you insert, update, and delete. A materialized view on a base table, partitioned table or view and define indexes on a materialized view. A materialized view replication can be stored in the same Database as its base table(s) or in a different Database.

Materialized views replication stored in the same Database as their base tables can improve query performance through query rewrites. Query rewrites are particularly useful in a data warehouse environment.

Partitioning is a powerful capability to help manage unwieldy table's indexes and materialized views replication. Table and index partitioning allows for data to be divided into smaller pieces to be managed and accessed as individual units.

Each partition can have its own unique storage attributes allowing the various partitions to be placed in different table spaces on different storage devices. By utilizing partitioning, the performance of certain queries and maintenance tasks can be vastly improved over using the traditional method of table and index design. Partitioning allows a table, index or index-organized table to be sub-divided into smaller pieces. Each piece of Database object is called a partition. A materialized view replication contains a complete or partial copy of a target **master** from a single point in time. The target master can be either a master table at a master site or a master materialized view at a materialized view site. A **master materialized view** is a materialized view that functions as a master for another materialized view. A **multitier materialized view** is one that is based on another materialized view, instead of on a master table [4].

TABLE_NAME	TABLESPACE_NAME	CLUSTER_NAME	IOT_NAME	PCT FREE	PCT USED	INI TRA
MLOG\$.CBIT_RANGEHASH_GP_4	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GP_3	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GP_2	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GP	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GLOBA3	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GLOBA2	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GLOBA1	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_GLOBA	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_LOCAL3	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_LOCAL2	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_LOCAL1	CBIT			60	30	
MLOG\$.CBIT_RANGEHASH_LOCAL	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GP_4	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GP_3	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GP_2	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GP	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GLOBA3	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GLOBA2	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GLOBA1	CBIT			60	30	
MLOG\$.CBIT_RANGELIST_GLOBA	CBIT			60	30	

Figure 2.1: shows the list of materialized views replication in remote Database.

2.2 Multimaster Replication

Two types of master replication: single master replication and multimaster replication. Multimaster replication includes multiple master sites, where each master site operates as an equal peer.

In single master replication, a single master site supporting materialized view replication provides the mechanisms to support potentially hundreds or thousands of materialized view sites. A single master site that supports one or more materialized view sites can also participate in a multiple master site environment, creating a hybrid replication environment(combination of multimaster and materialized view replication).

Materialized views can be based on master tables at master sites or on materialized views at materialized view sites. When materialized views are based on materialized views, a multitier materialized view environment. In such an environment, materialized views that have other materialized views based on them are called master materialized views.

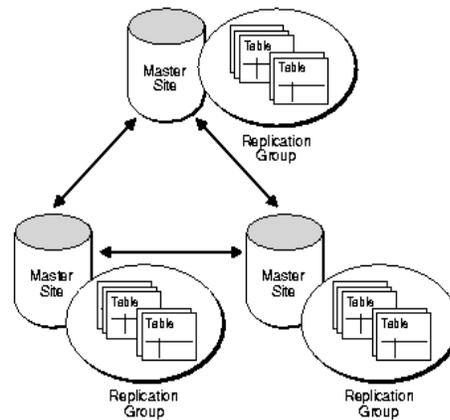


Figure 2.2: shows the multimaster replication

Multimaster replication (also called peer-to-peer or *n*-way replication) enables multiple sites, acting as equal peers, to manage groups of replicated Database objects. Each site in a multimaster replication environment is a master site, and each site communicates with the other master sites. Applications can update any replicated table at any site in a multimaster configuration.

Database servers operating as master sites in a multimaster environment automatically work to converge the data of all table replicas and to ensure global transaction consistency and data integrity.

2.2.1 Single Master Replication

A single master site can also function as the target master site for one or more materialized view sites. Unlike multimaster replication, where updates to a single site are propagated to all other master sites, materialized views update only their target master site. Conflict resolution is handled only at master sites or master materialized view sites. Materialized view replication can contain complete or partial replicas of the replicated table.

Master Sites

A master site can be both a node in a multimaster replication environment and the master for one or more materialized view sites in a single master or multimaster replication environment. The replicated objects are stored at the master site and are available for user access.

Master Definition Site

In a multimaster replication environment, one master site operates as the master definition site for a master group. This particular site performs many of the administrative and maintenance tasks for the multimaster replication environment.

Each master group can have only one master definition site, though the master definition site can be any of the master sites in the multimaster environment. Additionally, the master definition site can be changed to a different master site if necessary. A single master site supporting materialized view replication is by default the master definition site.

2.2.2 Benefits of Multimaster Replication

From a very basic point of view, replication is used to make sure that data is available when and where you need it. The following sections describe several different environments that have different information delivery requirements. replication environment may have one or more of the following requirements[6].

Failover

Multimaster replication can be used to protect the availability of a mission critical Database. For example, a multimaster replication environment can replicate data in Database to establish a failover site should the primary site become unavailable due to system or network outages. Such a failover site also can serve as a fully functional Database to support application access when the primary site is concurrently operational. Net to configure automatic connect-time failover, which enables Net to fail over to a different master site if the first master site fails. Configure automatic connect-time failover in tnsnames. file by setting the FAILOVER_MODE parameter to on and specifying multiple connect descriptors. For more information about configuring connect-time failover

Load Balancing

Multimaster replication is useful for transaction processing applications that require multiple points of access to Database information for the following purposes:

- Distributing a heavy application load
- Ensuring continuous availability
- Providing more localized data access

Applications that have application load distribution requirements commonly include customer service oriented applications.

3. Simulation Results:

VIEW REPLICATION

Initially all possible data retrieving and manipulating operations will be examined on Range, List, Composite Range – List and Composite Range – Hash partitioning techniques.

3.1 Full scan operation on materialized view replication

Full scan operation has to be examined in all partitioned techniques by gradually inserting the data in all materialized views replication. In the graph given below x-axis shows different types of partitioning techniques and gradual massive increment of data is noted as first phase , second phase ,third phase and fourth phase respectively and y-axis shows the cpu cost in m-seconds.

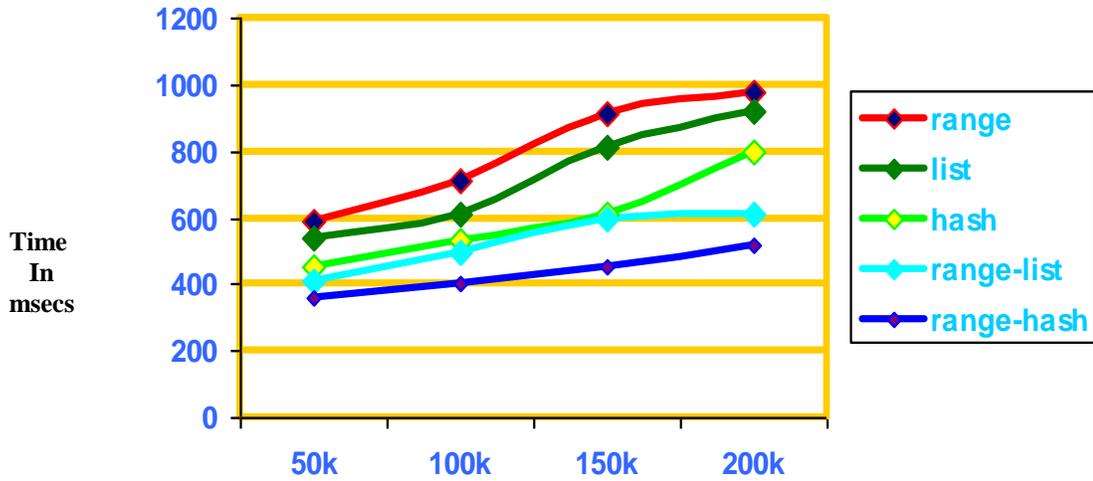


Figure 3.1 shows the response times of full scan operation.

The full scan operation graph shows the response times of different partitioning techniques when a full scan operation is applied. It has been observed that at first phase of full scan , range-list and range-hash performance is good when compared to range, list and hash partitioning techniques and when data is gradually growing , the performance of hash , range-list and range-hash are improved .Composite Range-hash is unchangeable even after massive insertion of data during second , third and fourth phases. In overall, the performance of Range-hash is better compared to other partitioning techniques when full scan operations are performed. Any full scan operations or partial scan operations with out using index objects performs well when underlying materialized views are partitioned with either Composite Range-List or Range-Hash partitioning techniques.

3.2 Full scan operation on Multi master replication

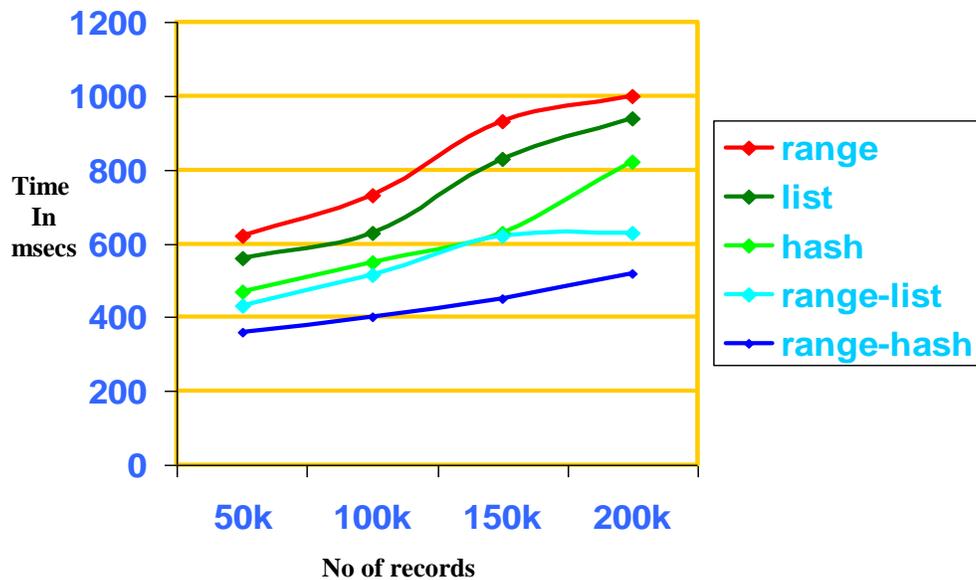


Figure 3.2 shows the response times of full scan operation.

The above graph shows the response times of different partitioning techniques when a full scan operation is applied. In overall, the performance of Range-hash is better compared to other partitioning techniques when full scan operations are performed. And range poor performance.

4. Conclusion

On the partitioned Databases using each of the algorithms materialized view replication strategy is applied and refreshment time is recorded. Similarly multimaster replication strategy is applied and refreshment times are recorded. From the results, it is observed that the performance of hash partitioning algorithm is poor when regular updates are performed. The composite range-hash partitioning and composite range-list partitioning algorithms showed better performance in both materialized view replication and multimaster replication strategies.

5. References

- [1] Data base advanced replication and management A.P.I Reference
- [2] Oracle Partitioning and QCO Space Management by Chuck Sodowsky, Dan Sherwin, Lenka Vanek and Thomas Szekely
- [3] J. A. Blakeley, N. L. Martin. Join Index, Materialized View Replication, and Hybrid Hash-Join: A Performance Analysis Proc. IEEE Int'l. Con\$ on Data Eng. Los Angeles,
- [4] A. Gupta, I.S. Mumick, Maintenance of Materialized Views: Problems, Techniques, and Applications, IEEE Data Eng. Bulletin, Special Issue on Materialized Views.
- [5] Materialized Views in Oracle (8.1.5/8i/9i/10g/11g) by Jangan pernah berhenti belajar
- [6] Oracle Master Replication, Performance and Scalability, Metalink article

Author



A. PRAMOD KUMAR

received the Bachelor's degree from M.V.S.R Engineering college affiliated to OSMANIA UNIVERSITY Hyderabad 2006. received the Masterr's degree from Chaitanya Bharathi Institute of Science and Technology affiliated to OSMANIA UNIVERSITY Hyderabad 2009 Computer Science & Engineering department.. I'm currently Working as a Assistant professor ACE Engineering College in Information Technology Department affiliated to JNTUH University.

Vehicular Number Plate Recognition Using Edge Detection and Characteristic Analysis of National Number Plates

¹Bharat Raju Dandu, ²Abhinav Chopra
Amity University

ABSTRACT:

In this work, we propose a framework that uses a camera installed at roadside to detect the vehicle number plate. A typical video component requires several adjustments after image has been stored such as enhancement of image, localization of number plate, separating each character and recognition of each character. In this work we extract the vehicle number plate from our image and then recognize it based on the characteristics of number plates in different countries. We use Sobel edge detection for plate localization, template matching and fuzzy logic for recognition. We make use of the characteristics of the vehicle number sequences to further enhance performance. Vehicle number plate recognition can be used to decrease human effort by making systems automatic.

General Terms: Computer Vision.

Keywords: Number plate, Sobel algorithm, vehicle and camera

1. INTRODUCTION

Number plate recognition is a technology based on computer vision which involves image processing, fuzzy logic and many other techniques.

Applications of Vehicle Number Plate Recognition(VNPR) can vary from automatic parking facilities, automatic fuel pumping for registered vehicles, restricted areas reserved for VIP members, automatic toll sensors, vehicles involved in theft and other imaginative forms as shown in Figure 1^[1]. Unfortunately there exist various factors which restrict us from recognizing a license plate such as weather, lighting, visual occlusion, placement of number plate, speed of vehicle, damaged plates, angles of camera, color of number plate and different fonts used for the characters.^[2]

These factors can be overcome in some ways so that our results are more accurate. With correct evaluation and analysis of the results of recognition of number plate it is possible to improve the accuracy of the number plate



Figure 1: Vehicle number plate recognition systems

recognition system. In India there has been a lot of work done on number plate recognition. It is possible to recognize the characters of Indian number plate based on the specifications of Indian number plates.

DL 01 C 1234

Considering the standard specifications of Indian Number Plates: Here 'DL' that is the first 2 characters stands for the state code, '01' that is a two digit number code which stands for the regional transport office where the car has been registered. 'C' stands for car and similarly 'S' for scooter and 'P' for public transport vehicles. Finally there comes a four digit car number ('1234'). After the 4 digit car number reaches '9999' the next vehicle number becomes

DL 01 C 9999 → DL 01 C A 0001 → DL 01 C A 9999 → DL 01 C B 0001 → DL 01 C AA 0001

2. PROPOSED TECHNIQUE

Vehicle number plate recognition follows a structural pattern for the recognition of characters. Figure 2 gives a schematic diagram of the of the VNPR system.

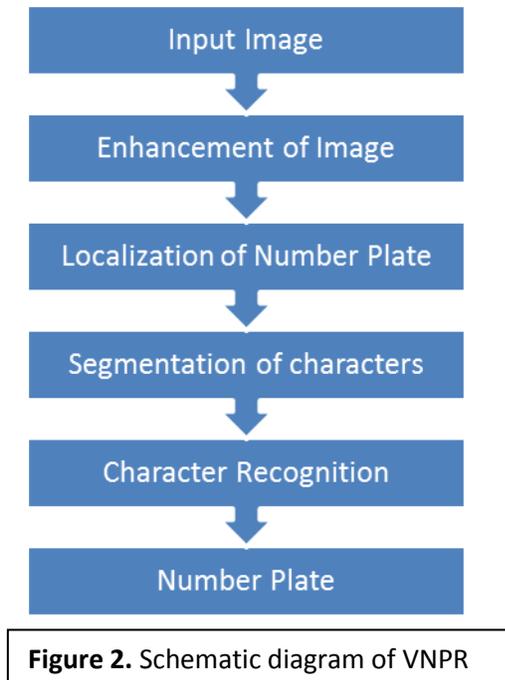
First of all we improve the quality of image by using familiar image processing techniques such as the following

2.1 Conversion to gray scale:

The RGB 24 bit color pixels are converted into 8 bit gray value as shown in figure 3(b).

2.2 Median Filtering

The non-linear filter changes the gray value of the pixels to the median of the gray value of surrounding pixels as seen in figure 3(c). We use a 3x3 mask and get the corresponding gray value of each pixel using the 8 neighbouring pixels which results in removal of noise with benefits^[7]. Median filtering gives advantages such



as no reduction in contrast since output values are its neighbourhood values, it does not shift the boundaries which is a contrast dependent problem and less sensitive than the mean values

2.3 Histogram Equalization

Histogram Equalization is used to enhance the contrast of the image for better functioning.^[3]

2.4 Edge Detection

After enhancement of image we perform extraction of vehicle number plate by localization of number plate region using Sobel edge detection and fuzzy logic as shown in figure 3(d). Sobel operator^[3] has a 3x3 convolution kernel. One kernel is the other rotated by 90 degree.

Masks used by Sobel operator are

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

The kernels respond to vertical and horizontal edges and the characters in the number plate have vertical edges of uniform nature at regular intervals. These characters are also equally distant to the number plate edge. These features help us to locate the number plate within the image.

After edge detection we perform smoothing of the image. Smoothing is often used to reduce noise within an image or to produce a less pixelated image. Most smoothing methods are based on low pass filters. It is done to reduce the no. of connected components so that the objects can be identified more efficiently (Figure 3(e)).

Finally we perform the Optical Character Recognition (OCR) clubbed with the fuzzy logic of character analysis of Indian Vehicle Number Plates which can be seen in [3]. However for accurate OCR, one of the key factors is the image quality. Hence it is important that for a fast travelling vehicle to be captured, the camera should have a short shutter time and high illumination for better image which avoids blurring of the image. One of the best illumination effects provided by a camera is an IR flash which gives very good results. We have used the template based OCR technique over the feature based OCR technique because the template based technique is more sensitive to font and size of the characters. Here the characters are correlated with the set of stored templates. The complete proposed technique that has been implemented can be seen in figure 3.

3. Characteristic Analysis

This system is designed on the basis of national number plate specifications and hence has characteristic features which can be taken into analysis^[1,4]:

- The background color of the plate is white or yellow and the characters are written in black.^[5]
- Characters ‘E’, ‘F’, ‘I’, ‘O’, ‘Q’, ‘V’, ‘X’ and ‘Z’ are not used in the first 2 characters in a number plate considering only national numbers and not the few exceptions such as numbers of diplomats, etc.
- The 5th character is mostly ‘C’, ‘S’ or ‘P’.
- The last 4 characters are always numbers. Hence alphabets need not be taken into account. This helps us in differentiating confusing matches between characters like ‘Z’ and ‘2’, ‘B’ and ‘8’, etc.
- The number plates in common have 10-12 characters at present.

Finally to calculate the efficiency of our number plate, we use the following parameters.

$$\text{Recognition Rate} = \frac{\text{Correct characters}}{\text{Total characters}}$$

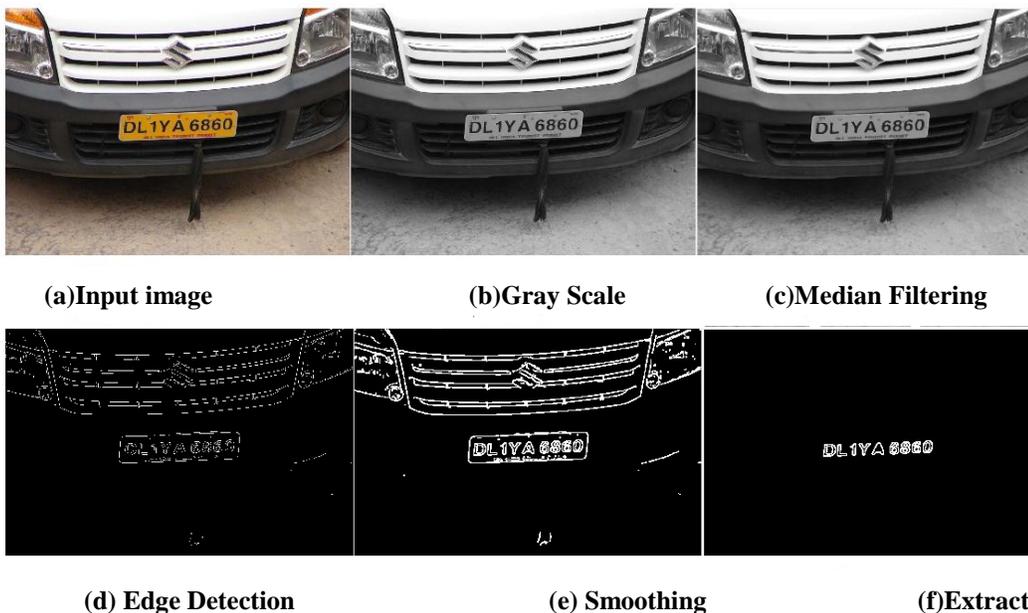


Figure 3: Proposed Technique

$$\text{Rejection Rate} = \frac{\text{Rejected characters}}{\text{Total characters}}$$

$$\text{Error Rate} = \frac{\text{Incorrect Characters}}{\text{Total characters}}$$

The above mentioned characteristics of vehicle number plates helps us recognize the number plates with more accuracy and makes the proposed system faster to recognize number plates using fuzzy logic. We tested the proposed system for 100 images of license plates. The system could not recognize 10 of these images due to bad camera angles, bad illumination and blurred images. The common misconceptions during character recognition can be seen in Table 1 and the common misconceptions that have been eliminated through our characteristic analysis are shown in Table 2. As a result, the images showed recognition rate of 95.1%, error rate of 1.3%, and rejection rate of 3.6

Character on the number plate	Character Recognized as
D	O
M	H
6	8
8	0

Table 1: Common misconceptions that exist

Characters on number plate	Correct output found with no confusion with characters below
B	8
O	0
Z	2
S	5
Z	7

Table 2: Misconceptions removed through characteristic analysis and table holds true vice versa.

4. Conclusion

In this paper we introduced a method of extracting specific features of the national number plates and increase its accuracy and speed in recognizing number plates. We have covered numerous number of the vehicle number plates in this paper but other type of number plates exist where such a character analysis cannot be applied using our fuzzy logic of character analysis. For such systems we have to make a few changes to our system. Hence we have focused our system on producing more accurate results and with lesser response time to the most common specifications of Indian vehicle number plates.

Acknowledgement

Authors are very thankful to C.S.E. Department of Amity University(Noida) for granting us the permission to carry out the work and providing necessary facilities during the progress of the work.

References

1. Jung, K., K.I. Kim, A.K. Jain, May 2004. Text Information Extraction in Images and Video: A survey. *Journal on Pattern Recognition*. Vol. 37, No. 5. pp: 977-997.
2. Kwasnicka, H., B. Wawrzyniak, 2002. Symposium on Methods of Artificial Intelligence (AI-METH 2002) Gliwice, Poland. License Plate Localization and Recognition in Camera Pictures.
3. R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Pearson Education Asia, 2002.
4. Nijhuis, J.A.G., M.H. ter Brugge, H. Hettema, L. Spaanenburg, 1995. Proceedings of the 5th Aachener Fuzzy Symposium, Aachen, Germany. pp: 147-153. License Plate Recognition by Fuzzy Logic.
5. Lee, S.W., D.J. Lee, H.S. Park, 1996. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, pp: 1045-1050. A New Methodology for Grayscale Character Segmentation and Recognition.
6. V. S. L. Nathan, Ramkumar. J, Kamakshi. P. S, "New approaches for license plate recognition system," ICISIP 2004, p.p. 149-152.
7. *Digital Image Processing, 2/e* By Gonzalez pg 263.
8. Danian Zheng, Yannan Zhao, Jiixin Wang, "An efficient method of license plate location", *Pattern Recognition Letters* (2005), Volume 26, Issue 15, November 2005, Pages 2431-2438.
9. J. Cano and J. C. Perez-Cortes, "Vehicle License Plate Segmentation in Natural Images", *Lecture Notes on Computer Science* 2652, F.J. Perales et al., Ed. Springer-Verlag, 2003, pp. 142-149.
10. Shigueo Nomura, Keiji Yamanaka, Osamu Katai, Hiroshi Kawakami, Takayuki Shiose, "A novel adaptive morphological approach for degraded character image segmentation", *Pattern Recognition*, Vol. 38, Issue 11, November 2005, pp.1961 - 1975.

Comparative Analysis of Image Registration using SIFT and RANSAC method

Riddhi J Ramani¹, N.G. Chitaliya²

¹ Department of Electronics and communication Engineering,SVIT Vasad-388306

²Assistant prof. in Department of Electronics and communication Engineering,SVIT Vasad-388306

ABSTRACT

Image registration is a prerequisite step prior to image fusion or image mosaic. It is a fundamental image processing technique and is very useful in integrating information from different sensors, finding changes in images taken at different times, inferring three- dimensional information from stereo images, and recognizing model-based objects. As a large dimension of the traditional SIFT descriptor and its complex algorithm and improved algorithm of SIFT is presented which can reduce the dimension and also can improve the time saving and complexity reducing and to improve the accuracy of matching, RANSAC is applied for removal the wrong matching points

Keywords— Image Registration, SIFT, RANSAC

I. INTRODUCTION

Image registration is multi spectral. Satellite image is a crucial problem for remote sensing applications, and remains challenging because of the inherent nonlinearity in intensity changes [1].

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and by different sensors. It geometrically aligns two images - the reference and sensed images [2].

Image processing which are possibly able to visualize objects inside the human body, are of special interest. Advances in computer science have led to reliable and efficient image processing methods useful in medical diagnosis, treatment planning and medical research. In clinical diagnosis using medical images, integration of useful data obtained from separate images is often desired. The images need to be geometrically aligned for better observation. This procedure of mapping points from one image to corresponding points in another image is called Image Registration. It is a spatial transform. The reference and the referred image could be different because were taken [3].

- At different times
- Using different devices like MRI, CT, PET, SPECT etc (multi modal).
- From different angles in order to have 2D or 3D perspective (multi temporal).

II. IMAGE REGISTRATION ALGORITHM

SIFT

A method to enhance the recognition of spatially distributed features, based on the scale invariant features transform is called SIFT method.

The Scale Invariant Feature Transform (SIFT) can detect and extract feature points which are invariant to changes in illumination, image noise, rotation, scaling, and small changes in viewpoint.

After features are extracted from images, the initial matching process can begin. Feature vectors from one image are compared to those from the other image in pairs. To give robustness, feature vectors from the other image which are within a certain similarity threshold are preserved. As a result, bad matches appear.

In image processing applications one class of problems sensitive to input variations is that involving image registration. It requires the transform estimation from just the image data, and the quality of the estimation can depend on the number and positions of the feature point's identified within the images. The reasons behind this are largely related to basic geometry: even simple relationships will be better served by input points that are as far apart as possible within the image (providing larger differences and reducing the impact of discretisation), while a nonlinear relationship such as radial lens distortion may have only a negligible impact over a large central part of an image but could be a very Significant influence on features selected nearer to the edges. Broadly speaking, a 'good' (wide- ranging, fairly uniform, and not too clustered) distribution of features is an important factor in ensuring the quality of registration [4].

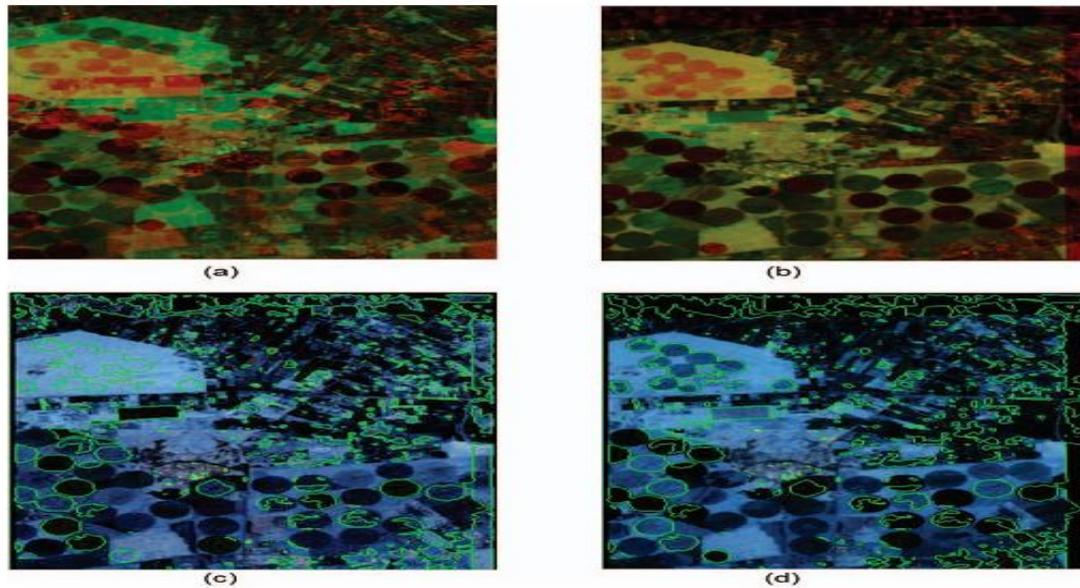


Figure 1 - Results on a sample LANDSAT5 image,(a) The two images before registration,(b) The two images after registration,(c) Changes in first image,(d) Changes in second image.

Experiment Result

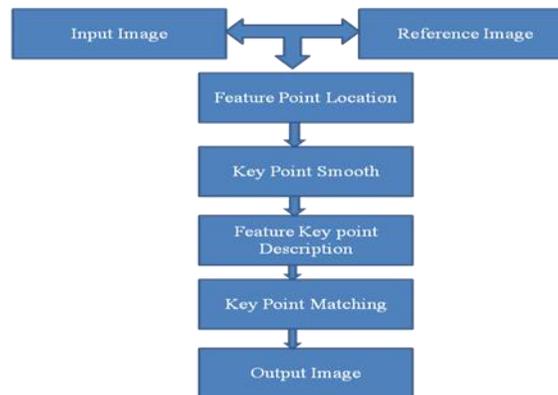


Figure 2- Flow chart of proposed method

The Experiment Algorithm Steps

- 1) Feature point location: It is used for to locate similar points between two images.
- 2) Key point smooth: Passed easily define points.
- 3) Feature Key point Description: Computing the gradient magnitude & orientation at each image sample point in a region around the feature point description.
- 4) Matching: When key point description was created computing minimum Euclidean distance for the invariant description for vector as similarity between two images.

Simulation Result

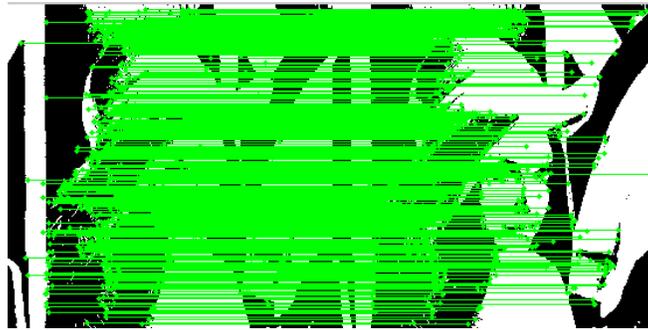


Figure 3- Output of lena Images



Figure 4-Experiment of Onion Result

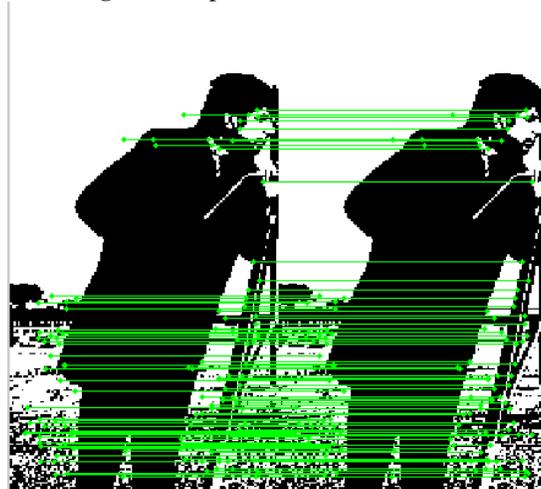


Figure 5-Experiment of Camera Man Result

Uses

- 1) Image alignment (Homo Graphy, fundamental matrix)
- 2) 3D reconstruction (e.g. Photo Tourism)
- 3) Motion tracking
- 4) Object recognition
- 5) Indexing and database retrieval
- 6) Robot navigation

RANSAC

RANSAC (RANdom SAMple Consensus) is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. It is a non-deterministic algorithm in the sense that it produces a reasonable result only with a certain probability, with this probability increasing as more iteration are allowed. The algorithm was first published by Fischler and Bolles in 1981.

RANSAC is a re sampling technique that generates candidate solutions by using the minimum number observations (data points) required to estimate the underlying model parameters. As pointed out by Fischler and Bolles unlike conventional sampling techniques that use as much of the data as possible to obtain an initial solution and then proceed to prune outliers, RANSAC uses the smallest set possible and proceeds to enlarge this set with consistent data points [5].

The Algorithm Steps

- (1) Select the randomly the no of points required to determine the images.
- (2) Determine the how many points from the set of all points fit with a images.
- (3) Select the optimum parameters according to the final parameters output.

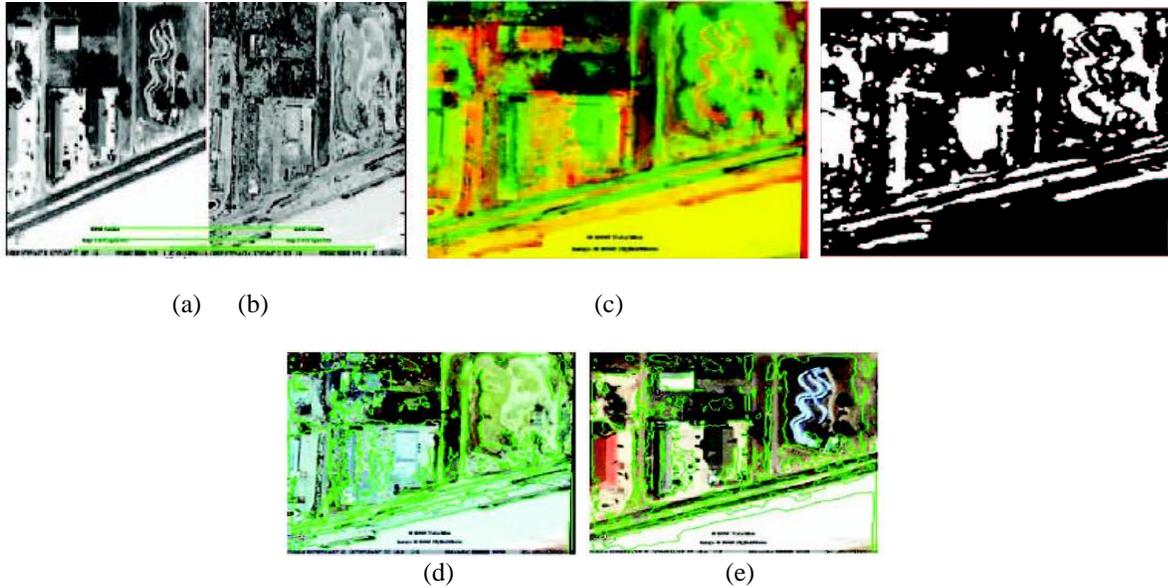


Figure 6 - Sample of results (a) Matched points after RANSAC, (b) The two images after registration, (c) The change mask, (d) Changes in first image, (e) Changes in second Image [6].

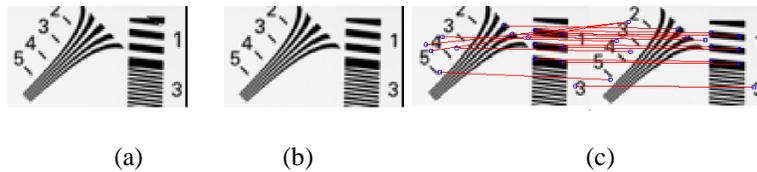


Figure 7 – (a) Input Image, (b) Reference Image, (c) RANSAC Point Matching [7].

Simulation Result More than 2 Input Images

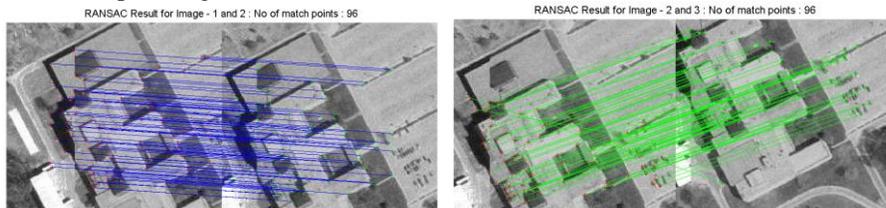


Figure 8 – Input image and reference image with output **Figure 9–Reference image and reference image with output**

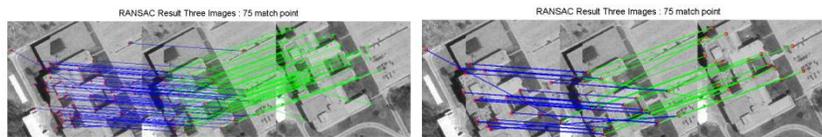


Figure 10– Input image and reference and reference images with output **Figure11 – Input image and reference and reference images with output**

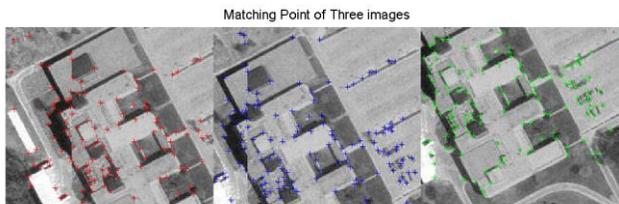


Figure 12 – Input image and reference and reference images with output

RANSAC Result for Image - 1 and 2 : No of match points : 38



RANSAC Result for Image - 2 and 3 : No of match points : 38



RANSAC Result Three Images : 33 match point

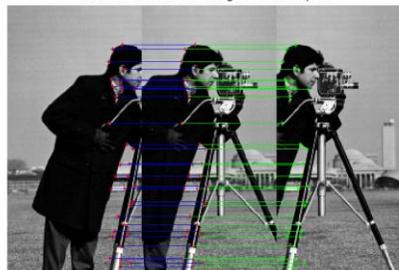


Figure 13 – Input image and reference output **Figure 14** – reference and reference reference and reference images with **Figure 15**– Input image and output

A comparison of SIFT & RANSAC Methods

SIFT (Scale Invariant Feature Transform)	RANSAC (RANDOM SAMPLE CONSENSUS)
The Scale Invariant Feature Transform (SIFT) can detect and extract feature points which are invariant to changes in illumination, image noise, rotation, scaling, and small changes in viewpoint. After features are extracted from images, the initial matching process can begin.	This method checks the number of elements of the input feature point data set which are consistent with the model just chosen.
Feature vectors from one image are compared to those from the other image in pairs. To give robustness, feature vectors from the other image which are within a certain similarity threshold are preserved. As a result, bad matches appear.	RANSAC repeats the two steps within a specified threshold until it finds the maximum number of elements within a model. It then selects this model and rejects mismatches.
SIFT idea was that by dividing the image into smaller windows, we would be able to match more features between the window and object.	In the transform fits a certain number of matches, it is considered a 'good' transform. Any points that do not match are discarded and then the transformation is recalculated using these new points.
SIFT is lengthy process involves multiple steps	In RANSAC when reinserted window it gives much better result with very short process.
SIFT is time consuming and costly method.	RANSAC is much more efficient than the Hough transform when SIFT Features are used.
Experiments shows that global localization can be achieved with just the current frame data in feature-rich environments, with the distinctive SIFT features.	Nevertheless, for symmetric environments or when there is a lack of features, global localization with the current frame may be uncertain and the robot should rotate or move around.

iii Conclusion

These Algorithms recover good registered images for more than two input images using SIFT and RANSAC methods.

Both algorithms are also applicable when there is matching point between two images.

Image registration algorithm using SIFT and RANSAC method take few minutes to register images. So the delay execution is time is very less comparatively to other methods i.e. one minute This algorithms are applicable to stereo images to find feature points.

References

- [1] Sang Rok Lee, "A *Coarse-to-Fine* Approach for Remote-Sensing Image Registration Based on a Local Method" International journal on smart sensing and intelligent system, vol.3, no.4, December 2010.
- [2] Barbara Zitova, Jan Flusser, "*Image* registration methods: a survey" Received 9 November 2001; received in revised form 20 June 2003; accepted 26 June 2003.
- [3] Medhav V.Wyawahare, Dr.Pradeep M.Patil, and Hemant k.abhyanka "Image Registration Techniques: An overview", Vishwakarm Institute of technology, pune, Maharashtra state, India. Vol.2, No3, September 2009.
- [4] R. Song, J. Szymanski "Well-distributed SIFT features" ELECTRONICS LETTERS 12th March 2009 Vol. 45 No. 6.
- [5] Konstantinos G. Derpanis, "Overview of the RANSAC Algorithm" May 13, 2010.
- [6] R. Song, J. Szymanski "Well-distributed SIFT features" ELECTRONICS LETTERS 12th March 2009 Vol. 45 No. 6.
- [7] Haidawati Nasir, Vladimir Stankovic, and Stephen Marshall, "IMAGE REGISTRATION FOR SUPER RESOLUTION USING SCALE INVARIANT FEATURE TRANSFORM, BELIEF PROPAGATION AND RANDOM SAMPLING CONSENSUS", 18th European Signal Processing Conference (EUSIPCO-2010) Aalborg, Denmark, August 23-27, 2010.

Assessment of Radiation Emission from Waste Dumpsites in Lagos State of Nigeria

¹Olubosedede. O, ²Akinngbe .O.B., ³Adekoya O.

^{1,2}Department of Physics and Electronics

Adekunle Ajasin University, Akungba-Akoko, Ondo State, Nigeria

³Yaba College of Technology, Lagos State, Nigeria

Abstract

This paper takes a look at the total radiation emanating from waste dumpsites in two cities of Lagos state Nigeria. This was achieved using a radiation survey meter (RADALERT50) to measure the radiation exposure rate in micro sievert per hour (μSvhr^{-1}). Readings were taken by placing the detector at gonad level i.e. about 1 meter above the ground level in five sampling locations; this was done at an interval of 5meters away from the point of reference up to 30 meters. The results obtained revealed that the annual absorbed dose rate measurements taken inside the five dumpsites are $29.80\mu\text{Svhr}^{-1}$, $28.05\mu\text{Svhr}^{-1}$, $19.29\mu\text{Svhr}^{-1}$, $17.53\mu\text{Svhr}^{-1}$ and $15.78\mu\text{Svhr}^{-1}$. This is far lower than the average of $70\mu\text{Svhr}^{-1}$ recommended by UNESCO on effect of Atomic Radiation.

Keywords: Radiation Emission, Waste dumpsites, absorbed dose rate.

Introduction

Wastes constitute an environmental and public health nuisance in major cities all over the world. Thus, governments consider waste management as an essential social service whose budgetary provision is made in line with population projections (Eja *et.al*, 2010, P. 110-123). Hazards posed by such dumpsite are not only in term of odour and presence of disease causing micro-organism, but can arise from the radiation emanating from such dumpsite (Ojoawo *et.al*, 2011, P.661-666). Various radioactivity measurement have shown the existence of traces of radionuclide in books (Imtiaz *et.al*, 2005, P. 169-174) and in the staple food consume in Nigeria (Jibri *et.al.*, 2007, P.53-59, Eyebiokin *et al*, 2005, P.187-191). It has also been established that vegetation and environmental fields in Nigeria contain traces of radionuclides (Akinloye and Olomo, 2005, P. 219-225). All these, are contained in the domestic waste which are indiscriminately dumped on open fields (Ojoawo *et.al*, 2011, P.661-666), farms soils (Jibri *et.al*, 2011, P.1039-1049), Quarry sites (Odunaike *et.al* 2008, P.174-176), rivers (Farai and Oni, 2002, P.94-97), well and boreholes (Jibri *et.al*, 2010, P.291-297), industries (Iyang *et.al* 2009, P.97-100) and even on road sides and mechanic workshops (Nworgu *et.al*, 2011, P. 801-805). In addition, industrial waste that are liable to contain traces of radionuclide are also dumped indiscriminately. Consequently, the radionuclide content in the waste dumpsite, if not properly managed emit mixed radiation to the environment (Eja *et.al*, 2010, P. 110-123). Radiation emission characterization of waste dumpsites (Odunaike *et.al*, 2008, P.174-176) and measurement of radiation level in refuse dumps (Odunaike *et.al*, 2008, P.174-176) shows the level and long term effects of these radiations if not properly monitored. The 2011 Fukushima Daiichi nuclear disaster displaced thousand of people and its effects is still been felt even in places far from the site (Tanabe, 2011, P. 1135-1139).The proper monitoring and evaluation of the radiation emanating from dumpsites in order to provide accurate data as part of environmental monitoring research for proper assessment of radiation exposure rate of the metropolis motivated this study.

Materials and Methods

This paper measures and calculates the absorbed dose of radiation in some selected dumpsites in Igando and ikotun area of Lagos state. In igando, we have the dumpsite I and dumpsite II, while in Agodo we have the dumpsites III, dumpsite IV and dumpsite V. The effects of the radiation on those living around the sites is also considered. The measurement is achieved by using a portable radiation survey meter with serial number 22205 to measure the dose rate or exposure in mixed field. The operating range for measurement can be obtained for subdivisions of mR/hour, $\mu\text{Sv}/\text{hour}$, CPM (Count per minute), CPS (Count per second) depending on the level of radiation. The meter uses 9V battery. The radiation exposure rate measurement was carried out using a survey meter device as described above, this meter record dose rate in micro sievert per hour (μSvhr^{-1}). Seventeen readings were taken altogether from five sampling locations, and every 5 meters away from the first two locations at Igando and Ikotun in Lagos state. The detector was placed at gonad level i.e. about one meter above ground level for effective detection, the detector was switched on to absorb radiation for a few seconds and the highest stable point was recorded. The procedure was repeated at each location and three readings in micro sievert per hour (μSvhr^{-1}) were recorded at each location in which an average value in micro sievert per hour (μSvhr^{-1}) was determined. This was converted to annual absorbed dose rate in micro sievert per year (μSvyr^{-1}).

Study Area

The study was carried out in Lagos state, with Seventeen readings taken. Five readings were taken from Igando and Ikotun dumpsite, and twelve readings were taken for every 5 meters away from Igando dumpsite in Lagos state. Lagos state is an administrative division created on May 27, 1967 and is located in the southwestern part of the country and has Ikeja as its capital.

Lagos state lies within latitude 6° and 35°N and longitude 3° and 45°E, with population of about 17,553,924 (Lagos state Social Security Exercise 2006 Census). Lagos state covers an area of approximately 3,475.1km². Lagos state is divided into five administrative division which are further divided into 16 local government areas. These local government are: Agege, Alimosho, Ifako-ijaye, Ikeja, Kosofe, Mushin, Oshodi-isolo, shomolu, Apapa, Eti-osa, Lagos island, Lagos mainland, Surulere, Ajeromi-ifelodu, Amuwo-odofin, Ojo, Badagry, Ikorodu. The sampling locations are Igando dumpsite (dumpsite I, dumpsite II) in Igando, Lagos state, and Agodo dumpsite (dumpsite I, dumpsite II, dumpsite III) in Ikotun, Lagos state. Also readings were taken for every 5 meters away from Igando dumpsite

Annual dose equivalent helps to determine the absorbed dose rate per year (μ Sv/yr). The relation below is used to calculate the annual dose equivalent as used to calculate the annual dose equivalent above.

$$D = \sigma \times \mu \times 24 \times 365.25 \text{ (}\mu\text{Sv/yr)} \dots\dots\dots(1) \text{ (Marilyn } et al., 1995, P. 296)$$

Where: D = Annual absorbed dose rate in micro sievert per year.

σ = Absorbed dose rate in micro sievert per hour.

μ = Occupancy factor, 0.2

mean annual absorbed rate is calculated using the equation (2)

$$\bar{X} = \frac{\sum x}{n} \dots\dots\dots(2)$$

Where: \bar{X} = Mean annual absorbed dose rate in micro sievert per year

\sum = Summation

x = Annual absorbed dose rate in micro sievert per year

n = Number of locations

Results And Discussion

Zero meters as indicated in the tables shows the measurement taken inside the dumpsites, other measurement are taken at the specified distances away. From table 2.0, the annual absorbed dose rate inside dumpsite I is 29.80 μ Svyr⁻¹, the mean annual absorbed dose rate from 5 meters to 30 meters away from the dumpsite is 21.04 μ Svyr⁻¹. The graph 1.0 shows that, the farther the distance away from the dumpsite, the lesser the value of the absorbed dose rate. The amount of radiation generated inside dumpsite I with annual absorbed dose rate 29.80 μ Svyr⁻¹ is higher than that generated inside dumpsites II, III, IV and V with annual absorbed dose rate of 28.05 μ Svyr⁻¹, 19.29 μ Svyr⁻¹, 17.53 μ Svyr⁻¹ and 15.78 μ Svyr⁻¹ respectively as seen from tables 3.0 to 6.0, the high value obtained in dumpsite I compared to the other four dumpsites is attributed to the high level of waste generated due to high population, this area is used for residential and commercial purposes. From table 3.0, the annual absorbed dose rate inside dumpsite II is 28.05 μ Svyr⁻¹, the mean annual absorbed dose rate from 5 meters to 30 meters away from the dumpsite is 16.95 μ Svyr⁻¹. The graph 2.0 follows the same trend as that of the figure I except that the values are lower due to the congestion in the area of figure I. Also, the amount of radiation generated inside dumpsite II with annual absorbed dose rate 28.05 μ Svyr⁻¹ is higher than that generated inside dumpsites III, IV and V with annual absorbed dose rate of 19.29 μ Svyr⁻¹, 17.53 μ Svyr⁻¹ and 15.78 μ Svyr⁻¹ respectively. As seen from tables 3.0 to 5.0, the annual absorbed dose rate for dumpsite III, IV and V is 19.29 μ Svyr⁻¹, 17.53 μ Svyr⁻¹, and 15.78 μ Svyr⁻¹ respectively. The low value obtained from the three dumpsites mentioned above is attributed to the low level of waste generated in this area, due to low population. The values obtained are lower compared to the result obtained from Obed *et al* (2005, P.305-312), with an average of 24.5 μ Svyr⁻¹ as reported in situ gamma spectroscopic measurement. All values discussed above is comparable, though lower than both result obtained from the investigation of radiation level in waste dumpsite across Lagos metropolis, Nigeria, with a mean annual absorbed dose rate 21.8 μ Svyr⁻¹ reported by Odunaike *et al.* 2008, P. 174-176) and that obtained from (Obed *et al.* 2005, P. 305-312) with an average of 24.5 μ Svyr⁻¹ as reported in situ gamma spectroscopic measurement. However, world's average is 0.06 μ Gyh⁻¹ approximately 70 μ Svyr⁻¹ (Ademola, 2008, P. 93-99), this is higher than the result obtained in this study.

Conclusion

The radiation dose rate characterization of the waste dumpsites in two cities (Igando and Ikotun) of Lagos Metropolis and some meters away from two of the dumpsites has been investigated in this study, and a mean annual absorbed dose rate

obtained from the five dumpsites is $29.80\mu\text{Svyr}^{-1}$, $28.05\mu\text{Svyr}^{-1}$, $19.25\mu\text{Svyr}^{-1}$, $17.53\mu\text{Svyr}^{-1}$, and $15.78\mu\text{Svyr}^{-1}$ respectively. Despite the fact that all levels of ionizing radiation are hazardous to human health (Imtiaz *et al.*, 2005, P. 169-174) the exposure level of the emitted radiation on the populace of the two cities is low when compared with Nigeria and World average which is $70\mu\text{Svyr}^{-1}$. Hence, fear of serious health hazards arising from the exposure to radiation emanating from these dumpsites should not be entertained. However, focus should be on the proper management of the waste generated in the city to prevent outbreak of mutation, cancerous diseases and also to prevent the radiation emitted from this dumpsites from getting into exposed foods sold in this areas. It will be advantageous to improve the environment and conserve natural resources; these can be achieved by locating waste dumpsites far away from residential areas, discouraging builders not to build houses near dumpsites. Waste to wealth approach should be incorporated by the government as well as carrying out regular investigations to monitor the level of radiation emission from dumpsites in the city to avoid high level of radiation emission outbreak in subsequent years.

References

1. Ademola, J.A, (2008). Exposure to high background radiation level in the tin minning area of Jos Plateau, Nigeria. J. Radiol. Prot, 28: 93-99.
2. Akinloye, M.K. and J.B. Olomo (2005). The radioactivity in some grasses in the environment of nuclear research facilities located within the OAU, Ile-Ife, Nigeria. Nig. J. Phys., 17S: 219-225.
3. Eja, M.E, Alobi, N.O, Ikpeme, E.M, Ogri, O.R and Iyang, A.O (2010). Environmental and public health-related assessment of solid waste management in Uyo, Akwa Ibom State, Nigeria. World Journal of Applied Science and Technology Vol.2. No. 1 (2010). 110-123.
4. Eyebiokin, M.R., A.M. Arogunjo, G. Oboh, F.A. Balogun and A.B. Rabi, (2005). Activity concentrations and absorbed dose equivalent of commonly consumed vegetable in Ondo State, Nigeria. Nig. J. Phys., 17S: 187-191.
5. Farai, I.P and Oni, O.M (2002). Natural radionuclide concentrations in aquatic species and absorbed dose equivalents to the dwellers of the coastal areas of Nigeria. Nig. Journal of Phys. 14, 2002 94-97.
6. Imtiaz, M.A., B. Aleya, A.S. Molla and M.A. Zaman (2005). Measurement of radioactivity in books and calculations of resultant eye doses to readers. Health Phys., 88: 169-174.
7. Inyang, S.O, Inyang, I.S and Egbe, N.O (2009). Radiation exposure levels within timber industries in Calabar, Nigeria. J Med Phys. 34(2): 97-100. doi: 10.4103/0971-6203.51937.
8. Jibiri, N.N., I.P. Farai and S.K. Alausa, (2007). Activity concentration of Ra-226, Th-228 and K-40 in different food crops from a high background radiation area in Bisichi Jos Plateau State, Nigeria. Radiat. Environ. Biophys., 46: 53-59.
9. Jibiri, N.O, Amakom, C.M and Adewuyi, G.O (2010). Radionuclide Contents and Physicochemical Water Quality Indicators in Stream, Well and Borehole Water Sources in High Radiation Area of Abeokuta, Southwestern Nigeria. J. Water Resource and Protection, Vol. 2, 291-297 doi:10.4236/jwarp.
10. Jibiri, N.O, Alausa, S.K, Owofolaju, A.E and Adeniran, A.A (2011). Terrestrial gamma dose rates and physical-chemical properties of farm soils from ex- tin mining locations in Jos-Plateau, Nigeria. African Journal of Environmental Science and Technology Vol. 5(12), pp. 1039-1049.
11. Marilyn, E.N. and G.Q. Maguire Jr., (1995). Radiation Protection in the Health Sciences. 1st Edn. World Scientific Publishing, Singapore, pp: 296. DOI 10.4028/www.scientific.net/AMR.367.801
12. Nworgu, O.D, Osahon, O.D and Obinyan, F.E (2011). Measurement of Gamma Radiation in Automobile Mechanic Workshops in an Area of Benin City, Nigeria. Advanced Materials Research. Volume 367. 801-805.
13. Obed, R.I., I.P. Farai and N.N. Jibiri, (2005). Population dose distribution due to soil radioactivity concentration levels in 18 cities across Nigeria. J. Radiol. Prot., 25: 305-312.
14. Odunaike, R.K, S.K. Alausa, O.A. Oyebanjo, G.C. Ijeoma and A.O. Alo, (2008). Measurement of radiation level in refuse dumps across lagos metropolis, Southwestern Part of Nigeria. Environ. Res. J., 2: 174-176.
15. Ojoawo. S, Agbede. O and Sangodoyin. A (2011). On the Physical Composition of Solid Wastes in Selected Dumpsites of Ogbomosoland, South-Western Nigeria. Journal of Water Resource and Protection, Vol. 3, 661-666 doi:10.4236/jwarp.2011.39076.
16. Tanabe. F (2011). Analysis of Core Melt Accident in Fukushima Daiichi-Unit 1 Nuclear Reactor Journal of Nuclear Science and Technology, volume 48, issue 8, pages 1135 to 1139.

Table 1.0: Locations at which the measurements were taken.

DUMPSITES	LOCATION
DUMPSITE I	IGANDO, LAGOS.
DUMPSITE II	IGANDO, LAGOS.
DUMPSITE III	AGODO, IKOTUN EGBE, LAGOS.
DUMPSITE IV	AGODO, IKOTUN EGBE LAGOS.
DUMPSITE V	AGODO, IKOTUN EGBE LAGOS.

Table 2.0: Radiation measurement for Dumpsite I

LOCATION NUMBER	DISTANCES (D)	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	AVERAGE ABSORBED DOSE RATE (μSvhr^{-1})	ANNUAL ABSORBED DOSE RATE (μSvyr^{-1})
1	0 METERS	0.017	0.016	0.017	0.017	29.80
2	5METER AWAY	0.016	0.016	0.017	0.016	28.05
3	10METERS AWAY	0.014	0.013	0.014	0.014	24.54
4	15METERS AWAY	0.012	0.012	0.011	0.012	21.04
5	20METERS AWAY	0.011	0.010	0.011	0.011	19.29
6	25METERS AWAY	0.010	0.010	0.010	0.010	17.53
7	30METERS AWAY	0.010	0.009	0.009	0.009	15.78

TABLE 3.0: Radiation measurement for Dumpsite II

LOCATION NUMBER	DISTANCES (D)	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	AVERAGE ABSORBED DOSE RATE (μSvhr^{-1})	ANNUAL ABSORBED DOSE RATE (μSvyr^{-1})
8	0 METERS	0.016	0.016	0.015	0.016	28.05
9	5METER AWAY	0.016	0.015	0.015	0.015	26.30
10	10METERS AWAY	0.012	0.013	0.012	0.012	21.04
11	15METERS AWAY	0.010	0.010	0.010	0.010	17.53
12	20METERS AWAY	0.009	0.009	0.010	0.009	15.78
13	25METERS AWAY	0.008	0.007	0.007	0.007	12.27
14	30METERS AWAY	0.005	0.005	0.004	0.005	8.77

TABLE 4.0: Radiation measurement for Dumpsite III

LOCATION NUMBER	DISTANCES (D)	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	AVERAGE ABSORBED DOSE RATE (μSvhr^{-1})	ANNUAL ABSORBED DOSE RATE (μSvyr^{-1})
15	0 METERS	0.011	0.010	0.011	0.011	19.29

TABLE 5.0: Radiation measurement for Dumpsite IV

LOCATION NUMBER	DISTANCES (D)	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	AVERAGE ABSORBED DOSE RATE (μSvhr^{-1})	ANNUAL ABSORBED DOSE RATE (μSvyr^{-1})
16	0 METERS	0.010	0.009	0.011	0.010	17.53

Table 6.0: Radiation measurement for Dumpsite V

LOCATION NUMBER	DISTANCES (D)	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	ABSORBED DOSE RATE (μSvhr^{-1})	AVERAGE ABSORBED DOSE RATE (μSvhr^{-1})	ANNUAL ABSORBED DOSE RATE (μSvyr^{-1})
17	0 METER	0.010	0.008	0.009	0.009	15.78

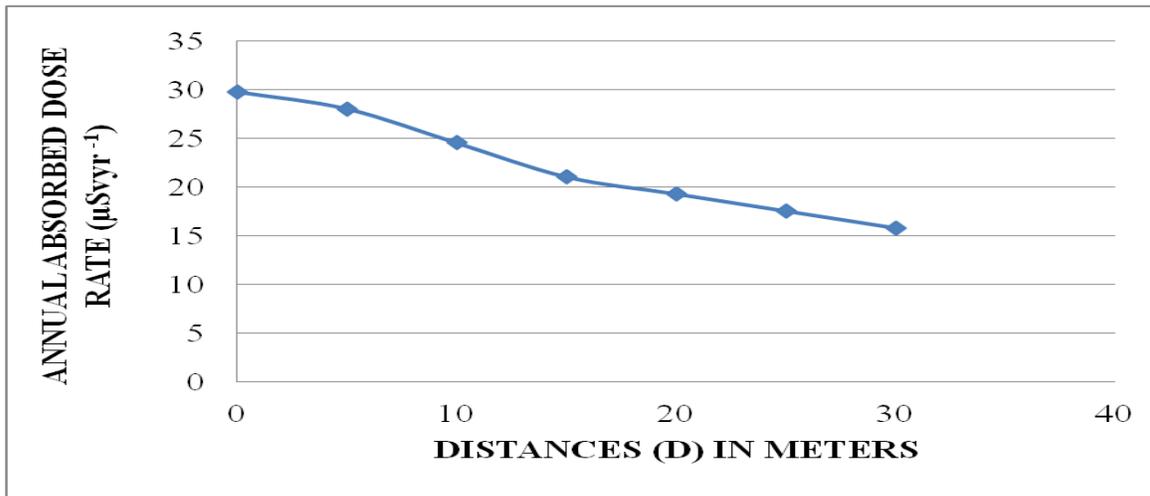


Figure 1.0: Graph of annual absorbed dose rate against distance for dumpsite I

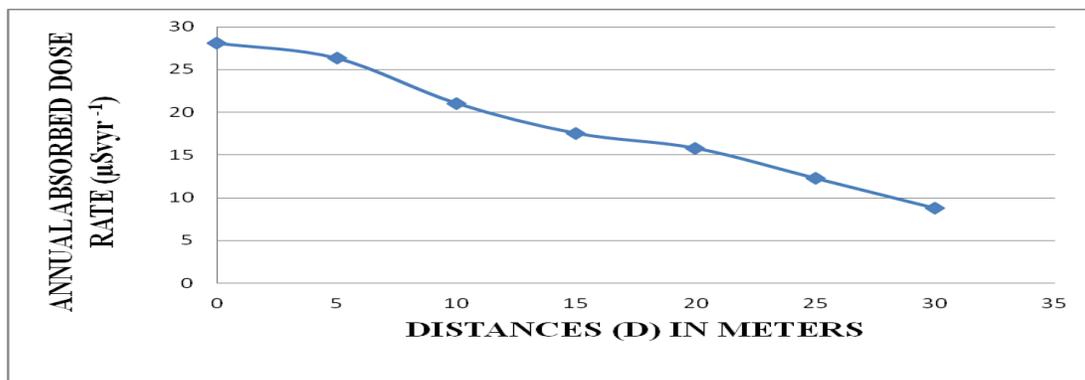


Figure 2.0: Graph of annual absorbed dose rate against distance for dumpsite II

Figure 3.0: Map of the study area



Performance Evaluation of different α value for OFDM System

Dr. K.Elangovan

Dept. of Computer Science & Engineering
Bharathidasan University
Trichirappalli

Abstract:

Orthogonal Frequency Division Multiplexing (OFDM) has recently been applied in wireless communication systems due to its high data rate transmission capability with high bandwidth efficiency and its robustness to multi-path delay. Fading is the one of the major aspect which is considered in the receiver. In this paper the Performance Evaluation of α (0.05, 0.005 and 0.0005) values for OFDM System using LMS Algorithm. These different α value is considered in this work and their performances are statistically compared by using computer simulations.

Keywords: OFDM, BPSK and QPSK modulations and LMS Algorithm

1.Introduction

Orthogonal frequency division multiplexing (OFDM) is a modulation and multiple access technique that has been explored for over 20 years [1]. Only recently has it been finding its way into commercial communications systems, as Moore's Law has driven down the cost of the signal processing that is needed to implement OFDM based systems. OFDM, or multitone modulation, is presently used in a number of commercial wired and wireless applications. On the wired side, it is used for a variant of digital subscriber line (DSL). For wireless, OFDM is the basis for several television and radio broadcast applications, including the European digital broadcast television standard, as well as digital radio in North America. OFDM is also utilized in several fixed wireless systems and wireless local area network products.

All modern mobile wireless systems employ a variety of techniques to combat the challenges of the wireless channel[2]. Some techniques are more effective than others, with the effectiveness depending on both the air interface and system architecture approach taken to satisfy the requirements of the services being offered. As mobile systems evolved from analog to digital, more sophisticated signal processing techniques have been employed to overcome the wireless environment. These techniques include diversity, equalization, channel or error correction coding, spread spectrum, interleaving, and more recently, space time coding[3]. Diversity has long been used to help mitigate the multipath induced fading that results from users' mobility. The simplest diversity technique, spatial diversity, involves the use of two or more receive antennas at a base station, separated by some distance, say on the order of five to ten wavelengths. The signal from the mobile will generally follow separate paths to each antenna. This relatively low cost approach yields significant performance improvement by taking advantage of the statistical likelihood that the paths are not highly correlated with each other. When one antenna is in a fade, the other one will generally not be [3].

A striking result is that the bit error probability of quadrature phase shift keying (QPSK) is identical to binary phase shift keying (BPSK), but twice as much data can be sent in the same bandwidth. Thus when compared to BPSK, QPSK provides twice the spectral efficiency with exactly the same energy efficiency. Similar BPSK, QPSK can also be differently encoded to allow non coherent detection[4].

Phase Shift Keying is a digital modulation scheme that conveys data by changing, or modulating, the phase of a reference carrier signal. Any digital modulation scheme uses a finite number of distinct signals to represent digital data. PSK uses a finite number of phases, each assigned a unique pattern of binary bits. Usually, each phase encodes an equal number of bits. Each pattern of bits forms the symbol that is represented by the particular phase. BPSK and QPSK are four major Modulation Techniques used in most of the applications.

2. Basic OFDM Block Diagram

The figure 2.1. shown below is the basic block diagram of the OFDM transmitter and receiver.

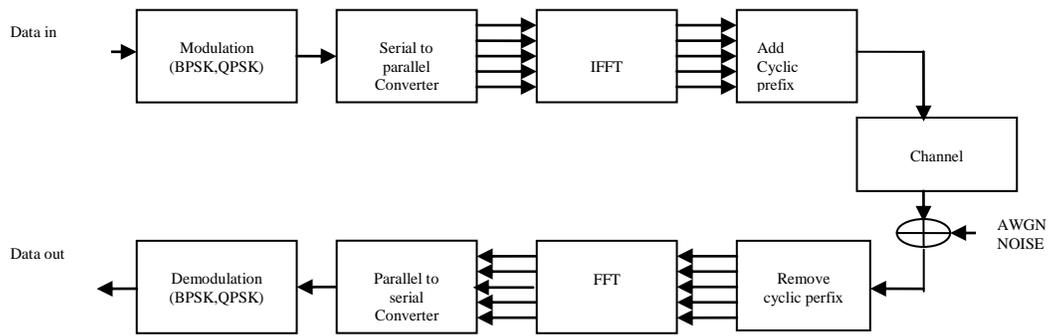


Fig. 2.1. OFDM Block Diagram.

3.Modulation Of Data

The data to be transmitted on each carrier is to be modulated using any modulation technique depending upon the user requirements. If a single bit is to be transmitted over a modulated symbol BPSK can be used. To transmit two data bits per symbol, QPSK can be made use of. In case of transmission of three bits per modulated symbol, 8- Phase Shift Keying (8-PSK) is used.

In Binary Phase Shift Keying (BPSK), the phase of a constant amplitude carrier signal is switched between two values according to the two possible signals m_1 and m_2 corresponding to binary 1 and 0, respectively. Normally, two phases are separated by 180° . If the sinusoidal carrier has an amplitude A_c and energy per bit $E_b = \frac{1}{2} A_c^2 T_b$, then the transmitter BPSK signal is either

$$S_{BPSK}(t) = (2E_b/T_b)^{1/2} \cos(2\pi f_c t + \theta_c) \quad 0 < t < T_b \text{ (for binary 1)} \quad (3.1)$$

or

$$S_{BPSK}(t) = -(2E_b/T_b)^{1/2} \cos(2\pi f_c t + \theta_c) \quad 0 < t < T_b \text{ (for binary 0)} \quad (3.2)$$

The BPSK signal is equivalent to a double sideband suppressed carrier amplitude modulated waveform, where $\cos(2\pi f_c t)$ is applied as a carrier, and the data signal $m(t)$ is applied as the modulating waveform.

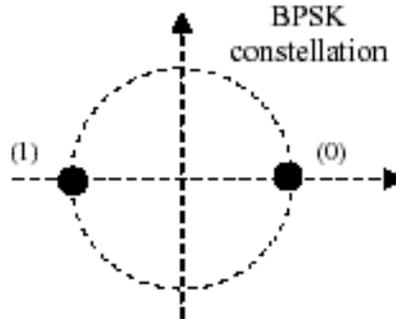


Fig. 3.1. Mapping bits into BPSK constellation.

Quadrature phase shift keying (QPSK), has twice the bandwidth efficiency of BPSK, since two bits are transmitted in a single modulation symbol. The phase of the carrier takes on one of four equally spaced values, such as 0° , 90° , 180° and 270° , where each value of phase corresponds to a unique pair of message bits. The QPSK signal for this set of symbol states may be defined as,

$$S_{QPSK}(t) = (2E_s/T_s)^{1/2} \cos(2\pi f_c t + (i-1)\pi/2) \quad 0 < t < T_s \quad i=1,2,3,4 \quad (3.3)$$

Where T_s is the symbol duration and is equal to twice the bit period.

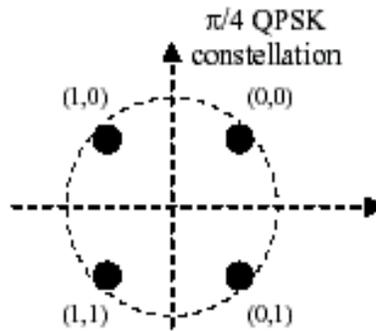


Fig. 3.2. Mapping bits into QPSK constellation.

A striking result is that the bit error probability of QPSK is identical to BPSK, but twice as much data can be sent in the same bandwidth. Thus when compared to BPSK, QPSK provides twice the spectral efficiency with exactly the same energy efficiency. Similar to BPSK, QPSK can also be differently encoded to allow non coherent detection.

4.Channel Estimation using LMS Algorithm

Channel estimation has been studied using various algorithms by an assortment of researchers[5]. In this work Channel estimation is done with the help of LMS algorithms. These techniques have shown better results than the conventional techniques. The suitability of the proposed model has been investigated and quantified. We present the results of simulations and analysis, in next chapter.

A most robust equalizer is the LMS equalizer where the criterion used is the minimization of the Mean Square Error (MSE) between the desired equalizer output and the actual equalizer output. Using the notations given below, the LMS algorithm can be readily understood.

New weights = Previous weights + $(\alpha) \times$ (Previous error) \times (Current input vector)

Where

Previous error = Previous desired output – Previous actual output

where α is constant and the constant may be adjusted by the algorithm to control the variation between the filter weights on successive iterations. The block diagram of LMS equalizer is shown.

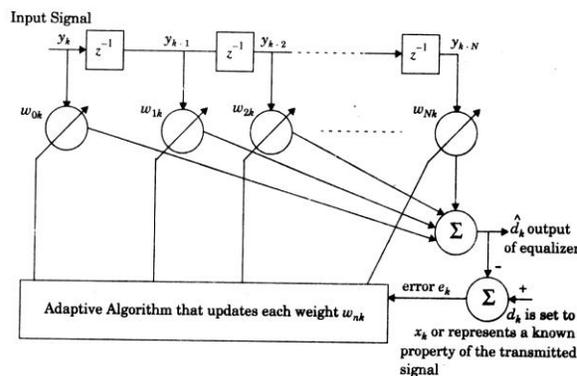


Fig. 4.1 LMS Equalizer.

Referring to the figure 3.1, the prediction error is given by

$$e_k = d_k - \hat{d}_k = x_k - \hat{d}_k \tag{4.1}$$

and

$$e_k = x_k - y_k^T w_k = x_k - w_k^T y_k \tag{4.2}$$

To compute the mean square error $|e_k|^2$ at time instant k, the mean square error function is given by

$$\xi = E[e_k * e_k] \tag{4.3}$$

The LMS algorithm seeks to minimize the mean square error given in the equation (4.3).

For a specific channel condition, the prediction error e_k is dependent on the tap gain vector w_N , so the MSE of an equalizer is a function of w_N . Let the cost function $J(w_N)$ denote the mean squared error as a function of tap gain vector w_N . In order to maximize the MSE, it is required to set the derivative of the equation (4.4) to zero.

$$\frac{\partial}{\partial w_N} J(w_N) = -2p_N + 2R_{NN}w_N = 0 \tag{4.4}$$

By simplifying the above equation we get,

$$R_{NN}\hat{w}_N = p_N \tag{4.5}$$

Equation (4.5) is a classic result, and is called the normal equation, since the error is minimized and is made orthogonal (normal) to the projection related to the desired signal X_k . When equation (4.5) is satisfied, the MMSE of the equalizer is

$$J_{opt} = J(\hat{w}_N) = E[x_k x_k^*] - P_N^T \hat{w}_N \tag{4.6}$$

To obtain the optimal tap gain vector \hat{w}_N , the normal equation in (4.5) must be solved iteratively as the equalizer converges to the acceptably small value of J_{opt} . There are several ways to do this and many variants of the LMS algorithm have been built upon the solution of equation (3.6). One obvious technique is to calculate

$$\hat{w} = R_{NN}^{-1} P_N \tag{4.7}$$

However, inverting a matrix requires $O(N^3)$ arithmetic operations. Other methods such as Gaussian elimination and Cholesky factorization require $O(N^2)$ operations per iteration. The advantage of these methods which directly solve equation (4.7) is that only N symbol inputs are required to solve the normal equation. Consequently, a long training sequence is not necessary.

In practice, the minimization of the MSE is carried out recursively, and may be performed by the use of the stochastic gradient algorithm. This is more commonly called the least mean square (LMS) algorithm. The LMS algorithm is the simplest equalization algorithm and requires only $2N+1$ operation per iteration. The filter weights are updated by the updated equations given below. Letting the variable n denote the sequence of iterations, LMS is computed iteratively by

$$\hat{d}_k(n) = w_N^T(n) y_N(n) \tag{4.8a}$$

$$e_k(n) = x_k(n) - \hat{d}_k(n) \tag{4.8b}$$

$$w_N(n+1) = w_N(n) - \alpha e_k^*(n) y_N(n) \tag{4.8c}$$

where the subscript N denotes the number of delay stages in the equalizer, and α is the step size which controls the convergence rate and stability of the algorithm.

The LMS equalizer maximizes the signal to distortion ratio at its output within the constraints of the equalizer filter length. If an input signal has a time dispersion characteristic that is greater than the propagation delay through the equalizer, then the equalizer will be unable to reduce distortion. The convergence rate of the LMS algorithm is due to the fact that there is only one parameter, the step size α that controls the adaptation rate. To prevent adaptation from becoming unstable, the value of α is chosen from

$$0 < \alpha < 2 / \sum_{i=1}^N \lambda_i \tag{4.9}$$

where λ_i is the i th eigen value of the covariance matrix R_{NN} . Since

$$\sum_{i=1}^N \lambda_i = y_N^T(n) y_N(n) \tag{4.10}$$

5. Simulation Results

Fig. 5.1, Fig. 5.2, show the signal to noise ratio (SNR) vs Bit Error Rate (BER) curve for different modulation techniques with AWGN. It is noteworthy that the BER value keeps on decreasing with a considerable increase in SNR value, in Additive White Gaussian Noise (AWGN) Channel environments.

Fig. 5.1 shows BER values for different SNR values, for BPSK modulation which delves that while SNR increases the BER is being decreased. At the transmitting point of $0db$ SNR value, it has $10^{-1.3}$ BER value and at the receiving point of $10db$ SNR value, it has $10^{-5.8}$ BER value. The similar value of SNR is considered for all modulation techniques.

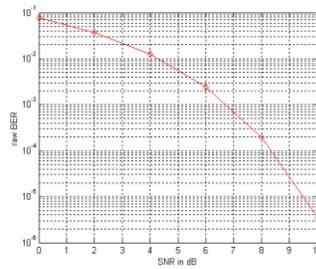


Fig.5.1.BPSK with AWGN.

Fig. 4.2 shows BER values for different SNR values, for QPSK modulation which delves that while SNR increases the BER is being decreased. At the transmitting point of 0db SNR value, it has $10^{-0.5}$ BER value and at the receiving point of 14dB SNR value, it has $10^{-6.3}$ BER value.

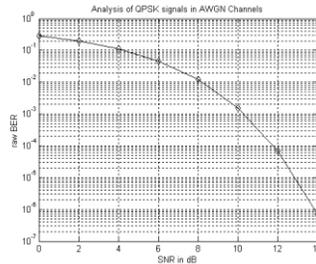


Fig. 5.2. QPSK with AWGN.

From the above two figures, we can infer that QPSK has lesser Bit Error Rate than BPSK.

Comparison of BPSK and QPSK modulation techniques

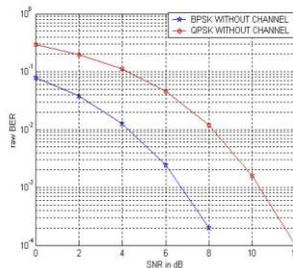


Fig.5.3 Comparative graph for QPSK and BPSK with AWGN.

Table 5.1 gives a comparison of BPSK and QPSK techniques. QPSK modulation have more BER compared to BPSK modulation.

Table 5.1. Comparison of BPSK and QPSK techniques.

S. No	Modulation Technique	SNR (at transmitter)	BER (at Transmitter)	SNR (at Receiver)	BER (at Receiver)
1	BPSK	0dB	$10^{-1.3}$	10dB	$10^{-5.8}$
2	QPSK	0dB	$10^{-0.5}$	14dB	$10^{-6.5}$

Comparison of different α value using LMS algorithms

Here, we compare the predictions of the channel coefficients using LMS algorithm. Before going to the analysis, we first applied LMS algorithm to predict Channel coefficient by using different α values. It is found that the prediction is better when we take α to be 0.0005 and hence we take $\alpha = 0.0005$ for comparing predicted and original values of the channel coefficients for each of LMS algorithms. Hence we can infer that $\alpha = 0.0005$ is more effective than the other $\alpha = 0.05$ and $\alpha = 0.005$.

Fig. 5.4 shows the original and predicted values of channel coefficients using LMS algorithm with $\alpha = 0.05$. Separate comparisons are made for the real and imaginary parts of the channel coefficients. It is observed that the difference between the actual and predicted values obtained using LMS algorithm is significant and hence we can infer that the tracking may not be efficient if we use LMS algorithm.

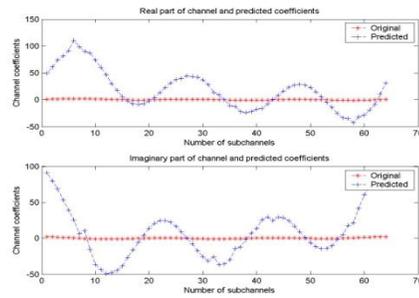


Fig. 5.4 Original and predicted channel coefficients for LMS with $\alpha= 0.05$.

Fig. 5.5 gives the plot for predicted SNR and BER values, using LMS algorithm, for BPSK modulation on wireless channel with OFDM. Here, the value of α is taken as 0.05. It is observed that the achievable BER is varying dramatically with respect to perfect α value.

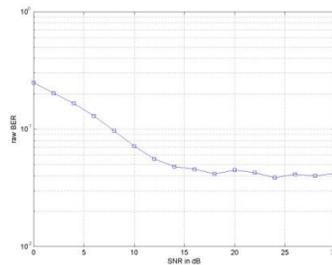


Fig. 5.5 SNR vs BER for LMS with $\alpha= 0.05$.

Fig. 5.6 shows the original and predicted values of channel coefficients using LMS algorithm with $\alpha = 0.005$. Separate comparisons are made for the real and imaginary parts of the channel coefficients. It is observed that the difference between the actual and predicted values obtained using LMS algorithm is significant and hence we can infer that the tracking may not be efficient if we use LMS algorithm.

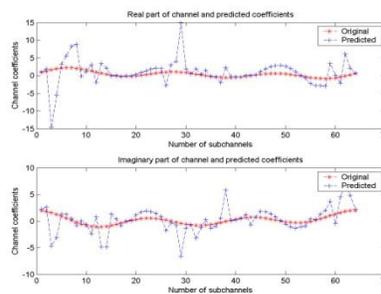


Fig. 5.7 Original and predicted channel coefficients for LMS with $\alpha= 0.005$.

Fig. 5.8 gives the plot for predicted SNR and BER values, using LMS algorithm, for BPSK modulation on wireless channel with OFDM. Here, the value of α is taken as 0.005. It is observed that the achievable BER is varying dramatically with respect to perfect α value.

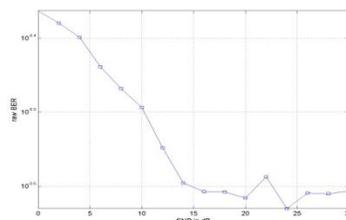


Fig. 5.8 SNR vs BER for LMS with $\alpha= 0.005$.

Fig. 5.9 shows the original and predicted values of channel coefficients using LMS algorithm with $\alpha = 0.005$. Separate comparisons are made for the real and imaginary parts of the channel coefficients. It is observed that the difference between the actual and predicted values obtained using LMS algorithm is significant and hence we can infer that the tracking may not be efficient if we use LMS algorithm.

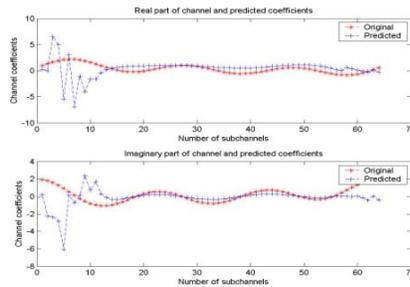


Fig. 5.9 Original and predicted channel coefficients for LMS with $\alpha= 0.0005$.

Fig. 5.10 gives the plot for predicted SNR and BER values, using LMS algorithm, for BPSK modulation on wireless channel with OFDM. Here, the value of α is taken as 0.0005. It is observed that the achievable BER is varying dramatically with respect to perfect α value.

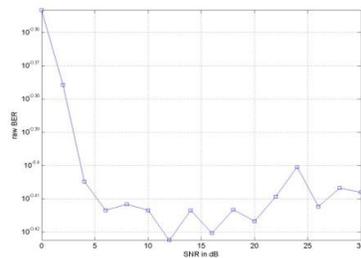


Fig. 5.10 SNR vs BER for LMS with $\alpha= 0.0005$.

5. Conclusion

In this paper the Performance Evaluation of α (0.05, 0.005 and 0.0005) values for OFDM System using LMS Algorithm. These different α value is considered in this work and their performances are statistically compared by using computer simulations. Here, we compare the predictions of the channel coefficients using LMS algorithm. Before going to the analysis, we first applied LMS algorithm to predict Channel coefficient by using different α values. It is found that the prediction is better when we take α to be 0.0005 and hence we take $\alpha = 0.0005$ for comparing predicted and original values of the channel coefficients for each of LMS algorithms. Hence we can infer that $\alpha = 0.0005$ is more effective than the other $\alpha = 0.05$ and $\alpha = 0.005$.

Reference

1. **Ahn, J., Lee, H. S.(1993)**, “Frequency Domain Equalisation of OFDM signals over Frequency Nonselective Rayleigh Fading Channel”, *Electronics Letters*, Vol. 29, No.16, pp. 1476 – 1477. **Bernard Sklar**, Rayleigh Fading Channels in Mobile Digital Communication Systems (Part I and II), July 1997.
2. **P. Schramm and R. Muller(1998)**, “Pilot symbol assisted BPSK on Rayleigh fading channels with diversity: Performance analysis and parameter optimization,” *IEEE Transaction on communication*, vol. 46, no. 12, pp. 1560–1563.
3. **Cimini, L.J.(1985)**, “Analysis and simulation of digital mobile channel using orthogonal frequency division multiplexing”, *IEEE Transaction on communication*, Vol.33.no.7, pp 665-675..
4. **Rappaport,T.S.,(2009)** , “Wireless Communication”, Prentice Hall.
5. **S.Galih, R.Karlina, F.Nugroho, A.Irawan, T.Adiono and A. Kurniawan(2009)**, “High mobility data pilot Based channel estimation for downlink OFDMA system based on IEEE 802.16e standard “,2009 international conference on Electrical Engineering and Information 5-7 Selangor, Malaysia.

Image Segmentation Using Active Contour Model

Abhinav Chopra¹, Bharat Raju Dandu²

^{1,2}Amity University Noida Sec-125,
Uttar Pradesh 201 301, India

Abstract:

Image segmentation is one of the substantial techniques in the field of image processing. It is vastly used for medical purposes, tracking growth of tumor for surgical planning and simulation. Active contours or snakes are used extensively for image segmentation and processing applications, particularly to locate object boundaries. Active contours are regarded as promising and vigorously researched model-based approach to computer assisted medical image analysis. However, its utility is limited due to poor convergence of concavities and small capture range. This paper shows the application of an external force that largely solves both problems. This external force is called gradient vector flow (GVF). Using several examples to show that, GVF because of its large capture range moves snakes into boundary concavities.

Keywords: Active contour models, edge detection, gradient vector flow, image segmentation, snakes

1. Introduction

Breaking an image into its constituents is called segmentation, which basically means separating the background and foreground. Active contours [1] or snakes provide an effective way of segmentation [2] of curves defined within the image domain that can move under the influence of external and internal forces. These forces are defined such that the snake will shrink wrap to an object boundary. This method is widely used in many applications, including motion tracking, edge detection and segmentation.

Active contours are extensively used in the field of digital image processing to find the contour of an object by forming a snake around its boundary. It involves running a low level image processing task such as canny edge detection which is mostly unsuccessful because often the edge is not continuous, i.e. the image might be noisy or there might be concavities present in the image. Hence Active Contours are an advanced way of image segmentation as it adds certain properties to the image before performing segmentation which makes the process of locating the boundary comparatively easier.

The major problem that was encountered in the traditional snake model was that the initial contour must in general, be close to the true boundary or else it will predict an incorrect result as shown in figure 1. Most of the methods that are proposed to solve the above problems are ineffective in solving this issue and end up creating more difficulties.

In this paper we present a class of external force for active contour model that answers both the problems mentioned above. This external force, which we call gradient vector flow (GVF)[5], is computed as a diffusion of the gradient vectors of a gray-level or binary edge map derived from the image. The resultant field has a large capture range and forces active contours into concave regions, hence solving both the problems. Since the external forces cannot be written as the negative gradient of a potential function, GVF snake is different from all other snake models used before. The major advantages of GVF snake over the other snake models are its intensity for initialisation and its ability to move into concavities. The GVF snake can be initialized far from the boundary since it has a large capture range. And unlike pressure forces, the GVF snake does not require prior knowledge about when to shrink or expand towards the boundary.

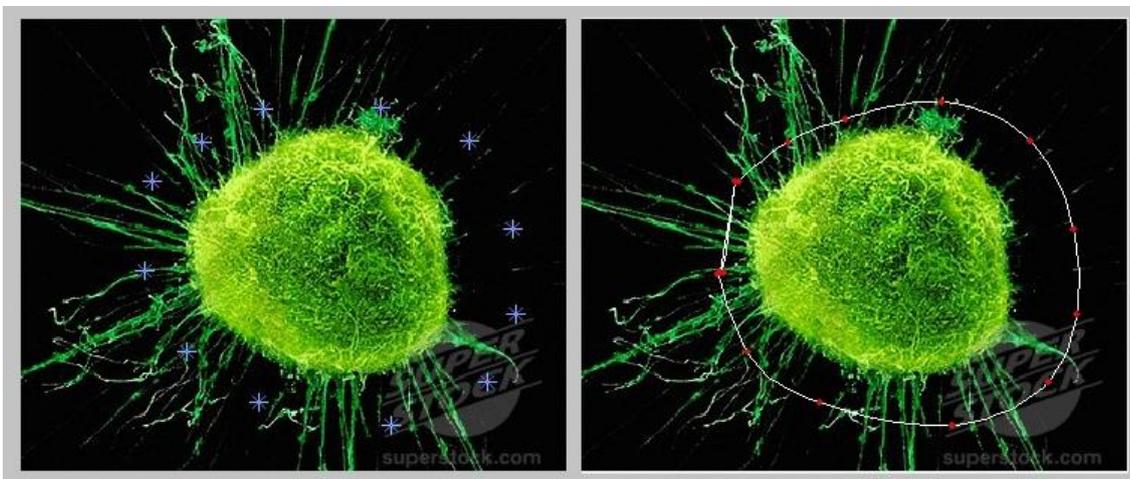


Fig1. Shows that major drawback of traditional active contour model i.e. small capture range

2. Literature Survey

2.1. Original Snake Model by Kass

The concept of active contours was introduced by Kass, in the seminal paper “Snakes: Active Contour Models” [Kass-88]. The paper was extremely influential and has since then been a major topic for research. As we have already discussed a snake is a parametric curve which tries to move into a position where its energy is minimized. Kass et al. introduced the following energy functional for calculating the snake energy.

$$E_{snake} = E_{internal} + E_{external} + E_{constraint}$$

The snake energy consists of three terms. The first term E_{int} represents the internal energy of the snake while the second term E_{img} denotes the image forces, the last term E_{con} gives rise to external constraint forces. The sum of the image forces E_{img} and the external constraint forces E_{con} is also simply known as the external snake forces, denoted by E_{ext} .

Internal Energy (E_{int}) depends on the intrinsic properties of the curve and is the sum of elastic energy and bending energy

$$E_{int} = E_{elastic} + E_{bending} = \int_s \frac{1}{2} (\alpha |v_s|^2 + \beta |v_{ss}|^2) ds$$

External energy (E_{ext}) of the contour is derived from the image so that it takes on its smaller values at the function of interest such as boundaries. Define a function $E_{image}(x,y)$ so that it takes on its smaller values at the features of interest, such as boundaries

$$E_{ext} = \int_s E_{image}(v(s)) ds$$

3. Gradient Vector Flow Model (GVF)

As discussed earlier one of the major problems faced during the implementation of active contour model was the poor convergence of this snake, this is because the forces point horizontally in opposite direction. Another weakness of the traditional snake model is that it has a limited capture range as seen in figure 1, this can be explained by the simple theory that the magnitudes of the external forces die out quite rapidly away from the object boundary. The boundary localization will become less accurate and distinct. To overcome this problem we introduce the GVF snake model.

The gradient vector flow snake is used in order to increase the capture range and improve the snakes ability to move into boundary concavities. The capture range of the original snake is generally limited to the vicinity of the desired contour. Furthermore the original snake has problems with moving into concave regions e.g. moving into the concave region of an U-shaped object, the gradient vector flow snake handles these problems by introducing a new external force. By minimizing an energy function we can derive a new vector field by using this external. We call these vector field as gradient vector flow fields

Gradient Vector Flow: - The GVF field is defined to be a vector field

$$V(x,y) = (u(x, y), v(x, y))$$

Force equation for GVF snake is, $\alpha v_{ss} - \beta v_{ssss} + V = 0$

$V(x,y)$ is defined such that it minimizes the energy functional,

$$E = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 dx dy$$

$f(x,y)$ is the edge map of the image.

The above equations are solved iteratively using time derivative of u and v . These equations provide further intuition behind the GVF formulation. We note that in the homogenous region the second term in both regions is zero because the gradient of $f(x, y)$ is zero.

After the application of the GVF snake model we can see that the snake can move towards object that are too far away from the boundary. Secondly, that now snakes can move into boundary concavities. Hence we can conclude that both the problems of the traditional snake models have been successfully resolved

4. Greedy Snake Algorithm

A greedy algorithm makes locally optimal choices, hoping that the final solution will be globally optimum. It is a feature extraction technique widely used in image segmentation. It works like a elastic band being stretched around an object and then being released. Initial points defined around feature to be extracted explicitly defined and the Pre-defined number of points are generated.

Points are calculated by an **Iterative Process**:

- Energy function for each point in the local neighborhood is calculated
- The point is moved to the next point with lowest energy function
- This process is repeated for every point
- Iteration is done until termination condition met
- Defined number of iterations
- Stability of the position of the points

The final step in the iteration of the greedy snake algorithm consists of checking whether the number of points moved in the iteration is below the threshold. This is used as a stopping criterion as the snake is presumed to have reached minimum energy when most of the control points have stopped moving.

5. Results and Discussion

The application of GVF to the traditional active contour model introduced by Kass significantly improves the quality and accuracy of image segmentation. The major problem that was encountered was that of the small capture range, this is eliminated by using the external force supplied by the gradient vector

In Fig 2 and Fig 3 we can see the different stages of segmentation for a cancer cell. First step involves localization of the object i.e. forming a rough outline along the boundary of the object. Next it moves through an iterative process and finds the boundary of the tumor

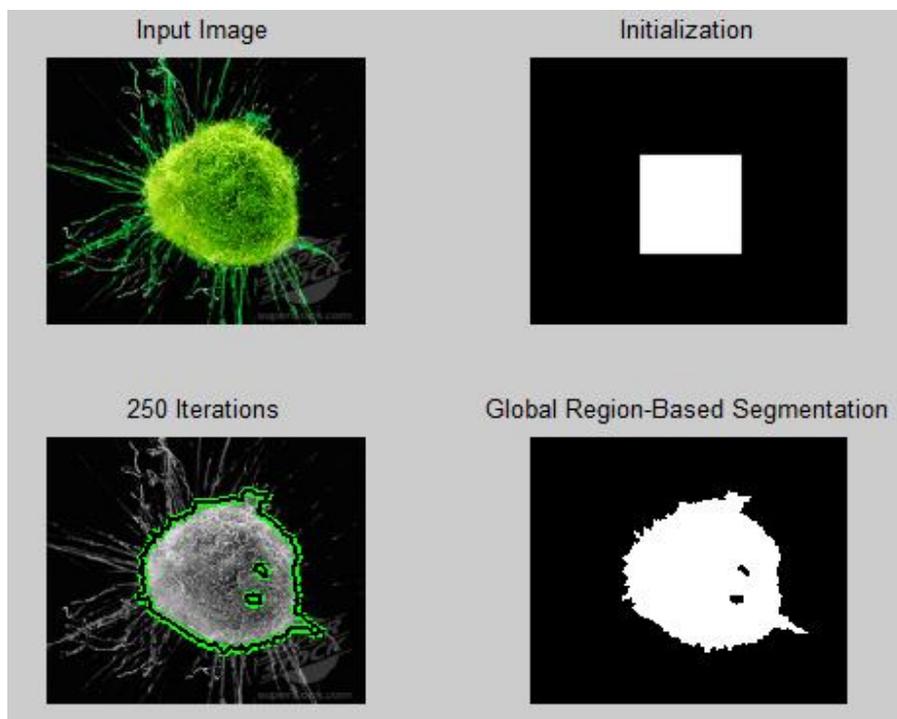


Fig 2. Stages of image segmentation for a cancer cell present in the lung

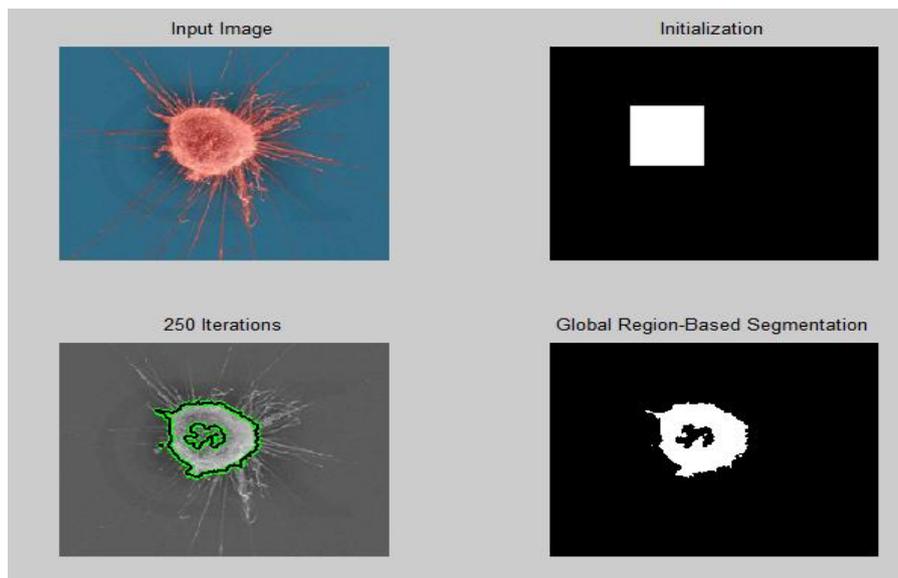


Fig 3. Stages of image segmentation of breast cancer cell

6. Conclusion

The complexity and variability of anatomic structures, poses many challenging problems for computer vision community in search for a good segmentation model. Snakes have proven to be very attractive approach and have produced good results in many medical imaging applications. We have successfully applied the Gradient Vector Flow (GVF) model. This new external force model for active contours successfully allows convergence to boundary concavities and also provides a large capture range for the snake. GVF model provides a collective way of treating visual problems that were till now treated differently. We can also conclude that although the GVF snake model is slower than the traditional snake model it is much more accurate. However, continued development, improvement and refinement of these methods is an important research area with the goal of producing automated, accurate and robust segmentation models

7. References

- [1] W.F. Ames. Numerical Methods for Partial Differential Equations. Academic Press, 1977.
- [2] R. Aris. Discrete Dynamic Programming. Blaisdel, 1964.
- [3] D.R. Bailes and C.J. Taylor. The Use of Symmetry Chords for Expressing Grey Level Constraints. In Proc. British Machine Vision Conference, pages296-305. Springer, 1992.
- [4] D.H. Ballard. Generalizing the Hough Transform to Detect Arbitrary Shapes. Pattern Recognition, 13:111-122, 1981.
- [5] R. Bellman. Dynamic Programming. Princeton University Press, 1957.
- [6] A. Blake and M. Isard. Active Contours. Springer, 1998.
- [7] G.J. Borse. Numerical Methods with Matlab. PWS, 1997.
- [8] J.P Boyd. Chebyshev and Fourier Spectral Methods. Springer, 1985.
- [9] W. Cheney and D. Kincaid. Numerical Mathematics and Computing. Brook-s/Cole, 1985.
- [10] J.H. Chuang. Potential-Based Approach for Shape Matching and Recognition. Pattern Recognition, 29:463-470, 1996.
- [11] P.G. Ciarlet and J.L. Lions. Handbook of Numerical Analysis, Volume 5. El-sevier, 1997.
- [12] L.D. Cohen. NOTE On Active Contour Models and Balloons. Computer Vision and Image Processing: Image Understanding, 53:211-218, March 1991.
- [13] L.D. Cohen and I.Cohen. Finit-Element Methods for Active Contour Models and Baloons for 2-D and 3-D Images. IEEE Transactions on Pattern Analysis and Machine Intellingence, 15:1131-1147, November 1993.
- [14] C. de Boor. A Practical Guide to Splines. Springer, 1978.
- [15] A. Chakraborty et.al. Deformable boundary finding influenced by region homogeneity. Available at the web site of the Yale University, March 1994.
- [16] A.A. Amini et.al. Using Dynamic Programming for Solving Variatinal Problems in Vision. IEEE Transactions on Pattern Analysis and Machine Intellingence, 12:855-867, September 1990.
- [17] A.K. Jain et.al. Object Matching Using Deformable Templates. IEEE Transactions on Pattern Analysis and Machine Intellingence, 18:267-278, March1996

Sbpgp Security Model Using Iodmrp

Meenakshi Mehla¹, Himani Mann²

¹C.S.E, Kurukshetra University,

JMIT(Radaur), Haryana, India

²C.S.E, Kurukshetra University,

Kurukshetra, Haryana, Pin-136118, India

Abstract

Today's world is mobile and using ad hoc network. Routing is the reactive on-demand philosophy where routes are established only when required. Security is one of the most important concepts in ad hoc networks. So different strategies for security are suggested. The study here proposes a theory in this paper based on PKI with IODMRP. The study should help in making protocols more robust against attacks and standardize parameters for security in routing protocols. PKI, PGP and SPGP plays the vital role in terms of security. It is easy to manage the security of a fixed network but for a mobile and dynamically changing network it is very cumbersome. Thus in this current paper we are focus on the security with Public key infrastructures and its various types that can help to maintain the security in the Mobile adhoc network.

Keywords: IODMRP MANETS, PKI, PGP, SPGP

1. Introduction

Mobile Ad-Hoc Networks are autonomous and self-configuring wireless systems. MANETs consist of mobile nodes that are free to move in and out of the network. These node can be mobile phone, system etc. Mobility affects the power indulgence of the nodes in a MANET and has great impact on its security. IODMRP is a more efficient multicast routing protocol It chooses partial forwarding nodes to relay packets, the number of which is decided by probabilistic forwarding algorithm. Security is provided by use of PKI, PGP and SBPGP. SBPGP has better performance than others. Security is main concern in Manets because nodes are mobile.

1.2 Security threat

The two broad classes of network attacks are active attacks and passive attacks.

Passive Attack: An attack in which an unauthorized party gains access to an asset and does not modify its content (i.e., eavesdropping). Passive attacks can be either eavesdropping or traffic analysis (sometimes called traffic flow analysis). These two passive attacks are described as

- **Eavesdropping:** The attacker checks message contents on transmission line.
- **Traffic analysis:** The attacker, in a more subtle way, gains intelligence by monitor the transmissions for

patterns of communication. A considerable amount of information is contained in the flow of messages between communicating parties.

Active Attack: in this attack an unauthorized person modifies the message contents, files, etc. such type of attacks are easily detectable but can't be prevented. Active attacks may take the form of one of four types masquerading, replay, message modification, and denial-of-service (DoS). These attacks are summarized as:

- **Masquerading:** The attacker imitate an authorized user and thereby gains certain unauthorized rights
- **Message modification:** The attacker alters a legitimate message by deleting, adding to, changing, or reordering it.
- **Denial-of-service:** The attacker prevents or prohibits the normal use or management of communications facilities.

1.3 Related Work

Many security models have been used PKI, PGP. every model gives different types of performance matrices. Maqsood Razi and Jawaid Quamar proposes paper on hybrid Cryptography Model for Managing Security in Dynamic Topology of MANET. Kamal Kumar Chauhan¹, Amit Kumar Singh Sanger², Virendra Singh Kushwah³ proposes Securing On-Demand Source Routing in MANETs. They conclude a protocol to secure on demand source routing in MANETs that fulfills the security requirements. Our protocol uses one way hash function to maintain the integrity of message. Therefore, deletion of a node from or any kind of modification in route control packet can be detected. Using Public Key Cryptography (PKC), nodes can negotiate the session key for secure communication that fulfills the requirement of confidentiality. Source, destination and intermediate nodes in route list authenticate others

nodes by verifying signature. Security analysis results show that protocol provides the security against many attacks such as reply attack, rushing attack, IP spoofing and man in the middle attack. Our protocol is based on Public Key Cryptography. Asymmetrical algorithms require more calculation than the symmetrical algorithms. Therefore, it consumes much battery power than protocols based on symmetric algorithms.

2. Proposed Technique

There are a number of proposed solutions for security authentication and key management in MANET. Proposed authentication architecture for MANET, describing the formats of messages, together with protocols which achieve authentication as in the architecture can accommodate different authentication schemes. One quite useful approach to the problem comprises PGP-based schemes

2.1 PGP-Based Solutions

The 'Public Key Infrastructure' (PKI) is the most scalable form of key management. Several different PKI techniques exist, such as SPKI, PGP and X.509. Various forms of these PKI techniques have been proposed for use in ad-hoc networks. Ref. [9] on security architecture proposes the use of a group-oriented PKI for large group formation. The leader of the group acts as a 'Certificate Authority' (CA), which issues group membership certificates. These are said to be SPKI-style certificates. They certify that the public key in the certificate belongs to a group member. However, this is not useful for two-party communications or non group-oriented tasks. On self-organized public key certificate management works like PGP [9], which allows users to create, store, distribute, and revoke their public keys without the help of any trusted authority or fixed server. This system does not assign specific missions to a node or subset of nodes (i.e. all the nodes have the same role). In this system, like in, users' public and private keys are created by the users themselves. It is assumed that each honest user owns a single mobile node. Hence the same identifier is used for the user and the other node (i.e. both being denoted by u). Unlike in PGP, where certificates are mainly stored in centralized certificate repositories, certificates in proposed system are stored and distributed by the nodes in a fully self-organized manner. Each certificate is issued with a limited validity period and therefore contains its issuing and expiration times. Before a certificate expires, its issuer issues an updated version of the same certificate, which contains an extended expiration time. This updated version is called the certificate update. Each node periodically issues certificate updates, as long as its owner considers that the user-key bindings contained in these certificates are correct. In this system, key authentication is performed via chains of public-key certificates in the following way: When a user u wants to obtain the public key of another user v , he / she acquires a chain of valid public-key certificates such that

1. The first certificate of the chain can be directly verified by u , by using a public key that u holds and trusts (e.g. her own public key).
2. Each remaining certificate can be verified using the public key contained in the previous certificate of the chain.
3. The last certificate contains the public key of the target user v .

In this system, the certificate revocation is an important mechanism. It enables two types of certificate revocation: explicit and implicit. The issuer explicitly revokes a certificate by issuing a revocation statement and by sending it to the nodes which stored the certificate in question. The implicit revocation relies on the expiration time contained in the certificates. Every certificate whose expiration time passes is implicitly revoked; this second mechanism is straightforward, but requires some loose time synchronization of the nodes. The quest for security in MANET led a PGP type PKI. In PGP, any node can issue a certificate and as such it allows a completely distributed architecture, apart from the central repository, which holds these certificates. It proposes a scheme to avoid the need for a central repository of certificates in the PGP system. This scheme involves each node keeping mini-repositories, which hold all the certificates the node issues and all the certificates issued on it. When nodes A and B meet, they merge their mini-repositories. The repositories are constructed according to the 'Shortcut Hunter algorithm'. This algorithm constructs repositories such that two nodes merging repositories have a high probability of finding a chain of certificates between them if one exists. This scheme is useful in a civilian environment where delegation of trust through a number of nodes is acceptable. Let the notation $A \rightarrow B$ mean that A trusts B . Then what the implications $A \rightarrow B, B \rightarrow C, C \rightarrow D$ and $D \rightarrow E$ signify is that A chooses to trust E i.e. $A \rightarrow E$. An alternative approach is to use a Certificate Authority (CA) to issue certificates. A CA is a third party trusted by all in the system, which effectively eliminates the need for a repository of certificates. Rather than finding a certificate linking $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$, one simply recovers the certificate $A \rightarrow E$. As such, the CA can be seen as a one-hop shortcut through the web of trust. The problem with this is the CA must be trusted by all and becomes a single point of failure in the event of an attack..

2.2 THE SB-TRUST MODEL

In PGP's "web-of-trust" model [9], each entity manages its own trust based on direct recommendation and seeks to further quantify the notions of trust and recommendation it uses a seniority-based (SB) trust model which is as follows. Trust management and maintenance are distributed in both space (k) and time (T) domains in the SB-model. Thus SB-model describes a seniors-securing approach to node authentication in MANET. In other words, the parameter T characterizes the time-varying feature of a trust relationship, while k signifies the number of senior nodes required to work as CA. An entity is trusted if any k trusted available senior entities claim so within a certain time period T . Once a node is trusted by its senior group, it is globally accepted as a trusted node. Otherwise, if the seniors distrusted an entity then it is regarded as untrustworthy in the entire network. If a node cannot find k senior nodes in certain network, it may roam to meet more nodes or wait for new senior nodes to move in.

2.3 CONSTRUCTION OF SB-PGP MODEL

In this work, we apply the SB-model for issuing PGP type certificate. Let us consider a MANET, to be established, for instance, in a conference where people having mobile nodes communicate with one another having insecure wireless channel. I assume N mobile nodes, and N may be dynamically changing as mobile nodes join, leave, or fail over time. Among them, some of the nodes that joined in the beginning are considered as senior nodes and later joining nodes are considered junior nodes but the size of senior nodes group may increase dynamically and sequentially according to the size of network. Besides, N is constrained if there may be a large device population otherwise not.

Specifically, for the model construction, we make the following assumptions:

- Each node has a unique nonzero ID and a mechanism to discover available senior member nodes of the network.
- Communication with senior nodes is more reliable compared with junior nodes of the networks.
- Mobility is centralized by a maximum node moving speed S_{max} .
- Each senior node is equipped with some local detection mechanism to identify Misbehaving nodes among its surrounding nodes, e.g. those proposed in [6, 1].

All nodes are maintaining the seniority table like routing table. Two nodes having off line certificate holder are used to centralize. Thus SB-PGP model describes a seniors-securing approach for issuance of PGP type certificate to a node & authentication in MANET. in which two or more (up to k) senior node are collectively sign a PGP type certificate and issue it to a newly incoming node after satisfying its information in T time. In other words, the parameter T characterizes the time-varying feature of a trust relationship, while k signifies the number of senior nodes required to sign on PGP type certificate or to work as CA. An entity is trusted if any k trusted available senior entities then it is globally accepted as a trusted node, Otherwise, untrustworthy for the entire network. The architecture of the model resulting from these assumptions is given in the following section.

2.4 ARCHITECTURE OF SB-PGP NETWORK

Consider a SB-trust model and introduce the PGP type certification design, which is based on the de facto standard RSA. Now what is the structure of group of senior nodes working as CA. To see this, consider a network environment which does not follow a hierarchical or centralized control and fixed infrastructure and all member of the network are equivalent in terms of status. In this model functionality of the CA is performed by two or more senior most nodes of the network. These senior nodes collectively sign on the certificate of a new node, after satisfying themselves about its information. PGP type certificate is signed by more than one node. The size of CA nodes increases dynamically. Initially we divide our ad-hoc networks nodes in two groups, senior group SN and junior group JN. The size of senior group increases dynamically. Let

$$SN = \text{ceiling}(N \times M\%) + 1 \quad (1a)$$

Where SN = (set of senior nodes in senior group) N = (total number of nodes in ad-hoc network) (1b) SCA = (set of nodes required for CA functionality)

$$M = (\text{variable } \%)$$

Notice that M can change according to security level required in the networks. If M increases then the size of the senior group increases and availability of the networks also increases. However, security of the network decreases, because if the seniority number of a node is lower down, then its confidence level is also down.

$$SCA = \text{ceiling}(SN \times K\%) + 1 \quad (2a)$$

Where

$$SCA = (\text{number of senior most nodes required in the network for CA})$$

$$SN = (\text{senior-most nodes})$$

$$K = (\text{Variable } \%) \quad (2b)$$

K : Depends on M. K can change according to security level required in the networks. If K increases then

the number of nodes require for CA also increases and security of the network increases but availability of the CA of network decreases. Here SCA is number of senior most nodes require for CA to sign on the certificate for new reliable node. The signature procedure by each senior node of CA is done sequentially[9].

Again, notice that the junior group consideration involves a dynamic topology, which is proportional to the network size and senior group size. Consequently, the size of junior group (JN) will grow with the difference of growth in total number of nodes (N) of the network being considered and the growth in size of senior group (SN), which results in the following equation $JN=N-SN$

3. METHODOLOGY

3.1 Simulation Environment

Simulations are done to compare these routing protocols. Simulator ns-2 is used for performance comparison. The network simulator ns-2 developed by the VINT research group at University of California at Berkeley in 1995 . The network simulator NS2 is a discrete event network simulation. Ns is a discrete event simulator targeted at networking research. Ns provides substantial support for simulation of TCP, routing, and multicast protocols over wired and wireless (local and satellite) networks.

NS2 is used to simulate the proposed algorithm. It work on network layer and inform about link breakage. The implementation of the protocol has been done using C++ language in the backend and TCL language in the frontend. TCL(Tool Command Language) is compatible with C++ programming language.

Interpretation is based upon two files trace files and nam files are to be generated. Network Animator (.nam) file, records all the visual events that happened during the simulation. Trace files (.tr), records the entire network event that occur during the simulation. And file is post analyzed with the help of awk scripts..

Table 2 : Simulation Parameter

<i>Parameter</i>	<i>value</i>
Simulation Time	50 Sec
No. of Nodes	50
Traffic Type	CBR
Pause Time	10 Sec
Maximum X-Y coordinate value	1000 M
Packet Size	512 byte
MAC Protocol	802.11
Mobility Model	Random Waypoint
Routing Protocol	IODMRP
Observation Parameters	EED, Throughput, PDF

3.2 Performance Metrics:

The estimation of performance of AODV, OLSR and TORA is done on the basis of following Performance metrics:

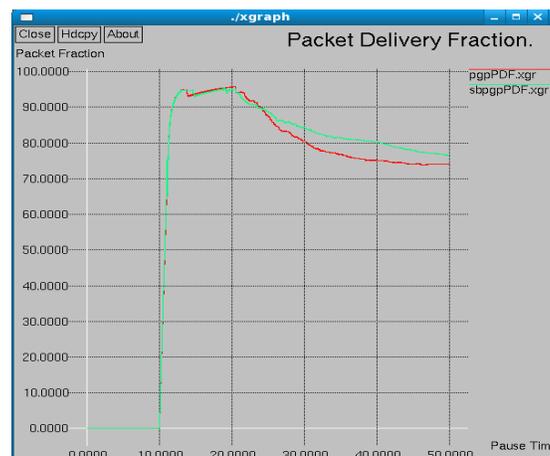
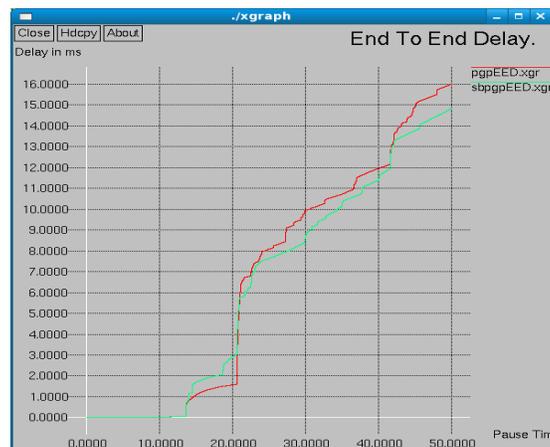
- **Packet Delivery Ratio:** It is the ratio of the packets received by destination to those generated by the sources. CBR traffic type is used by source. It specifies the packet loss rate, which limits the maximum throughput of the network. The routing protocol which have better PDR, the more complete and correct. This reflects the usefulness of the protocol. And provide good performance.

$$\text{Packet Delivery Ratio} = (\text{Received Packets} / \text{Sent Packets})$$

- **End to End Delay:** Average end-to-end delay is the average time it taken by the packet to reach to destination in seconds.
- **Throughput:** No. of packet passing through the network in a unit of time. It is measure in kbps.

4. Results

Throughput for IODMRP:



5. Conclusion

In this current thesis work we have described the design of secure techniques namely PGP and SBPGP with IODMRP protocol. It has been observed from the previous paper that SBPGP is giving the better security as

compared to the other techniques of PKI model. Current study is performed for comparison analysis for SBPGP model with PGP security models. From the above graph results and averaging it is found that SBPGP is giving better results and gives better security.

6 Acknowledgement

First and foremost, I would like to express my sincere gratitude to my guide Mrs. Meenakshi Mehla, Assistant Professor, Computer Science and Engineering Department for immense help, guidance, stimulating suggestions, and encouragement all the time with this thesis work. This work would have not been possible without her support. She always provided a motivating and enthusiastic atmosphere to work with; it was a great pleasure to do this thesis under her supervision.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction, to help me stay calm in the oddest of times and keep moving at times when there was no hope.

Himani Mann

REFERENCES:

- [1] Kimaya Sanzgiri, Daniel LaFlamme and Bridget Dahill, “Authenticated Routing for Ad hoc Networks” ,Refinements and extensions to IEEE ICNP 2002.
- [2] Svein Johan Knapskog, “New Cryptographic Primitives (Plenary Lecture)”,7th Computer Information System and Industrial Management Applications, IEEE 2008.
- [3] Yue Ai and Fuwen Pang, “Improved PKI Solution for Mobile Ad Hoc Networks”, IEEE 2010.
- [4] Venkatesan Balakrishnan and Vijay Varadharajan, “ Designing Secure Wireless Mobile Ad hoc Networks”, 19th International Conference on Advanced Information Networking and Applications, IEEE 2005.
- [5] G Varaprasad and P. Venkataram, “The Analysis of Secure Routing in Mobile Ad Hoc Network”, International Conference on Computational Intelligence and Multimedia Applications, IEEE 2007.
- [6] Nilesh P Bobade and Nitiket N Mhala, “ Performance Evaluation of Adhoc On Demand Distance Vector in Manets with varying Network size using NS-2 Simulation”, International Journal on Computer Science and Engineering (IJCSE) Volume 02 , August, 2010.
- [7] Geetha Jayakumar and Gopinath Ganapathy, “Performance Comparison of Mobile Ad-hoc Network Routing Protocol”, International Journal of Computer Science and Network Security (IJCSNS), Volume 07, November 2007.
- [8] Kamarudin Shafinah and Mohammad Mohd Ikram, “File Security based on Pretty Good Rivacy (PGP) Concept”,www.ccsenet.org/cis, Computer and Information Science, Volume 04, July 2011.
- [9] Maqsood Razi, Jawaid Quamar, “A Hybrid Cryptography Model for Managing Security in Dynamic Topology of MANET” IEEE 2008.
- [10] Antonio Vincenzo Taddeo, Alberto Ferrante, “A Security Service Protocol for MANETs”, IEEE 2009.
- [11] Q. Wang and W.C. Wong, “A Robust Routing Protocol for Wireless Mobile Adhoc Networks”, IEEE 2002.

Design of Neural Architecture in 0.35 μ m Technology Using Analog VLSI.

¹ Mr.Maulik B.Rami , ² Prof.H.G.Bhatt , ³Prof.Y.B.shukla

¹PG student(EC), Electronics and Communication Department S.V.I.T., Vasad-Gujarat.

^{2, 3}Associate professor, Electronics and Communication Department, S.P.C.E., Visnagar-Gujarat.

Abstract

Artificial neural network are attempts to mimic, at least partially, the structure and function of brains and nervous systems. The human brain contains billion of biological neurons whose manner of interconnection allow us the reason, memorize and compute. Advances in VLSI technology and demand for intelligent machines have created strong resurgence of interest in emulating neural system for real time applications. Such an artificial neural network can be built with help of simple analog components like MOSFET circuits and basic circuits with help of operational amplifier. This paper gives information about neuron behavior and how it takes intelligent decision.

Index Terms - Introduction, Architecture of neuron, Analog Neural components, Simulation Results, Conclusion& future work.

I. INTRODUCTION

Neural networks are used when there is no algorithmic solution to a problem or a problem is too complicated to be solved by known algorithms Neural networks can be used when the definition of the problem does not exist, but the samples of inputs and corresponding outputs are available.if we are ever going to understand intelligence and develop artificially intelligent machines or computers, we need to study the brain and its neurons, and how neurons work together to solve problems.It is not useful consider neural networks to solve problems for which the analytical solution can be easily found and implemented. In that case, the corresponding neural implementation will be generally larger and less accurate than the direct algorithmic solution.[2]

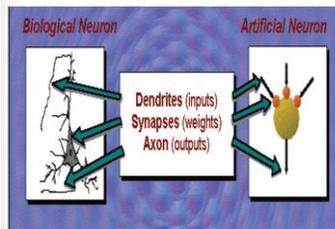


Fig.1 Comparison Of Neurons[7]

Here we compare a biological neuron with a typical artificial neuron. Similarly both take inputs, use weights and generate an output.Neural Networks are composed of a large number of computational nodes operating in parallel. Computational nodes, called neurons, consists in processing elements with a certain number of inputs and a single output that branches into collateral connections, leading to the input of other neurons.[2]

Normally they perform a nonlinear function on the sum (or collection) of their input. The neurons are highly interconnected via weight strengths these interconnections are typically called synapses and control the influence of neurons on the others neurons.[2]

The synaptic processing is typically modelled as multiplication between a neuron outputs and synaptic weight strengths.[3] Each neuron's output level depends therefore on the outputs of the connected neurons and on the synaptic weight strengths of the connections.[3]

Digital technology has advantages of mature fabrication techniques, weight storage in RAM, and arithmetic operations exact within the number of bits of the operands and accumulators. Digital chips are easily embedded into most applications. Digital operations are usually slower than in analogue systems, specially in the weight x input multiplication, and analogue input must first be converted to digital.[2] Analog neural networks can exploit physical properties to perform network operations and thereby obtain high speed and densities. Analog design can be very difficult because of the need to compensate for variations in manufacturing, in temperature, etc Creating an analog synapse involves the complication of analog weight storage and the need for a multiplier being linear over a wide range. Hybrid design attempts to combine the best of analog and digital techniques. The external inputs/outputs are digital to facilitate integration into digital systems, while internally some or all of the processing is done in analog. There are three types of neural networks.

1. Non-learning neural networks:
2. Off-chip learning networks:

3. On-chip learning networks:[3]

On-chip learning networks is the neural network performs both the feed forward phase and the learning one The advantages are the high learning speed due to the analog parallel operations and the absence of the interface with a host computer for the weight update. On-chip learning networks are suited to implement adaptive neural systems, i.e. systems that are continuously taught while been used.

II. ARCHITECTURE OF NEURON

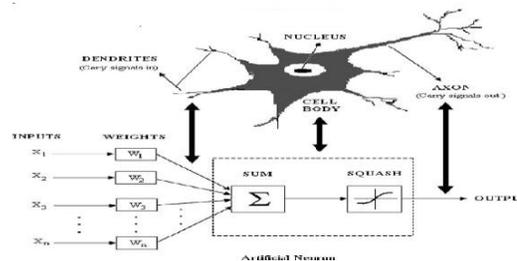


Fig.2 Architecture of Neuron.

Neural network produces a forecast by taking the weighted average of the predictors. This is just what a regression equation does Where neural networks go further is to layer this procedure The first layer of weighted averages (regression equations) produces a “hidden layer” of intermediate forecasts.

These forecasts are then used as predictors in another regression equation to produce the final forecast Note that It is possible to have more than one hidden layer, each one taking the forecasts of the previous layer as it’s predictors Each layer is made up of nodes Each node takes the weighted average of the predictors generated by the previous layer The first layer is an “input layer” consisting of the predictors It feeds the first layer of averaging nodes Each of these averaging nodes feeds a corresponding sigmoid to produce a forecast You can stack as many of these forecasting layers as you want, each one taking the forecasts of the previous layer as its predictors The final forecast layer produces the forecast of the quantity you were seeking. This type of neural network is called “feed forward” because the information feeds forward from the predictors, through the layers, and on to the final prediction It is said to be “fully connected” because every node in the one layer connects to every node in the layer above and below it Feed forward networks with sigmoid squashing functions are sometimes called “perceptrons”

$$u_i = \sum_j w_{ij}x_j \tag{1}$$

Where x_j s the j th predictor, w_{ij} is the weight for that predictor for node i , and u_i is the weighted average coming out of the i th node These weighted averages are then “squashed” by a non-linear sigmoid (s-shaped) function in order to prevent the occurrence of extreme values

$$y_i = \frac{1}{1 + e^{-u_i}} \tag{2}$$

Where, y_i is the forecast generated by the i th node

III. ANALOG NEURAL COMPONENTS

The inputs to the neuron as shown in figure 2 are multiplied by the weight matrix, the resultant output is summed up and is passed though an neuron activation function (NAF). The output obtained from the activation function is taken through the next layer for further processing. The multiplier block, adder block and the activation function model the artificial neural network. Blocks to be used are as follows

1. Multiplication block
2. Adders
3. NAF (Neuron Activation Function)block

A. Analog Multiplier

Here I have used Gilbert multiplier is named for Barrie Gilbert who designed the circuit in 1968 .The circuit combines diode-connected transistors, current mirrors, summing junctions, and differential pairs to multiply two differential signals. Consider two differential pairs that amplify the input by opposites gain

$$v_{out1}/v_{in} = -g_m R_d \tag{3}$$

$$v_{out2}/v_{in} = g_m R_d \tag{4}$$

$$V_{out} = V_{out1} + V_{out2} = A_1 V_{in} + A_2 V_{in}, \quad (5)$$

where A_1 and A_2 are gain which are controlled by V_{count1} and V_{count2} , respectively. If I_1 is zero then $V_{out} = +g_m R_d V_{in}$. V_{count} is used to vary currents monotonically. if $V_{count1} - V_{count2}$ is large then V_{out} will be most positive or most negative. $V_{out} = V_{in} * f(V_{count})$ and $f(V_{count})$ is Taylor expansion. $V_{out} \propto V_{in} V_{count}$ which is multiplication of two inputs.

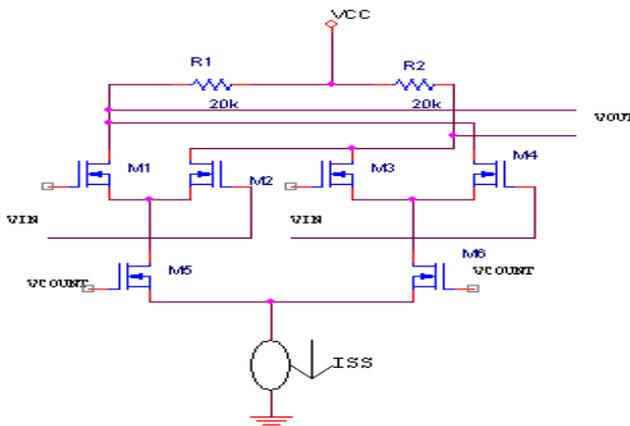


Fig.3 Gilbert Multiplier.

As shown in figure.3 six N-MOSFET are used in designing of Gilbert Multiplier. Designing of MOSFET's are based on 0.35um technology. For given lengths W/L ratios are given in Tabel I.

TABEL I. W/L RATIO OF GILLBERT CELL.

MOSFET(M)	W(μm)
1	1.4
2	1.4
3	1.4
4	1.4
5	0.2
6	0.2

B. Operational Amplifier.

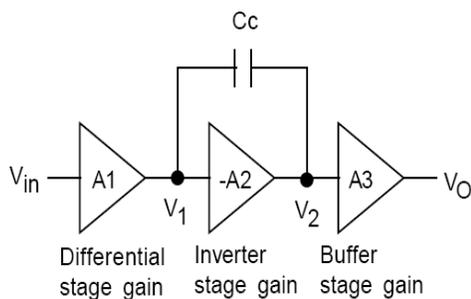


Fig.4 General Block Diagram Of OPAMP[8]

This OPAMP is a two-stage OPAMP where the first stage is a differential amplifier whose differential current output is mirrored into the next stage and converted to a single ended output through circuitry very similar to the synapse circuit. The outputs of the synapses can easily be summed. The summation is done by connecting all current outputs together. The summed current then must be converted to a voltage by a current to voltage (IV) converter.

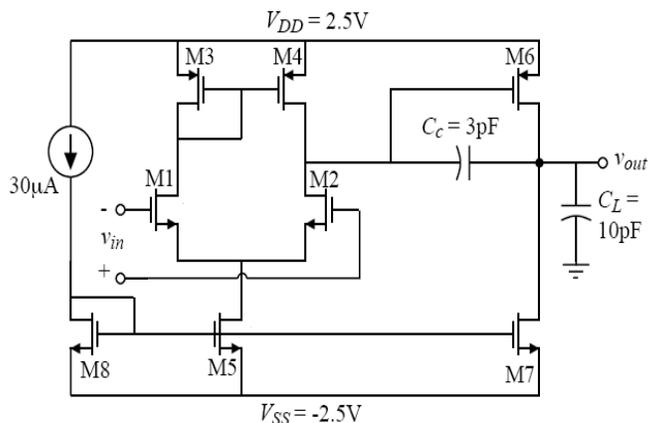


Fig.5 Schematic Of Two stage OPAMP.

TABEL II. W/L RATIO OF OPAMP.

MOSFET(M)	W/L(µm)	W(µm)
(1,2)	3	1.05
(3,4)	15	5.25
(5,8)	4.5	1.58
(6)	94	32.9
(7)	14	4.92

C. Opamp As Adder.

Adder can be design using operational amplifier. So it is necessary to implement op-amp. As shown in figure, opamp is designed with eight MOSFETs. In which five are N-MOSFET and 3 are P-MOSFET'S are designed with L=0.35µm length.

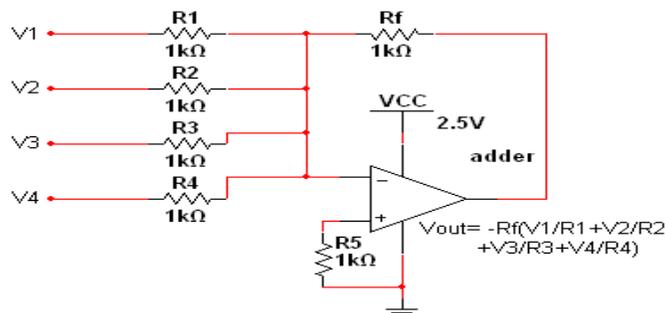


Fig.6 OPAMP As Adder Circuit.

TABEL III. SPECIFICATION OF ADDER.

Adder		
V1	0.25V	-2.0 volt
V2	0.3V	
V3	0.5V	
V4	1.0V	

D. Neuron Activation Function Block.

Neuron activation function designed here is tan sigmoid. The design is basically a variation of the differential amplifier with modification for the differentiation output. The same circuit should be able to output the neuron activation function and the differentiation of the activation function. Here three designs are considered for NAF. NAF can be operated in three regions.

(1) Linear function with adjustable threshold

The output voltage of the circuit is determined by $V_o = V_i(R_2/R_1) - V_{ref}(1+R_2/R_1)$ and threshold is determined by $q = V_{ref}(R_1+R_2)/R_1$. Lower set point is determined by $V_{ref}(R_1+R_2)/R_2 - V_{cc}(R_1/R_2)$ and Upper set point is determined by the $V_{ref}(R_1+R_2)/R_2 + V_{cc}(R_1/R_2)$.

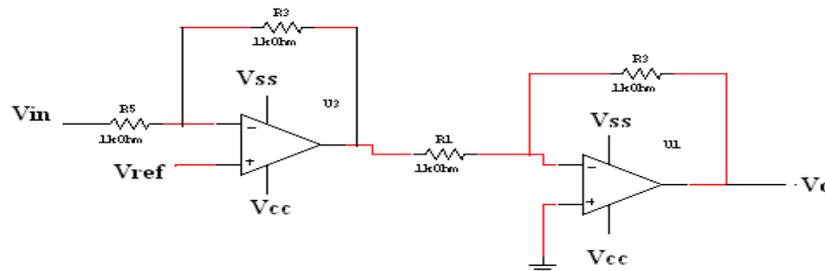


Fig.7 Schematic of Linear function with adjustable threshold

TABEL IV. PARAMETER VALUES.

Parameter Values	
R_1	1000k
R_2	2500k
R_3	1000k
V_{ref}	1 volt

(2) Sigmoid function with fixed gain control.

Another non linear function widely used is sigmoid function as shown in figure.

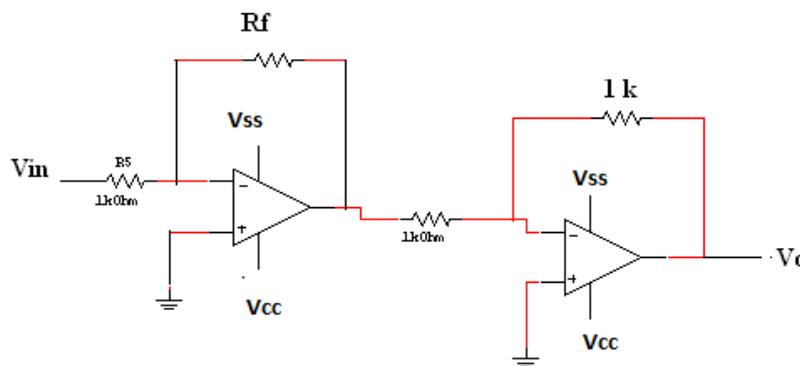


Fig.8 Schematic Of Sigmoid Function With Fixed Gain Control

This is same as Linear threshold function only V_{ref} is made to ground, so Upper limit as well as lower limit will be symmetrical around axis as shown its transfer function.

TABEL V. PARAMETER VALUES.

Parameter Values	
R_s	1000k
R_f	2500k
R_3	1k
V_{ref}	Ground

(3) Step Function

This is also same as sigmoid function but the difference is only in Upper Limit and Lower Limit is almost same. As shown in its Transfer Function it is at zero voltage.

E. Design of Neural architecture.

Artificial Neuron is combination of above developed and simulated blocks. Here, two multipliers with weight and input neuron are developed. Multiplier outputs are differentiated and amplified then applied to adder inputs. Input of multiplier is applied and amplified with respective weight and results of multiplier outputs and adder output can be identified with simulated results. Sigmoid function or Neuron activation function without derivative is used to limit the output of neuron within high and low logic. Result is shown in Fig.19.

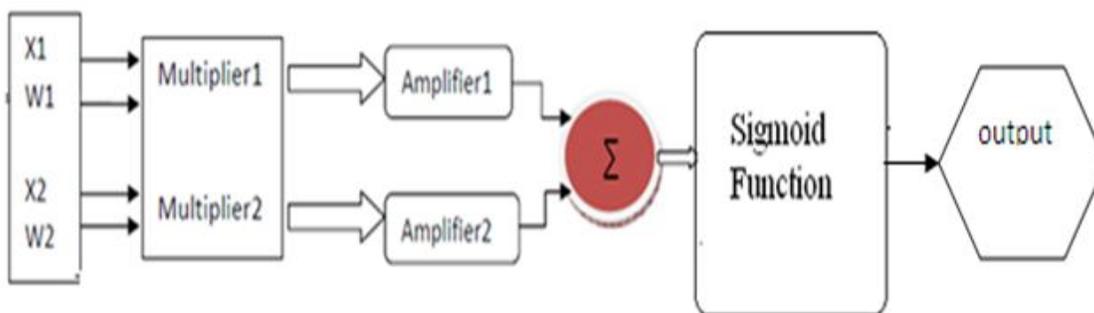


Fig.9 Design of Neural Architecture

IV. SIMULATION RESULTS.

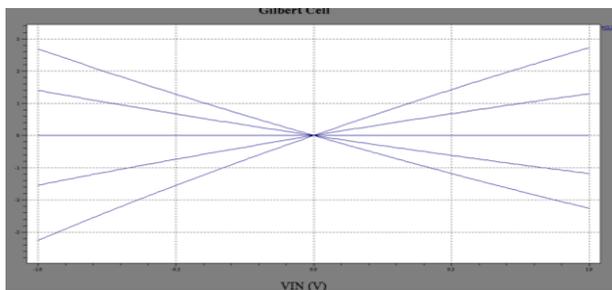


Fig.10 DC Transfer Characteristics

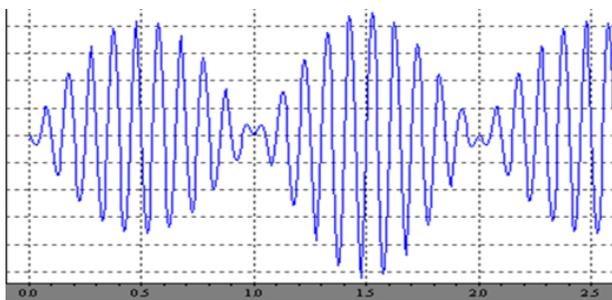


Fig.11 Modulated Waveform for Multiplier.

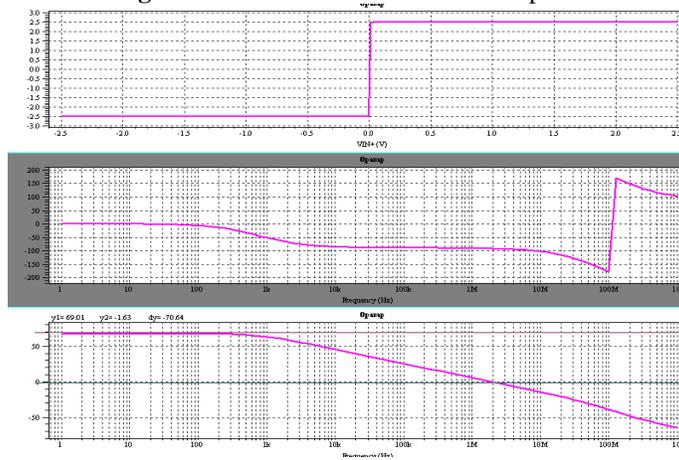


Fig.12 Transfer characteristic ,Gain and frequency response Of OPAMP.

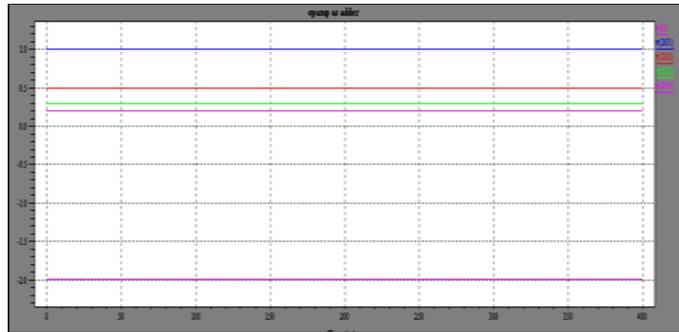


Fig.13 Adder Circuit and Simulation Result.



Fig.14 Transfer Function Of Linear Function With Adjustable Threshold



Fig.15 Transfer Function Of Sigmoid function with Fixed Gain Control

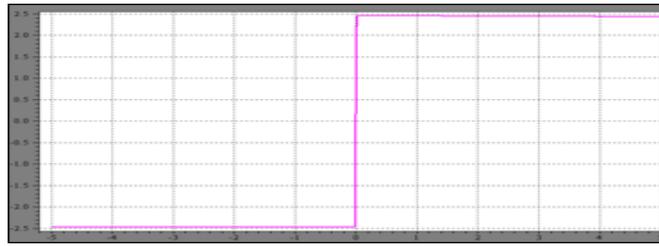


Fig.16 Transfer Function Of Step Function.

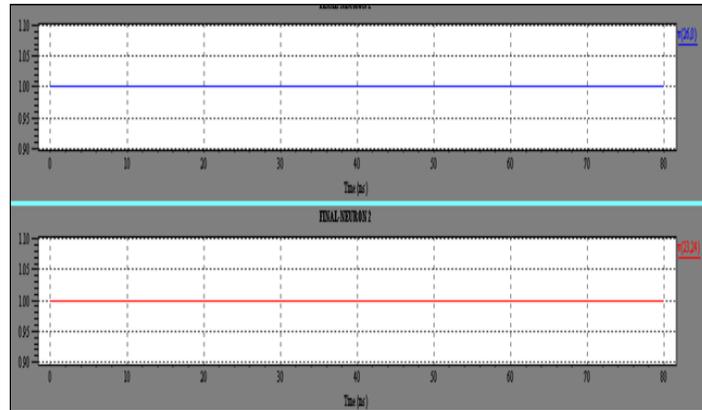


Fig.17 Input signal with weight signal for multiplier 1

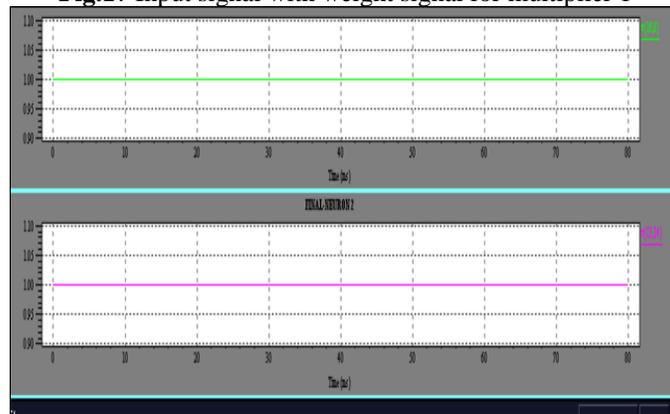


Fig.18 Input signal with weight signal for multiplier 2

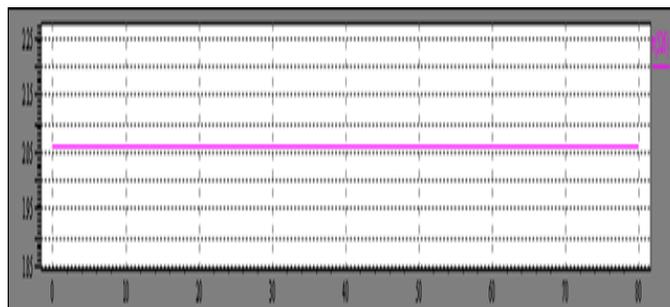


Fig.19 Output of Neural architecture

CONCLUSION & FUTUREWORK

Neural architecture is widely used in Real world interface & it can be implemented using different methods but we choose analog VLSI because it is very fast compared to digital VLSI & no need of A/D or D/A converter. One important is that it can be directly interface with physical sensors & actuators so used in pattern recognition like ECG, image processing and many applications.

Here we designed analog neural components like (Gillebert cell) multiplier or mixer, CMOS OPAMP, adder using OPAMP, activation function sigmoid, step function & linear threshold function with help of simulation TSPICE Software. After design of neural components I have combined all this blocks and designed neural architecture and got the simulation results.

In future we will design winner takes all circuit which is used to identify maximum active input than combine all neural components with WTA circuit and apply neural architecture with WTA for X-RAY image segmentation.

REFERENCES

Journals:

- [1] B. Gilbert, "A precise four quadrant multiplier with subnanosecond response," *IEEE*
- [2] Ismet Bayraktaroglu, "Circuit Level Simulation Based Training Algorithms For Analog Neuralnetworks", Electrical and Electronic Engineering, M.S. Thesis, 1996
- [3] Bo Gian Marco, "Microelectronic Neural Systems: Analog Vlsi For Perception And Cognition", Thesis November 1998
- [4] Bose N. K. Liang P., "Neural Network Fundamentals with graphs, algorithms and Application", Tata McGraw hill, New Delhi, 2002, ISBN 0-07-463529-8
- [5] Cyril Prasanna Raj P. , "Design and Analog VLSI Implementation of Neural Network Architecture for Signal Processing"
- [6] P.E. Allen, "Cmos Analog Circuit Design", 2005
- [7] A PPT On Fundamentals Of Neural Network.

Books:

- [8] Razavi Behzad, "Design of Analog CMOS Integrated Circuits", Tata McGrawhill, New Delhi, 2002, ISBN 0-07-052903-5
- [9] Mohammad Rashid "Pspice Using Orcad for Circuits and Electronics, PHI learning
- [10] R. Jacob Baker, Harry W. Li and David E. Boyce "CMOS Circuit Design, Layout, and Simulation" Department of Electrical Engineering Microelectronics Research Center.
- [11] Douglas R. Holberg, Phillip E. Allen, "CMOS Analog circuit design OXFORD University Press

Combining Multimedia Building Blocks In Image Analysis

P. Shanmugam¹, Dr. C. Loganathan²

¹ Associate Professor, Department of Mathematics, Kongu Engineering College,
Perundurai – 638052, India ² Principal, Maharaja Arts and Science College, Coimbatore – 641407, India

Abstract

Messages are given as text needs to be coded to avoid loss of secrecy in any transaction. This has been attempted through any media requires tough crypto analysis and can be handled through known ciphers. This job gets better results while combining the text message along with other building blocks of multimedia. In this paper a novel approach to insert a message on an image and both of them are passed in a media using elliptic curve crypto systems. Caution is made in maintaining the quality of the vehicle image carrying the text. We have presented the algorithm and illustrated through standard images used in image analysis.

Key Words: Elliptic curve cryptography, Encryption, Image analysis, Multimedia, Text conversion.

1. Introduction

The idea of information security lead to the evolution of Cryptography. In other words, Cryptography is the science of keeping information secure. It involves encryption and decryption of messages. Encryption is the process of converting a plain text into ciphertext and decryption is the process of getting back the original message from the encrypted text. Cryptography, in addition to providing confidentiality, also provides Authentication, Integrity and Non-repudiation

The use of elliptic curves in public key cryptography was independently proposed by Koblitz and Miller in 1985^[1] and since then, an enormous amount of work has been done on elliptic curve cryptography. Public key cryptography^[2] does not require any shared secret key between the communicating parties. Elliptic Curve Cryptography^[3] is emerging as an attractive public key cryptosystem. The attractiveness of using elliptic curves arises from the fact that similar level of security can be achieved with considerably shorter keys than in methods based on the difficulties of solving discrete logarithms over integers or integer factorizations. Elliptic Curve Cryptography^[4] provides an excellent solution not only for the data encryption but also for the secure key transport between two communicating parties. Understanding ECC needs full mathematical background on elliptic curves.^[5]

2. Elliptic Curve Cryptography

Elliptic curve cryptography starts with generating points of a curve. Using the real numbers^[6] for cryptography will cause a problem because it is very hard to store them precisely in computer memory and to predict how much storage will need for them. Also calculations over the real numbers are slow and inaccurate due to round-off error and the implementation of these calculations into cryptographic schemes require fast and precise arithmetic; thus elliptic curve groups over finite fields F_p and F_{2^m} are used.

2.1 Construction of Elliptic Curve Group

Consider the elliptic curve $E: y^2 = x^3 + ax + b \pmod{p}$ defined over F_p where p is prime number and $4a^3 + 27b^2 \pmod{p}$ is not 0. Then the set of $(p - 1)/2$ quadratic residues Q_p is obtained from the reduced set of residues $Z_p = \{1, 2, 3, \dots, p - 1\}$. Now, for $0 \leq x < p$, compute $y^2 = x^3 + ax + b \pmod{p}$ and determine whether the value of y^2 is in the set of quadratic residues Q_p . If so then there are two values in the elliptic group and if not so then the point is not in the elliptic group $E_p(a, b)$.^[7]

Example: Consider the elliptic curve $E_{23}(1,1): y^2 = x^3 + x + 1 \pmod{23}$ defined over the prime number $p = 23$ and $a = 1, b = 1$ for which $4a^3 + 27b^2 \pmod{p}$ is not 0. For the reduced set of residues $E_{23} = \{1, 2, 3, \dots, 21, 22\}$ the quadratic residues Q_{23} is obtained as follows.

Table-1: Quadratic residues Q_{23}

$x^2 \pmod{23}$	$(p-x)^2 \pmod{23}$	Q_{23}
$1^2 \pmod{23}$	$22^2 \pmod{23}$	1
$2^2 \pmod{23}$	$21^2 \pmod{23}$	4
$3^2 \pmod{23}$	$20^2 \pmod{23}$	9
$4^2 \pmod{23}$	$19^2 \pmod{23}$	16
$5^2 \pmod{23}$	$18^2 \pmod{23}$	2
$6^2 \pmod{23}$	$17^2 \pmod{23}$	13
$7^2 \pmod{23}$	$16^2 \pmod{23}$	3
$8^2 \pmod{23}$	$15^2 \pmod{23}$	18
$9^2 \pmod{23}$	$14^2 \pmod{23}$	12
$10^2 \pmod{23}$	$13^2 \pmod{23}$	8
$11^2 \pmod{23}$	$12^2 \pmod{23}$	6

The set of $(23-1)/2 = 11$ quadratic residues $Q_{23} = \{1, 2, 3, 4, 6, 8, 9, 12, 13, 16, 18\}$ For $0 \leq x < 11$,

Table-2: Generation of Elliptic Curve Points

x	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
y^2	1	3	11	8	0	16	16	6	15	3	22	9	16	3	22	10	19	9	9	2	17	14	22
$y^2 \in Q_{23}$	y	y	n	y	n	y	y	y	n	y	n	y	y	y	n	n	n	y	y	y	n	n	n
y_1	1	7		10	0	4	4	11		7		3	4	7				3	3	5			
y_2	22	16		13		19	19	12		16		20	19	16				20	20	18			

The points on the elliptic curve are

$$E_{23}(1,1) = \left\{ (0,1), (0,22), (1,7), (1,16), (3,10), (3,13), (4,0), (5,4), (5,19), (6,4), (6,19), (7,11), (7,12), (9,7), (9,16), (11,3), (11,20), (12,4), (12,19), (13,7), (13,16), (17,3), (17,20), (18,3), (18,20), (19,5), (19,18) \right\}$$

2.2 Arithmetic in an Elliptic Curve Group

The negative of the point at infinity is $-\infty$ and is equal to ∞ and the negative of a point $P = (x, y)$ is defined to be $-P = (x, -y \pmod{p})$.

Adding distinct points: If $P(x_P, y_P)$ and $Q(x_Q, y_Q)$ are distinct points on $E_p(a, b)$ such that $P \neq -Q$ then their sum $P + Q = R$ is given by

$$x_R = [\lambda^2 - x_P - x_Q] \pmod{p} \text{ and}$$

$$y_R = [\lambda(x_P - x_R) - y_P] \pmod{p} \text{ where}$$

$$\lambda = \frac{y_P - y_Q}{x_P - x_Q} \pmod{p}.$$

Doubling the point: If the y coordinate of P is zero, modulo p, then $P = -P$. Doubling of the point $P = (x_P, y_P)$ with $y_P \neq 0 \pmod{p}$, is defined as $2P = P + P = R$ and is given by

$$x_R = [\lambda^2 - x_P - x_Q] \pmod{p} \text{ and}$$

$$y_R = [\lambda(x_P - x_R) - y_P] \pmod{p} \text{ where}$$

$$\lambda = \frac{3x_P^2 + a}{2y_P} \pmod{p}.$$

2.3 Elliptic Curve Encryption

Elliptic curve cryptography can be used to encrypt plaintext message M into ciphertext. The plaintext message M is encoded into a point P_M from the finite set of points in the elliptic group, $E_p(a,b)$ by multiplying the corresponding ASCII value with the generator point G . The first step consists in choosing the generator point, $G \in E_p(a,b)$, such that the smallest value of for which $nG = O$ is a very large prime number. The elliptic curve group $E_p(a,b)$ and the generator G are made public.

Suppose that A wants to encrypt and transmit plaintext message M to B . Each A and B selects a private key and uses it to compute their public key. A selects a random integer $n_A < n$ as his private key and computes his public key $P_A = n_A G$ and makes it publicly available. Similarly B selects a random integer $n_B < n$ as his private key and computes his public key $P_B = n_B G$ and makes it publicly available. To encrypt the message point P_M corresponding to the message M , A chooses his private key n_A and B 's public key P_B to compute the ciphertext pair of points $C_M = [(n_A G), (P_M + n_A P_B)]$.

2.4 Elliptic Curve Decryption

After receiving the ciphertext pair of points C_M , B decrypts it to retrieve the plaintext message point P_M as follows:

First point, $n_A G$ is multiplied with his private key n_B to obtain $n_A n_B G$ and it is subtracted from the second point $P_M + n_A P_B = P_M + n_A n_B G$, the plaintext point P_M corresponding to the plain text message M is retrieved.

Example: Consider the Elliptic curve $E_{281}(1,4)$. Let the generator point $G = (0, 2)$. Then multiple of are $2G = (123, 178)$, $3G = (194, 239)$, etc and $311G = O$ - Point at infinity.

If A wants to encrypt the plaintext 'ENCRYPT' in which the first character is 'E'.

The ASCII value of the character is 101. Then $P_M = 101(0, 2) = (95, 173)$.

Let the private key of A be $n_A = 100$ then its public key is $P_A = n_A G = 100G = (254, 244)$.

Let the private key of B be $n_B = 200$ then its public key is $P_B = n_B G = 200G = (29, 265)$.

Then the ciphertext pair of points is

$$C_M = [(n_A G), (P_M + n_A P_B)]$$

$$C_M = [(254, 244), ((95, 173) + 100(29, 265))]$$

$$C_M = [(254, 244), ((95, 173) + (31, 151))]$$

$$C_M = [(254, 244), (269, 136)]$$

Upon receiving the ciphertext pair of points $C_M = [(254, 244), (269, 136)]$, B decrypt is as follows:

$$P_M + n_A P_B - n_B (n_A G) = (269, 136) - 200 \times (254, 244)$$

$$P_M + n_A P_B - n_B (n_A G) = (269, 136) - (31, 151)$$

$$P_M + n_A P_B - n_B (n_A G) = (269, 136) + (31, 130) \text{ since } -(31, 151) = (31, -151) = (31, 130)$$

$$P_M + n_A P_B - n_B (n_A G) = (95, 173) = P_M$$

Thus $(95, 173) = (\text{ASCII value of the character}) \times G$.

Applying discrete logarithm concept to get back the ASCII value of the character from which the character is retrieved.

3. Inserting Text In An Image

To hide the text inside the image, the ASCII value of the text and the pixel value of the image are converted into streams of 8-bit binary. Two pixel pairs of the image in two adjacent rows are used to hide one character of the text. The four least significant bits of the selected pixel value in two adjacent rows are replaced respectively by the four least significant bits and four upper significant bits of one character of the text. The modified 8-bit binary numbers are converted to decimal number which gives the pixel value of the image, after hiding the text. To hide each character of secret message, two pixels are needed. So the number of characters that can be hidden in $(n \times n)$ image is given by the equation: Number of characters $\leq (n \cdot n) \div 2 - n$. In this equation, n pixels are subtracted because the secret text is not set in the first row of the cover image^[8]. So start setting data from the second row of the cover image. The first row of the covered image is used to store specific data, like the position of last pixel in the covered image that contains secret data. Reconstruction of the secret text message is performed by reversing the process used to insert the secret message in the container image.

4. Image Encryption Using Elliptic Curve Cryptography

Elliptic Curve Cryptography, with proper choice of curve parameters, is used to encrypt images. The pixel value at each point of the image is used in place of ASCII value and encoded into pixel point. The same encryption and decryption method is adopted to compute the encrypted pair of points corresponding to the pixel value and to retrieve the pixel value from the encrypted pair of points.

5. RESULT

Camerman image is taken and the text 'image' is inserted in the image by using the method explained in section 3.



Camerman Original



Text inserted

157	157	157	157	157	...
158	158	158	158	158	...
159	159	159	159	159	...
158	158	158	158	158	...
156	156	156	156	156	...
...

Pixel value of Text inserted image

To encrypt the text inserted image, consider the elliptic curve $y^2 = (x^3 + x + 4) \text{ mod } 281$; that is $a = 1, b = 4$ and $p = 281$. Let the generator point $G = (0, 2)$ and note that $311G = O$, the point at infinity.

If A wants to encrypt the Text inserted image, A selects his private key $n_A < n$ and generate the public key $P_A = n_A \times G$ and produces the ciphertext pair of points $C_M = \{n_A G, P_M + n_A P_B\}$. After receiving the ciphertext pair of points C_M , B multiplies the first point, $n_A G$ with his private key n_B and the subtract this from the second point $(P_M + n_A P_B)$. Thus $(P_M + n_A P_B) - n_B(n_A G) = P_M$. The pixel value M of the image at the position (1, 1) is 157 and therefore $P_M = 157G = 157 \times (0, 2) = (243, 156)$.

Let A chooses his private key $n_A = 100$ randomly and the private key of B is $n_B = 200$ and so

$$P_B = n_B G = 200 \times (0, 2) = (29, 265).$$

Hence

$$C_M = \{100 \times (0, 2), (243, 156) + (100 \times (29, 265))\}$$

$$C_M = \{(254, 244), (243, 156) + (31, 151)\}$$

$$C_M = \{(254, 244), (23, 162)\}$$

All the Pixel values of the image are encrypted and the corresponding ciphertext pair of points are computed for the entire image.

After receiving the ciphertext pair of points $C_M = \{(254, 244), (23, 162)\}$, B computes

$$P_M + n_A P_B - n_B(n_A G) = (23, 162) - [200 \times (254, 244)]$$

$$P_M + n_A P_B - n_B(n_A G) = (23, 162) - (31, 151)$$

$$P_M + n_A P_B - n_B(n_A G) = (23, 162) + (31, -151)$$

$$P_M + n_A P_B - n_B(n_A G) = (23, 162) + (31, 130), \text{ since } -151 \text{ mod } 281 = 130$$

$$P_M + n_A P_B - n_B(n_A G) = (243, 156) = P_M$$

Applying discrete logarithm concept to get back, the pixel value of the image is retrieved. After retrieving all the pixel values, the text inserted image can be regenerated. From this reconstruction of the secret text message is performed by reversing the process used to insert the secret message in the container image.

6. Conclusion

In this paper, a novel idea has been developed from Elliptic Curve Cryptography. Instead of sending the encrypted ciphertext pair of points, of a plaintext message, from entity A to B, the plaintext is hidden within the image as described in section 4 and the pixel value at each point of the image is encrypted to a pair of points as described in section 5 and can be send from A to B with added security.

References

1. Koblitz. K, Elliptic Curve Cryptosystems, Mathematics of Computation, 48: 203-209, 1987.
2. Anoop. M S., Elliptic curve cryptography, an Implementation Guide
3. Prof. Jagdale. B. N., Prof Bedi. R. K., and Sharmishta Desai, "Securing MMS with High Performance Elliptic Curve Cryptography", International Journal of Computer Applications, 8(7): 2010.
4. Suneetha. Ch., Sravana Kumar. D and Chandrasekhar. A., "Secure key transport in symmetric cryptographic protocols using elliptic curves over finite fields", International Journal of Computer Applications, 36(1): 2011
5. Padma Bh, Chandravathi. D, Prapoorna Roja. P, "Encoding and Decoding of a Message in the Implementation of Elliptic Curve Cryptography using Koblitz's Method", International Journal on Computer Science and Engineering, 2(5): 1904-1907, 2010.
6. Kaabneh. K, Al-Bdour. H, "Key Exchange Protocol in Elliptic Curve Cryptography with no Public Point" American Journal of Applied Sciences, 2(8): 1232-1235, 2005.
7. Kefa Rabah, "Implementation of Elliptic curve Diffe-Hellman and EC Encryption Scheme" Information Technology Journal 4(2): 132-139, 2005,
8. Habes. A, "Information Hiding in BMP image Implementation, Analysis and Evaluation", Information Transmission in Computer Networks 6(1): 1-10, 2006.
9. Stallings. W, "Cryptography and Network Security", 4th edition, Prentice Hall, 2005

An Intrinsic Dislocation Density – Finite Element Formulation Of Metal Plasticity

*Njoroge K. D.¹, Mutuli S. M.², Kihui J. M.³

¹ Department of Mechanical Engineering, School of Mechanical, Manufacturing and Materials Engineering, Jomo Kenyatta University of Agriculture and Technology, P.O. Box 62000-00200, Nairobi, Kenya.

² Associate Professor, Department of Mechanical and Manufacturing Engineering, School of Engineering, University of Nairobi.

³ Associate Professor, Department of Mechanical Engineering, School of Mechanical, Manufacturing and Materials Engineering, Jomo Kenyatta University of Agriculture and Technology.

Research done at the Department of Mechanical and Manufacturing Engineering, School of Engineering, University of Nairobi, P.O. Box 30197-00100, Nairobi, Kenya.

Abstract:

A computational model was developed to simulate elastic and plastic behavior in Body Centered Cubic (BCC) metals and alloys. The model provided for simultaneous simulation of the micro and macro length scales and used periodicity to link the two length scales. The model was implemented in a 3dimensional framework giving rise to a finite element technique incorporating intrinsic dislocation information in the simulation of the material's behavior. The technique was validated by simulating loading over the elastic range and the immediate region beyond yield, of thin steel strips, and comparing the results to those obtained by conventional analysis. Stress-strain curves and slip plane percentage contribution factors were generated. Specifically the stress-strain curves generated upheld Hooke's law and demonstrated a definite yield plateau followed by material recovery after yielding.

Key words: Dislocation density, Multi scale, Percentage slip plane contribution

Notation:

- B - Dislocation mobility (Pa. sec.)
- b** - Burgers vector
- dl* - Length of dislocation segment (m)
- E_e - Young's Modulus for Loading Excluding dislocation information (N/m²)
- E_p - Young's Modulus for Loading Including dislocation information (N/m²)
- $E(\beta)$ - Pre-logarithmic energy coefficient (J)
- G - Shear Modulus (N/m²)
- R** - Relative position vector for interacting dislocation segments (m)
- T_e - Temperature (deg. K)
- t - Thickness (m)
- V - Representative volume (m³)
- W - Energy per unit volume (J/m³)
- ΔW_{pm} - Work done in dislocation motion (J)
- ΔW_{pf} - Work done in net dislocation formation (J)
- ΔW_{pi} - Work done as the dislocations interact (J)
- ΔW_{po} - Work done in overcoming the obstacles to dislocation motion (J)
- α - Constant of proportionality
- β - Dimensionless constant
- β_t - Outer radius of dislocation cylinder (m)
- ΔG - Gibbs free energy (J)
- ϵ_e - Strain tensor

- v - Glide velocity (m/s)
- θ - Grain boundary mis-orientation angle (radians)
- ρ - Dislocation density (m/m^3)
- ρ_o - Density of obstacles to dislocation motion (m^{-3})
- σ_{ij} - Stress tensor (N/m^2)

1. Introduction

Traditional use of numerical techniques for structural analysis applies constitutive relations that are based on average material properties [1]. While being convenient and sufficient for use in elastic deformation problems, they are however at best coarse approximations in their use in large strain problems and problems where size effects are significant [1]. The search for more accurate relations based on a physical and mechanistic foundation has resulted in the micro structural approach. The micro structural approach looks at materials using length scales in the order of 10^{-5}m to 10^{-10}m [2], and involves the monitoring of the movement of atoms and atomic defects that contribute to material deformation and is used to aggregate these contributions to predict macroscopic strain. A fully micro structural approach is however computationally costly due to the large number of degrees of freedom involved. A blend between the micro structural and macro structural analysis has been proposed and involves two or more length scales. The macro structural approach involves the use of length scales in the order of 10^{-3}m and above.

Multi scale material modeling is a modern technique in the field of structural mechanics that is capable of analyzing structures in at least two different length scales [2]. This method has evolved from the field of materials modeling where it was understood that a material's response is profoundly influenced by its microstructure. The need to understand the mechanisms responsible for material behavior and the development of microstructure-property correlations has driven the development of theory and simulation of material behavior. Several techniques have emerged differentiated by their treatment of the micro scale of the problem [2-6].

The commonly used constitutive relations are based on the macro scale approach which ignores the random occurring micro structure to enable the use of averaged material property constants giving reasonable results. The micro structural behavior is based on a phenomenological approach in which the material behavior is modeled along phenomena observed during empirical tests. Alternatively it is based on a mechanistic approach where true material deformation mechanisms are simulated and therefore a more realistic model of the material's micro structural response is obtained. Most previous research work has addressed the use of the phenomenological foundation [2] and opportunity exists to incorporate a mechanistic foundation with the obvious results of negating the need to assume any specific degradation phenomena in the deformation process.

Several researchers have contributed to the development of the so-called discrete dislocation techniques where a continuum finite element model is extended to include individual dislocations [7, 8]. Only their Burgers vectors, and their core locations through stress and displacement fields that they induce, are used to represent these dislocations. The discrete dislocation approach accurately describes the long-range interactions between dislocations but cannot be described as a truly atomic scale phenomenon.

The standard discrete dislocation formulation for the inhomogeneous problem of an elastic-plastic body containing an elastic inclusion was derived by Van der Gissen and Needleman [9]. This method models edge dislocations as line defects in an isotropic elastic material constrained to glide on a fixed slip plane. Long-range dislocation interactions occur through their continuum elastic fields while short-range interactions are governed by constitutive rules for dislocation nucleation, motion and annihilation. This method's capacity is limited by the computation of the polarization stress, which corrects for the presence of the inclusion, which dominates the finite element calculation when the number of dislocations becomes large.

A superposition technique of the discrete dislocation method, as developed by O'day et al. [10] resolved the problem associated with the polarization stress in the standard discrete dislocation method. In this approach the regions of elastic inhomogeneity and the area of the inclusion are handled separately and superimposed onto the fully elastic solution. The technique is therefore advantageous as it may be used in elastically in-homogeneous problems with large numbers of dislocations or many inclusion elements.

The use of 3D discrete dislocation method in the study of small scale plastic phenomena was reported by Zbib et al. [11]. This three dimensional continuum based finite element formulation for elastic-viscous-plasticity incorporates discrete dislocation simulation replacing the usual plasticity constitutive relations. The superposition principle is utilized in order to find the effects of boundaries on the dislocation movement and the multi-scale frame merges the two scales of micro-scale where plasticity is determined, to the continuum scale where the energy transport is based. This technique uses a two stage computation which makes the method computationally expensive. The evaluation of the dislocation forces requires very fine mesh particularly when dealing with dislocation-defect interaction [11]. As a result, to capture the effects of small defects, the dislocation segment size must be comparable to the size of the defect. This requirement adds a further computational load to the simulation process by making it necessary to make several simulations before appreciable macro strain is achieved. To achieve 0.3% strain over one

million iterations are required [11]. As a result application is limited to nano-technology. Opportunity exists for researchers to overcome these drawbacks and extend the technique to larger sized structures.

Atomistic to continuum modeling methods are either simultaneous or sequential in their implementation. The Quasi Continuum approach [5] is one method where simultaneous/concurrent solution of the governing equations at multi scales is employed. ‘Handshake’ algorithms are used to bridge neighboring regions of different resolution. Other such methods include the Three-scale Bridging technique by Broughton et al [12] and the Dynamic Bridging Scale method by Park et al [13]. These methods have been applied successfully in the solution of problems where the bulk material remains relatively homogeneous and the micro scale is confined to a relatively small volume. Sequential/decoupled methods do not involve direct matching of boundary conditions at multiple length scales. Instead lower length scale simulations generate information that is used as input data in higher length scales. Atomistic simulations of motion of a representative number of atoms were used by Hao et al [14] to generate bulk elastic and plastic properties while Spearot et al [15] used them to generate interfacial properties. Clayton et al [16] used atomistic simulations to generate displacement gradients which were transferred across length scales. These methods can be used where a small group of defect types are considered under periodic conditions which enable the micro scale simulation to as reasonably small as possible.

Several methods have been presented for the development of variational principles for materials with microstructure. Svendsen [17,18] presented a thermodynamic model for materials with internal degrees of freedom. The model is cast in a variational formulation founded upon a dissipation rate density based on a dissipation potential. Cermelli et al [19] presented a formulation of a variational principle consisting of an energy functional for a body containing isolated defects in terms of a regular function of the defect configuration. The method avails a basis for the study of a finite set of dislocations and is illustrated in the study a finite number of Volterra dislocations in a plane domain. Groma et al [20-22] proposed a variational method to calculate the stress field generated by a system of dislocations. The principle is restricted to the small deformation limit theory and the functional is sum of the elastic Gibbs free energy and a term representing the interaction of the defects and the stress field. The method allows for the study of a wide range of problems as illustrated in the study of core regularization problems and dislocation-solute interaction.

In all these approaches the variational methods provide systematic and compact technique to study the problems at hand while allowing relatively simple accounting of the physical interactions. These methods however focus on a global thermodynamic evolution based on small strain theory and do not differentiate various dislocation constitutive behaviour. A formulation for large strains and the capacity to account for different dislocation evolution phenomena needs to be developed. In addition a formalization that captures the interactions at the dislocation core and its integration into a meso scale dislocation density variational principle is required. An extended review of dislocation dynamics applied to multi scale techniques is found in [23].

2. Materials and Methods

2.1. Model Development

This work developed a framework that incorporated the dislocation information in a continuum finite element model, under quasi static conditions. The underlying objective was to formulate a finite element technique capable of single cycle micro-macro evaluation of elasto-plastic deformation of structures. It formed the basis of a dislocation density continuum approach that could be extended to a variety of materials based boundary value problems.

From the principle of conservation of energy, work done by external loads in plastic deformation at room temperature, W_p , is dissipated in the evolution of the micro structure consisting of dislocation fields in material matrix [24]. Consequently, W_p equals [24]

$$W_p = \int_V \Delta W_m dV \dots \dots \dots (1)$$

The work done in the evolution of the microstructure ΔW_m is a sum of the contributions of dislocation motion, dislocation formation and annihilation, dislocation interactions and work done against obstacles to dislocation motion as shown in Fig. 1, and may be summarized in the relation

$$\Delta W_m = \Delta W_{pm} + \Delta W_{pf} + \Delta W_{pi} + \Delta W_{po} \dots \dots \dots (2)$$

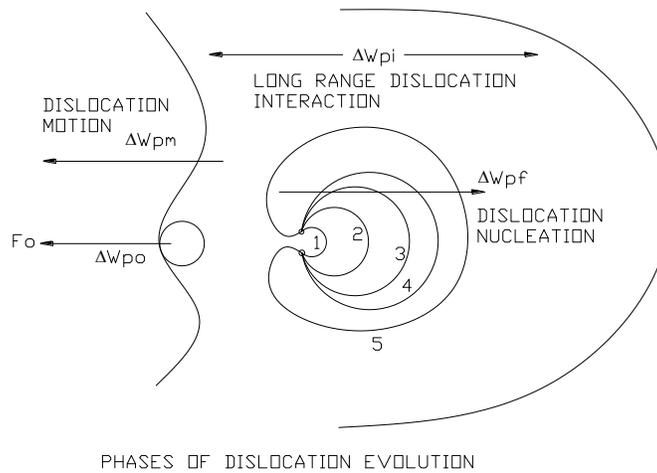


Fig. 1: Work done in dislocation evolution

Under adiabatic conditions, work is done at the expense of the strain energy of the total system. The above various contributions are developed as follows:

Work done in dislocation motion ΔW_{pm} is given by the relation [24]

$$\Delta W_{pm} = \int \int_{ul} [(b_p \cdot \sigma_{ij}) x dl_p] du \dots \dots \dots (3)$$

where

b_p is the burgers vector for segment 'p'

dl_p is the length of segment 'p'

du is the displacement of segment 'p'

Work done in net dislocation formation, ΔW_{pf} includes work in total dislocation formation less the work done in dislocation annihilation. This work is equal to the net change in elastic strain energy of the elemental volume with change in total length of dislocations contained and is given by the relation [24]

$$\Delta W_{pf} = \int \int_{ul} E(\beta) \ln \left(\frac{\alpha |\beta_t|}{|b_t|} \right) d(dl_t) \dots \dots \dots (4)$$

Consider two dislocation segments denoted "p" and "q" and separated by a vector R . Work done as the dislocations interact ΔW_{pi} is given by [24]

$$\Delta W_{pi} = -\frac{G}{2\pi} \int \int_{ul} \left(\frac{(b_p \cdot x b_q)(dl_p \cdot x dl_q)}{R} \right) + \frac{G}{4\pi} \int \int_{ul} \left(\frac{(b_p \cdot dl_p)(b_q \cdot dl_q)}{R} \right) + \frac{G}{4\pi(1-\nu)} \int \int_{ul} (b_p \cdot x dl_p) T \cdot (b_q \cdot x dl_q) \dots \dots \dots (5)$$

where $T = \frac{\partial^2 R}{\partial x_i \partial y_j}$

Work done in overcoming the obstacles to dislocation movement is ΔW_{po} . This work, done by slow moving dislocations, is modeled as a varying body force over the elemental volume acting normal to the dislocation line. The body force is determined by obtaining the activation energy for dislocation motion against these effects as a function of the number of discrete obstacles and summing over an elemental volume to give [24]

$$\Delta W_{po} = \int \int_{ul} \Delta G \cdot \rho_o \cdot dl \cdot du \cdot t + \int \int_{ul} B \cdot v \cdot dl_p \cdot du \dots \dots \dots (6)$$

The work allows for invariance of the micro structural properties of the material hence $\Delta G, \rho_0$ and t are constant. The physical interpretation of this constraint is that dislocation obstacle's dissipative strength and density, and the slip plane spacing are maintained constant. Having provided for slow moving dislocations, dislocation mobility and velocity are also assumed constant.

A dislocation length function

$$l = f(u) \dots \dots \dots (7)$$

is adopted and as a first approximation, the dislocation is modeled as a near circular boundary between the slipped and unslipped portions of the crystal about the slip plane. Consequently, the dislocation function is written as

$$l = 2\pi u \dots \dots \dots (8)$$

Using equations (3), (4), (5), (6) and (8) in (2), expanding the cross products in terms of the alternating tensor, using $(\mathbf{b}_p, \sigma_{ij}) = S_j$, and rewriting gives

$$\Delta W_m = \frac{1}{2\pi} \int \int_{u,l} [S_j dl_p \varepsilon_{jpk}] dl_k + \int \int_{u,l} E(\beta) \ln \left(\frac{\alpha |\beta_t|}{|b_t|} \right) n_t d(dl_t) - \frac{11G}{4\pi} \int \int_{u,l} \left(\frac{b_p b_q dl_p dl_q}{R} \right) + \frac{1}{2\pi} (\Delta G \cdot \rho_0 \cdot t + B \cdot v) \int \int_{u,l} dl_p dl_p \dots (9)$$

Equation (9) is the dislocation evolution energy functional for a micro element 'm', with the dislocation displacement expressed in terms of the dislocation segment length. This expression is integrated over the meso scale into the macro scale and combined with the elastic strain energy of a macro scale element to yield the following expression

$$\Pi_{pe} = \int_V (\varepsilon_e^T \sigma_e + u_e^T X_e) dV + \int_{\zeta} (u_e^T P_e) d\zeta + \int_V \Delta W_m \dots \dots \dots (10)$$

Equation (10) is the total energy functional for a macro element incorporating two variables, the macro scale displacement u_e and the micro scale dislocation segment length dl . Applying variational techniques in minimizing this relation with respect to the macro scale displacement and the dislocation length yields an equilibrium equation governing the evolution of elasto-plastic behavior of the material. This formulation employs the simultaneous form of multi scale technique as the solution of the micro scale and macro scale are handled simultaneously with periodicity used to couple the two length scales.

2.2. Implementation

Thin strips were subjected to loading cycles with loads varied from 0.04 to 3.0 times the estimated yield stress with equal load increments. The maximum induced stress and its corresponding induced strain over the length of the sample for each loading cycle were recorded. A constant obstacle density of $1 \times 10^{11} / m^3$ and a constant stored dislocation density of $5 \times 10^{12} m/m^2$ were used. Mobile dislocation densities were generated by a stress based function [25]

$$\rho = (\sigma_{max} / (0.4 * G * b))^2 \dots \dots \dots (11)$$

where the generated stress from the aprior loading cycle was used as an input. Dislocation nucleation was intrinsically accounted for in the increments to dislocation densities across the simulation cycles. Grain boundary dislocation density was generated by a quadratic function [25]

$$\rho_m = (\theta / b)^2 \dots \dots \dots (12)$$

dependant on the angular mis-orientation across the grain boundary which in turn was related to the generated stress.

$$\theta = \theta_{min} + \left((\theta_{max} - \theta_{min}) * \left(\frac{\sigma_{max}}{G} \right) \right) \dots \dots \dots (13)$$

Single dislocation density values were used for each micro scale cell and the interaction between the various dislocation families was expressly captured in the dislocation evolution energy functional. Grain boundaries were naturally generated as the boundaries of the micro structural cells with a wavelength of 50nm.

3. Results And Discussion

Simulations were carried out to evaluate the contribution of the {110}, {112} and {123} slip planes. A percentage slip plane contribution 'a-b-c' was defined where 'a' was the contribution of the {110} plane, 'b' was the contribution of the {112} plane and 'c' was the contribution of the {123} plane. These simulations were used to investigate the performance of the model in three areas. The first was where the {110} slip plane provided low percentage slip plane contribution contrary to observations from crystal geometry studies [24]. The second was where the {112} and {123} slip planes provided equal contributions again contrary to observations from crystal geometry studies [24]. The results obtained from these simulations showed that the

simulation resulted in scattered data leading to the inference that this model correctly demonstrated that low percentage contributions by the {110} plane and equal contribution by the {112} and {123} planes did not occur. The third was where the {110} slip plane was the main plane of activity and hence had the highest percentage slip plane contribution consistent with observations from crystal geometry studies [24].

The second validation criterion was generation of acceptable deformation curves. Fig. 2 shows curves incorporating dislocation information simulations for percentage slip plane contribution 70-25-5, 75-20-5, 80-15-5, 75-25-0, 80-20-0, 90-10-0 and 95-5-0.

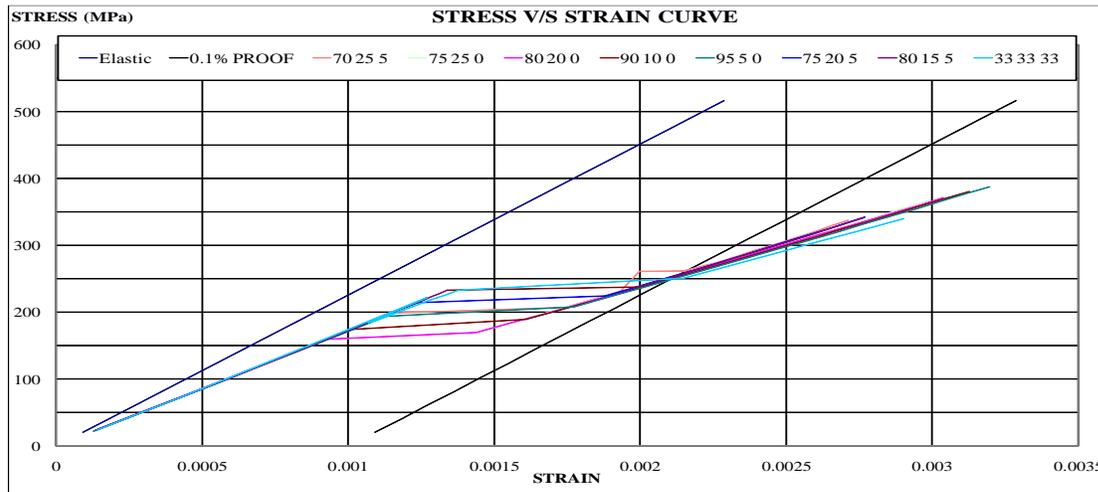


Fig 2: Stress/Strain curves for coarse mesh simulation for various slip plane percentage contribution

The elastic range of these curves consisted of straight lines. This indicated that the stress was proportional to the strain despite the incorporation of dislocation information, and Hooke’s law was upheld. Elastic constants $C_{11}=24.2$, $C_{12}=14.65$, $C_{44}=11.2$, $G=8.6$, Poisson’s ratio of 0.291 [24] were used as input data. Young’s Modulus obtained for these curves lay between 171-174 GPa which was 17% below the value of 206 GPa obtained from conventional analysis for low carbon steels [26]. This was attributed to the fact that crystal material properties used in the dislocation data were for perfect, stress free iron crystals grown in controlled conditions. These did not incorporate the effects induced by carbon content and other alloy constituents that strengthen typical steel materials at the macro level. The model therefore conformed to the expected behavior in the elastic region of loading with dissonance occurring only in the value of the Young’s Modulus.

In addition the curves showed a “plateau” yield at near constant stress as dislocation contribution become significant. The ability of the formulation to predict yield was considered a major achievement and this was consistent with existing knowledge that dislocation evolution was the principal contributor to plastic flow.

These results also showed that as the percentage contribution of the {110} slip plane was increased, the yield stress increased. This indicated that the existence of the {112} slip plane resulted in a weaker material matrix. This was consistent with the fact that the {112} slip plane is a twinning plane consisting of a material defect with lower slip activation energy levels [24]. In addition, the two-slip plane systems exhibited lower yield values than the 3-slip plane system. This suggested that the {123} slip plane required larger amounts of energy for activation of slip than the other two slip planes. This again was consistent with empirical evidence as the {123} slip plane is active only at elevated material temperatures which provide the additional slip activation energy required.

The post yield behavior for all the cases studied showed material recovery and a straight line post yield behavior of gradient 117-124 GPa. The recovery was attributed to the countering of the dissipative effects of the dislocations by the strengthening effects of the increased dislocation density. The simulations generated an almost constant gradient post yield curve. This was contrary to experimental observation and was attributed to the use of a constant stored dislocation density and a stress based dislocation density algorithm. It is proposed that future work focus on the development of dislocation density evolution algorithms.

Key to this formulation was the notion of the “percentage slip plane contribution”. Drawing from minimization concepts a plot of the Yield stress versus the {110} percentage slip plane contribution for 2 plane systems were prepared and is shown in Fig. 3. The results presented show a minimum yield stress of 159MPa at 75% {110} percentage slip plane contribution for the 2 slip plane system for all obstacles densities equal to or less than 10×10^{21} . At 10×10^{24} the yield stress dropped to 137MPa but at

the same slip plane contribution. This common minimum suggested that the representative Yield Stress was in the order of 137MPa to 159MPa at a {110} percentage slip plane contribution of the order of 75%.

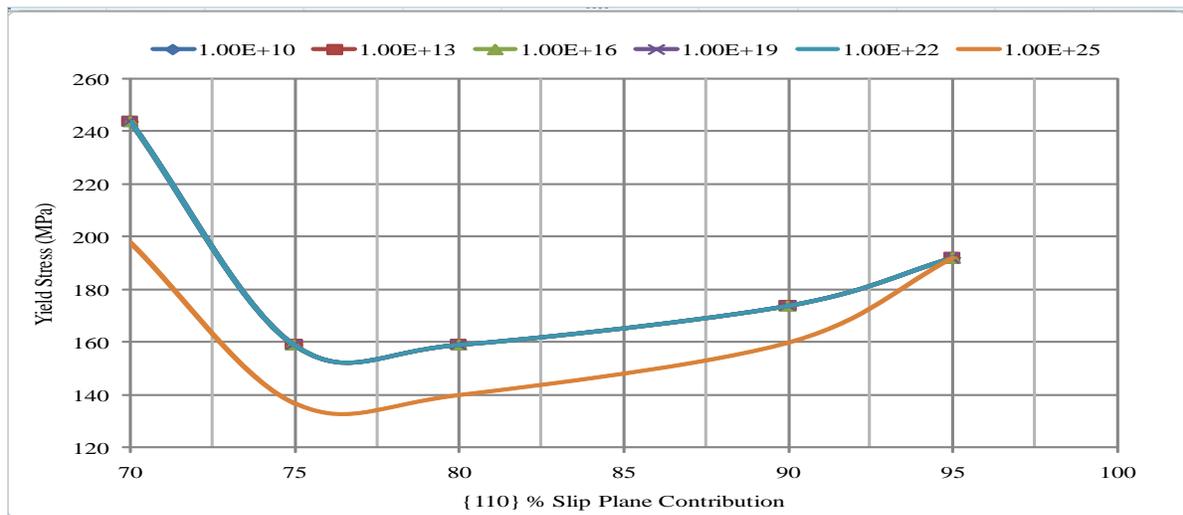


Fig 3: Yield Stress v/s {110} Slip plane percentage contribution For 2 Plane Slip System for Various obstacle densities

The formulation utilized a black box technique to superimpose dislocation information onto an elastic matrix thereby precluding the need to monitor individual dislocation segments. This was one of the principal advantages of the method as the simulation's demand for computational power approached that of standard elastic analysis with the only additional load existing in the micro element's material property computation. The 3D discrete dislocation formulation [11], executed a variable time step and simulation cycles were in the order of one million to achieve a strain of 0.3% for nano-scale samples. The method presented in this paper achieved similar strains in 25 cycles which took approximately 2 seconds for "coarse mesh" of 40 finite elements and 76 seconds for "fine mesh" of 240 finite elements.

It was found that the method was not constrained by large increases of dislocation segments which posed a computational challenge during the evaluation of the polarization stress for the standard discrete dislocation method. In addition, the method was not limited to nano-technology, again due to the computational challenge encountered as was 3D discrete dislocation method. The method did require the differentiation of the levels of initial plasticity across the specimen as did the superposition discrete dislocation method and this was conveniently achieved by specifying initial dislocation densities values at the macro scale. Traction boundary conditions were also handled at the macro scale.

4. Conclusion

This work detailed the development and evaluation of a variational technique that utilized a discrete dislocation formulation in a black box finite element technique. The dislocation information was contained in a scalar density function. The results confirmed that this minimization of total energy involving two length scales, and the correlation of dislocation segment lengths and dislocation densities was acceptable.

The results agreed with existing literature and it was observed that slip plane activity was shared between a {110} slip plane and the {112} slip plane. In addition the results were consistent with empirical findings that indicated that the {110} slip plane had higher slip activation energy than the {112} slip plane, and the {123} plane required the highest slip activation energy normally provided in high temperature loading.

The model correctly simulated elastic behavior and upheld Hooke's law. The model also correctly simulated material yielding and post yield material recovery. However the choice of dislocation density evolution algorithm was considered to be a critical component in the model's inability to simulate post yield strengthening.

The method successfully demonstrated the direct use of dislocation evolution information in an elastic finite element simulation to generate elastic behavior and elastic-plastic transition.

5. References

1. **D. L. McDowell, C. N. Paden Jr.** Future Research Trends In Metal Plasticity For Simulation of Metals Processing and Life Cycle Analysis. Georgia : Georgia Institute of Technology, 2004.
2. **F. F. Abraham, J. Q. Broughton, N. Bernstein, E. Kaxiras.** Spanning the Length Scales in Dynamic Simulation. Computers in Physics. 1998, Vol. 12, 6.
3. **P. Gumsch.** An Atomistic Study of Brittle Fracture: Towards Explicit Failure Criteria From Atomistic Modeling. Journal Materials Research. 1995, Vol. 10, 11, pp. 2897-2907.
4. **F. F. Abraham, J. Q. Broughton, N. Bernstein, D. Hess.** Dynamic Fracture of Silicon: Concurrent Simulation of Quantum Electrons, Classical Atoms and the Continuum Solid. MRS Bulletins. 2000, Vol. 25, 5, pp. 27-32.
5. **V. Shenoy, R. Miller, E. B. Tadmor, D. Rodney, R. Phillips, M. Ortiz.** An Adaptive Methodology for Atomic Scale Mechanics-The Quasi Continuum Method. Journal of Mechanical Physical Solutions. 1999, Vol. 47, pp. 611-642.
6. **M. S. Daw, M. I. Baskes.** Embedded Atom Method: Derivation and Applications to Impurities, Surfaces, and Other Defects in Metals. Physical Review B. 1984, Vol. 29, pp. 6443-6453.
7. **J. P. Hirth, M. Rees, H. Zbib.** Modeling of Deformation by a 3D Simulation of Multiple Curved Dislocations. Journal of Computer Aided Materials Design. 1996, Vol. 3, 1, pp. 164-166.
8. **L. P. Kubin, G. Canova.** Modeling of Dislocation Patterns. Scripta Metallurgica et Materialia. 1992, Vol. 27, 8, pp. 957-962.
9. **Van der Gissen, A. Needleman.** Discrete Dislocation Plasticity: A simple Planar Model. Model. Simul. Mater. Science Eng. 1995, Vol. 3, pp. 689-735.
10. **M. P. O'day, W. A. Curtin.** A Superposition Technique for Discrete Dislocation Plasticity. Brown University, USA. 2004.
11. **H. M. Zbib, M. Shehadeh, S. M. A. Khan, G. Karami.** Multiscale Dislocation Dynamics Plasticity. School of Mechanical and Materials Engineering, Washington State University. 2002.
12. **J. Q. Broughton, F. F. Abraham, N. Bernstein, E. Kaxiras.** Concurrent Coupling of Length Scales: Methodology and Applications. Physical Review B. 1999, Vol. 60, pp. 2391-2403.
13. **H. S. Park, E. G. Karpov, W. K. Lui, P. A. Klein.** The Bridging Scale Method for Two-dimensional Atomistic/Continuum Coupling. Philosophical Magazine. 2005, Vol. 85, pp. 79-113.
14. **S. Hao, W. K. Lui, B. Moran, F. Vernerey, G. B. Olson.** Multi-scale Constitutive Model and Computational Framework for the Design of Ultra-High Strength, High Toughness Steels. Computational Methods in Applied Mechanical Engineering. 2004, Vol. 193, pp. 1865-1908.
15. **D. E. Spearot, K. I. Jacob, D. L. McDowell.** Nucleation of Dislocations from [001] Bicrystal Interfaces in Aluminum. Acta Materialia. 2005, Vol. 53, pp. 3579-3589.
16. **J. D. Clayton, P. W. Chung.** An Atomistic to Continuum Framework for Nonlinear Crystal Mechanics Based on Asymptotic Homogenization. Journal of the Mechanics and Physics of Solids. 2006, Vol. 54, pp. 1604-1639.
17. **B. Svendsen.** On the Variational Formulation of Models for Materials with Microstructure. Proceedings of Applied Mathematical Mechanics. 2003, Vol. 3, pp. 220-221.
18. **B. Svendsen.** On Thermodynamical and Variational Based Formulations of Models with Inelastic Continua with Internal Length Scales. Computer Methods in Applied Mechanics and Engineering. 2004, Vol. 193, 48-51, pp. 5429-5452.
19. **P. Cermelli, G. Leoni.** Renormalized Energy and Forces on Dislocations. SIAM Journal on Mathematical Analysis. 2005, Vol. 37, 4, pp. 1131-1160.
20. **I. Groma, G. Gyorgyi, P.D. Ispanovity.** Variational Approach in Dislocation Theory. Department of Materials Physics, Eotvos University. Budapest : arXiv.org, 2009. pp. 1-19. 1105v2.
21. **I. Groma, G. Gyorgyi, B. Kocsis.** Debye Screening of Dislocations. Physical Review Letters. 2006, Vol. 96, 16, pp. 5503-4.
22. **I. Groma, P. D. Ispanovity.** Role of Elastic Anharmonicity in Dislocation Patterning. Physical Review B. 2007, Vol. 76, 5, p. 4120.
23. **S. Groh, H. M. Zbib.** Advances in Discrete Dislocations Dynamics and Multi Scale Modeling. Journal of Engineering Materials and Technology. 2009, Vol. 131, 4, pp. 1209.
24. **J. P. Hirth, J. Lothe.** Theory of Dislocations. 2nd. Ed. : John Willey & Sons, 1982.
25. **Argon, S. Ali.** Constitutive Equations in Plasticity. The MIT Press, 1975.
26. **E. C. Rollanson.** Metallurgy for Engineers. 4th. Ed. : Edward Arnold Publishers, 1973. p. 363.

Dynamic analysis of dislocation cores in the α -Fe lattice using the embedded atom model

**K. D. Njoroge^{1,2,5}, G. O. Rading^{1,5}, J. M. Kihuu^{2,5}, M. J. Witcomb^{3,5},
L. A. Cornish^{3,4,5}**

¹ Department of Mechanical and Manufacturing Engineering, School of Engineering, University of Nairobi, P.O. Box 30197-00100, Nairobi, Kenya.

² Department of Mechanical Engineering, School of Mechanical, Manufacturing and Materials Engineering, Jomo Kenyatta University of Agriculture and Technology, P.O. Box 62000-00200, Nairobi, Kenya.

³ DST/NRF Centre of Excellence in Strong Materials, University of the Witwatersrand Private Bag 3, WITS, 2050, Johannesburg, South Africa.

⁴ School of Chemical and Metallurgical Engineering, University of the Witwatersrand Private Bag 3, WITS, 2050, Johannesburg, South Africa.

⁵ African Materials Science and Engineering Network (AMSEN)

Research done at the Department of Mechanical and Manufacturing Engineering, School of Engineering, University of Nairobi, P.O. Box 30197-00100, Nairobi, Kenya.

Abstract

The Embedded Atom Method (EAM) was employed to study the structure of body centered cubic (BCC) dislocation cores. Core energies, number of nearest neighbour atoms, stress tensor components, resolved shear stresses and dynamic dislocation core stresses were calculated for four types of dislocation cores. A dynamic dislocation model was presented and a “path of least resistance” (POLR) mechanism suggested for the determination of the Peierls stress. It was concluded that a sequence of stress components acting on the dislocation core in a $\{110\}\langle 111 \rangle$ slip system were responsible for the proposed core atom motion resulting in the overall dislocation motion. A review of the resolved shear stress in the lattice was then used to collaborate the results of the dynamic dislocation model and the core atom motion mechanism model.

Key words: Embedded Atom Method, Dislocation, Peierls Stress, Body Centered Cubic

1. Introduction

The characterization of dislocation properties will frequently go beyond the simple continuum model and involve the lattice structure at the dislocation's core. The effects of the dislocation core have been found to be significant in the low temperature behavior of body centered cubic (BCC) metals [1], the plastic anisotropic behavior of ionic crystals [2, 3] and hexagonal close packed (HCP) metals [4], and the dislocation climb behavior in oxide crystals [5].

The Peierls-Nabarro model was the first successful model to account for the structure at the core of the dislocation [6]. A more accurate representation of the core configuration was later accomplished through lattice static or atomic models. Lattice static models are limited to the harmonic approximation, which is not strictly valid in the dislocation's core region. Atomic models are more suited to dealing with highly distorted lattice configurations [7]. In atomic models, the volume surrounding the core of the dislocation (region I) has been modeled as consisting of atoms interacting via a given non-linear potential, while the volume surrounding region I (region II) has been treated using the theory of elasticity. The power of the atomic models lies in their ability to accurately describe the core, which is dependent on the physical validity of the potentials used to describe the interaction among the atoms. A rigorous determination of the minimum energy with respect to the configuration of the dislocation core requires the use of quantum mechanical (QM) models. However, the highly distorted nature of the core severely limits the detail to which these interactions can be evaluated. The aim has been to find tractable approximate treatments, which are physically accurate over the range of distortions anticipated. The description of the core may then be used to predict the physical behavior of the dislocation in terms of its evolution and motion, and to determine the strain energy and the Peierls stress at which these physical actions take place.

The EAM has matured as a technique and is now applied in the study of material structures containing defects [8, 9, 10, 11]. Applications have been developed for the study of the structure of tilt boundaries [12, 13], phonon dispersion [14, 15], linear thermal expansion [15], point defects [14, 16, 17], and lattice dynamics [14, 18]. More recent research has focused on the modeling of potentials for application to industry specific problems. Potentials for studies on the effect of phosphorous on the embrittlement of nuclear reactor pressure vessels [19], the study of reactor pressure vessel steel thermal annealing [20], the generation of phase diagrams [1, 21], and the dissolution and diffusion of hydrogen in bulk α -Fe as well as binding of hydrogen to surfaces, vacancies and dislocations [22] are available. However, the simulation of the dynamic-spatial atomic

structure of dislocation cores, and its contribution to the behavior of dislocations viewed as an atomic scale problem, has not been comprehensively addressed. There exists opportunity to advance the development and application of potentials to model dislocation cores and dislocation families for the application to a wider range of industrial problems.

2. Method

This work generated the dislocated structure by the application of a set of displacement vectors to the points of a perfect lattice about a prescribed dislocation core centre. The dislocation line direction was set as $\frac{1}{2} [1\bar{1}1]$. The distortion was

introduced and the slip plane created by displacing the block of atoms on one side of a predetermined plane by half a Burgers vector. A transition zone was formed around the dislocation core where the first and second rows of atoms were displaced at 1/6 and 1/3 of the Burgers vector.

The direction of the Burgers vector was defined for the pure edge, pure screw and mixed dislocations as that direction, oriented such that together with the dislocation line vector, defined the dislocation type, and such that this direction resulted in a whole number of lattice displacements for the defined dislocation type. Consequently, two pure and two mixed dislocations were possible for the body centred cubic (BCC) lattice. These were the pure screw, the 35.26 degree screw, the 70.53 degree screw and the pure edge dislocations. The dislocation core was formed at the intersection of the slipped and unslipped block of atoms at the edge of the slip plane. The use of the different displacement vectors resulted in the generation of the different types of dislocation cores.

Simulation was carried out for static and dynamic conditions. The dynamic conditions were generated by moving the block of atoms on one side of the slip plane by predefined displacement vectors. These displacement vectors were set to a fraction of the Burgers vector for each dislocation type, within the slip plane.

This work applied the Embedded Atom Method (EAM) potential of Mendeleev et al. [23] to evaluate the cohesive strength and stress tensor of the four different types of dislocation cores. This potential has been tested and found to stabilize the non-degenerate dislocation cores on $\{110\}$ glide planes in agreement with experimental data for dislocation motion at low

temperatures [24]. The cohesive strength and the stress tensors for the dynamic simulations were calculated after each incremental displacement giving rise to a varying strength and stress profile.

Code was prepared on a Fortran 90 platform and the computation utilized a single scale atomic model with an extended "region I" of 1331 atoms in an "11 x 11 x 11" block of atoms to carry out the simulations. This was in variation to the traditional atomic model [7], as it eliminated the need to couple the two windows of resolution.

This paper presents results on the simulation of the behavior of the dislocation cores. Section 1 presents results from the review of the cohesive strength and stress tensor of the dislocation core at equilibrium lattice spacing. This is followed by Section 2, which presents a dynamic dislocation stress analysis.

3. Results And Discussions

3.1. Static Analysis

The cohesive strengths of the static dislocated lattices are given in Table 1. These results were obtained from simulations for the four types of dislocation cores and compared with the results for the perfect lattice. The results revealed a reduction of cohesive strength with the introduction of dislocations, consistent with empirical observations. However, the differences in cohesive strength were small (less than 1%), hence the results were only of qualitative value. Whereas the pure screw dislocation presented a higher cohesive strength per atom than the edge dislocation, indicating higher edge dislocation mobility consistent with empirical findings [25, 26], the 70.53 degree screw did not show a high cohesive strength, which was inconsistent with empirical findings, which report that the 70.53 degree mixed dislocation has exceptionally high Peierls stress [27].

Two reasons were suggested for these observations. The first was that these results were generated from simulations of a static lattice structure, providing a snapshot of the comparative cohesive strength of the equilibrium lattice. This state may not be that where the Peierls stress would be determined. The second was that the lack of gradient terms in the formulation of the cohesive strength may render the computation incapable of capturing the distortion gradient in the dislocation core. This suggestion was reasonable as it is observed that the embedding energy did not vary with the type of dislocation, while the pairing energy varied only marginally. The evaluation of the dynamic dislocation was used to explore the first assertion, while the use of a gradient sensitive electron density function was suggested as a test of the second assertion.

The cohesive energy per atom about the dislocation core was taken as a measure of the limiting stress required to overcome inter atomic bonding, to enable dislocation glide. This was then related to the macroscopic yield stress by the amalgamation of the various effects of the density of dislocations contained in the matrix. However, the above results illustrated the shortcomings of this approach, and the postulated hypothesis that a gradient-based EAM functional was more suitable, will be evaluated elsewhere. The use of dislocation dynamics for the direct computation of the Peierls stress was feasible and its

application has been supported by existing literature where empirical findings relate the macroscopic yield stress to the Peierls stress at low temperatures [28].

Table 1: Embedding and pair energy contributions to the cohesive energy for four types of dislocations (eV/atom).

Dislocation type	Embedding energy, $F(\rho)$	1 st Pair potential, $V(r_{ij})$	2 nd Pair potential, $V(r_{ij})$	3 rd Pair potential, $V(r_{ij})$	Cohesive energy / atom, E_c (eV)	No. of nearest neighbours	Burgers vector direction
Pure screw	-5.495193	1.48466	0	0	-4.01053	57	$[\bar{1} 1 \bar{1}]$
70.53° screw	-5.495193	1.48539	0	0	-4.009801	57	$[\bar{1} 1 0]$
35.26° screw	-5.495193	1.48466	0	0	-4.01053	57	$[\bar{1} 1 \bar{1}]$
Pure edge	-5.495193	1.48514	0	0	-4.010046	56	$[\bar{1} 1 2]$
Perfect lattice	-5.495382	1.48328	0	0	-4.012102	58	-

Simulations were also carried out to generate the stress tensor components at the dislocation core. These values represented the resulting stress due to lattice distortion forming the dislocation core. Six stress tensor components were generated with the remaining three obtained by application of principle of complimentary stress.

Table 2: Stress tensor components calculated for four types of dislocations (GPa.).

Dislocation type	σ_{11}	σ_{12}	σ_{13}	σ_{22}	σ_{23}	σ_{33}
Pure screw	1.4817	-6.2778	-5.1545	4.1319	-5.6076	3.7813
70.53 degree screw	-2.163	-7.6378	-5.5587	4.6258	-5.7451	2.6820
35.26 degree screw	1.4947	-6.2258	-5.1025	4.3527	-5.4128	4.0021
Pure edge	2.2437	-7.1664	-5.0673	1.9309	-5.8403	3.7959

Results presented in Table 2 revealed that the lowest stresses were the σ_{11} tensor component, which corresponded to $[100]$ direction, and the σ_{22} tensor component, which corresponded to the $[010]$ direction. The $\langle 100 \rangle$ directions were of interest as dislocations in these directions contribute substantially to the flow stress [29] and are formed by the interaction of two $\frac{1}{2}\langle 111 \rangle$ dislocations. The resulting stress components suggested that two or more stress components combined to generate the $\langle 110 \rangle$ type motion. Additionally, dislocation core reconstruction across a $\langle 110 \rangle$ zone axis was attributed to the similar combination of stress components. This was consistent to observations of dissociation of screw dislocation cores across adjacent $\langle 110 \rangle$ planes [30, 31, 32]. The intersection of non-coplanar dislocations and their interaction was enhanced by the presence of a $\langle 110 \rangle$ zone axis. In addition, edge dislocation core spreading to adjacent $\langle 112 \rangle$ planes has been reported [33]. It was proposed that dislocation core spreading and core-core interaction be explained by the analysis of combination of dislocation core stress components.

The precise motion of the individual atoms was not determined, but the values of the stress tensor components suggested that motion in the $[100]$, $[010]$ and $[001]$ directions was to be expected. This motion corresponded to atomic movement in different $\{110\}$ and $\{112\}$ planes belonging to the given $\langle 110 \rangle$ zone axis. The very nature of a 3-dimensional lattice suggested that out-of-plane motion was necessary for the “micro-cleavage” of the lattice to enable dislocation motion. Additionally, the presence of the $\langle 110 \rangle$ zone axis enabled the out-of-plane motion to occur at the lower stress levels

corresponding to the [100], [010] and [001] directions. This visualization for dislocation core evolution was consistent with observations of moderate non-planar spreading of dislocation cores in BCC lattice [32].

3.2. Dynamic Analysis

Peierls stress calculations have concentrated on the body centered cubic (BCC) metals because it was realized that their dislocation core effects contribute to a high Peierls stresses. Dislocation dissociation in BCC metals is localized to the core region (there is no stable stacking fault separating the partials) and, in the case of the $\langle 111 \rangle$ screw dislocation, it is non-planar. The BCC screw dislocation core is rather compact and this gives rise to a large Peierls stress [34].

Figures 1 to 5 present the variation of the lattice stress about the dislocation core as the dislocation moves rigidly along the direction of its Burger's vector. All the dislocations presented different sinusoidal stress component variation. These curves suggested that the movement of atom rows past rows of other atoms was synchronised, enabling the simulation to detect the variation. The positive and negative values were interpreted as tensile and compressive stresses respectively and it was noted that the direct stresses were positive (except for σ_{11} component where the troughs dip below zero) and the shear stresses are negative. It was expected that movement of atoms around the dislocation core resulting in dislocation motion would follow the path of least resistance (POLR) and would involve changes in direction.

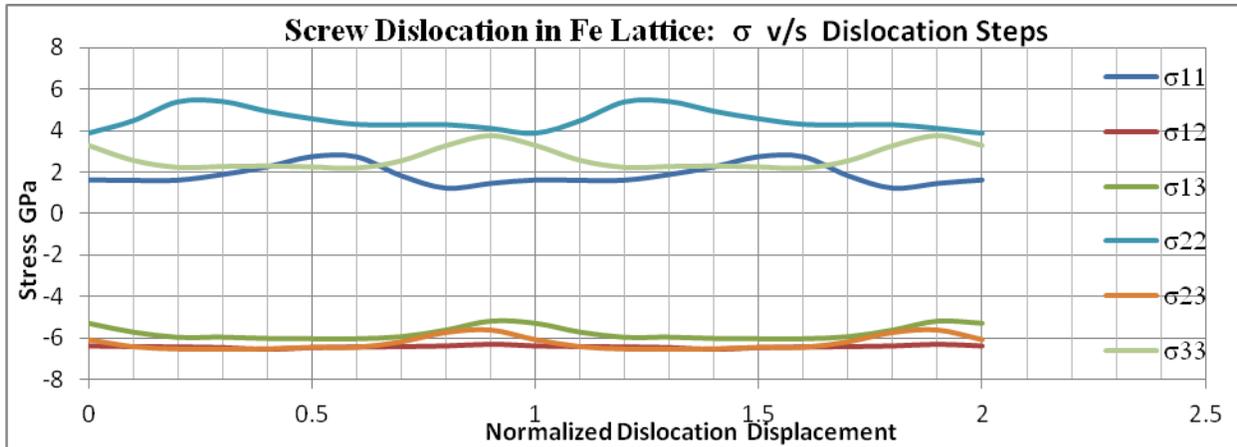


Fig 1: Variation of the stress generated as a pure screw dislocation in Fe lattice moves in the $[1\bar{1}1]$ direction.

The stress cycle for the pure screw is presented in Figure 1. The POLR was in the σ_{11} direction over most of the path. At some point, the curve for σ_{11} intersected that of σ_{23} and exceeded it for a short portion of the cycle. Over this period, σ_{23} was at its lowest value and provided the POLR, after which σ_{11} resumed as the POLR. This sinusoidal path oscillated between 1.26GPa and 2.27GPa. This compared well with published values of the Peierls stress of 1.2 to 1.8GPa from Chaussidon et al. [24] and 1.3 to 1.9GPa from Ventelon [35] for BCC Fe. The stress component σ_{11} corresponded to motion along the (011) plane, and involved displacement of atoms away from the (110) glide plane ahead of the dislocation core. The stress component σ_{33} corresponded to motion along the (110) plane, and involved displacement of atoms along the glide plane ahead of the dislocation core. The proposed sequence supported the hypothesis that dislocation motion occurred predominantly by the “peeling” of atoms around the dislocation core as a stress-enabled mechanism. The shear stresses were in the order of magnitude -6.44GPa, with smaller amplitude than the direct stresses. Additionally, the magnitude of the shear stresses was of the same order of magnitude as the peak direct stresses of 5.41GPa. These higher stress values were an indication of the stress required to generate catastrophic fracture of the lattice. Consequently, it was reasonable to infer that the ultimate tensile stress should be in the order of 2.8 times the yield stress. This result was obtained by comparing the maximum stress values related to catastrophic failure and those related to dislocation motion and hence material yielding.

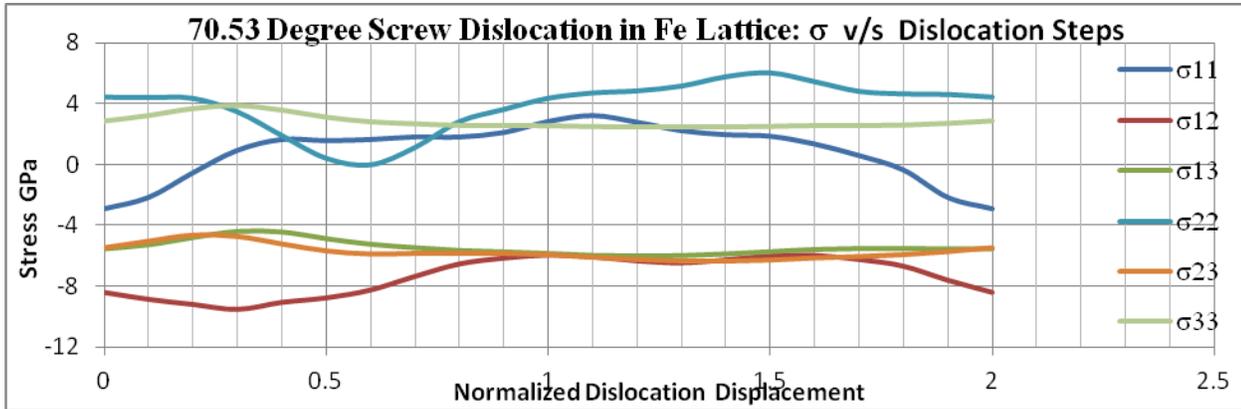


Fig 2: Variation of the stress generated as a 70.53 degree screw dislocation in Fe lattice moves in the $[\bar{1}10]$ direction.

The stress component cycles for the 70.53 degree screw are presented in Figure 2. The POLR was along the direction corresponding to the σ_{11} stress component. It was noteworthy that this stress component traversed the zero stress value, with a maximum positive value 3.21GPa and a maximum negative value of -2.91GPa. This curve intersected that of σ_{22} at 1.91GPa and σ_{23} at 2.54GPa at different periods of the cycle. The intersection occurred over a small interval, and the resulting change in POLR, for the atoms around the dislocation core traversed three directions. The larger maximum value of σ_{11} (though negative in sign) implied that a 70.53 degree screw was harder to move than a pure screw dislocation. In comparison to the results for the pure screw dislocation, the maximum direct stress increased to 6.06GPa, with a corresponding increase of the maximum shear stress value to -9.55GPa. The larger value of the maximum shear stress and the divergence in shear stress curves implied that a narrow band catastrophic ultimate stress would not result from the presence of 70.53 degree screw dislocations. These results were supported by empirical findings [27] that report that an exceptionally high Peierls stress is experienced by the 70.53 degree screw.

The stress component curves for the 35.26 degree screw are presented in Figure 3. The POLR was again along the direction corresponding to the σ_{11} stress component. This stress component oscillated between a maximum value of 2.76GPa and a minimum value of 1.27GPa. The POLR changed direction when the σ_{22} stress component curve intersected with the σ_{11} stress component curve at about 2.42GPa. The σ_{22} stress component corresponded to motion on the (011) plane, and involved movement of atoms away from the slip plane around the dislocation core. A maximum magnitude shear stress of -6.51GPa was noted, with a spread of the shear stress components similar to that of the pure screw dislocation. This implied that a narrow band catastrophic ultimate stress may result from the 35.26 degree screw dislocation.

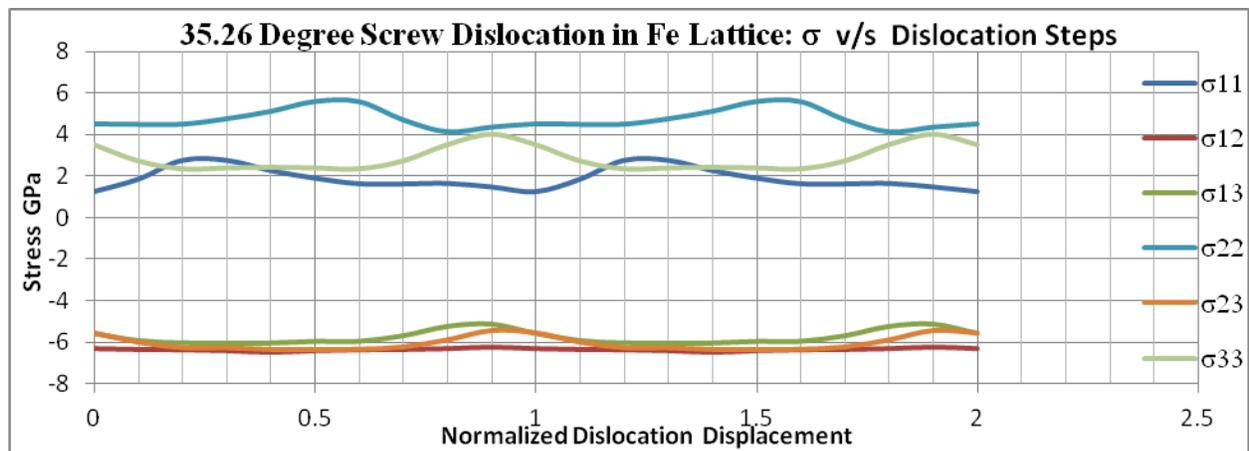


Fig 3: Variation of the stress generated as a 35.26 degree screw dislocation in Fe lattice moves in the $[\bar{1}11]$ direction.

Results on simulations for the pure edge dislocation presented in Figure 4 presented the most random curves for the four types of dislocations. Similar sinusoidal curves were obtained for the shear stress components. The POLR was again provided by the curve corresponding to the σ_{11} stress component, which spanned the range -0.848GPa to 2.45GPa. Along this curve, intersection with σ_{22} stress component curve at 1.93GPa and σ_{33} stress component curve at 2.4GPa occurred with corresponding change in direction into adjacent zone $\langle 110 \rangle$ axis planes. The availability of a larger number of atom motion directions, only replicated by the 70.53° screw, was envisaged as the possible mechanism for a greater mobility of the dislocation core. Additionally, this dislocation presented the lowest magnitude stress of -0.848GPa, which was considered to enhance dislocation motion over other types of dislocations. The spreading of the shear stress curves suggested that a narrow band catastrophic ultimate stress would not result from these dislocations.

The curves in Figures 1 to 4 demonstrate that the Embedded Atom Method (EAM) formulation applied to evaluate the Peierls stress had the required sensitivity to differentiate between the dislocation types. This supported the earlier assertion that dislocation dynamics based upon the EAM model presented a more complete description in material defect analysis. Specifically, it was suggested that the interaction of combinations of stress components resulted in the POLR, and therefore determined the Peierls stress profile. The lack of symmetry of the stress curves resulted from the effects of the twinning/anti-twinning asymmetry of the BCC lattice, and this was therefore more pronounced for the edge dislocation. The short range nature of the EAM potential resulted in a stress computation at the dislocation core only. It was proposed that a more complete approach would be to generate stress components due to dislocation cores at locations removed from the dislocation core. This would be invaluable in the study of dislocation interaction with other dislocations, obstacles and grain boundaries. Thus, there was need for the development of a longer range potential, specifically suited to dislocation interaction.

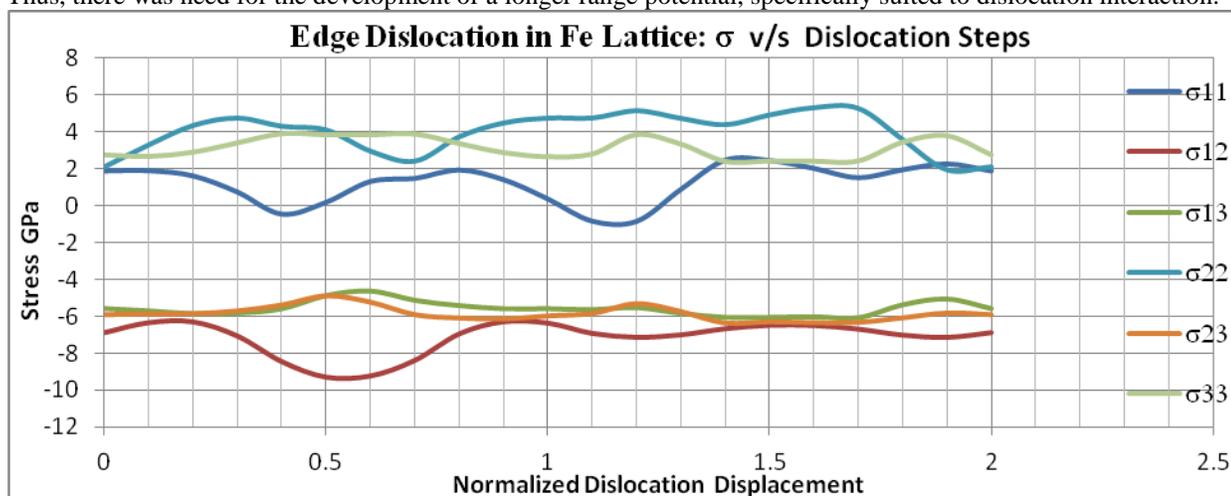


Fig 4: Variation of the stress generated as a pure edge dislocation in Fe lattice moves in the $[\bar{1}12]$ direction.

Schmid's law [36] states that glide on a given slip system is controlled by the resolved shear stress on that system known as the Schmid stress, and in a rate-independent formulation, glide commences when this stress reaches a critical value known as the critical resolved shear stress (CRSS). Implied is that plastic deformation should not be affected by non-glide stress tensor components and is largely dependent on the existence of planar dislocation dissociation mechanisms. For body centered cubic (BCC) lattice structures, stress tensor components in associated $\{110\}$ planes may act on dislocation cores spread into these associated planes resulting in non-glide stresses contributing to plastic deformation. Non-Schmid behaviour of BCC metals therefore results from the three way dislocation core dissociation or spreading on associated $\{110\}$ zone planes and is supported by atomistic calculations [26, 37, 38, 39]. However, a convergence of atomistic calculations and experimental results is lacking, with computations overestimating the Peierls stress [40, 41]. The resolved shear stress in the direction of the Burger's vectors is presented in Figure 5.

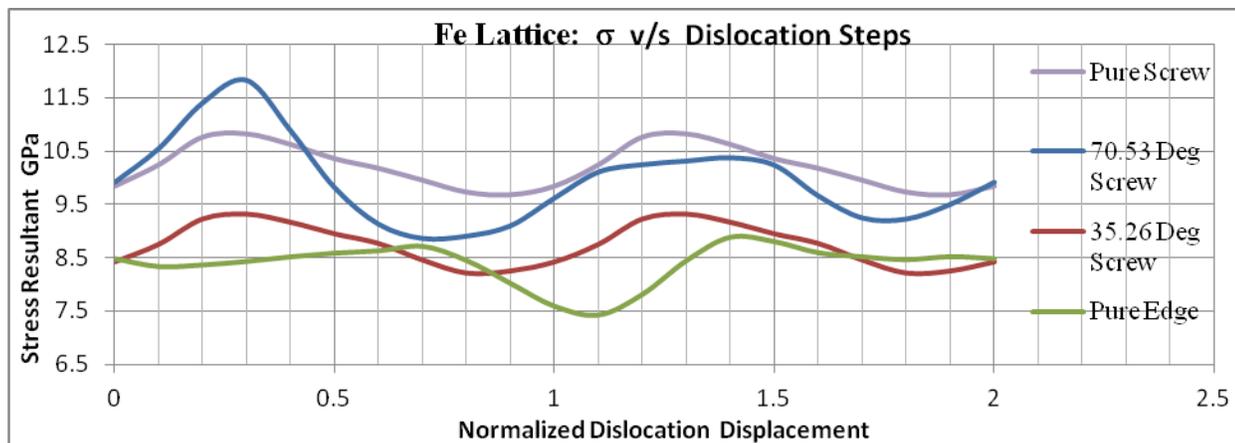


Fig 5: Variation of the resolved stress generated as a dislocation in Fe lattice moves in the direction of the burgers vector.

Figure 5 summarised the results of Figures 1 to 4 and revealed that the edge dislocation provided the least resolved shear stress over most of the dislocations motion. It was therefore inferred that the edge dislocation possessed the greatest mobility consistent with empirical findings [25]. The 70.53 degree screw dislocation presented a curve with a much higher amplitude and peak stress value. This was again consistent with empirical findings, which reported its high Peierls stress [27]. Also noteworthy was that the magnitude of the resolved shear stress was much higher than both experimental and calculated Peierls stress [35]. Therefore it is inferred that the Peierls stress was related to the active stress components, and not to the overall resolved shear stress.

4. Conclusions And Recommendations

The application of the Embedded Atom Method (EAM) to the evaluation of the cohesive strength and the stress tensor at the dislocation core, using a single static window of resolution, demonstrated the limitation of the technique and led to the proposal that dynamic dislocation analysis or a gradient-based functional be developed to enhance the method's sensitivity to distortions at the dislocation core. Stress components generated at equilibrium lattice conditions were used to show the potential relation between the evolution of the dislocation core and the specific displacements related to the stress components.

Analysis of the dislocation in motion was also presented and a "path of least resistance" POLR mechanism proposed to explain the evolution of the dislocation core. These results illustrated the contribution of different stress components to the motion of atoms about the core and supported the proposed "peeling" of atoms about the core. A comparison of the resolved shear stress in the lattice provided a alternative view to collaborate the results obtained from the analysis of the dynamic dislocation core and the proposed mechanism explaining the motion of the dislocation core atoms. Finally, it was concluded that the Peierls stress was related to the active stress components, and not to the overall resolved shear stress.

Acknowledgements

This research is funded by the African Material Science and Engineering Network (AMSEN) vote no. 500-661-382, a Carnegie-IAS Regional Initiative in Science and Education Network. This research is also supported by the Jomo Kenyatta University of Agriculture and Technology. This support is greatly appreciated.

References

1. G. Bonny, R. C. Pasianot, L. Malerba., Fitting Interatomic Potentials Consistent with Thermodynamics: Fe, Cu, Ni and their Alloys. *Philosophical Magazine*. 2009, Vol. 89, 34-36, pp. 3451-3464.
2. M. P. Puls. Liblice : s.n., 1983. Proceedings of "The Structure and Properties of Crystal Defects".
3. W. Skrotzki, P. Haasen. Hardening Mechanisms of Ionic Crystals on {110} and {100} Slip Planes. *Journal Physical Science*. 1981, Vol. 42, C3, pp. 119-148.
4. D. J. Bacon. Liblice : s.n., 1983. Proceedings of "The Structure and Properties of Crystal Defects".
5. T. Bretheau, J. Castaing, J. Rabier, P. Veyssiere. Dislocation Motion and High Temperature Plasticity of Binary and Ternary Oxides. *Advances in Physics*. 1979, Vol. 28, 6, pp. 835-1014.
6. F. R. N. Nabarro. Dislocations in a Simple Cubic Lattice. *Proceedings of the Physical Society*. 1947, Vol. 59, 2, pp. 256-272.

7. **M. P. Puls.** *Dislocation Modeling of Physical Systems.* [ed.] R. Bullough, G.S. Hartley, J.P. Hirth, M.F. Ashby. New York : Pergamon Press, 1981. p. 249.
8. **M. S. Daw, M. I. Baskes.** Embedded Atom Method: Derivation and Applications to Impurities, Surfaces, and Other Defects in Metals. *Physical Review B.* 1984, Vol. 29, pp. 6443-6453.
9. **F. Ercolessi, M. Parrinello, E. Tosatti.** Simulation of Gold In The Glue Model. *Philosophical Magazine A.* 1988, Vol. 58, pp. 213-226.
10. **S. M. Foiles, M. I. Baskes, M. S. Daw.** Embedded Atom Method Functions for the FCC Metals Cu, Ag, Au, Ni, Pd, Pt and their Alloys. *Physical Review B.* 1986, Vol. 33, 12, pp. 7983-7991.
11. **M. S. Daw, S. M. Foiles, M. I. Baskes.** The Embedded-Atom Method: A Review of Theory and Applications. *Materials Science Reports.* 1993, Vol. 9, pp. 251-310.
12. **M. Yan, M. Sob, G. J. Ackland, D. E. Luzzi, V. Vitek, M. Methfessel, C. O. Rodriguez.** Interatomic Forces and Atomic Structure of Grain Boundaries in Copper-Bismuth Alloys. *Physical Review B.* 1993, Vol. 47, 10, pp. 5571-5583.
13. **R. Besson, A. Fraczkiewicz, M. Biscondi.** An Empirical Many-body Potential for B2 Fe Al. *Journal De Physique IV.* 1996, Vol. 6, C2, pp. 47-52.
14. **G. J. Ackland, D. J. Bacon, A. F. Calder, T. Harry.** Computer Simulation of Point Defect Properties in Dilute Fe-Cu Alloy using a Many-body Interatomic Potential. *Philosophical Magazine A.* 1997, Vol. 75, 3, pp. 713-732.
15. **W. Zang, Q. Xie, X. Ge.** Interatomic Potentials Between Distinct Atoms from First Principles Calculation and Lattice Inversion Method. *Journal of Applied Physics.* 1997, Vol. 82, 2, pp. 578-582.
16. **C. Domain, C. S. Becquart.** Ab Initio Calculations of Defects in Fe and Dilute Fe-Cu Alloys. *Physical Review B.* 2001, Vol. 65, 024103, pp. 1-14.
17. **T. T. Lau, C. J. Forst, X. Lin, J. D. Gale, S. Yip, K. J. V. Vliet.** Many-Body Potential for Point Defects Clusters in Fe-C Alloys. *Physical Review Letters.* 2007, Vol. 21, 5501, pp. 1-4.
18. **R. C. Pasianot, L. Malerba.** Interatomic Potentials Consistent with Thermodynamics: The Fe-Cu System. *Journal of Nuclear Materials.* 2007, Vol. 360, pp. 118-127.
19. **G. J. Ackland, M. I. Mendelev, D. J. Srolovitz, S. Han, A. V. Barashev.** Development of Interatomic Potential for Phosphorous Impurities in alpha-Iron. *Journal of Physics: Condensed Matter.* 2004, Vol. 16, 27.
20. **G. Bonny, R. C. Pasianot, N. Castin, L. Malerba.** Ternary Fe-Cu-Ni Many-Body Potential to Model Reactor Pressure Vessel Steels: First Validation by Simulated Annealing. *Philosophical Magazine.* 2009, Vol. 89, 34-36, pp. 3531-3546.
21. **G. Bonny, R. C. Pasianot, L. Malerba.** Fe-Ni Many-Body Potential for Metallurgical Applications. *Modelling and Simulation in Materials Science and Engineering.* 2009, Vol. 17, 025010, pp. 1-13.
22. **A. Ramasubramaniam, M. Itakura, E. A. Carter.** Interatomic Potentials for Hydrogen in Alpha-Iron based on Density Functional Theory. *Physical Review B.* 2009, Vol. 79, 174101, pp. 1-13.
23. **M. I. Mendelev, S. Han, D. J. Srolovitz, G. J. Ackland, D. Y. Sun, M. Asta.** Development of Interatomic Potentials appropriate for Crystalline and Liquid Iron. *Philosophical Magazine.* 2003, Vol. 83, 35, pp. 3977-3994.
24. **J. Chaussidon, M. Fivel, D. Rodney.** The Glide of Screw Dislocations in BCC Fe: Atomistic Static and Dynamic Simulations. *Acta Materialia.* 2006, Vol. 54, 13, pp. 3407-3416.
25. **T. Imura, K. Noda, H. Matsui, H. Saka, H. Kimura.** *Dislocations in Solids.* Tokyo : University of Tokyo Press, 1985. p. 287.
26. **M. S. Duesbery.** On Kinked Screw Dislocations in the BCC Lattice - I. The Structure and Peierls Stress of Isolated Kinks. *Acta Metallurgica.* 1983, Vol. 31, 10, p. 17471758.
27. **V. Vitek, M. Yamaguchi.** Core Structure of Non-screw $1/2[111]$ Dislocations on (110) Planes in B.C.C. Crystals, II. Peierls Stress and the Effect on an External Shear Stress on the Cores. *Journal of Physical F. Metal Phys.* 1973, Vol. 3, 3, pp. 537-542.
28. **M. Tang, L.P. Kubin, G.R. Canova.** Dislocation Mobility and the Mechanical Response of BCC Single Crystals: A Mesoscopic Approach. *Acta Materialia.* 1998, Vol. 46, pp. 3221-3235.
29. **W. Carrington, K. F. Hale, D. McLean.** Arrangement of Dislocations in Iron. *Proceedings of the Royal Society of London, Series A, Mathematical and Physical Sciences.* 1960, Vol. 259, 1297, pp. 203-227.
30. **A. H. Cottrell, B. A. Bilby.** A Mechanism for the Growth of Deformation Twins in Crystals. *Philosophical Magazine.* 1951, Vol. 42, 329, pp. 573-581.
31. **F. C. Frank, J. F. Nicholas.** Stable Dislocations in the Common Crystal Lattice. *Philosophical Magazine.* 1953, Vol. 44, 358, pp. 1213-1235.
32. **J. B. Cohen, R. Hinton, K. Lay, S. Sass.** Partial Dislocations on the {110} Planes in the BCC Lattice. *Acta Metallurgica.* 1962, Vol. 10, 9, pp. 894-895.

33. **D. A. Terentyev, Y. N. Osetsky, D. J. Bacon.** Effects of Temperature on Structure and Mobility of the $\langle 100 \rangle$ Edge Dislocation in Body Centered Cubic Iron. *Acta Materialia*. 2010, Vol. 58, pp. 2477-2482.
34. **V. Vitek.** Core Structure of Screw Dislocations in BCC Metals: Relation to Symmetry and Inter Atomic Bonding. *Philosophical Magazine*. 2004, Vol. 84, 3-5, pp. 415-428.
35. **L. Ventelon.** *Core Structure of Screw Dislocations in Fe From First-Principles*. Department of Materials for Nuclear Energy at the Nuclear Energy Division, Service de Recherches de Métallurgie Physique. Yvette, France : Commissariat à l'énergie Atomique et aux Energies Alternatives, 2008. PhD Thesis.
36. **J. P. Hirth, J. Lothe.** *Theory of Dislocations*. 2nd. s.l. : John Willey & Sons, 1982.
37. **J. Chang, W. Cai, V.V. Butalov, S. Yip.** Dislocation Motion in BCC Metals by Molecular Dynamics. *Materials Science and Engineering A*. 2001, Vols. 309-310, pp. 160-163.
38. **K. Edagawa, T. Suzuki, S. Takeuchi.** Motion of a Screw Dislocation in a 2 Dimensional Peierl's Potential. *Physical Review B*. 1997, Vol. 55, 10, pp. 6180-6189.
39. **W. Xu, J. Moriarty.** Accurate Atomistic Simulations of the Peierl's Barrier and Kink-Pair Formation Energy for $\langle 111 \rangle$ Screw Dislocations in BCC Mo. *Computational Material Science*. 1998, Vol. 9, 3-4, pp. 348-356.
40. **H. J. Kaufmann, A. Luft, D. Schulze.** Deformation Mechanism and Dislocation Structure of High Purity Molybdenum Single Crystals at Low Temperatures. *Crystal Research Technology*. 1984, Vol. 19, 3, pp. 357-372.
41. **S. Takeuchi, K. Maeda.** Slip in High Purity Tantalum between 0.7 and 40 K. *Acta Metallurgica*. 1977, Vol. 25, 12, pp. 1485-1490.

Performance Enhancement in Mobile Computing Using Replicated Cache Agent

Meenakshi Mehla¹ and *Reena Dahiya²

^{1,2} C.S.E, Kurukshetra University,
JMIT (Radaur), Haryana, India
Kurukshetra, Haryana, Pin-136118, India

Abstract

With the rapid advances of wireless communications and portable computing devices, a new computing paradigm which is called "mobile computing" has evolved. In mobile computing environment, users carrying portable devices have access to data and information services regardless of their physical location or movement behaviour. Wireless LANs (WLANs); allow greater flexibility and portability than traditional wired local area networks (LAN). This environment is called MANET. In MANETs, however, there are many unsolved problems so far:

- (1) The mobile resource constrains such as limited wireless bandwidth and limited battery life.
- (2) The mobility of nomadic users and their computers.
- (3) The intermittent disconnection of mobile users and their computers, etc.

Therefore, the methods used in conventional distributed systems and wired systems are not appropriate in mobile computing systems. This research paper makes a comparison of replicated cache agent to that of normal cache agent. Simulation is used to compare the performance of mobile system using metrics like throughput, end-to-end delay & packet delivery fraction. NS2 (Network Simulator version2) is used as simulator here. With the help of an Agent-based cache strong consistent validate strategy, result shows that replicated cache agent improves not only battery life & connection but also the writing delay.

Keywords: WLANs, LANs, MANET, Replicated Agent, NS2.

1. Introduction

A technology that allows transmission of data, via a computer, without having to be connected to a fixed physical link is called mobile computing. Mobile voice communication is widely established throughout the world and has had a very rapid increase in the number of subscribers to the various cellular networks over the last few years. An extension of this technology is the ability to send and receive data across these cellular networks. This is the principle of mobile computing. Mobile data communication has become a very important and rapidly evolving technology as it allows users to transmit data from remote locations to other remote or fixed locations. This proves to be the solution to the biggest problem of business people on the move - mobility. In mobile computing environment, users carrying portable devices have access to data and information services regardless of their physical location or movement behaviour.

1.1 Requirements: The major requirements for a mobile computing environment are:

1. Facilities that permit mobile host to be connected from different access points and to stay connected while on movement.
2. A good bandwidth for fast reliable communication.
3. Since the battery life sets a limitation, we need some kind of facilities that will permit disconnection and save power without affecting the performance and reliability of the whole system.

1.2 Characteristics: Various important characteristics in a mobile computing environment are:

1. Non-symmetrical nature of communications between mobile hosts and fixed servers.
2. Long disconnection from mobile host to save power and to permit the user to work at his will.
3. The mobility of host also implies the availability of a virtual server that is available to provide an efficient service to the mobile host.

The question that always arises when a business is thinking of buying a mobile computer is "Will it be worth it?" In many fields of work, the ability to keep on the move is vital in order to utilize time efficiently. Efficient utilization of resources can mean substantial savings in transportation costs and other non-quantifiable costs such as increased customer attention, impact of on-site maintenance and improved intercommunication within the business.

1.3 Applications: Various applications of Mobile Computers are:

- **For Estate Agents:**

Estate agents can work either at home or out in the field. They can provide clients with immediate feedback regarding specific homes or neighbourhoods, and with faster loan approvals, since applications can be submitted on the spot. Therefore, mobile computers allow them to devote more time to clients.

- **Emergency Services:**

Information regarding the address, type and other details of an incident can be dispatched quickly, via a CDPD system using mobile computers, to one or several appropriate mobile units which are in the vicinity of the incident.

- **In Courts:**

When the opposing counsel references a case which they are not familiar, they can use the computer to get direct, real-time access to on-line legal database services, where they can gather information on the case and related precedents.

- **In Companies:**

Managers can use mobile computers in, say, and critical presentations to major customers. They can access the latest market share information. Therefore, mobile computers can leverage competitive advantages.

- **Stock Information Collation/Control:**

The use of small portable electronic databases accessed via a mobile computer would be ideal. Data collated could be directly written to a central database, via a CDPD network, which holds all stock information. This ensures that from the time that a stock count is completed, there is no inconsistency between the data input on the portable computers and the central database.

- **Credit Card Verification:**

At Point of Sale (POS) terminals in shops and supermarkets, when customers use credit cards for transactions, the intercommunication required between the bank central computer and the POS terminal, in order to effect verification of the card usage, can take place quickly.

- **Electronic Mail/Paging:**

Usage of a mobile unit to send and read emails is a very useful asset for any business individual, as it allows him/her to keep in touch with any colleagues as well as any urgent developments that may affect their work.

1.2 Agents:

Agent is a term used to describe autonomous, "intelligent" programs that can be deployed through a WAN, to provide services to a mobile user. An agent is an encapsulation of code, data and execution context that is able to migrate autonomously and purposefully within a computer network during the execution. An agent is able to react on external events. It may be persistent in the sense that it can suspend execution and keep local data in stable storage. After resuming activity an agent's execution is continued. Agent is an object it contains:

a. Cached data

b. Methods for access

c. State information

d. Consistency rules

e. Tolerance parameters for time and attributes

1.2.1 Properties: An agent has the following properties:

1. **Delegation:** A user or a program can delegate tasks to an agent and vest it with an authority (power of attorney) to act on their behalf by providing terms of reference and time dead-lines.
2. **Autonomy:** Agent can make its own decisions, based on rules, goals, policies and preferences set by its owner.
3. **Communication:** Agents can interact with host environment and owners.
4. **Co-operation:** Agents can cooperate with each other to achieve common goals.
5. **Equity:** Agents can support peer-peer model for distributed computing.
6. **Flexibility:** Agents can vary their roles-may act as clients or servers, observers, seekers.
7. **Intelligence:** Ability to reason and learn from interactions with other agents, owners and environment.
8. **Mobility:** Agents can move across networks and accomplish tasks assigned.

1.2.2 Types of agents:

a. Buyer agents(Shopping Bots):

Buyer agents travel around a network (i.e. the internet) retrieving information about goods and services. These agents, also known as 'shopping bots', work very efficiently for commodity products such as CDs, books, electronic components, and other one-size-fits-all products. **For example**, *Amazon.com* is a good example of a shopping bot. The website will offer you a list of books that you might like to buy on the basis of what you're buying now and what you have bought in the past.

b. User agents(Personal agents):

User agents are intelligent agents that take action on your behalf. In this category belong those intelligent agents that already perform, or will shortly perform, the following tasks:

1. Check your e-mail, sort it according to the user's order of preference, and alert you when important emails arrive.
2. Play computer games as your opponent or patrol game areas for you.
3. Assemble customized news reports for you. There are several versions of these, including News hub and CNN.
4. Find information for you on the subject of your choice.
5. Fill out forms on the Web automatically for you, storing your information for future reference.
6. Scan Web pages looking for and highlighting text that constitutes the "important" part of the information there.
7. Discuss topics with you ranging from your deepest fears to sports.
8. Facilitate with online job search duties by scanning known job boards and sending the resume to opportunities who meet the desired criteria.
9. Profile synchronization across heterogeneous social networks.

c. Monitoring-and-surveillance agents(Predictive Agents):

They are used to observe and report on equipment, usually computer systems. The agents may keep track of company inventory levels, observe competitors' prices and relay them back to the company, watch stock manipulation by insider trading and rumors etc. **For example**, NASA's Jet Propulsion Laboratory has an agent that monitors inventory, planning, and scheduling equipment ordering to keep costs down, as well as food storage facilities. These agents usually monitor complex computer networks that can keep track of the configuration of each computer connected to the network.

d. Data mining agents:

Data mining agent uses information technology to find trends and patterns in an abundance of information from many different sources. The user can sort through this information in order to find whatever information they are seeking .A data mining agent operates in data warehouse discovering information. A 'data warehouse' brings together information from lots of different sources. "Data mining" is the process of looking through the data warehouse to find information that you can use to take action, such as ways to increase sales or keep customers who are considering defecting.

1. 3 Working of Mobile Agents:

A mobile agent consists of the program code and the program execution state (the current values of variables, next instruction to be executed, etc.). Initially a mobile agent resides on a computer called the home machine. The agent is then dispatched to execute on a remote computer called a mobile agent host (a mobile agent host is also called mobile agent platform or mobile agent server). When a mobile agent is dispatched the entire code of the mobile agent and the execution state of the mobile agent is transferred to the host.

The host provides a suitable execution environment for the mobile agent to execute. The mobile agent uses resources (CPU, memory, etc.) of the host to perform its task. After completing its task on the host, the mobile agent migrates to another computer. Since the state information is also transferred to the host, mobile agents can resume the execution of the code from where they left off in the previous host instead of having to restart execution from the beginning. This continues until the mobile agent returns to its home machine after completing execution on the last machine in its itinerary.

1.3.1 The life cycle of a mobile agent:

1. The mobile agent is created in the Home Machine.
2. The mobile agent is dispatched to the Host Machine A for execution.
3. The agent executes on Host Machine A.
4. After execution the agent is cloned to create two copies. One copy is dispatched to Host Machine B and the other is dispatched to Host Machine C.
5. The cloned copies execute on their respective hosts.

6. After execution, Host Machine B and C send the mobile agent received by them back to the Home Machine.
7. The Home Machine retracts the agents and the data brought by the agents is analysed. The agents are then disposed.

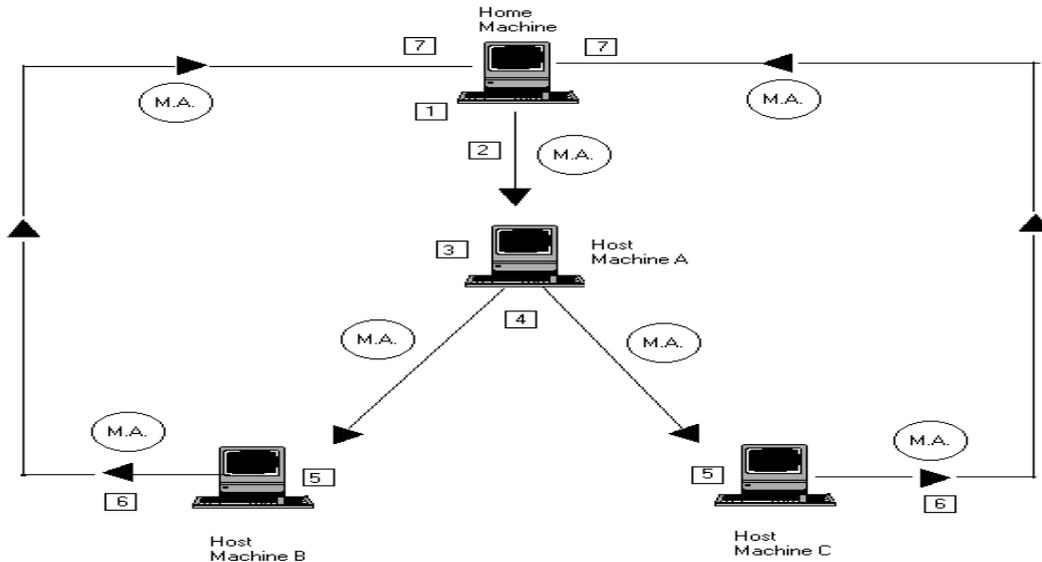


Figure 1 working of Mobile Agent

1.3.2 Events: From this we observe that a mobile agent experiences the following events in its life cycle:

- a. **Creation:** a brand new agent is born and its state is initialized.
- b. **Dispatch:** an agent travels to a new host.
- c. **Cloning:** a twin agent is born and the current state of the original is duplicated in the clone.
- d. **Deactivation:** an agent is put to sleep and its state is stored on a disk of the host.
- e. **Activation:** a deactivated agent is brought back to life and its state is restored from disk.
- f. **Retraction:** an agent is brought back from a remote host along with its state to the home machine.
- g. **Disposal:** an agent is terminated and its state is lost forever.

1.4 Advantages of Mobile Agents:

1.4.1 Reduction in Network Load:

The interactions in a distributed system are often achieved using communication protocols. An alternative to using communication protocols is the use of mobile agents. Mobile agents are dispatched to the remote hosts containing the data. The agents perform the computations at the remote hosts and return back with the results. Since computations are moved to the data storage location instead of moving data to the computing location, network load is reduced.

1.4.2 Overcome Network Latency:

Controlling many systems through a network involves significant delays, which are not acceptable for critical real time systems. To overcome this problem, mobile agents can be directly dispatched from the central controller in the manufacturing plant to the real time systems.

1.4.3 Protocol Encapsulation:

However, protocols evolve over a period of time and new features such as better security may be introduced in the protocol. It is a cumbersome task to upgrade the protocol code at all locations in the distributed system. The mobile agent code can encapsulate the protocol. When a protocol is upgraded, only the mobile agent has to be altered.

1.4.4 Asynchronous and Autonomous Execution:

Mobile agents operate asynchronously. Once a mobile agent is dispatched from the home machine, the home machine can disconnect from the network. The mobile agent executes autonomously without the intervention of the home machine.

1.4.5 Fault Tolerance:

Mobile agents react dynamically and autonomously to the changes in their environment, which makes them robust, and fault tolerant.

1.5 Related Work

Various proposals have been defined for avoiding the writing delay due to call back algorithm. Ichiro Satoh [2] described that when a task is assigned to mobile agents, those suitable to perform the task need to be selected according to, not only their application-specific behaviours, but also their mobilities. The focus is on the development of execution platforms and applications for mobile agents and not on methodologies for selection of mobile agents. Seungsang Sun [6] various migration schemes for mobile agents. The agents play a key role to manage a huge amount of tasks instead of human being in ubiquitous computing environments. Jinghua Wang [9] presented an improved cache invalidation schema based on mobile Agents. An asynchronous invalidation broadcast technology and the control messages which are broadcast by MSS periodically is introduced. MSS maintains for each MC a data structure that stores the latest time stamp for each data item cached by MC. The improved schema can maintain client cache consistency better and improve the speed of execution of transaction. Chu Yan [10] observed that in mobile computing environments, the limited bandwidth of wireless communications, the weak capability of mobile computing equipment and the valuable power make it difficult for mobile devices to be maintained with the network all the time. Cache technique caches part of data in the clients to reduce the query of database server, so it solves the problem to some extent.

2. Methodology

2.1 Agent-based cache strong consistent validate strategy

The key issue in cache algorithm is how to maintain consistency of cached data. We studied the caching consistent checking problems which aroused by data query operations from clients and write operations on server. Utilizing *predicate-based* method to represent cache data structure, which described a data record as a string (called predicate) including '0', '1' and '*'. Predicate-based scheme could reduce networks communications cost and execute subtraction sign '-' operation and intersection sign '^' operation to compute 'remainder predicate' in practically reasonable complexity. Here, remainder predicate is the cache data record which should be updated. Figure 2 shows the architecture of proposed agent-based cache validation scheme.

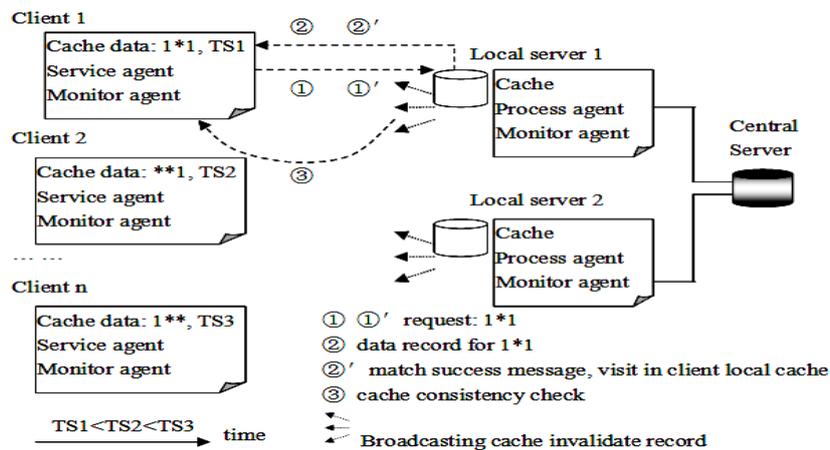


Figure 2: The architecture of our agent-based cache Validation scheme

Figure 3 shows the proposed architecture of 'Client/Agent/Server' for mobile computing. Local server encapsulates MSS's functions.

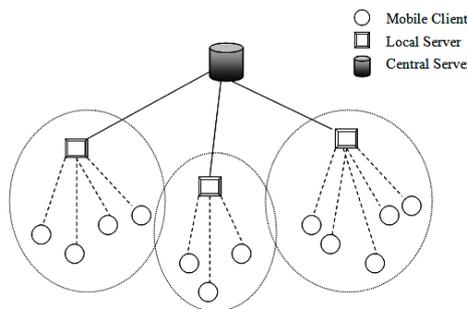


Figure 3: our proposed client/agent/server architecture

To remove writing delay and to improve the performance we propose methodology/research has been Come out which removed the drawbacks of existing call back method, introduces an agent technology, and extends conventional Client/Agent into Client/Agent/Server.

Main objective of the agent cache system is:

- 1) We will implement the previous model of the agent cache in the NS2
Then we will Implement the new agent cache system and compare the following parameters

- 1) Throughput
- 2) End to end delay
- 3) Round Trip time

Purposed scheme that extend the Existing system are as under:

- 1) We will introduce replication agent system into previous work , to make it much sophisticated
- 2) With addition of replication agent , we need to prepare a system that will handle the data replication cost and data accessing cost in the Mobile computing
- 3) In addition to this we will discard the Malicious attack in this system which will lead to provide the quality of service in the mobile computing

2.2 Performance Metrics:

The estimation of performance of mobile computing system by replicating the agents can be estimated by following metrics:

a. Packet Delivery Ratio

Packet delivery ratio is defined as the ratio of data packets received by the destinations to those generated by the sources. This performance metric gives us an idea of how well the agent is performing in terms of packet delivery at different speeds using different traffic models.

b. End to End Delay: end-to-end delay refers to the time taken for a packet to be transmitted across a network from source to destination.

$$d_{\text{end-end}} = N [d_{\text{trans}} + d_{\text{prop}} + d_{\text{proc}}]$$

Where

$d_{\text{end-end}}$ = end-to-end delay

d_{trans} = transmission delay

d_{prop} = propagation delay

d_{proc} = processing delay

N = number of links (Number of routers + 1)

Note: we have neglected queuing delays.

c. Throughput: The system throughput or aggregate throughput is the sum of the data rates that are delivered to all terminals in a network. The throughput can be analysed mathematically by means of queuing theory, where the load in packets per time unit is denoted arrival rate λ , and the throughput in packets per time unit is denoted departure rate μ .

3. Result and Discussion

3.1 Throughput: Network throughput is the average rate of successful message delivery over a communication channel. This data may be delivered over a physical or logical link, or pass through a certain network node. The throughput is usually measured in bits per second (bit/s or bps), and sometimes in data packets per second or data packets per time slot. The system throughput or aggregate throughput is the sum of the data rates that are delivered to all terminals in a network.

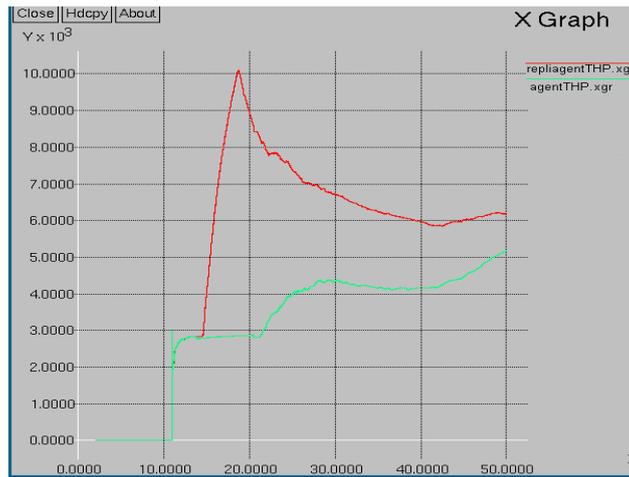


Figure 4: Throughput graph between Pause time and KPBS

As it is clear from the graph that the throughput data packets are higher in Replicated agent as compared to simple agent cache system.

3.2 **Average end-to-end delay of data packets** — This includes all possible delays caused by buffering during route discovery latency, queuing at the interface queue, retransmission delays at the MAC, and propagation and transfer times. The average end-to-end delay of packet delivery is higher in Agent Cache as compared to replicated agent system.

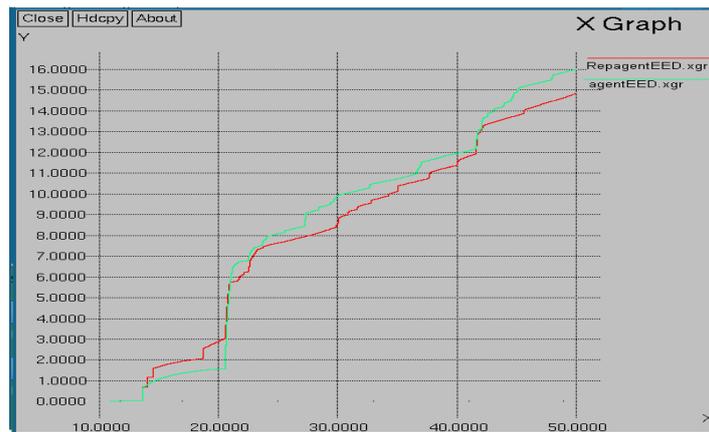


Figure 5: end to end Delay graph between Pause time and Delay time in milli sec

3.3 **Packet delivery fraction** — the ratio of the data packets delivered to the destinations to those generated by the CBR sources. The delivery fraction is higher in Replicated agent as compared to agent cache.



Figure 6: Packet Delivery fraction graph between pause time and packet fraction

4 Conclusion:

In mobile computing existing call-back arithmetic has the drawback of 'writing delay' due to which server must wait for invalidation confirmed messages from every client. And client's caching data need be revalidated again when any one appears. In this paper, based on wireless access network, I propose a technology of agent-cache which is fit for a mobile computing environment with the architecture of Client/Agent/Server'. I use the artifice of agent-cache to reduce traffic and take some proper control to avoid dirty data been read by one. It also includes that user data can be updated in time by virtue of validating the consistency of data among the client, agent and server. A new Callback Validating Policy is presented by me for mobile environment because the current arithmetic for this is believed with the problem of writing delay, in which the cache data need be revalidated again when any client appears.

5 Acknowledgement

First and foremost, I would like to express my sincere gratitude to my guide Mrs. Meenakshi Mehla, Assistant Professor, Computer Science and Engineering Department for immense help, guidance, stimulating suggestions, and encouragement all the time with this thesis work. This work would have not been possible without her support. She always provided a motivating and enthusiastic atmosphere to work with; it was a great pleasure to do this thesis under her supervision.

Most importantly, I would like to thank my parents and the almighty for showing me the right direction, to help me stay calm in the oddest of times and keep moving at times when there was no hope.

Reena

References

- [1] Parineeth M Reddy, "Mobile Agents", General Article, Year 2002, pp. 35-43.
- [2] Ichiro Satoh, "Building Reusable Mobile Agents for Network Management", IEEE, Year August 2003, vol. 33, no. 3, pp. 350-357.
- [3] Stefan Pleisch, Andre Schiper, "Fault-Tolerant Mobile Agent Execution", IEEE Transactions on Computers, vol. 52, no. 2, Year 2003, pp. 209-222.
- [4] Wenyu Qu, Hong Shen, "Some Analysis on Mobile-Agent Based Network Routing", IEEE International Symposium on Parallel Architectures, Algorithms and Networks, Year 2004, pp. 1-6.
- [5] Gabor Vilmos Szabó, Diploma. Eng. in Comp. Sc., Swiss Federal Institute of Technology, ETH Zurich {DISS. ETH No. 16207, 2005 © IEEE}
- [6] Seungsang Sun, Hyun-Su Jang, Youn-Woo Kim, and Young Ik Eom, "An Efficient Migration Scheme for Mobile Agents in Ubiquitous Environments", IEEE International Symposium on Information Technology Convergence, Year 2007, pp. 110-114.
- [7] Amanpreet Kaur, Sukhdeep Kaur, "Role of Mobile Agents In Mobile Computing", National Conference on Challenges & Opportunities in Information Technology, March 23, 2007, pp.68-71.
- [8] Khaleel Ur Rahman Khan, Rafi U Zaman, A Venugopal Reddy, "A Three-Tier Architecture for Integrating Mobile Ad Hoc Network and the Internet using a Hierarchical Integrated Routing Protocol", IEEE International Conference on Advanced Computer Theory and Engineering, Year 2008, pp. 518-522 Rubing Liang, Qiong Liu., "A Technology of Agent-Cache Used in Mobile Computing Environment", World Congress on Computer Science and Information Engineering, Year 2009, pp.120-124
- [9] Jinghua Wang, Zhiyun Zhou, Minhra Liu, "An improved cache invalidation schema based on mobile Agents", IEEE, Year 2010, vol. 1, pp. 7-9.
- [10] Chu Yan, Zhang Jianpei and Zhao Chunhui, "Data Cache Strategy Based on Colony Algorithm in Mobile Computing Environment", IEEE Fourth International Conference on Internet Computing for Science and Engineering, Year 2010, pp. 235-238.

Oscillation Test Methodology for Built-In Analog Circuits

Ms. Sankari.M.S and Mr.P.SathishKumar

Department of ECE, Amrita School of Engineering, Bangalore, India

Abstract

This article aims to describe the fundamentals of analog and digital testing methods to analyze the difficulties of analog testing and to develop an approach to test the analog components in a mixed signal circuit environment. Oscillation based, built-in self-test methodology for testing analog components in mixed-signal circuits, in particular, is discussed. A major advantage of the OBIST method is that it does not require any complex response analyzers and test vector generators which are costly. Furthermore, since the oscillation frequency is considered to be digital it can be easily interfaced to test techniques dedicated to the digital part of the circuit under test (CUT). OBIST techniques show promise in detecting faults in mixed signal circuits and requires little modification of the CUT to improve the fault coverage. Extensive simulation results on some sample analog benchmark circuits are described in Spice format.

Keywords- System on chip (SOC), built-in Self test (BIST), Oscillation based built-in self test(OBIST), Circuit Under Test(CUT), Design for Testability (DFT).

1. Introduction

Testing is a critical technology in the semiconductor production process. IC test is used for debugging, diagnosing and repairing the sub-assemblies in their new environment. The test should be designed to indicate the desired perfection. The objective is to realize through detailed testing, that the manufactured products are free from defects. It may ultimately help in increasing the product yield and reducing the product cost. The VLSI realization process has a distributed form of testing. Requirements and specifications are audited, design and tests are verified, and fabricated parts are tested. The broad specifications of analog circuits require detailed and long performance tests as well. This results in lengthy time consuming and very expensive test procedures. These factors have resulted in ample research being channeled in the direction of mixed signal testing. Researchers are now seeking to combine both analog and digital-circuits testing either by applying digital signals to drive analog circuits or by using analog signals to drive digital circuits. The applications of analog and mixed-signal, embedded-core-based, system-on-chip in recent years have motivated system designers and test engineers to direct their research to develop methodologies in effective very large-scale integrated circuits and systems testing. Mixed signal hardware systems have digital cores, very often interconnected with analog filters, analog and digital converters for digital processing. Testing is done to detect defects and diagnosis to determine the root cause of the defects. This would help in finding points in the process to be altered. If the product fails even after a detailed, rigorous and exhaustive test procedure, then pitfalls in the specification, design, or the fabrication process is suspected. The IC fabrication process involves photolithography, printing, etching, and doping steps. It is difficult to achieve 100% perfection in these steps while fabricating. Minor imperfections do creep in, leading to failures in the operation of the individual ICs. In the case of mixed signal ICs, the performance will not be satisfactory. The main sources of test difficulties in digital and analog circuits are also different. The size and complexity in digital circuits remain a measure of test difficulty whereas in analog and mixed-signal circuits, the behavior of circuit signals are much more important than circuit sizes. A major problem in the analog and mixed-signal-circuit testing is in defining the line of demarcation between a fault-free and faulty circuit, which leads to uncertainty in quantification of the product yield. In mixed signal circuits, imperfection in the form of small capacitance like parasitic, between the traces, produce a significant parameter variation altering the circuit behavior., Many parts of SOC are not accessible due to SOC complexity and limited test pins .To improve monitoring and control, test buses and scan chains are used by DFT methodology. Here, analog signals get degraded in long wire transmission resulting in erroneous test results and reduced production yield. One vital solution to this problem is by introducing Built-in self-test mechanism that permits a machine to test itself. The test generation, test application and response verification, all are accomplished through a Built-in hardware, which allows different parts of the chip to be tested in parallel. The test hardware can be placed near to the digital portion of the SOC which reduces the test time and eliminates the usage of external test t equipment. This concept originated with the idea of including pseudorandom number generator and cyclic redundancy check on the IC. BIST functionality is incorporated with the system level design. Fault models from the digital domain, like stuck-at faults, stuck on, stuck open, bridging faults, digital DFT and test techniques cannot be directly extended to analog circuits. Typically, analog circuits are tested by verifying against specifications. Test inputs can be generated easily for this straightforward method. Analog circuits have detailed, extensive specifications. Checking all the specifications is time consuming and expensive. Analog circuit signal monitoring is reduced in a mixed signal system. Since the testing of mixed signal circuit is complex, some effort is to be taken to lower the cost of test, particularly for the analog portions of mixed signal circuits. Almost every mixed-signal

integrated system contains circuits such as operational amplifiers, filters, oscillators, PLLs, etc. In this paper, an approach to develop a test methodology based on oscillation based built-in self-test is investigated. During the test mode, all these circuits (CUTs) could be transformed to an oscillator so as to make it testable by connecting some additional circuitry (a feedback network). In the process, the defined fault model and test algorithms considered to test the circuits in the mixed signal SOC environment and simulation results are provided for the selective benchmark circuits.

2. Test Methods

DESIGN FOR TESTABILITY (DFT)

DFT refers to hardware design styles or added hardware that reduces test generation complexity and test application cost. The basic concept of design for testability is given in Fig. 1. The test generation complexity increases exponentially with size of the chip.

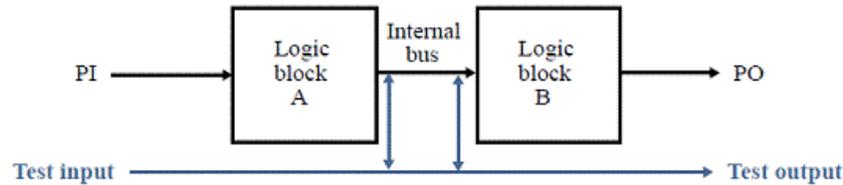


Fig .1. Design for Testability

GENERAL BIST ENVIRONMENT

It is a design process that provides the capability of solving many of the problems encountered in testing analog, mixed-signal or digital systems. Test generation, test application and response verification is through Built-in hardware. It allows different parts of the chip tested in parallel thereby reducing the required testing time. It eliminates the necessity for external test equipment. BIST circuitry is located in the digital portion of the mixed-signal circuitry to minimize area overhead. The basic principle of BIST is explained in the Fig.2.

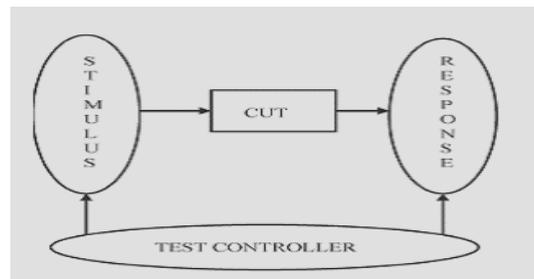


Fig .2 BIST Environments

A **built-in self-test (BIST)** or **built-in test** is a mechanism that permits a machine to test itself. Engineers design BISTs which ensure high reliability and reduced repair cycles. In integrated circuits, BIST is used to make faster, less-expensive manufacturing tests. The IC has a function that verifies all or a portion of the internal functionality of the IC. A BIST mechanism is provided in advanced field bus systems to verify its functionality. It reduces test-cycle duration.

3. Building of an Oscillator

The way to design a sinusoidal oscillator from the transfer function is to connect the output terminal of the filter to the input terminal. The basic requirements for oscillation are a signal feedback from the output to the input of proper phase and sufficient amplitude. The design equations of an oscillator are determined by analyzing the denominator of the transfer equation of the circuit. The poles of the denominator of the characteristic equation, or the zeros of $T(s)$, determine the time-domain behavior and stability of the system. The magnitude and phase equations of an oscillator must also be analyzed. If the magnitude of the loop-gain is greater than one and the phase is zero, the amplitude of oscillation will increase exponentially.

The process of building general oscillators is different than that of building oscillators for testing purposes. In designing general oscillators, well-defined, stable oscillation frequency and amplitude are required. But an oscillator that is built from conversion of CUT is designed such that the variation of the components in CUT can be detected by measuring the oscillation frequency and amplitude.

$$H(s) = V_o(s) / V_i(s)$$

$$= (a_2s^2 + a_1s + a_0) / (s^2 + b_1s + b_0) \dots \dots \dots (1)$$

where, $b_1 = \omega_0/Q$, and $b_0 = \omega_0^2$, ω_0 is the pole frequency, and Q is the pole quality factor that determines the distance of the poles on $j\omega$ -axis in the s -plane. An infinite Q locates the poles on the $j\omega$ -axis, and this can cause the circuit to oscillate. Therefore, in order for a filter to oscillate, the quality factor must be increased. First the Circuit component to be tested is converted into an oscillatory circuit by adding external circuitry in the feedback path so as to place a pair of poles of the system on the imaginary axis causing the system to become marginally stable causing oscillations. The concept of building of an Oscillator is explained in Fig. 3.

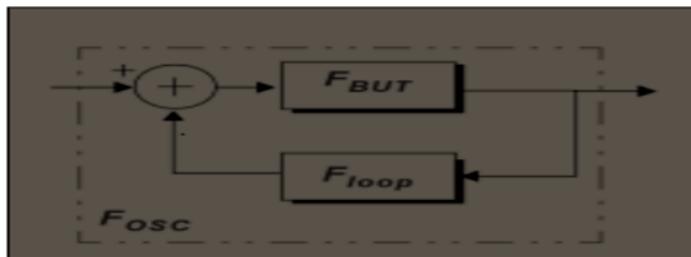


Fig 3.Concept of an building an Oscillator

4. General OBIST procedure

The test signals are sent into the system and the output from the normal mode and the faulty mode of the system are compared to create fault coverage. The General OBIST procedure is explained in Fig .4. Various faults are injected into the system such as removing a part of the circuit; Stuck faults etc. to cover the overall fault range.

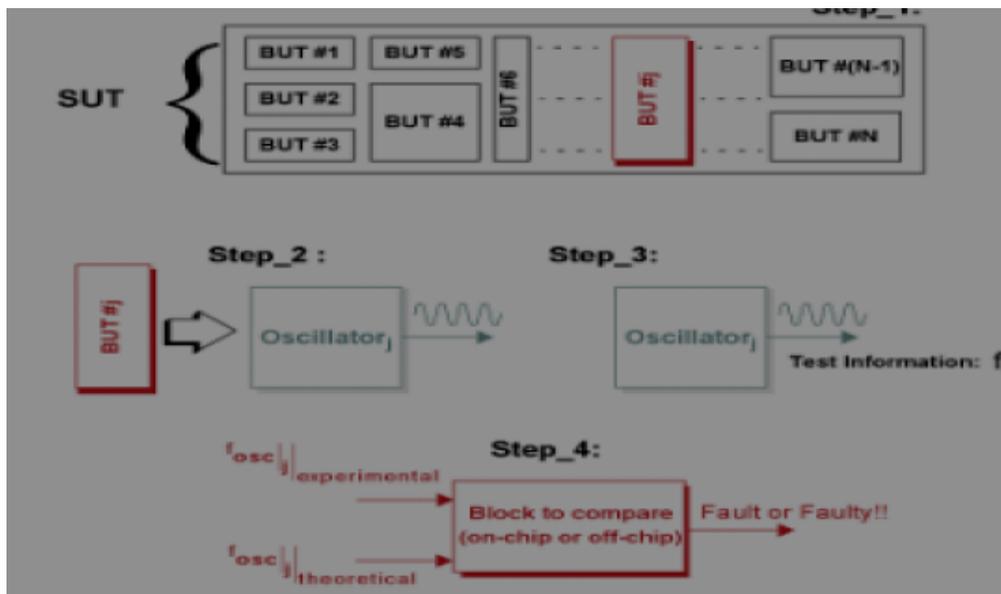


Fig .4. General OBIST Procedure

5. Concept of OBIST Strategy

A complex analog circuit is portioned into functional building blocks such as op-amps, filters, comparators, PLL etc. or a combination of these blocks. Each building block is converted into an oscillator by adding the proper circuitry in order to achieve sustained oscillation. The oscillation parameters are then evaluated. A faulty circuit is detected from a deviation of its oscillation parameters under fault free conditions. The oscillation parameters are independent of the CUT type and analog testing. The block diagram of OBIST strategy is explained in Fig .5.

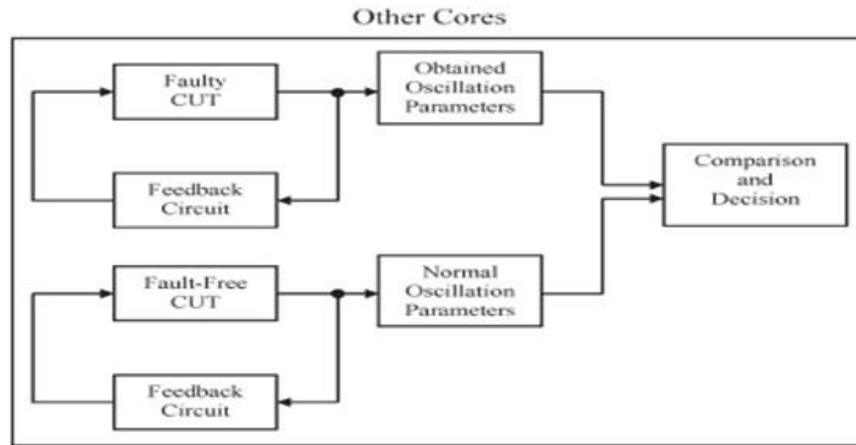


Fig .5. Block Diagram of OBIST Strategy

6. Analog Fault Modeling

Fault models for analog and mixed signal circuits can be classified into two categories. They are hard faults and parametric faults. A catastrophic fault is analogous to the stuck-at fault model in the digital domain where the terminals of the component can be stuck open or stuck-short. Parametric faults are deviations in component parameters that cause performance overshoot beyond acceptable limits. It is caused by statistical fluctuations in the manufacturing process. Catastrophic faults are introduced by random defects and results in failures in components. For example, dust particles on a photolithographic mask can cause either short or open in circuits or it may create large deviations of CUT parameters such as aspect ratio, threshold voltage change in a MOS transistor. These faults can be modeled as below in Fig.6. Stuck open faults are hard faults in which the component terminals are out of contact with the rest of the circuit. These faults can be simulated by adding high resistance in series. A Stuck- short fault is a short between the terminals of the component

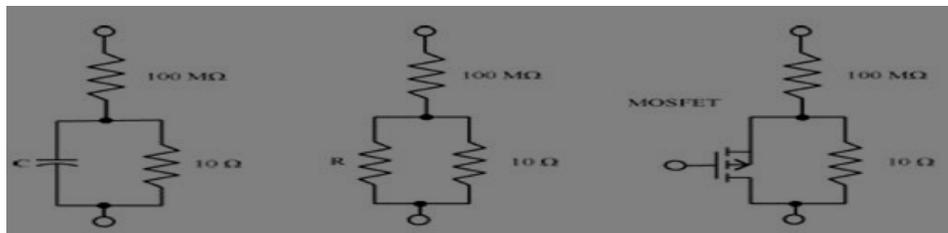


Fig .6.Stuck open and stuck short Fault Models for capacitor, resistor and MOSFET

7. Test Procedure

Parametric and catastrophic faults are injected into the circuit under test. These faults being injected into nominal circuit description are described in Spice format. By using Spice simulator transient and frequency domain response is evaluated. Fig.7.explains the test procedure for Oscillation based Built-in Self-test methodology. The different steps of the procedure are briefly given below.1. The fault free circuit is converted into an oscillator and simulated, and its test parameters are derived.

1. A fault list was derived from the CUT (circuit net list).
2. The faulty net list is generated (through fault injection).
3. A simulation was done for the faulty CUT.
4. The fault detection was completed on comparing the faulty-output measurements with fault-free test parameters.
5. The circuit fault coverage was calculated.

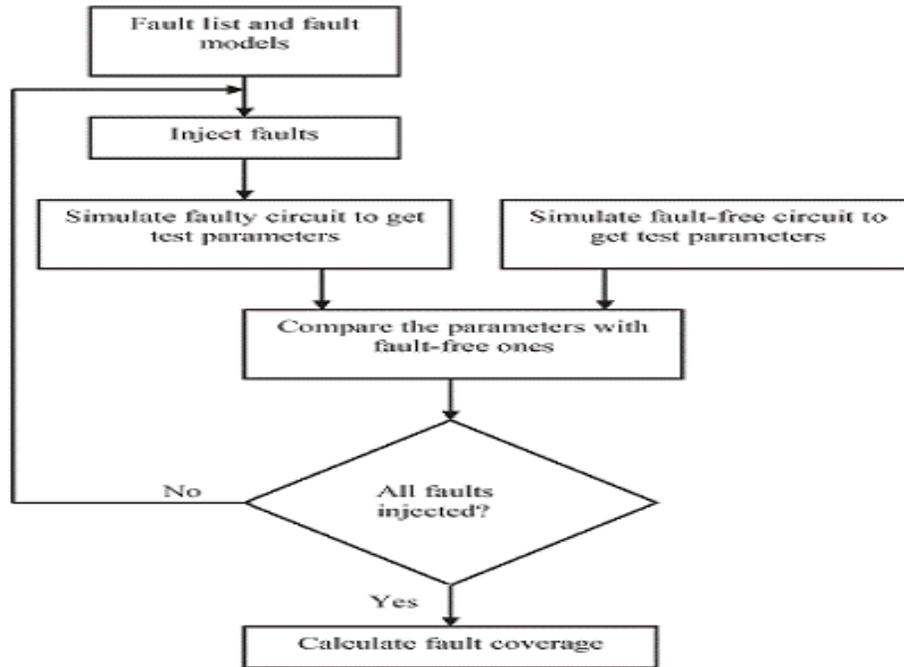
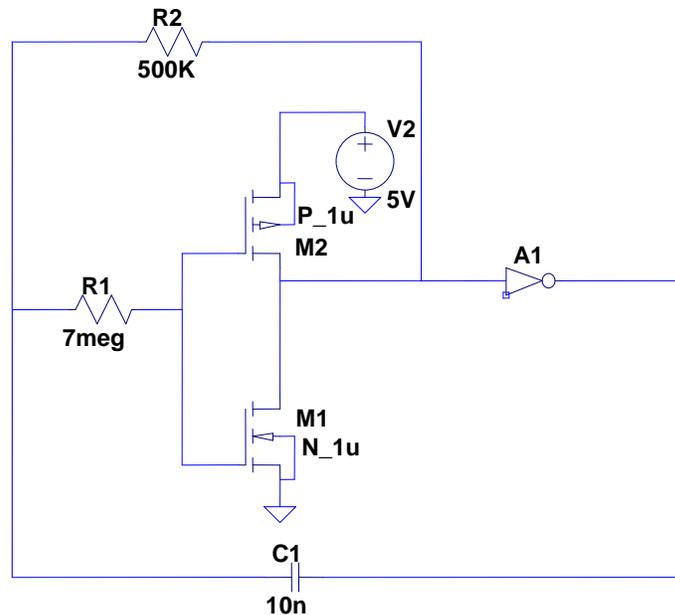


Fig.7. Test Procedure based on OBIST approach

8. Experimental Results

Purely analog ICs usually consist of amplifiers, comparators, PLL; filters etc., The test parameters specified by designers can be gain, signal to noise ratio, amplitude, power gain, phase shift, gain margin and phase margin and so on. Efficient ways to test the op –amps and filters are desired because of their importance in analog systems. In order to test any circuit with oscillation based method, first the circuit under test must be converted to an oscillator by adding extra circuitry as a feedback. If the circuit is faulty, converted circuit either won't oscillate or the response parameters of oscillation will differ from fault free condition. The proposed OBIST methodology for calculating the oscillation frequency is explained below (Fig. 8) by considering CMOS inverter as an example.



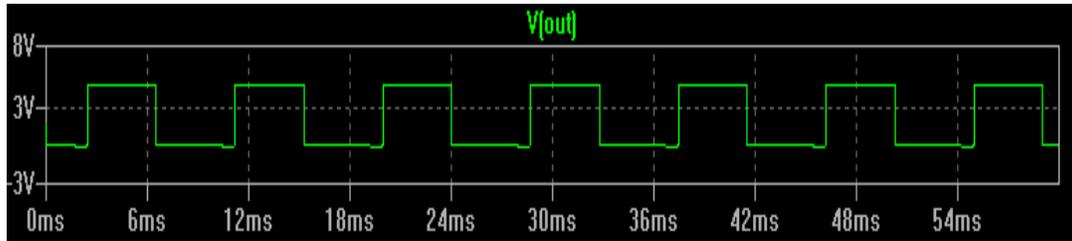


Fig.8.CMOS inverter as an oscillator and the output signal of CUT in test mode

By choosing $R1 = 7M\Omega$, $C1 = 10\text{ nF}$, and $R2 = 500K\Omega$, the oscillation frequency is calculated as $f_{osc} = 90.9\text{ Hz}$. The Fig. 6. depicts the fault model for MOS transistors, where the value for the parallel resistor R_p is 100Ω , which emulates stuck short fault, and the series resistor R_s has a value of $100M\Omega$ that emulates stuck-open fault. The output signal of the CUT in SPICE Simulation is shown in the Fig .8

Next, we consider the catastrophic (hard) fault injection in CUT in test mode. Injection of the below four faults specified in the CUT (inverter) in the test mode causes stopping of the oscillation of the circuit. If any one of the transistor is short or open, the output remains high or low according to the type of the fault. All the simulation results obtained below agree with this hypothesis.

Qn stuck-short

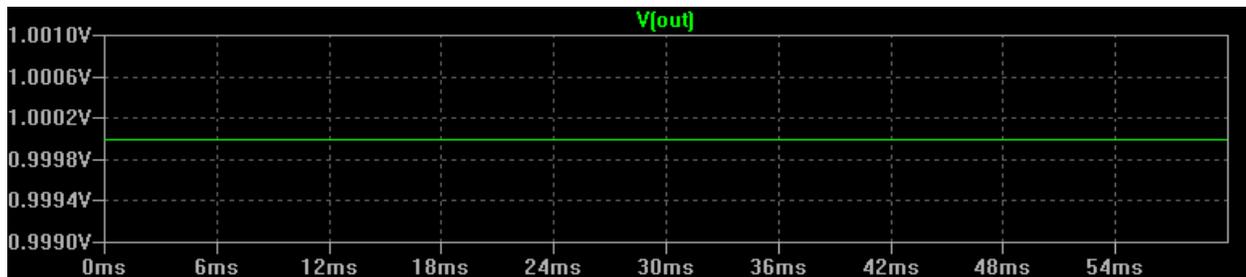


Fig.9. Output of CUT when Qn is stuck-short

We consider the stuck-short fault of NMOS Qn. Since the input of the second inverter shorted to ground node, the output will always be high. The same result has been obtained in SPICE simulation. The Simulation result is shown in Fig.9.

Qp stuck-short

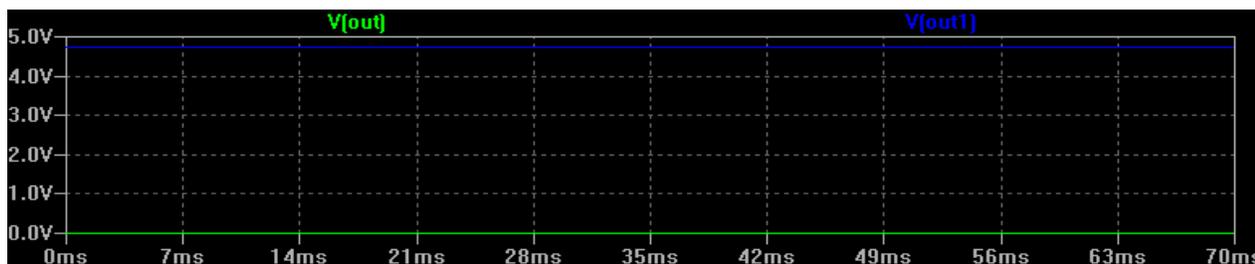


Fig.10. Output of CUT when Qp is stuck-short

A short fault is injected in PMOS transistor Qp. Since the input of second inverter is shorted to VDD, the output will always be zero. The same result has been obtained in SPICE simulation. The simulation result of output of CUT when Qp stuck- short is shown in Fig.10.

Qp stuck-open

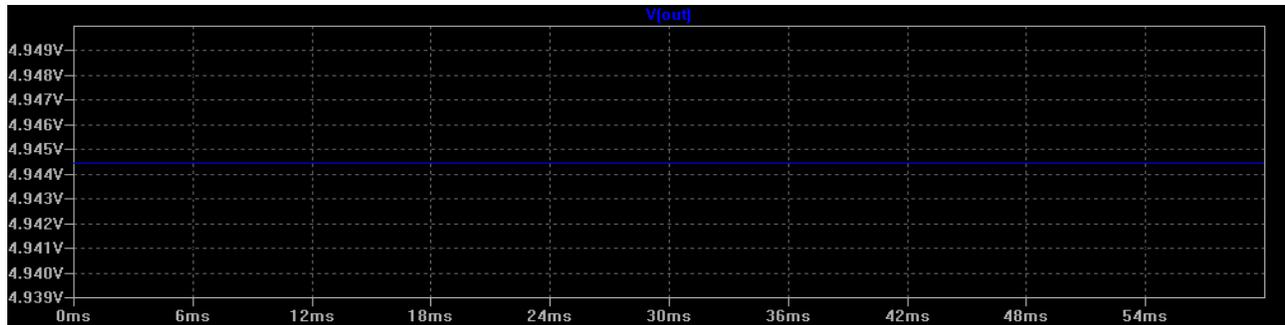


Fig.11. Output of CUT when Qp is stuck-open

A stuck- open fault is injected in Qp transistor, such that, it is disconnected from the remaining part of the circuit. The input of the second inverter is connected to the ground node. The output of second inverter turns high (near to 5V). A similar result is obtained with SPICE simulation. The Simulated result is shown in Fig.11.

Qn stuck-open

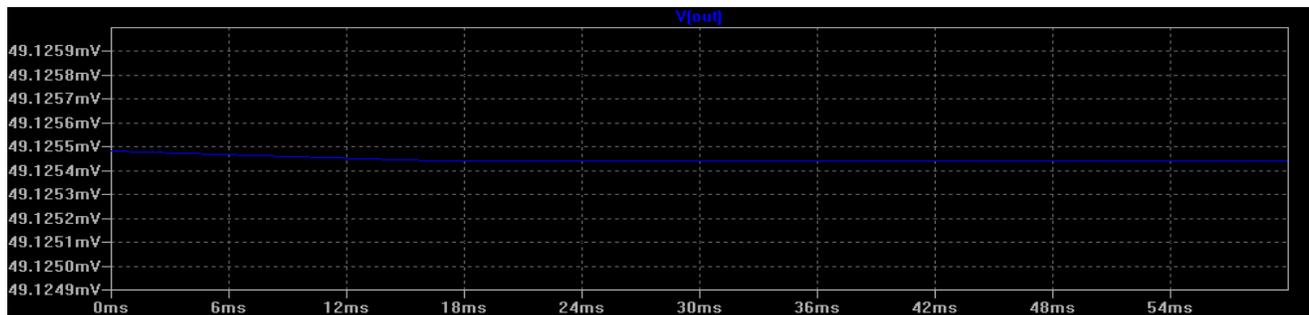


Fig.12. Output of CUT when Qn is stuck-open

A CMOS inverter is converted into astable oscillator using one more inverter and one RC feedback. Qn is disconnected from the remaining part of the circuit. The input of the second inverter goes high from VDD. So the output of that inverter goes low. The same result obtained in SPICE Simulation. The simulated result is shown in Fig.12

FET Oscillator in test mode

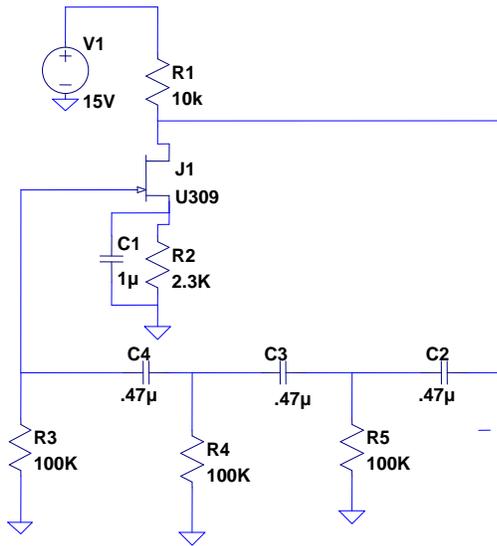


Fig.11. FET Oscillator in test mode

Table 1.The FET Oscillator faults and fault coverage analysis

Faults	Amplitude	BW (Hz)	Power BW (Hz)
R1 short	14.9V	268.50	42.360
R2 short	92.7mV	269.76	42.340
R5 short	7.70 V	268.36	42.603
C2 short	7.07V	270.24	42.699
C4 short	8.00V	126.90	19.860
R4 short	7.77V	268.36	42.600
R3 short	7.775V	267.36	42.603
R3 open	7.767V	268.36	42.600
Fault free	14.80V	107.10	18.804

Table.1 presents the results obtained by spice simulation for FET oscillator when the hard faults are injected. FET Oscillator produces oscillation frequency without fault as $f_{osc} = 1.45$ KHz. By injecting short faults in the circuit, the FET operates in cutoff region. When the FET is open, it is disconnected from the circuit does not oscillate. Fault is detected when the power bandwidth is out of a tolerance band, defined as $\pm 5\%$ of the fault free power bandwidth to get nominal fault coverage.

Fault free power BW= 18.804Hz.

Threshold band range =19.1742-17.8638Hz.

Here, all the possible faults present in the circuit have not been considered. For the faults considered, 100% fault coverage was achieved.

Single Op-amp oscillator

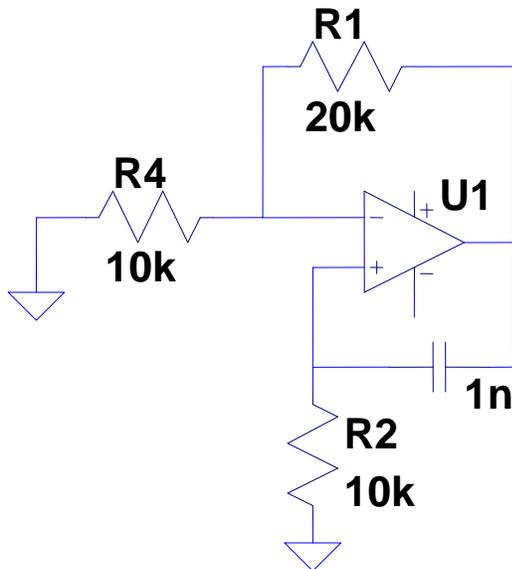


Fig.12. Single Op-amp Oscillator

As given in Fig.12. an op-amp is converted to an oscillator by adding both positive and negative feedback. In a fault free case, the oscillation frequency is $f_{osc} = 9 \text{ KHz}$. The results of parametric analysis of single Op-amp oscillator is displayed in time domain and frequency domain Fig.13 and Fig.14. This oscillation frequency significantly varies with any hard or parametric faults injected into the circuit. To improve the fault coverage, both amplitude and frequency measurements are to be considered. The parametric test analysis on this circuit, when analyzed indicated significant frequency deviation in the value of R and C. Changing different components in different parts of the circuit shows different effect on the output response. For example, changing the value of capacitor in the circuit from 1n to 100n resulted in a drastic change in the output frequency value. It has been shown in Fig.15. This was completely away from the desired range of frequency. It confirms that different components have different sensitivity.

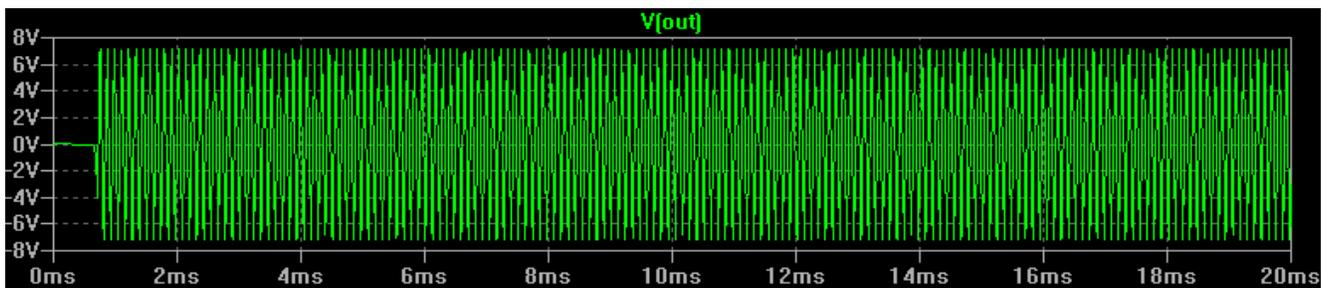


Fig.13. Simulation result in time domain for a single Op-amp oscillator

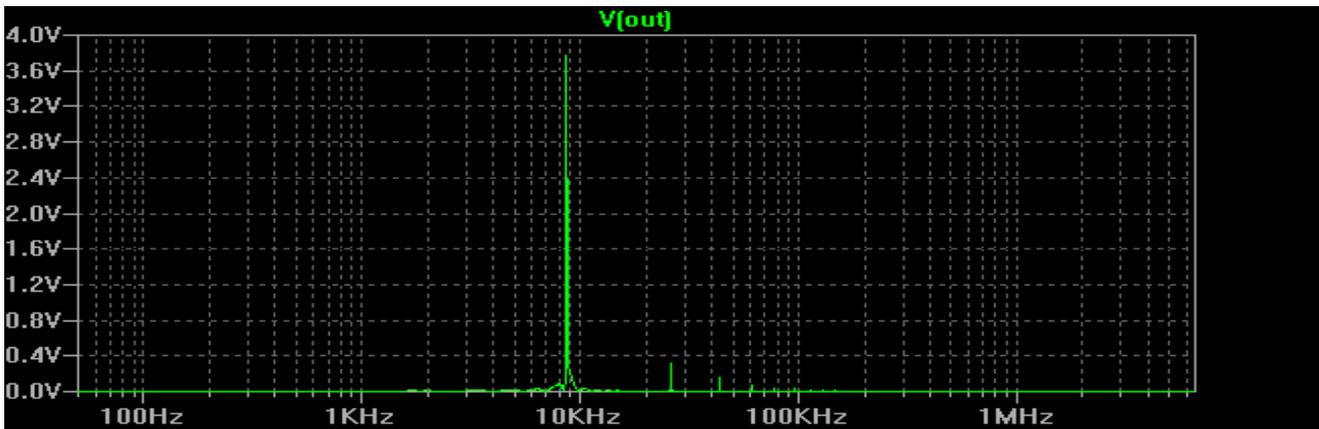


Fig.14.Simulation result in frequency domain for a Single Op-amp oscillator

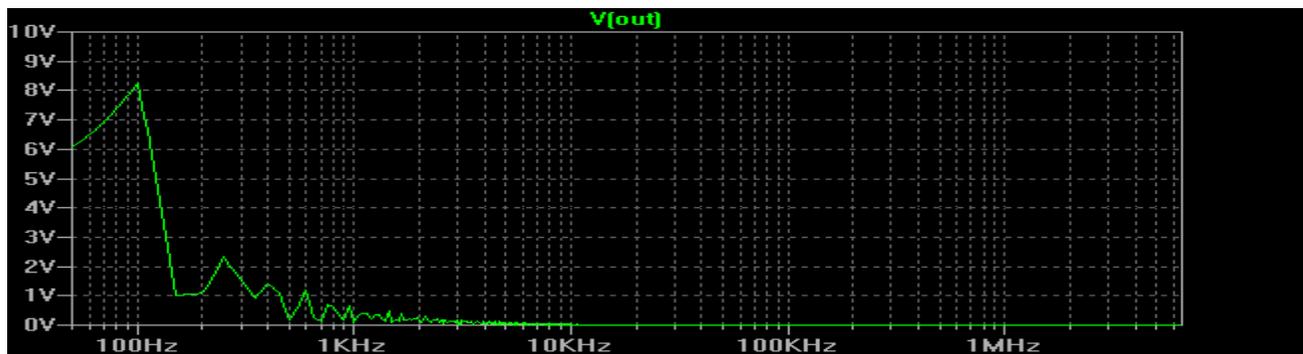


Fig.15.Simulation result for C1 parametric

9. Conclusion

The OBIST method has been effectively employed in testing and mixed signal circuits in embedded core based SOC environments. There are many ways to ensure proper functioning of a designed circuit, but this built-in hardware approach has proven to be one of the most reliable method. This oscillation based technique is implemented where the hard and parametric fault models are defined for fault coverage evaluation. By using Amplitude and Frequency measurement together we have seen improvement in fault coverage. Some sample benchmark circuits have been simulated and studied. The results have been verified in spice Simulation Software. The detection of faults in fault coverage shot up when the TIME-DOMAIN output was converted to FREQUENCY-DOMAIN and compared with the same $\pm 5\%$ threshold values. A very high coverage of 100% was achieved as well. The OBIST methodology has been applied for many circuits like analog to digital converter, filters, Dual tone multifrequency Detector, Switched Capacitor circuits and even MEMS systems.

References

- [1] M. J. Ohltz, "Hybrid built-in self-test (HBIST) structure for mixed analog/digital integrated circuits", in Proc. Eur. Test Conference, 51-57, 1991.
- [2] Karim Arabi and Bozena Kaminska, "Oscillation-Test Methodology for Low-Cost Testing of Active Analog Filter", IEEE Transactions and Instrumentation and Measurement, August 1999.
- [3] S. R. Das, "Self-testing of embedded cores-based systems with built-in hardware", *Proc. Inst. Electrical. Eng.—Cir. Dev. Syst.*, vol. 152, no. 5, 539-546, March 2005.
- [4] S. Sedra and K. C. Smith, *Microelectronic Circuit*, 2nd ed. New York: Holt, Rinehart, and Winston, 1987
- [5] Sunil.R.Das, j et.al, June, "Testing Analog and Mixed -signal circuits with Built-in Hardware-A New Approach", IEEE Transactions On Instrumentation And Measurement, vol. 56, no. 3, 840-852, June 2007.

A new approach Data hiding in 2D data matrix and tilt correction algorithm

KIMMY GHANAIYA¹ GAGANDEEP KAUR²

¹Student, M.Tech Electronics and Communication Section, YCOE, Talwandi Sabo, Guru Kashi Punjabi University Campus, Patiala

²Asst Prof., Electronics and Communication Section, YCOE, Talwandi Sabo, Guru Kashi Punjabi University Campus, Patiala

Abstract- Today barcodes are very much popular and present all around us. Not only on groceries in the super-market, but also in industrial manufacturing. Barcodes are designed for automated reading and interpretation. This paper presents two algorithm, first for data hiding in 2D data matrix provides a high level security and second for decoding purpose, independent of orientation of data matrix in any direction. This method also provides de noising of data matrix images captured by a webcam in real time YUY2_640x480 resolution. Experimental results shows this technique provides more accurate results as compared to previous methods with improved speed.

Keywords- Blind source separation, coordinate detection and slope detection, rotation detection.

I. Introduction

Bar codes are very much popular and reliable method of encoding of machine readable information on different products and services. These codes can be one dimensional or two dimensional. Now a day's technology is very vast, with the invention of camera phones this technique becomes quite popular because camera phones provide us optical imaging system. Two dimensional bar code is developed from the traditional one dimensional bar codes. It provides us more information capacity rather than 1D codes[1]. Data matrix two dimensional bar codes are of two types of ECC000-140 and ECC200. Each Data Matrix two-dimensional bar code is formed by some square blocks according to specific rules[2]. Here a new pattern for 2D data matrix is designed. The four sides of the bar code are surrounded by the specific graphics, naming Finder Pattern. First 8 bits are 11111111 here that will help in alignment.

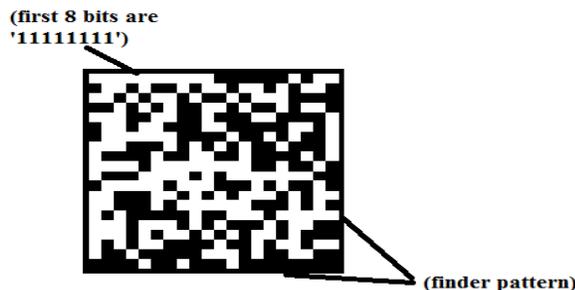


Figure 1. 2D Data Matrix

II. Encoding of data

Security is first requirement, for this purpose an encryption technique is used here technique called blind source separation. By using this encryption technique an unauthorized person will not be able to decode data. The size of data matrix can be changed according to our requirement. Here 'i' represent the maximum length of data string. The dimensions of data matrix can be varied according to your need with a finder pattern of 10x10. The size of square blocks is 20x20. The data is encoded within this finder pattern with the help of black and white blocks.

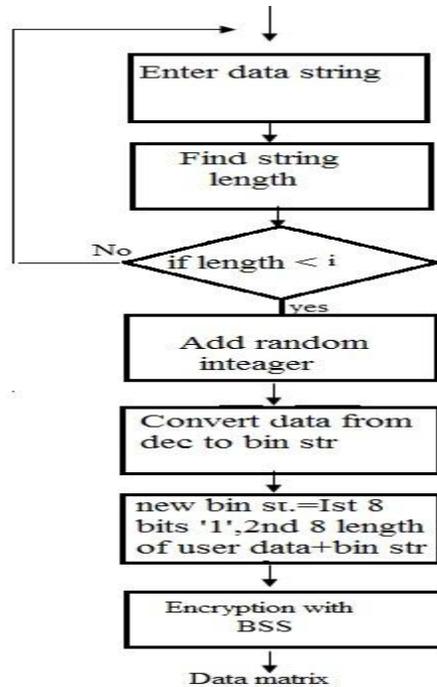


Figure 2. Encoding process

III. Decoding process

To extract data from data matrix, firstly image of that data matrix is to be decoded. Experimental results show that if images are tilted, it is impossible to be decoded correctly[3]. Some the traditional two-dimensional bar code correction algorithms, such as: algorithm based on projection method, Hough transform method, Fourier transform method are there[4]. But these are quite complex and more time consuming. This algorithm has the ability to correct the position of data matrix aligned in any direction. Angle of orientation will be detected with the help of slope finding. The objective of this paper is to solve the rotation problem, noising and decoding image correctly.

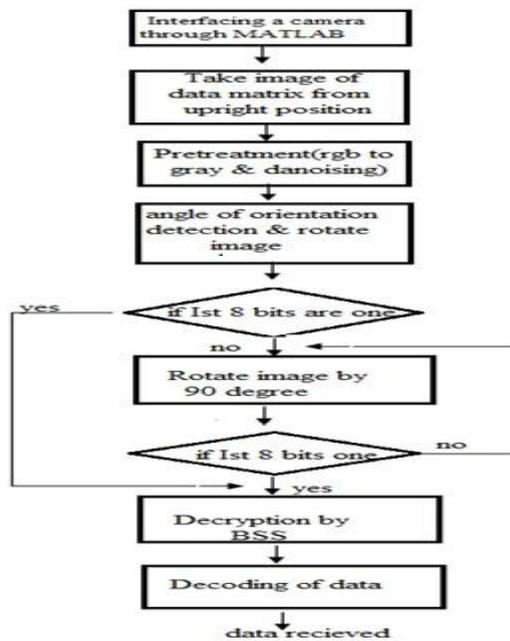


Figure3. Decoding process

any direction. Angle of orientation will be detected with the help of slope finding. The objective of this paper is to solve the rotation problem, denoising and decoding images correctly.

IV. Image processing

This approach analyzes the image captured by camera using MATLAB software. Image processing techniques are used in this approach for the detection of angle of orientation and resizing the images. This approach requires the camera to focus on the data matrix. Camera should be exactly at upright position of data matrix while capturing the image. The distance between data matrix and camera can vary. For this purpose it requires precise camera focus adjustment

V. Angle of orientation

To detect orientation of image ,coordinates of image have to find[3], with the help of these co-ordinates slope will be detected. Equation of line is given below

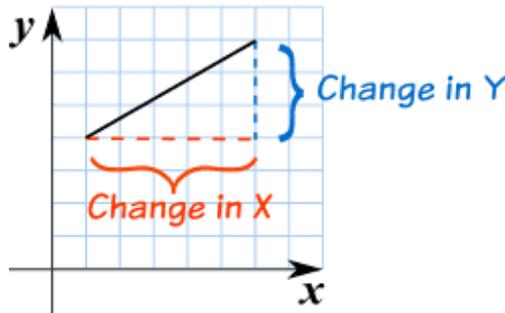


Fig. 4 Slope detection

$y=mx+c$ (1)

here m represents slope

$m=dy/dx$ (2)

After finding angle of orientation image will be rotated to that angle by using imrotate command in MATLAB.

VI. Cutting of image

Second step is cutting of image, only required portion of image will be cut from the whole image. With the help of imresize command in MATLAB image can be resized to original dimensions.

VII. Decoding of data matrix

If first 8 bits of data matrix are ‘11111111’ decryption process will start , it will convert binary string into original binary string, then it will be decoded into original information, otherwise image will be rotated to an angle of 90 until first 8 bits will be‘11111111’.

VIII. Rotation process

After cutting image with the help of slope finding rotation of image will be processed to find the actual data Matrix. This will be done with the help of following procedure.

Detection of first 8 bits	Rotation of image
if bits not equal 11111111	Rotate image anticlockwise
If bits equal to 11111111	Rotation stop, start decoding

Table1-Bits detection vs rotation

IX. Conclusion

A new method is proposed here which is a hybrid model for data hiding in data matrix and provides decoding algorithm which is independent of image orientation in any direction and de noising is one of the main advantages of this. It runs at 640×480 resolution for real time scenarios. For practical implementation dimensions of data matrix image is 420×420 includes finder pattern of 10×10 and size of squares is 20×20. This proposed system provides an accuracy of 97.6%. Our experiments show that the proposed system produces accurate results with improved speed.

X. Reference

- [1] Yongyue Yang , Guang Zhu and Peng Wang, “Reading system for 2D code based on machine vision”, ICEMI ,2009
- [2] Zhi Liu, Herong Zheng Wenting Cai, “ Research on Two- Dimensional code positioning approach based on Convex hull algorithm”, International Conference on Digital Image Processing, 2009
- [3] Bo Chen, Zhi Liu, Xiaohui Xu, “ The Skew correcting method for two dimension bar codes based on Least Square Methods” International conference on Digital Image Processing , 2009
- [4] Yinghong Liang, Zhiyan Wang, “ A Skew Detection Method for 2D Barcodes Images in High Noise Conditions Based on Projection Method” , Micro computer Information, 22(8) 2006.
- [5] Yuanyuan Chen, Pengfei Shi,” Image Recognition in 2-D bar code”[J]. Measurement & Control Technology,25(12) : 17-19,2006
- [6] Da Yang, Yu Shenglin, “Encoding and decoding of 2D code by Computer”[J].Data Acquisition and Processing, Vol.18 No.3,2003
- [7] Shuhua Wang, Zuo Li, Shijie Cai, Detection for Skew of Document Image Based on Least Square Method. Computer Application and Software, 2001.

Estimation of the Periods of Occurrence of Spread- F Over Ouagadougou

Tomwa, A. C.

Physics and Electronics Department, Adekunle Ajasin University,
Akungba Akoko, Ondo State, Nigeria.

Adimula, I. A.

Department of Physics, University of Ilorin, Ilorin, Nigeria

Abstract

Spread F data from the ionogram Collected in Ouagadougou (west Africa) were used to investigate the rate of occurrence of medium-scale irregularities in the electron concentration in the F-region of the ionosphere at all the seasons of the year. It was observed that at both equinox and solstices irregularities occur mostly at night, which is the same period at which scintillation also occur, when data were considered on an hourly basis.

Occurrence of spread F was also studied during magnetic storms and during quiet periods. It was shown that spread- F is observed during disturbed period as well as quiet periods. At equinox, the values of the thickness of spread-f is seen to be higher for the quiet periods than the disturbed days throughout, but at solstices it is only in the evening period that the thickness for the quiet period is more than that of the disturbed periods.

Keywords: ionosphere, ionogram, irregularities, equinox, solstices, quiet and disturbed periods.

1.0 Introduction

Spread F data obtained from a typical Ionogram showing records of the ionospheric conditions indicated by the relationships between the frequency of radio pulse emitted upward and the vertical height of echoes reflected from the ionosphere from Ouagadougou (Geo latitude 12.37°N, Geo longitude 358.47°N) are used for my study. Spread F occurs mainly in the dawn hours during magnetic storm (disturbed period) and during quiet periods (Soicher et al 1997). Data for the months of January, April, July and October are used for my analysis each representing different seasons of the year.

Spread f is an indication of the occurrence of irregularities in the f2 region (Dieminger et al 1995). It is also referred to as the ionogram signature of the f-region irregularities in the ionosphere. The f-region of the ionosphere is the ionized region above the E-layer.

The ionosphere is under constant agitation by a variety of external influences which causes the irregularities in this region. Apart from this irregularities in the ionosphere the agitation is either caused by scintillation or ionospheric storm.

Spread f phenomenon is particularly severe near the magnetic equator and in the auroral zones. In middle and high latitudes these irregularities appear to occur in patches with horizontal extent up to several hundreds of kilometers which drift with speeds of around the 100m/s.

The most likely time of occurrence is between 2100 and 0100LT (Dieminger et al 1995).

The F-region of the ionosphere is subdivided into f1 and f2 region. The f1 region denotes that part of the atmosphere between heights of about 140km to about 220km. The temperature in the f1 region increases rapidly with heights from around 500K at 140km to over 1000K at 220km.

The data used for my study is from the ionogram recorded from Ouagadougou (Geo latitude 12.37°N, Geo longitude 358.47°N) for different seasons of the year i.e. equinox and societies. The ionogram is scaled on an hourly basis between the hours of 1900 and 0700LT. The ionogram signature on the ionosonde is considered from the two ends of the spread referred to in this work as the bottom and topside of the spread.

Occurrence statistic of spread F(over Ouagadougou) which is the word-wide probability distribution of spread F per local time. Spread – F occurrence was recorded only between 1900 and 0700 LT. In summer (solstices) higher night time occurrence was observed at low and high solar activity.

In Equinox (winter) higher percentage occurrence takes place during low solar activity especially in the later part of the night (Saksena 1996).

2.0 Methodology

The values of the height of occurrence of spread F (Hmin and Hmax) and the corresponding frequencies of occurrence were obtained from the ionogram data obtained in Ouagadougou using an ionosonde.

Also the number of times spread F occur (rate of occurrence) per day were also determined on an hourly basis mainly by observation. The case of fig Ig is a typical case of which spread does not occur.

3.0 Definition Of Parameters Used.

- i. **Fmin** – This is the minimum frequency at which spread F occur. It indicate the frequency at the bottom side of the spread and is measures in Megahertz (MHz)
- ii. **Fmax** – This is the maximum frequency at which spread F occur. It indicate the frequency at the topside of the spread and is also measured in Megahertz (MHZ).
- iii. **Hmin** – This is the minimum height at which spread F occur. i.e the average height at the bottom side of the spread. It is measured in kilometer (km).
- iv. **Hmax** – This indicate the maximum height at which the spread occur. It is the average height at the topside of the spread and is also measured in kilometers (km).
- v. **Bo** – This parameter indicate the thickness (in km) of spread F at the bottomsides of the spread.
- vi. **B1** – This also indicate the thickness of spread F at the topside of the spread. It is also measured in kilometers (km).

The graphs of the average values of Hmin and Hmax are plotted against the local time (LT) for months of January, and October representing two seasons of the year. From these graphs the average minimum and maximum heights at which spread F occur can be determined. Also the time at which the heights is greatest is also determined.

Also, the graph of the rate at which spread F occur was plotted against the local time per month. The time at which the occurrence is greatest can easily be determined from the graphs.

For the quiet and disturbed periods the five most disturbed days and the five most quiet days of the months were considered. The periods at which the Ap index is less than 26 is considered to be a quiet day and disturbed if otherwise.

Graphs of their average values of heights (Hmin and Hmax) are plotted against the local time. The graph shows the periods when the values of Hmin or Hmax become lower or higher for quiet and disturbed periods. Also the graph of B1 against local time for both quiet and disturbed periods are plotted.

**Table 3.1 Ionogram readings with time.
JANUARY**

Fmin (MHz)	Bo (km)	Hmin(km)	Fmax (MHz)	B1 (km)	Hmax(km)	LT	Date
1.65469	25	271.875	5.7495	71.875	329.6875	0:00	04-01-91
1.65469	81.25	278.125	5.71845	68.75	296.875	1:00	
1.70011	28.125	248.4375	6.69084	31.25	312.5	2:00	
1.70934	25	243.75	6.61877	150	400	3:00	
1.71862	28.125	229.6875	5.35867	37.5	325	4:00	
1.92561	25	237.5	4.45755	46.875	310.9375	5:00	
1.8241	37.5	253.125	3.70796	65.625	395.3125	6:00	
2.34009	18.75	296.875	7.49667	53.125	384.375	7:00	
1.76579	31.25	421.875	6.72717	43.75	568.75	19:00	
1.8241	56.25	428.125	6.7637	59.375	579.6875	20:00	
1.8241	75	403.125	6.69084	150	512.5	21:00	
1.84397	37.5	318.75	8.35419	87.5	443.75	22:00	
1.75625	34.375	267.1875	8.49102	115.625	360.9375	23:00	

APRIL

Fmin (MHz)	Bo (km)	Hmin(km)	Fmax (MHz)	B1 (km)	Hmax(km)	LT	Date
1.56746	12.5	296.875	11.4369	143.75	484.375	0:00	10/4/1991
1.45303	9.375	270.3125	6.7637	18.75	346.875	1:00	
1.53388	15.625	260.9375	8.9635	59.375	439.0625	2:00	
1.5256	12.5	256.25	9.01217	46.875	410.9375	3:00	
1.48485	12.5	262.5	8.9635	34.375	398.4375	4:00	
1.8241	25	153.125	8.49102	31.25	390.625	5:00	
2.01086	9.375	298.4375	8.49102	25	362.5	6:00	
3.55075	12.5	262.5	11.5614	28.125	395.3125	7:00	
1.90486	15.75	435.875	5.7495	18.75	534.375	19:00	
1.86405	87.5	565.625	5.71845	34.375	701.5625	20:00	
1.95714	103.125	623.4375	4.31504	146.875	639.0625	21:00	
1.89458	43.75	487.5015	5.7495	115.625	632.8125	22:00	
1.85398	131.25	396.875	6.7637	78.125	542.1875	23:00	

JULY

Fmin (MHz)	Bo (km)	Hmin(km)	Fmax (MHz)	B1 (km)	Hmax(km)	LT	Date
1.87417	43.75	440.625	4.73114	121.875	535.9375	0:00	6/7/1991
1.78502	12.5	371.875	4.70559	103.125	454.6875	1:00	
1.49291	9.375	301.5625	5.71845	62.5	456.25	2:00	
1.75625	9.375	282.8125	5.5657	50	428.125	3:00	
1.66368	9.375	276.5625	5.78072	84.375	457.8125	4:00	
1.72795	9.375	270.3125	5.97163	87.5	406.25	5:00	
2.25304	18.75	296.875	6.80043	18.75	334.375	6:00	
3.55075	12.5	250	9.3098	18.75	350	7:00	
1.8241	12.5	334.375	6.69084	12.5	475	19:00	
1.88434	131.25	478.125	4.70559	106.25	590.625	20:00	
1.86405	43.75	418.75	4.48175	153.125	651.5625	21:00	
1.93606	65.625	514.0625	3.95691	215.625	585.9375	22:00	
1.84397	15.625	410.9375	4.73114	56.25	468.75	23:00	

OCTOBER

Fmin (MHz)	Bo (km)	Hmin(km)	Fmax (MHz)	B1 (km)	Hmax(km)	LT	Date
1.80445	15.625	264.0625	8.35419	131.25	375	0:00	1/10/1991
1.65469	15.625	242.1875	8.44516	93.75	371.875	1:00	
1.65469	18.75	246.875	8.39955	128.125	482.4125	2:00	
1.67271	18.75	246.875	6.80043	25	362.5	3:00	
1.65469	12.5	287.25	6.94934	18.75	390.625	4:00	
1.73733	21.875	242.1875	7.14007	109.375	401.5625	5:00	
2.02178	9.375	276.5625	7.74425	25	337.5	6:00	
3.06776	9.375	267.1875	11.4369	40.625	357.8125	7:00	
1.68179	109.125	564.0625	7.14007	75	740.625	19:00	
1.97846	96.875	592.1875	5.84367	243.75	656.25	20:00	
1.92561	84.375	545.3125	7.10151	212.5	625	21:00	
2.20477	84.375	407.8125	7.10151	178.125	698.4375	22:00	
1.65469	37.5	381.375	6.7637	175	509.375	23:00	

3.2.1 Results of Occurrences Rate.

The following deductions were inferred through general out look or sighting of the graph of number of occurrence (N) against local time for the various mouths of the year. Figures 3a-3d

+

(i) **January: figure 3a**

For the month of January, spread F occurs throughout the period of 1900 LT to 0700LT with the maximum occurrence at the 2100 LT when it occurs for 30 days of the month. The minimum number of occurrence of spread f is at 0700LT when it occurs for only 6 days of the month.

(ii) **April: figure 3b**

Here spread of also occurs from 1900LT to 0700LT with the highest occurrence at the 2000LT where it occurs for 29 day of the month. The least occurrence is at 0700LT when it occurs for just 3 days of the month.

(iii) **July: figure 3c**

Here spread f occurs between 2000LT and 0600LT. Spread f occurrence reaches its maximum at 0100LT and 0300LT when it occurs for 28 days of the month. Its least occurrence is at 0500LT when it occurs for 14 days of the month. It is important to note that here there is no spread f beyond 0500LT.

(iv) **October: figure 3d**

Spread f occurs between 1900LT and 0700LT for the month of October. At 2000LT the number of occurrence reaches its peak at 25days of the month. At 0700LT the least number of occurrence of 5 days is observed.

3.2.1 Summary of Results of Occurrence Rate.

It is pertinent to note from the above inferences that spread f occurs most at between hours of 2000LT and 0100LT. This Unique observation conform with what was suggested by Dieminger et al 1995 on his study of spread f.

At equinoxes (April, October). Spread f occurs throughout the period of 1900 to 0700LT with the spread reaching its maximum at 29 days for April. Also here at 1900LT the occurrence rate is as high as 23 days for October.

At solcitics (January, July) spread f does not necessarily occur throughout the periods of 1900 to 0700LT. For July especially spread f occur between 2000LT and 0500LT and disappear before and beyond this hours. Here, we have the highest number of occurrence of 30 days at 2000LT for January.

Result of the Height of Occurrence.

The minimum and maximum heights of occurrence of spread f were also observed and the following inference deduced from the graphs.

Table 3.2a: Occurrence Statistics for January

NO OF OCCURRENCE (N)	LOCAL TIME (LT)	Jan
8	19	
29	20	
30	21	
27	22	
23	23	
26	0	
23	1	
23	2	
24	3	
21	4	
22	5	
17	6	
6	7	

Table 3.2b: Occurrence Statistics for April

NO OF OCCURRENCE (N)	LOCAL TIME (LT)	April
4	19	
29	20	
28	21	
28	22	
27	23	
27	0	
25	1	
26	2	
24	3	
16	4	
8	5	
7	6	
3	7	

Table 3.2c: Occurrence Statistics for July

NO OF OCCURRENCE (N)	LOCAL TIME (LT)	July
0	19	
24	20	
24	21	
26	22	
27	23	
28	0	
28	1	
28	2	
28	3	
18	4	
14	5	
0	6	
0	7	

Table 3.2d: Occurrence Statistics for October

NO OF OCCURRENCE (N)	LOCAL TIME (LT)	Oct
23	19	
25	20	
22	21	
22	22	
21	23	
21	0	
21	1	
20	2	

18	3
16	4
13	5
16	6
5	7

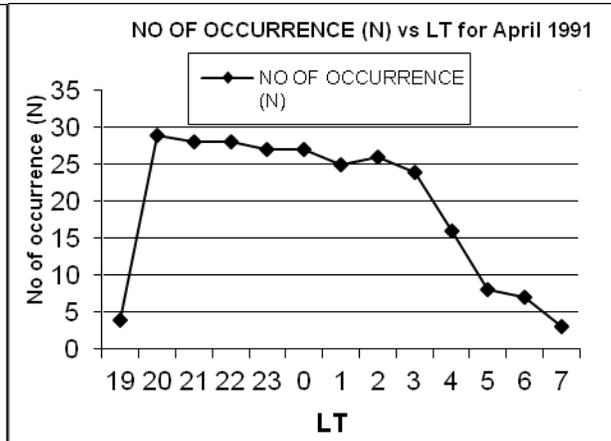
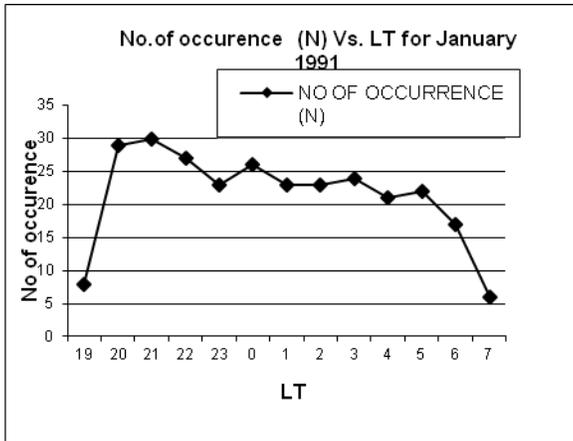


Fig. 3a: Number of occurrence vs LT for January

Fig. 3b: Number of occurrence vs LT for April

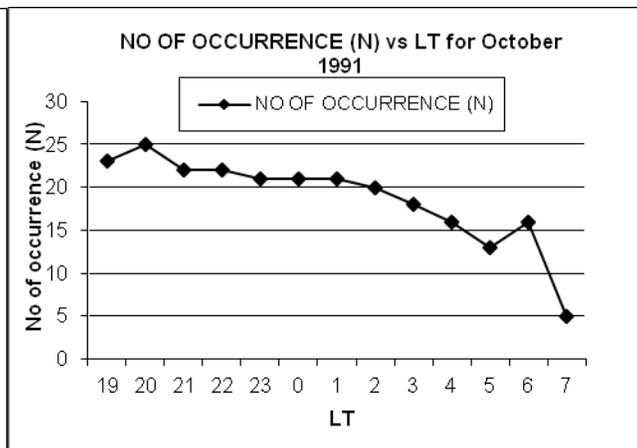
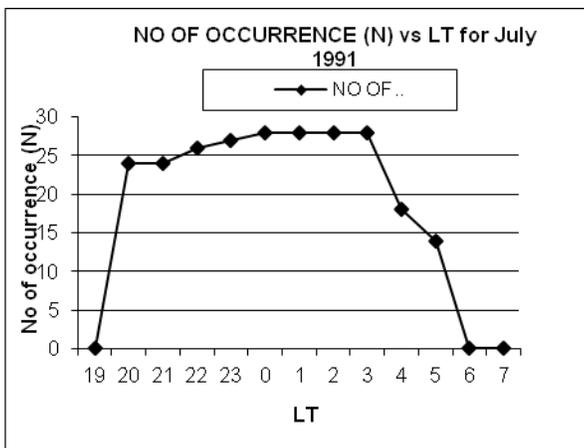


Fig. 3c: Number of occurrence vs LT for July

Fig. 3d: Number of occurrence vs LT for October

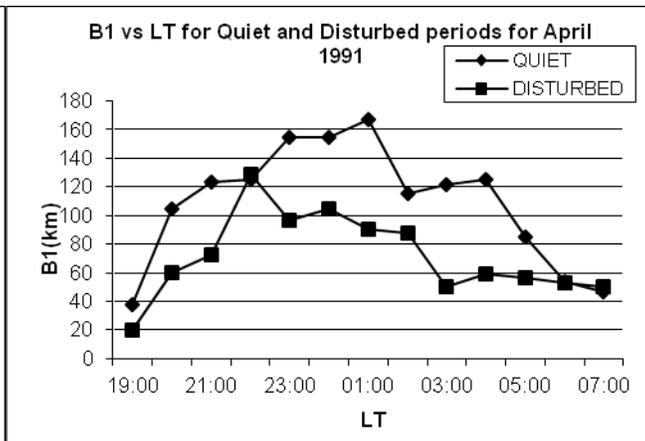
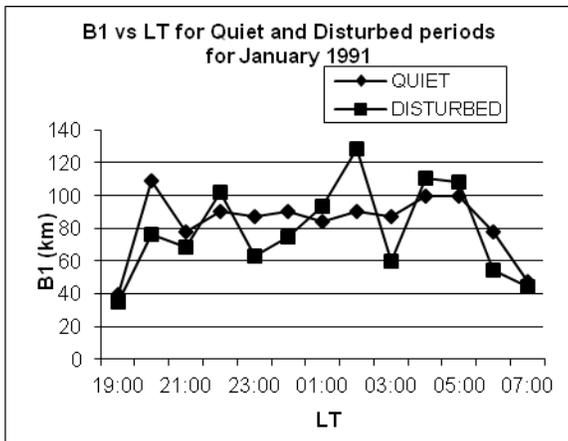


Fig. 3e: B1 vs. LT for Quiet and Disturbed periods

Fig. 3f: B1 vs. LT for Quiet and Disturbed periods

For January

for April

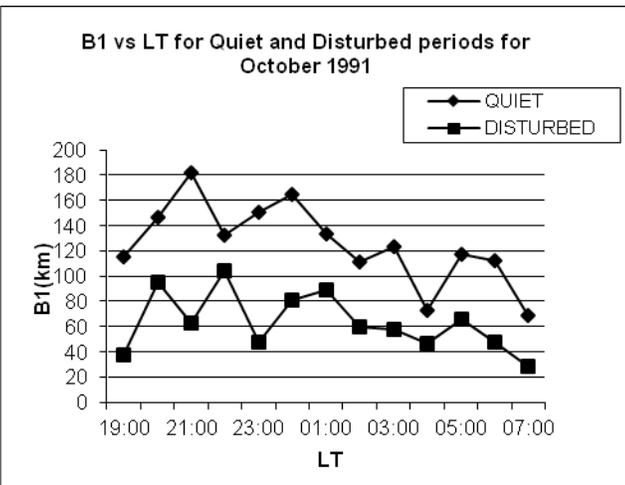
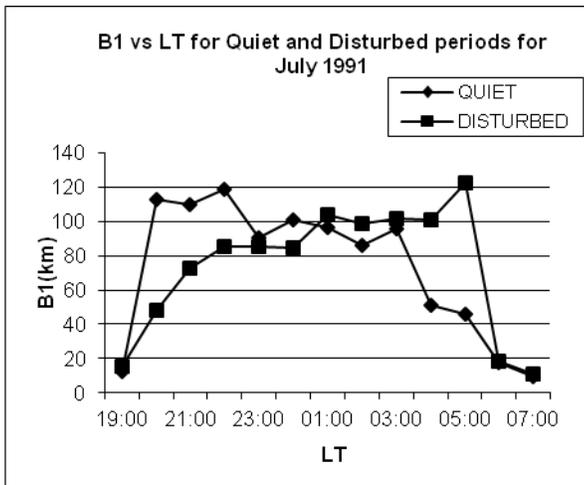


Fig. 3g: B1 vs. LT for Quiet and Disturbed periods For July

Fig. 3h: B1 vs. LT for Quiet and Disturbed periods for October

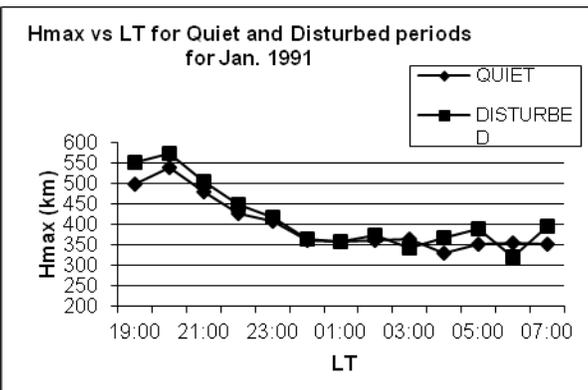
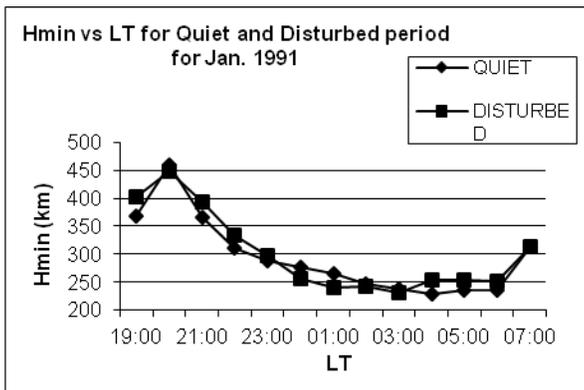


Fig. 3i: Hmin vs. LT for Quiet and Disturbed period For January

Fig 3j: Hmax vs. LT for Quiet and Disturbed period for January

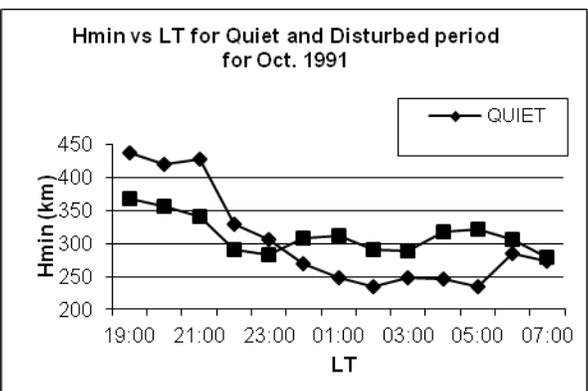
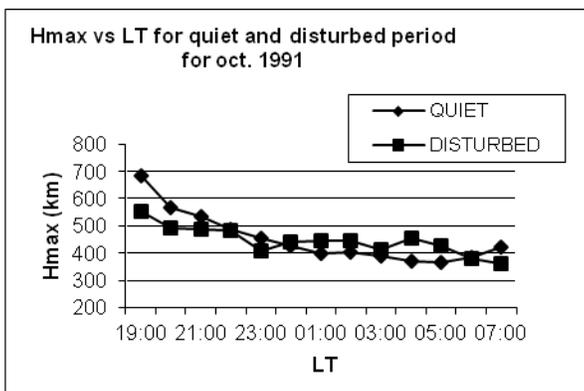


Fig. 3k: Hmin vs LT for Quiet and Disturbed period for January

Fig 3l: Hmax vs LT for Quiet and Disturbed period for January

i. January: figures 3i and 3j

The maximum height of occurrence (Hmax) is greater at the 2000LT with the height of 580.996km and least at 0400LT at the height of 340.5242km. Hmin (minimum height of occurrence) is also greater at the 2000LT with the height of 449.74798km and least at 0400LT with the height of 237.19758km. The curve of the variation of Hmin with the LT is similar in pattern.

ii October: figures 3k and 3l

For the month of October Hmax reaches its peak value at 1900LT at a height of 653.8996km and least at 0500LT at the height of 371.2379km. Hmin also has its highest value at the 1900LT with the value of 442.7520km and least value at 0200LT with the height of 248.67742km.

The curves of Hmin, Hmax with the local time also follow the same trend.

According to Diemenger 1995, spread F which is the region of maximum electron density and occurs mostly between the height of 220 and 500km.

The height range of both the Hmin and Hmax of the spread fall within this ranges of height except for some exceptional cases where the maximum height is slightly above this 500km. This observed height ranges can thus be said to conform with the theoretical values.

At equinoxes (April, October) the maximum height of occurrence is as high as ~650km, while for solstices (January, July) it is as high as ~580km. The reason for the higher value of Hmax at equinox than solstices is due to the fact that equinox is more magnetically disturbed than solstices. Hence, the every likelihood of vertical movement of electrons in the F2 region in the upper direction. This conform with the observation of Adeniyi J.O et al 1998 in the study of the effect of magnetic storm on the bottomside profile parameters.

4.0 Summary of result of quiet and disturbed periods

The interesting thing is that spread F is observed both at very quiet and disturbed period. This observation conform with what Soicher H. et al observed in their study of spread of during quiet and disturbed period.

At both equinox and solstices the variation in the height of occurrence of disturbed and very quiet period were less in the day time but in the evening and night time they exhibit a very wide variation this also conform with theoretical observation made by Soicher in his study.

4.1 Result of thickness of occurrence (B_1)

The graph of thickness of occurrence against Local time was also plotted for both very quiet period and magnetically disturbed period. Very interesting results were obtained;

(i) January: figure 3e

Spread F is very thick in the evening time for very quiet period from i.e from 1960 to 2300LT the night to the early morning the thickness of the spread is more for the magnetically disturbed period than the quiet periods.

(ii) April: figure 3f

Here, the thickness of spread f is more for the quiet period that the disturbed period throughout i.e both in the evening and night periods till the early hour of the day.

(iii) July: figure 3g

Between the hours of 1900 to midnight the thickness of spread is higher for the quiet period and that of the disturbed period becomes higher than that of very quiet period for from 0100LT upward.

(iv) October: figure 3h

The thickness of spread f is always more for the quiet period than the magnetically disturbed period. It is also important to note here that the graph follow the same pattern.

4.2 Summary of result for thickness of Occurrence for quiet and disturbed period B_0 and B_1

The most interesting thing to note here is that at equinox (April, October) the values of the thickness of spread f is higher for the quiet period that the disturbed days throughout. Whereas, at solstices (January, July) it is only in the evening period that the thickness of the quiet period is more than that of the disturbed period.

The difference in the thickness of the spread in the f2 layer at quiet and disturbed period is due to movement of ionization in the vertical direction. From the study of Adeniyi and Redicella on the effect of magnetic storm on the thickness of

spread in the F₂ layer we were made to understand that the main effect of magnetic storm on the thickness of spread is observed only in the day time i.e between 0700 and 1700LT. This is the reason why the thickness for the quiet period is always greater than that of the disturbed periods especially at equinox.

Conclusion

The result of this work can be concluded that spread F, which is the irregularities in the F region is a phenomenon that occur mostly in the night time, the quiet and disturbed magnetic conditions for all the seasons of the year ie at both equinoxes and solcities. In most cases there is no spread F occurrence in the day time. The average height of occurrence of spread –f as in this work is between 200 and 600km into the ionosphere.

Between the hours of 2100 and 0100LT is when spread – F is thickest as shown in his study. It is important to note here that the thicker the spread, the longer radio signal stays in the ionosphere before returning back to the earth receiver. Except in some cases the thickness of spread F for the quiet period (B₁) is always more than that for the disturbed period especially in the evening periods.

References

1. Abdu, M.A et al, Equatorial Spread. F Occurrence statistics in the American Longitude: Diurnal, Seasonal and Solar Cycle variations. J. adv. Space Res. Vol. 22 pp 851 – 854 (1998).
2. Adeniyi J.O and S.M. Radicella, “The effect of magnetic storm on the Bottom side profile parameters B₀, and B₁. Proceedings of IRI Task force activity, pp (2001).
3. Berkey F.T. and G.H. Stonehocker; A comparison of the height of the maximum electron density of the F₂ layer from real height analysis and estimates based on M(3000)F₂. J. atmos. And teriestrial physics. Vol. 51 pp 873 – 879 (1989).
4. Buonsanto, M.J. A review of Ionospheric storm. Space science reviews, vol. 88 pp 563 – 601, (1999)
5. David, P and J. Voge propagation of waves pp. 157-197 (1969).
6. Dieminger W., G.K. Hartmann and R. Leitinger (eds). (1995) the upper atmosphere
7. Ionospheric station information Bulleton (Jan 1970). Ionospheric Network Advisory Group (INAG).
8. Massey H.S.W. and R.L.F Boyd. The upper atmosphere.
9. Mc Graw – Hill, Encycloledia of science and technology vol. 9 pp 439-442.
10. Mc Graw –Hill, Encyclopedia of science and technology vol. 15 pp 92-98
11. Oyinloye J.O et al. Handbook on radio propagation for tropical and subtropical countries.
12. Rodger, A.S and J. Aarons. Studies of Ionospheric F-region irregularities from geo-magnetic mid-latitude conjugate regions, J. atmospheric and Terrestrial Physics vol. 50 pp 63 –72(1988).
13. Saksena, R.C. Occurrence statistics of spread – F over Indian subcontinent. J. adv space Res. Vol. 18 pp 99-102 (1996).
14. Soicher H, Gorman, E.E and O.V Weitsman, Spread –F during quiet and disturbed periods at mid latitudes. J. adv space Res. Vol 20 pp 2199 – 2202 (1997)
15. Wakai, N; H. Ohyama and T. Koizumi, Manual of Ionispheric propagation (1987).
16. Xin – yu Hnang and Shun-Rong Zhang, sporadic-E. spread F and some ionospheric oscillatory phenomena on the far east. J. adv. Space Res. Vol 18 pp 645 – 648 (1996).

Application Specific Quality of Service (QoS) Centric Parameters Simulation in Wireless Mobile Ad hoc Networks

Padmashree S¹, Manoj P B²,

4thsem M.Tech(DEC), AMC Engineering College¹, Asst Prof. AMC Engineering College²,

Abstract— Wireless mobile ad hoc networks is a collection of mobile nodes communicating with each other without any existing infrastructure networks via wireless links. In this paper, modelling and simulation of Ad hoc On-demand Distance Vector(AODV) Routing Protocol is performed. Study and analysis of parameters for Vehicular Communication application is considered where a very highly mobile ad hoc networks scenario will be created. Simulation is performed in ns2 (Network Simulator 2) for Quality of Service (QoS) parameters such as throughput and delay. Throughput variation also results due to various other factors such as jitter, number of mobile node variation and so on. The simulations are then plotted on x-graph.

Keywords— Wireless mobile ad hoc networks, AODV, DSDV, Quality of Service (QoS) parameters, ns2.

I. INTRODUCTION

Mobile adhoc networks (MANET) is a collection of wireless mobile nodes communicating with each other using wireless links without any existing network infrastructure as shown in Fig.1. There is no base station (BS) or access point (AP) in adhoc network. Each node in the network acts as a router, forwarding data packet for other nodes. Each node is free to roam about while communicating with others. The path between each pair of the nodes may have multiple links, and the radio between them can be heterogeneous. This allows an association of various links to be a part of the same network. The routing and resource management are done in a distributed manner in which all nodes coordinate to enable communication among them. Routing of data in adhoc network is a challenging task due to rapidly changing topology and high mobility of the nodes. The responsibilities for organizing and controlling the network are distributed among the terminals themselves.

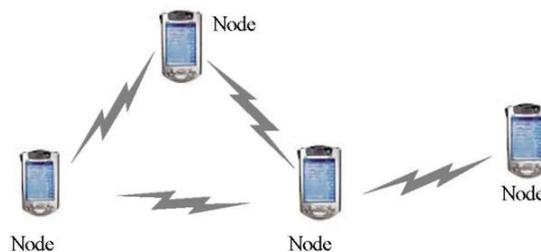


Fig. 1 Mobile Ad hoc Networks Structure

In ad hoc networks, some pairs of terminals may not be able to communicate directly with each other and relaying of some messages is required, so that they are delivered to their destinations. Such networks are often referred to as multi-hop or store-and-forward networks. The network topology is dynamic, because the connectivity among the nodes may vary with time due to node departures, new node arrivals, and the possibility of having mobile nodes. Hence, there is a need for efficient routing protocols to allow the nodes to communicate over multi-hop paths in a way that does not use the network resources than necessary.

Mobile hosts and wireless networking hardware are becoming widely available and extensive work has been done recently in integrating these elements into traditional networks such as the Internet. Often sometimes, mobile users may want to communicate with each other without the barriers of fixed infrastructure like fixed backbones or confined within a certain area. For example, a group of students may want to communicate with each other to share some lecture notes, assignments and so on; crisis management services, applications, such as in disaster recovery, where the entire communication infrastructure is destroyed and resorting communication quickly is crucial; The military can track an enemy tank as it moves through the geographic area covered by the network; Bluetooth, which is designed to support a personal

area network by eliminating the need of wires between various devices, such as printers and personal digital assistants; vehicular communications where each vehicle is equipped with a communication device will be a node in the adhoc networks for applications such as collision warning, road sign alarms, traffic information and so on.

II. APPLICATION: VEHICULAR COMMUNICATION

Vehicular communication is possible through wireless networks in which vehicles and roadside units are the communicating nodes. VANETs (Vehicular Ad hoc Networks) are highly mobile wireless ad hoc networks and will play an important role in public safety communications and commercial applications. Routing of data in VANETs is a challenging task due to rapidly changing topology and high speed mobility of vehicles. Fig. 2 shows Vehicular communication scenario in a Vehicular ad hoc networks where each of the vehicles which have mobility either is communicating with infrastructure network or in adhoc manner with another vehicle to exchange information. The growing mobility of people and goods incurs high societal costs: traffic congestion, fatalities and injuries. At the same time, vehicles have increasingly effective driver assistance and protection mechanisms.

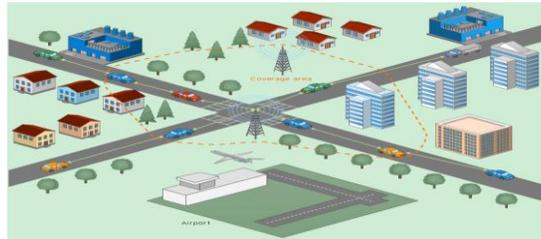


Fig. 2 Vehicular Communication Scenario

Radio communication based solutions can operate beyond the line of sight constraints of radar and vision solutions, and they can enable cooperative approaches. Vehicles and infrastructure cooperate to perceive potentially dangerous situations in an extended space and time horizon. Vehicles will be equipped with novel computing, communication and sensing capabilities and user interfaces. These will support a spectrum of applications that enhance transportation safety and efficiency, but also provide new or integrate existing services for drivers and passengers. Wireless transmission and medium access technologies adapted to vehicular communication environment are the primary enabling technology.

A. VANETs Characteristics

The fundamental characteristics of Vehicular Ad hoc Networks are as follows [5]:

1. High computational capability: Operating vehicles can afford significant computing, communication and sensing capabilities.
2. Highly dynamic topology: Vehicular network scenarios are very different from classic ad hoc networks. In VANETs, vehicles can move fast. It can join and leave the network much more frequently than MANETs. The topology in VANETs changes much more frequently.
3. Mobility: Vehicles tend to have very predictable movements that are (usually) limited to roadways. The movement of nodes in VANETs is constrained by the layout of roads. Roadway information is often available from positioning systems and map based technologies such as GPS. Each pair of nodes can communicate directly when they are within the radio range.
4. Potentially large scale: vehicular networks can in principle extend over the entire road network and so include many participants.
5. Partitioned network: Vehicular networks will be frequently partitioned. The dynamic nature of traffic may result in large inter vehicle gaps in sparsely populated scenarios and hence in several isolated clusters of nodes.
6. Network connectivity: The degree to which the network is connected is highly dependent on two factors: the range of wireless links and the fraction of participant vehicles, where only a fraction of vehicles on the road could be equipped with wireless interfaces.

B. Related Works

Several protocols have been defined by researchers for vehicular networks. New protocols are also required with the improving technology. The history of VANET routing begins with the traditional MANET routing protocols.

Kakkasageri et al [4] compared AODV and DSR with Swarm intelligence routing algorithm and have shown that AODV and DSR has less performance than swarm intelligence routing algorithm in VANET.

Jerome Haerri et al [6] evaluated the performance of AODV and OLSR for VANET in city environment, in their study all the characteristics are handled through the Vehicle Mobility Model. Their study showed that OLSR has better performance than AODV in the VANET, as the performance parameters that they used have less overhead on the network as compared to OLSR.

Performance analyses of traditional ad-hoc routing protocols like AODV, DSDV and DSR for the highway scenarios have been presented in [3], and the authors proposed that these routing protocols are not suitable for VANET.

Their simulation results showed that these conventional routing protocols of MANET increase the routing load on network, and decrease the packet delivery ratio and end to end delay.

In this paper, modelling and simulation of Ad hoc On-demand Distance Vector (AODV) Routing Protocol and Destination-Sequenced Distance-Vector Routing (DSDV) Protocol is performed and parameters like jitter, number of mobile nodes are varied and the throughput variation is analysed.

I. ROUTING CONSIDERATIONS

Since several hops may be needed for a packet to reach its destination, routing protocols are needed. There are two main functions of the routing protocols. First is finding the routes from the source to the destination and the second one is to deliver the messages or data packets from the source to the destination. The second process is more complicated as it also requires some control mechanism to assure the proper delivery of data packets so that the packets are not lost in the way. This process may also use a variety of protocols and data structures. In this paper, routing is carried out through AODV protocol.

A. AODV

AODV (Ad hoc On-demand Distance Vector) Routing protocol is a reactive ad hoc routing protocol. It creates a new route only when necessary. Route discovery in AODV is by using Route Request (RREQ) and Route Reply (RREP). As long as the new route remains valid AODV would not act. When it has expired or becomes invalid AODV will only react when the route is needed again. AODV protocol is composed of two mechanisms: Route Discovery and Route Maintenance. One important feature in AODV is the use of sequence numbers to avoid counting to infinity problem. By maintaining sequence numbers in each node for each destination, the problem can be avoided. [8]

If there is no information in the routing table of the source then a route request is broadcasted. A node that receives the route request and has an up-to-date path to the destination will return it to the source and all nodes on the return path will update their routing tables. If no valid path is present in an intermediary node then the request is rebroadcasted. If a node receives multiple paths to the destination the one with the highest utility is chosen.

AODV specifies two different ways in which a link break can be detected. Either all nodes regularly broadcast a 'hello' message to its one-hop neighbours, which makes it possible for them to verify the link operation, or it is detected by a link signalling mechanism when the link is used. When a link break is detected the end nodes (source and destination) are informed and it is up to them to find a new path.[9]

B. DSDV

DSDV [9] (Destination-Sequenced Distance-Vector Routing) is a table driven routing protocol that is an enhanced version of the distributed Bellman-Ford algorithm. In all table driven protocols each node maintains a table that contains the next hop to reach all destinations. To keep the tables up to date they are exchanged between neighbouring nodes at regular intervals or when a significant topology changes are observed.

In DSDV, each node maintains a routing table which is constantly and periodically updated (not on-demand) and advertised to each of the node's current neighbours. Each entry in the routing table has the last known destination sequence number. Each node periodically transmits updates, and it does so immediately when significant new information is available. The data broadcasted by each node will contain its new sequence number and the following information for each new route: the destinations address the number of hops to reach the destination and the sequence number of the information received regarding that destination, as originally stamped by the destination. No assumptions about mobile hosts maintaining any sort of time synchronization or about the phase relationship of the update periods between the mobile nodes are made. Following the traditional distance-vector routing algorithms, these update packets contain information about which nodes are accessible from each node and the number of hops necessary to reach them. Routes with more recent sequence numbers are always the preferred basis for forwarding decisions. Of the paths with the same sequence number, those with the smallest metric (number of hops to the destination) will be used. The addresses stored in the route tables will correspond to the layer at which the DSDV protocol is operated. Operation at layer 3 will use network layer addresses for the next hop and destination addresses, and operation at layer 2 will use layer-2 MAC addresses [12].

II. NETWORK SIMULATOR

After setting up the platform, software named NS-2 was set up on it which was used for all the analysis and simulation work apart from other tools used. NS-2 is the de facto standard for network simulation. Its behaviour is highly trusted within the networking community. It is developed at ISI, California, and is supported by the DARPA and NSF.

NS-2 is an object oriented simulator, written in C++, with an OTcl (Object Tool command language) interpreter as a frontend. This means that most of the simulation scripts are created in Tcl (Tool command language). If the components have to be developed for NS-2, then both Tcl and C++ have to be used. NS-2 uses two languages because any

network simulator, in general, has two different kinds of things it needs to do. On the one hand, detailed simulations of protocols require a systems programming language which can efficiently manipulate bytes, packet headers, and implement algorithms that run over large data sets. For these tasks run-time speed is important and turn-around time (run simulation, find bug, fix bug, recompile, re-run) is less important. On the other hand, a large part of network research involves slightly varying parameters or configurations, or quickly exploring a number of scenarios. In these cases, iteration time (change the model and re-run) is more important. Since configuration runs once (at the beginning of the simulation), run-time of this part of the task is less important. [10]

NS-2 meets both of these needs with two languages, C++ and OTcl. C++ is fast to run but slower to change, making it suitable for detailed protocol implementation. OTcl runs much slower but can be changed very quickly (and interactively), making it ideal for simulation configuration.

The simulator supports a class hierarchy in C++, and a similar class hierarchy within the OTcl interpreter. The two hierarchies are closely related to each other; from the user's perspective, there is a one-to-one correspondence between a class in the interpreted hierarchy and one in the compiled hierarchy. The root of this hierarchy is the class TclObject. Users create new simulator objects through the interpreter. These objects are instantiated within the interpreter, and are closely mirrored by a corresponding object in the compiled hierarchy. The interpreted class hierarchy is automatically established through methods defined in the class TclClass. User instantiated objects are mirrored through methods defined in the class TclObject. There are other hierarchies in the C++ code and OTcl scripts, these other hierarchies are not mirrored in the manner of TclObject. [11]

III. QUALITY OF SERVICE (QoS)

Quality of service (QoS) is the performance level of a service offered by the network to the user. A network or a service provider can offer different kinds of services to the users. Various parameters have to be resolved for an application. The parameters which affect the user end include delay, security, minimum bandwidth, jitter, maximum packet loss rate and so on. QoS does not create bandwidth, but manages it so it is used more effectively to meet the wide range of application requirements. The goal of QoS is to provide some level of predictability and control beyond the current Internet Protocol best effort services. Quality of service protocols uses a variety of complementary mechanisms to enable deterministic end-to-end data delivery.[12]

To support QoS, the link state information such as delay, bandwidth, cost, loss rate, and error rate in the network should be available and manageable. However, getting and managing the link state information in MANETs is very difficult because the quality of a wireless link is apt to change with the surrounding circumstances. Furthermore, the resource limitations and the mobility of hosts make things more complicated. The challenge is to implement the complex QoS functionality with limited available resources in a dynamic environment [13].

IV. SIMULATION RESULTS

The evaluation is performed considering the scenario of 100 nodes being static and attaining mobility. Also evaluation is performed for 200 nodes scenario. Table 1 gives the values of various parameters listed below.

TABLE I
GENERAL SIMULATION PARAMETERS

Sl. No.	Parameters	Values
1	Traffic type	Constant bit rate
2	Number of nodes	100, 200 nodes
3	Speed	50msec
4	Pause time	0.5msec
5	Simulation time	10, 20msec
6	Network Area	1000x1000 meters
7	Routing Protocol	AODV
8	MAC Protocol	802.11a
9	Packet size	500bits/sec
10	Queue length	50
11	Simulator	ns-2.32

The throughput obtained for the 100 node in a static scenario is plotted in x-graph as shown below which is time vs. bits per second in Fig. 3.

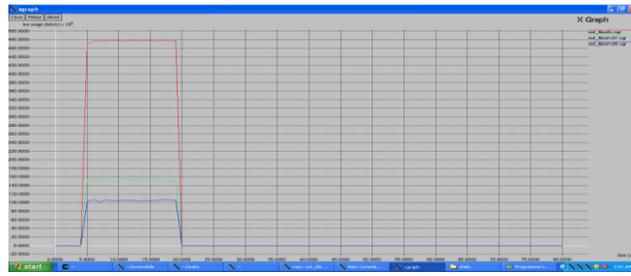


Fig. 3 Throughput plotted in x-graph for all 100 nodes in static condition

The performance metrics of throughput for the 100 nodes with mobility being simulated is shown in Fig. 4 below in the x-graph.



Fig. 4 Throughput plotted in x-graph for all 100 nodes with mobility

Simulation performed in similar way for 200 nodes also gives the results and is plotted in the x-graph and also network animator is used to view the animation.

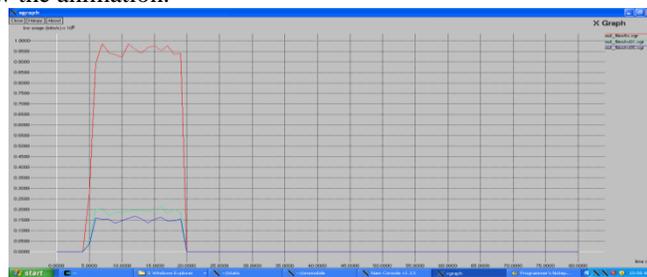


Fig. 5 X-graph for 200 mobile nodes verifying throughput

Fig. 5 shows the x-graph for 200 mobile nodes scenario being simulated for throughput performance metrics. The network animator is used to show animation for the same scenario and the node movement here can be observed. The network animator (NAM) is a Tcl/TK based animation tool for viewing network simulation traces and real world packet trace data. The first step to use NAM is to produce the trace file. The trace file should contain topology information, e.g., nodes, links, as well as packet traces. Usually, the trace file is generated by NS-2. During an NS-2 emulation, user can produce topology configurations, layout information, and packet traces using tracing events in ns2 [10]. When the trace file is generated, it is ready to be animated by NAM. Upon startup, NAM will read the trace file, create topology, pop up a window, do layout if necessary and then pause at the time of the first packet in the trace file. Through its user interface, NAM provides control over many aspects of animation. Fig. 6 shows the NAM for 200 nodes simulation below.

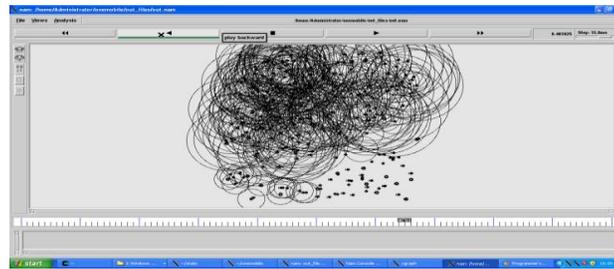


Fig. 6 NAM representing 200 mobile nodes involved in packet transmission

Simulation is performed by jitter variation which varies the throughput for the scenario. The Fig. 7 below shows the throughput for the jitter variation in 100 mobile node conditions. Jitter being set to 0.5ms and the overall simulation time considered is 10ms.

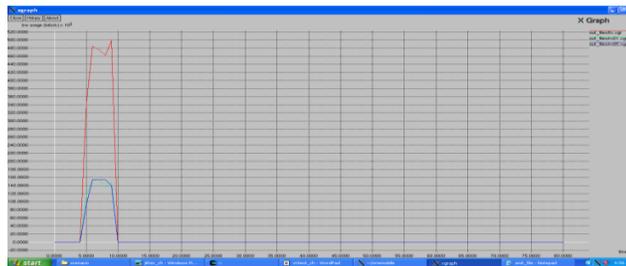


Fig. 7 X-graph of throughput for jitter variation of 0.5ms for 100 mobile nodes

The parameter delay verified for the mobility conditions are also plotted on x-graph in Fig. 8 which on x-axis will be the time and on y-axis is the delay variation taking place with respect to the time in milli second condition. The delay variation also shows how important it is to know the delay happening in the packet transmission. Thus the simulation and analysis will help to understand the concept of study of various parameters to meet the user requirements.

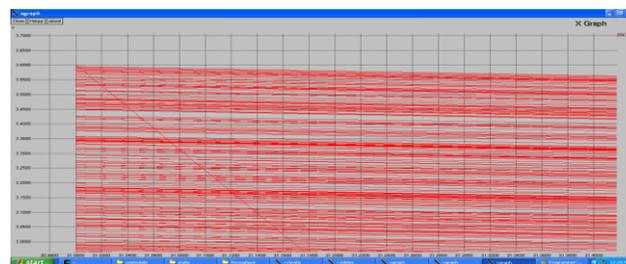


Fig. 8 X-graph for delay variation with respect to time

Vii. Conclusion

A systematic study and analysis of all the aspects of wireless mobile adhoc networks is carried out. Vehicular communication is considered and the comparison is made for transmission packet ratio. The applications have a range of QoS requirements in terms of the granularity of determinism and the level of guarantee, there are also a variety of services and protocols available. The ns2 simulator is used to simulate all the results of performance metrics like throughput, delay and jitter and the various results are plotted on x-graph.

REFERENCES

- [1] Li Layuan, Li Chunlin, YaunPeiyan, Performance evaluation and simulations of routing protocols in adhoc networks, Computer communications 30 (2007), 1890-1898.
- [2] PanosPapadimitratos, EPFL, Vehicular communication system, IEEE Communication Magazine, Nov 2009.
- [3] Xiong Wei, Li Qing-Quan, "Performance evaluation of data disseminations for vehicular ad hoc networks in highway scenarios", <http://www.isprs.org/congresses/proceedings/1pdf/174.pdf>
- [4] Manvi, S.S., Kakkasageri, M.S., Mahapurush, C.V., "Performance Analysis of AODV, DSR, and Swarm Intelligence Routing Protocols In Vehicular Ad hoc Network Environment" In International conference on future Computer and Communication., pp. 21-25, April. 2009.
- [5] K.Prasanth, K.Duraiswamy, K.Jayasudha&C.Chandrasekar, "Reducing Packet Transmission Delay in Vehicular Ad Hoc Networks using Edge Node Based Greedy Routing "

- [6] Jerome Haerri, FethiFilali, Christian Bonnet, "Performance Comparison of AODV and OLSR in VANETs Urban Environments under Realistic Mobility Patterns", www.eurecom.fr/
- [7] Manvi, S.S., Kakkasageri, M.S., Mahapurush, C.V., "Performance Analysis of AODV, DSR, and Swarm Intelligence Routing Protocols In Vehicular Ad hoc Network Environment" In International conference on future Computer and Communication., pp. 21-25, April. 2009.
- [8] Noor Azlan Ahmad, Shamala K. Subramaniam, Jalil Md. Desa, "Increasing Packet Delivery in Ad Hoc On-Demand Distance Vector (AODV)Routing Protocol", Proceedings of the International Conference on Computer and Communication Engineering 2008 May 13-15, 2008 Kuala Lumpur, Malaysia.
- [9] I.Vijaya, PinakBhusan Mishra, "Influence of Routing Protocols in Performance of Wireless Mobile Adhoc Network", 2011 Second International Conference on Emerging Applications of Information Technology
- [10] Carnegie Mellon, Computer Science Department. 15-744 Spring 2007 "Problem Set 2".
- [11] Kevin Fall and KannanVaradhan, "The ns Manual (formerly ns notes and documentation)".
- [12] Krishna Gorantala, "Routing in Mobile Ad-hoc Network", Umea University, Sweden, June-2006. "PDCA12-70 data sheet," Opto Speed SA, Mezzovico, Switzerland.
- [13] C.Siva Ram Murthy and B.S.Manoj, Adhoc wireless networks-Architecture and protocols, Pearson publications.

BIOGRAPHY



Padmashree S. pursuing final year M.Tech in Digital Electronics and Communication, completed B.E in Telecommunication Engineering in 2008. Research interests are in cellular networks, ad hoc networks, digital signal processing, and communication technology. Paper presentations have been given at National Conferences in the field of mobile ad hoc networks.



Mr. P. B. Manoj received his Masters degree in Information technology from Visvesvaraya Technological University, Belgaum, Karnataka, India. His research interests include mobile wireless networks, sensor networks, parallel and clustering computing, and performance modeling and evaluation. He has authored/co-authored several technical papers in the areas of computer networking, performance evaluation, and parallel and distributed computing. He is currently an Assistant Professor in the Department of Electronics and Communication Engineering, AMC Engineering college, VTU, Karnataka.

The Implementation OF Prosthetic Index Finger Based On EMG Signals

¹Amanpreet Kaur, ²Gagandeep Kaur

Student, M.Tech Electronics and Communication Section, YCOE, Talwandi Sabo, Guru Kashi
Punjabi University Campus, Patiala
postal address- VPO Jhallian Kalan Distt. Ropar Pin code 140001

^bAsst Prof., Electronics and Communication Section, YCOE, Talwandi Sabo, Guru Kashi
Punjabi University Campus, Patiala

Abstract

The main goal of this paper is to provide an integrated design of the artificial index finger and to present the result of human like behavior. The big advantages of the prosthetic index finger are their incredible small size, volume and weight, their low cost .This paper describes our implementation of one finger of a future biomechatronic hand and remote control .This index finger captures the muscle activity like human index finger. The purpose of this paper is to illustrate the methodology for EMG signal analysis to provide efficient and effective ways of understanding the signal and its nature.

Keywords: Index Terms-EMG, Hilbert Transform, Neural Network.

1. Introduction

Small electrical currents are generated by muscle fibres prior to the production of muscle force. These currents are generated by the exchange of ions across muscle fibre membranes, a part of the signaling process for the muscle fibres to contract. The signal called the Electromyogram (EMG) can be measured by applying conductive elements or electrodes to the skin surface, or invasively within the muscle. Surface EMG is the more common method of measurement, since it is non-invasive and can be conducted by personnel other than Medical Doctors, with minimal risk to the subject. Measurement of surface EMG is dependent on a number of factors and the amplitude of the surface EMG signal (sEMG) varies from the μV to the low mV range.

There are many applications for the use of EMG. EMG is used clinically for the diagnosis of neurological and neuromuscular problems. It is used diagnostically by gait laboratories and by clinicians trained in the use of biofeedback or ergonomic assessment. EMG is also used in many types of research laboratories, including those involved in biomechanics, motor control, neuromuscular physiology, movement disorders, postural control, and physical therapy.

In this paper, we propose and develop a new prosthetic index finger based on the EMG signals. The proposed system uses EMG signals to realize a control of the prosthetic index finger.



Fig.1.Prosthetic index finger

2. The prosthetic index finger

This prosthetic index finger used in **the bionic man** machine control system and remote control is shown in fig. 1.The prosthetic index finger is almost same size as an adult's index finger .This prosthetic index finger is driven by the stepper motor .This prosthetic index finger adapts the same shape when the human finger open and close. In the base joint of index finger, we use one electrode .

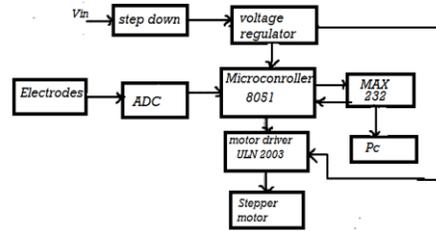


fig.2 The system of the control circuit

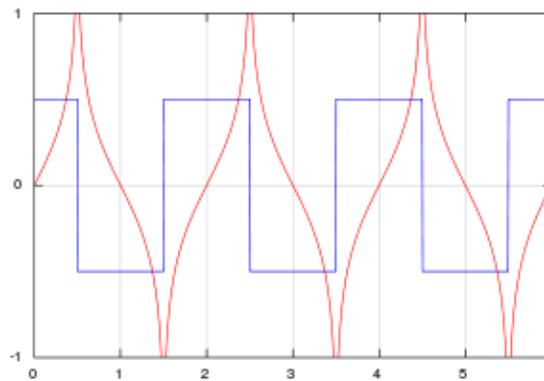
3. Control circuit system

The control circuit system of the prosthetic index finger is presented in fig.2. Microcontroller 8051 is used, which will drive the prosthetic index finger using the stepper motor. The communication with PC is realized by max232. Emg signals from electrode communicate with motor controller through 8051 microcontroller. The complete control circuit board implement on the PCB .EMG signals are used as the control signals to control the prosthetic index finger. These signals are measured from the operator’s forearm muscles when the operators contract their muscles to control their finger motion.

4. Methodology

In mathematics and in signal processing, the Hilbert transform is a linear operator which takes a function, $u(t)$, and produces a function, $H(u)(t)$, with the same domain. The Hilbert transform is named after David Hilbert, who first introduced the operator in order to solve a special case of the Riemann–Hilbert problem for holomorphic functions. It is a basic tool in Fourier analysis, and provides a concrete means for realizing the harmonic conjugate of a given function or Fourier series. Furthermore, in harmonic analysis, it is an example of a singular integral operator, and of a Fourier multiplier. The Hilbert transform is also important in the field of signal processing where it is used to derive the analytic representation of a signal $u(t)$.

The Hilbert transform was originally defined for periodic functions, or equivalently for functions on the circle, in which case it is given by convolution with the Hilbert kernel. More commonly, however, the Hilbert transform refers to a convolution with the Cauchy kernel, for functions defined on the real line \mathbb{R} (the boundary of the upper half-plane). The Hilbert transform is closely related to the Paley–Wiener theorem, another result relating holomorphic functions in the upper half-plane and Fourier transforms of functions on the real line.



In this fig. The Hilbert transform, in red, of a square wave, in blue.

The Hilbert transform of u can be thought of as the convolution of $u(t)$ with the function $h(t) = 1/(\pi t)$. Because $h(t)$ is not integrable the integrals defining the convolution do not converge. Instead, the Hilbert transform is defined using the Cauchy principal value (denoted here by p.v.) Explicitly, the Hilbert transform of a function (or signal) $u(t)$

is given by
$$H(u)(t) = \text{p.v.} \int_{-\infty}^{\infty} u(\tau)h(t-\tau) d\tau = \frac{1}{\pi} \text{p.v.} \int_{-\infty}^{\infty} \frac{u(\tau)}{t-\tau} d\tau,$$

provided this integral exists as a principal value. This is precisely the convolution of u with the tempered distribution $\text{p.v.} 1/\pi t$ (due to Schwartz (1950); see Pandey (1996, Chapter 3)). Alternatively, by changing variables, the principal value integral can be written explicitly (Zygmund 1968, §XVI.1) as

$$H(u)(t) = -\frac{1}{\pi} \lim_{\varepsilon \downarrow 0} \int_{\varepsilon}^{\infty} \frac{u(t+\tau) - u(t-\tau)}{\tau} d\tau.$$

When the Hilbert transform is applied twice in succession to a function u , the result is negative u :

$$H(H(u))(t) = -u(t),$$

provided the integrals defining both iterations converge in a suitable sense. In particular, the inverse transform is $-H$. This fact can most easily be seen by considering the effect of the Hilbert transform on the Fourier transform of $u(t)$ (see Relationship with the Fourier transform, below).

For an analytic function in upper half-plane the Hilbert transform describes the relationship between the real part and the imaginary part of the boundary values. That is, if $f(z)$ is analytic in the plane $\text{Im } z > 0$ and $u(t) = \text{Re } f(t + 0 \cdot i)$ then $\text{Im } f(t + 0 \cdot i) = H(u)(t)$ up to an additive constant, provided this Hilbert transform exists.

5. Conclusions

EMG signal carries valuable information regarding the nerve system. So the aim of this paper was to give brief information about EMG and reveal the methodology to analyze the signal. This study clearly points up the various types of EMG signal analysis techniques so that right methods can be applied during any clinical diagnosis, biomedical research, hardware implementations and end user applications. A prosthetic finger system based on the EMG signals is proposed and developed in this paper. In order to achieve simplicity, lightweight, dexterity underactuated self adaptive theory is adopted to decrease the weight. The feature of the control system is that it uses a VLR based Hilbert transform two EMG patterns. We use this prosthetic finger for remote control and for those people who has no index finger. The experimental results showed that using Hilbert transform has high recognition ability even for several samples of each motion.

References

- [1]. Basmajian JV, De Luca CJ (1985) Muscles Alive. Their Function Revealed by Electromyography. Williams & Wilkens, Baltimore.
- [2]. A Five-fingered Underactuated Prosthetic Hand System*
Jingdong Zhao, Li Jiang, Shicai Shi, Hegao Cai Hong Liu, Gerd Hirzinger
Robotics Institute of Harbin Institute of Technology Institute of Robotics and mechatronics
Harbin Institute of Technology German Aerospace Center, DLR
Harbin, 150001 Heilongjiang, P.R.China DLR, 82230 Wessling, Germany
- [3]. Techniques of EMG signal analysis: detection, processing, classification and applications
M.B.I. Raez,¹ M.S. Hussain,¹ and F. Mohd-Yasin¹ ¹Faculty of Engineering, Multimedia University. 63100 Cyberjaya, Selangor. Malaysia. Corresponding author.M.B.I. Raez, Faculty of Engineering, Multimedia University. 63100 Cyberjaya, Selangor. Malaysia. Email: mamun.raez@mmu.edu.my Received October 4, 2005; Revised January 9, 2006; Accepted January 18, 2006.
- [4]. <http://en.Wikipedia.org/wiki/Electromyography,2010>
- [5]. edge.rit.edu/edge/P08027/public/IRB/Papers/intro_EMG.pdf

Design of an Intelligent SMS Based Remote Metering System for AC Power Distribution to HT and EHT Consumers.

**¹Mrs. Mahalakshmi N, ² Mr. Krishnaiah Paramesh,
³Ms. Elavarasi E,**

¹Department of ECE,
AMC Engineering College, Bangalore, Karnataka,
INDIA - 560 083.

²Asst Prof., Department of ECE, AMC Engineering College, Bangalore, Karnataka,
INDIA - 560 083.

³Executive Engineer,
KPTCL, Bangalore, Karnataka,
INDIA - 560 001.

Abstract

Electrical distribution utilities are facing problems due to high energy losses that amount to 8% of the total generation. 4% of the losses are unaccounted. The problem is mainly associated with sub-transmission and distribution networks. This paper presents a unique solution to devise a unit which is tamperproof, cost effective, fast, accurate and remote metering at any level of the distribution system. The system helps to access accurate and sufficient data from metering devices to measure the electrical parameters, eliminating the use of energy meters and human intervention. The software application provides a real time parameter like line voltage, line current, power factor, true power, apparent power and reactive power from remote substations. Remote metering system is implemented using microcontroller based mixed signal circuitry.

Keywords: Global System for Mobile communication, General Packet Radio Service, Remote Metering, Short Message Service.

1. Introduction

The conventional billing system for electricity, gas, or water is that an assigned person visits each house and read the meter readings manually. Then the collected meter readings are used for bill calculation. This manual process can become very time consuming and tiresome. It can cause human error and can open an opportunity for corruption done by the human meter reader. Thus the billing system can become inaccurate and inefficient. The recent advances in the field of information technology have made the exchange of information fast, secured and accurate.

The digital revolution caused the rapid drop of digital devices such as computers and telecommunication devices. Communication networks like the internet, GSM networks, etc. are available almost all the countries of the world. In the work presented here, a technique has been developed to read electricity meter readings from a remote server automatically using the existing GSM networks for cellular phones. This technique can be applied for gas or water meters as well. The meters send the meter readings like kilo-watt-hour (kWh), voltage, current, bill, etc. by SMS to a central server. The SMS based data collection can be done very quickly and efficiently. Data can be collected after any desired time interval such as hourly, daily, weekly, or monthly basis. As there is no human intervention in the entire process, there is no chance of human error and corruption. Also, unwanted weather conditions like heavy snow, rain, storm, etc. will not hamper on collecting data as long as the GSM networks are stable. By applying complex encryption algorithms on the data SMS, data security can be ensured.

Traditional meter reading, which depends on manual reading on site at a fixed date, has many deficiencies in real-time and accuracy. GPRS automatic meter reading system is a gradually developed intelligent system combined with the power meter reading characteristics, it can collect the real time data of industrial and residential electricity, and then sends the data to the remote management through GSM/GPRS network. By the system, it can not only save the human resources, but also can supply accurate and reliable meter reading data for the power enterprises to do load prediction and control, so that the overall economic efficiency of the power industries can be improved and the management level progressed.

2. Materials & Methods

2.1 Working Principle

The working principle can be explained as follows. The system helps utilities to access accurate and sufficient data from metering devices installed at a remote substation. The software application provides real time parameters like line voltage, line current, power factor, Apparent Power (VI), Reactive Power (VI-R), True Power (VI-A) from a remote substation and the same is transmitted to a centralized station. The centralized station receives data to be processed from the distribution system through a GPRS link. This module collects data from a network of metering devices installed at incoming and outgoing feeders within the substation or from the feeders at different locations.

2.2 Functional Block Diagram

The proposed hardware is realized with various building blocks based on functional aspects. The functional block diagram with all the blocks is depicted in Fig 2.1.

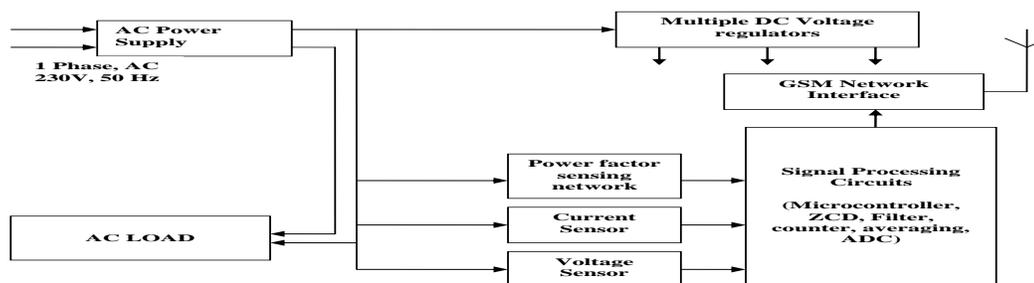


Fig. 2.1: Functional Block Diagram.

3. Hardware Design & Development

3.1 The Power Supply

The power supply unit is designed to deliver $\pm 12V @ 500mA$, $+5V @ 1A$ and $+3.8V @ 150mA$ to power up the whole electronics. The power supply is realized using fixed voltage linear regulator ICs 7812, 7912 and 7805 and a variable voltage regulator LM 317 along with appropriate filter capacitors.

The various requirements for the project are:

1. $+12V/-12V 500mA$ for all OPAMPS.
2. $+5V DC$ for Hall Effect Current Transducers.
3. $+3.8V DC$

3.2 The voltage Sensor

An input AC of 230V and current of 5A cannot be measured directly using a microcontroller. The high input voltage needs to be conditioned in a way suitable for the microcontroller. This is because, the microcontroller is a semi conducting device that operates at low voltages. The voltage sensor comprises of a step down transformer that steps down the 230V AC to 3V. This is rectified using a full wave rectifier, filtered using a capacitor and reduced or scaled down to a voltage approximately equal to 2.5v which represents 230v AC. The circuit is designed such that the microcontroller is able to measure a supply voltage of $230v \pm 20\%$ variations in the supply mains.

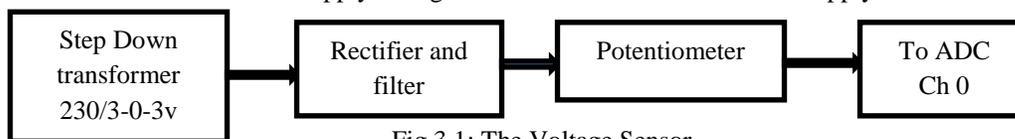


Fig.3.1: The Voltage Sensor.

The Current Sensor

Current sensors are used to sense and measure current variations. They give an output voltage proportional to the current under measurement. The current sensor used is a Hall Effect current transformer which is a transducer.

The Hall Effect is explained by the Lorentz force principle. When a charge moves in a direction perpendicular to an applied magnetic field, it experiences a force defined by the Lorentz Law, "The direction of this force is perpendicular to the direction of propagation of the charge and that of the external magnetic field".

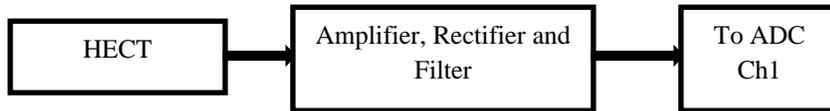


Fig. 3.2: The current Sensor.

Power Factor Sensing Network

Power factor is defined as the Cosine of the phase angle between voltage and current.

Power factor = Cosine ϕ or Cos ϕ

It is also defined as the ratio of true power to apparent power or volt amperes. It is also the ratio of resistance to impedance.

True power in an AC circuit is given by:

$$P = V I \cos \phi$$

V = RMS value of voltage

I = RMS value of Current

Cos ϕ = Power factor

The phase angle between voltage and current is computed by the micro controller by accounting the pulse width of the signal applied as input. The pulse width represents the time difference between the voltage and current signals at the output of the sensors. The AC voltage and current wave forms are clipped by rectifiers and squared through the ZCD. The two diodes used act as a bi directional clipper. These signals are applied to a two input AND gate to obtain the pulse equivalent of the phase angle differences between voltage and current.

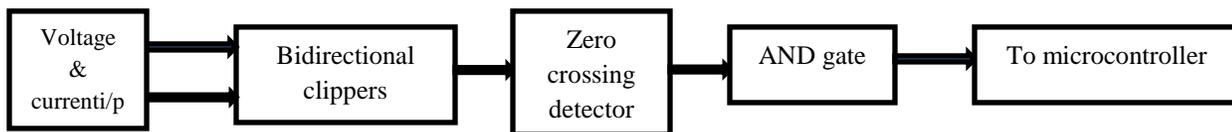


Fig. 3.3: Power Factor Sensing Unit

Microcontroller, ADC and GSM Module Interface Circuit

The analog to digital converter samples the voltage and current inputs at regular intervals and converts them into digital form. The 8-bit output is converted by the microcontroller into actual values.

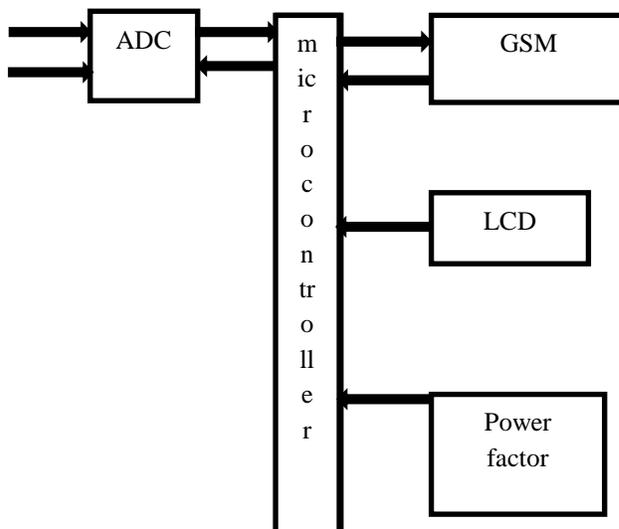


Fig. 3.4: Microcontroller, ADC and GSM Module Interface Circuit

3.4.1 Analog to Digital Converter

Out of 8 channels first two channels of ADC are used for the measurement of Voltage and current signals. The channel is selected by flipping the LSB of the selection line of ADC through the microcontroller.

The analog inputs corresponding to the voltage and current are applied to channel 0 and channel 1 of the ADC. The clock frequency for the ADC operation from the astable multi vibrator is 1.6 kHz. The two analog inputs are selected by the ADC for conversion, one at a time.

The successive approximation frequency of the ADC is set externally with a help of an RC network and IC 7400 as shown which is configured to work as an astable multi vibrator. The clock frequency for the ADC is given by:

3.4.2 The Serial Communication Circuit

In this project mode 1 is used. That is, standard UART mode which is 10 bit full duplex transmits and receive mode.

Transmitted data is sent as a start bit, eight data bits and a stop bit.

BAUD RATE: If standard baud rates are desired, then an 11.0592 megahertz crystal should be selected. To obtain the standard baud rate of 9600 hertz, the setting of TH1 will be as follows.

$$TH1 = 256d - (2/32 \times 11.0592 \times 10^6 / 12 \times 9600d) = 253.000d$$

TH1 = 0FDh

3.4.3 LCD Interface

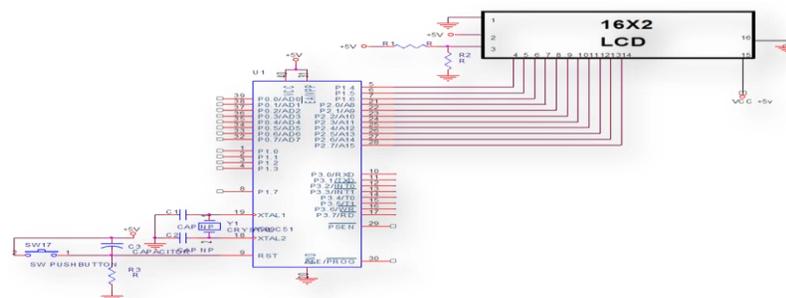


Fig. 3.5: LCD Interface.

To display data on the LCD the following steps are executed:

Initialize the LCD with a set of command words. The command words are sent on the D0–D7 data lines with RS = 0, R/W = 0 and a high-to-low pulse on the enable ‘E’ pin.

The ASCII value of the character to be displayed (say 41H to display ‘A’) is sent on the D0–D7 data lines with RS = 1 (data register), R/W = 0 and a high-to-low pulse on the ‘E’ pin.

4. Software

4.1 Flow Chart

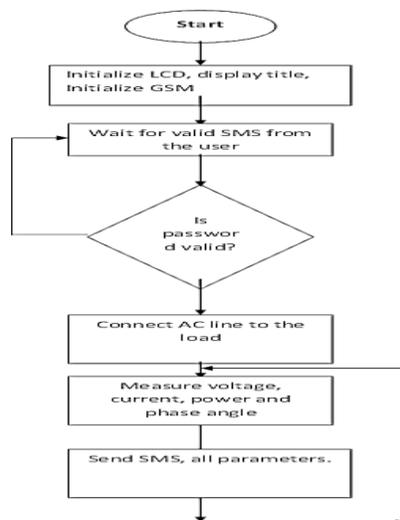


Fig. 4.0: Flow Chart

4.2 Algorithm

1. Initialize LCD and display title of project.
2. Initialize GSM into text mode i.e. "SMS" mode, and wait for password SMS.
3. After receiving password SMS, connect AC 230 line to load.
4. Select current channel and get digital value for current, Calibrate digital value into ampere and display on LCD.
5. Select voltage channel, and get digital value for voltage, Calibrate to voltage and display on LCD.
6. Measure pulse width of phase angle signal and calibrate to angle.
7. Calculate power, active power, reactive power and display on LCD.
8. Repeat steps from 4 to 7 for fixed number of times, and send all details to the base station mobile.

5. Experimental Setup, Test Results and Analysis

A technique for remotely reading the power drawn by the load using short message service is developed and tested. The existing global system for mobile communication networks is used for sending and receiving short messages. The prototype has been designed, developed and experimented. The functioning of the prototype was tested by creating conditions in the laboratory similar to the conditions in which it is expected to operate. The load curve is a peak-valley curve indicating the variation in power demand during different hours of the day. The load is normally resistive between 6.00 P.M. and 9.00 A.M. It becomes complex between 9.00 A.M. and 6.00 P.M. due to industrial loads. These conditions are simulated by subjecting the prototype to resistive and complex loads and the performance is evaluated.

5.1 Test Setup

The circuit schematic of the experimental setup to evaluate the performance of the hardware is as shown in Fig. 5.1. The setup includes a voltmeter, ammeter, the prototype of the remote metering system, resistive, inductive and capacitive loads and a global system for mobile communication network receiver. The voltmeter and ammeter are required to compare the input voltage and current with the output readings of the prototype on the liquid crystal display.

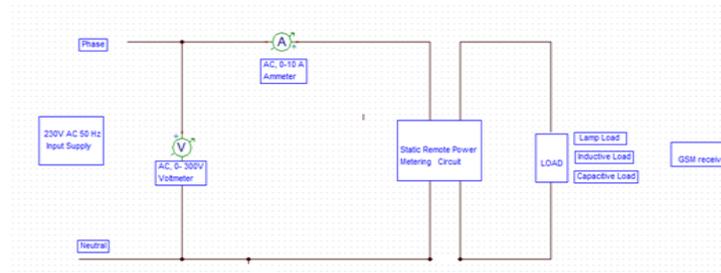


Fig. 5.1: Laboratory Setup to test SMS based Remote Metering System.

A photograph of the experimental setup is shown in Fig. 5.2. It consists of the hardware unit, loads and the display unit.

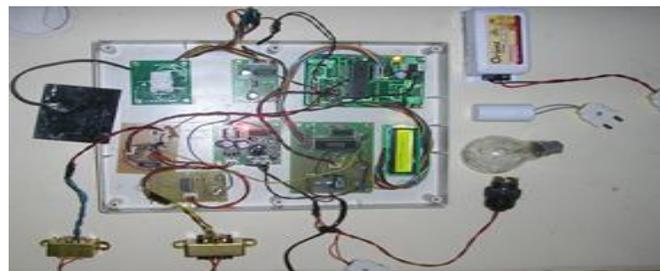


Fig. 5.2: Experimental setup.

Steps involved in testing the performance of the proposed system are:

- The system is initialized and zero reading on the liquid crystal display is ensured.
- A load is connected at the prototype output.
- The magnitudes of voltage and current on the voltmeter and ammeter are compared with those on the display.
- The parameters phase angle, true power, reactive power, apparent power computed by the microcontroller is also available on the display.

- At every 30th measurement as set by the program, an SMS is sent by the microcontroller to the initialized mobile number.
- The results of true power and reactive power are tallied with the power ratings of the load.

Fig. 5.3 shows the display when the GSM was initialized on no load, on a lamp load, sending an SMS to the initialized mobile and displaying the various parameters.

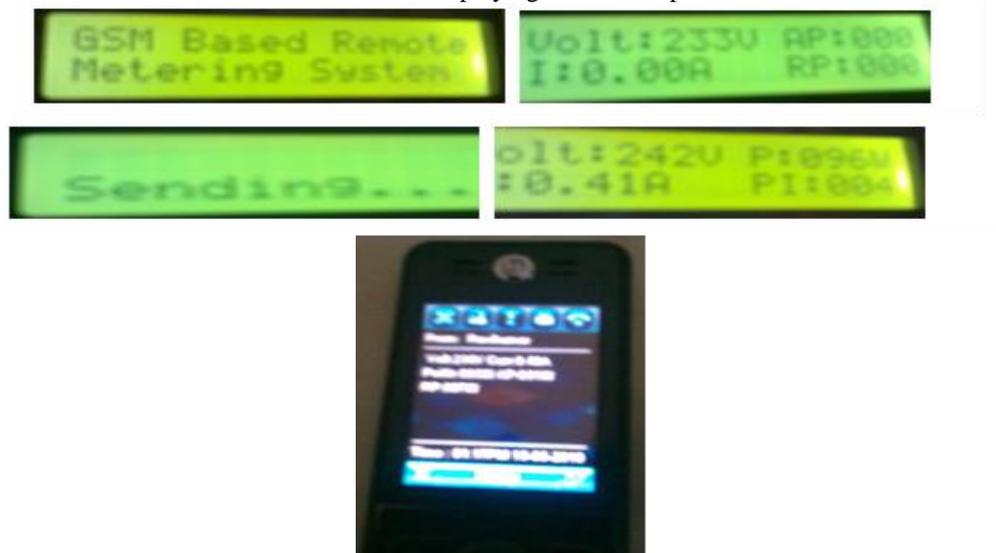


Fig. 5.3: Test Results

5.2 Results and Discussion

The measured parameters and their values for eight case studies conducted at different hours of the day are listed in Tables. It is observed that for a lamp load which is resistive the phase angle difference between voltage and current in all the cases is within 10°. The true power is larger and complies with the rating of the lamp. For complex loads comprising of resistor – inductor (R–L) and resistor – capacitor (R–C), the reactive powers have considerably increased.

Case Study 1

Table 5.1 Measured Parameters and their values.

Sl. No.	Load Type	HECT Output			Voltage sensor DC o/p (V)	LCD & SMS Readings						ZCD AND gate o/p (ms)
		CT o/p (V)	Rectifier o/p (V)	Amplifier o/p (V)		Input voltage (V)	Input Current (A)	Phase Angel (Φ)	P=VIcosΦ (w)	Q=VI sinΦ (VAR)	S=VI (VA)	
1	No Load	V _Q =2.2 (DC)	0	0	5.8	226	0	0°	0	0	0	Noise
2	R Load	0.16	0.5	6.0	5.8	225	0.44	9°	98	16	99	0.5
3	L Load	0.28	0.8	9.0	5.6	232	0.69	76°	38	155	160	4.2
4	RL Load	0.4	1.2	10.0	5.0	236	0.9	22°	196	80	213	1.2
5	RC Load	0.16	0.8	6.4	5.2	226	0.52	72°	36	111	117	4.0

Case study 2

Table 5.2 Measured Parameters and their values.

Sl. No.	Load Type	HECT Output			Voltage sensor DC o/p (V)	LCD & SMS Readings						ZCD AND gate o/p (ms)
		CT o/p (V)	Rectifier o/p (V)	Amplifier o/p (V)		Input voltage (V)	Input Current (A)	Phase Angel (Φ)	P=VIcosΦ (w)	Q=VI sinΦ (VAR)	S=VI (VA)	
1	No Load	V _Q =2.2 (DC)	0	0	5.6	224	0	0°	0	0	0	Noise
2	R Load	0.18	0.5	6.0	5.6	225	0.47	8°	104	15	105	0.45
3	L Load	0.28	0.8	9.0	5.5	230	0.7	72°	50	153	161	4.0
4	RL Load	0.4	1.2	9.8	4.8	235	0.92	22°	200	81	216	1.2
5	RC Load	0.16	0.8	6.2	5.2	228	0.51	70°	40	109	116	3.9

Case study 3

Table 5.3 Measured Parameters and their values.

Sl. No.	Load Type	HECT Output			Voltage sensor DC o/p (V)	LCD & SMS Readings						ZCD AND gate o/p (ms)
		CT o/p (V)	Rectifier o/p (V)	Amplifier o/p (V)		Input voltage (V)	Input Current (A)	Phase Angel (Φ)	P=VIcos Φ (w)	Q=VI sinΦ (VAR)	S=VI (VA)	
1	No Load	V _Q =2.2 (DC)	0	0	5.6	223	0	0°	0	0	0	Noise
2	R Load	0.17	0.5	6.0	5.6	234	0.46	9°	105	17	107	0.5
3	L Load	0.28	0.8	8.9	5.5	232	0.68	76°	38	153	158	4.2
4	RL Load	0.4	1.2	9.8	5.0	237	0.91	24°	197	87.8	216	1.3
5	RC Load	0.16	0.8	6.4	5.2	227	0.52	72°	36	112	118	4.0

Graphs

Graphs of variations in voltage, current, phase angle and true power for the same load in different case studies is as shown in figures 5.4 to 5.15. The tabulated values are tables 5.9 to 5.12. From the graphs it is observed that the variations in the parameters measured using the prototype is due to the fluctuations in supply voltage.

Table 5.4: Variation in measured parameters at every 30th measurement (Resistive Load)

Sl. No.	Measurement No.	Voltage in Volts	Current I (A)	Phase Angle° (f)	Power (W)
1	30	225	0.44	9	98
2	60	225	0.47	8	104
3	90	234	0.46	9	105
4	120	226	0.49	8	109
5	150	232	0.44	7	101
6	180	228	0.45	8	101
7	210	230	0.46	8	104
8	240	229	0.45	9	102

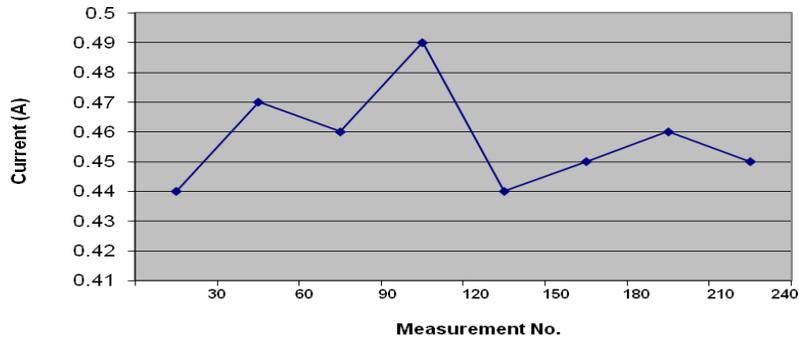


Fig. 5.4: Graph of Variations in Current at every 30th measurement.

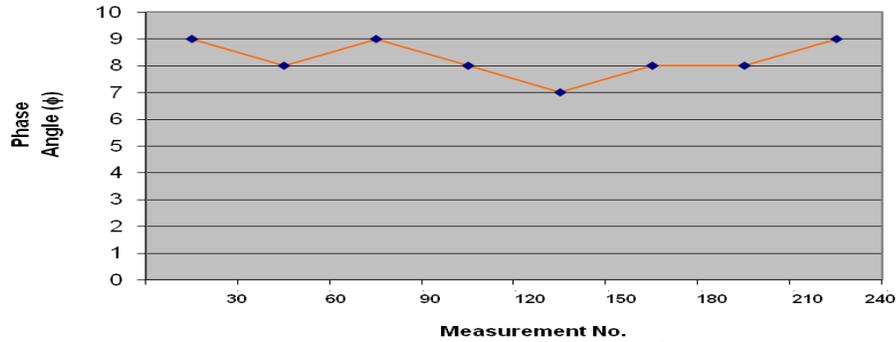


Fig. 5.5: Graph of Variations in Phase Angle at every 30th measurement.

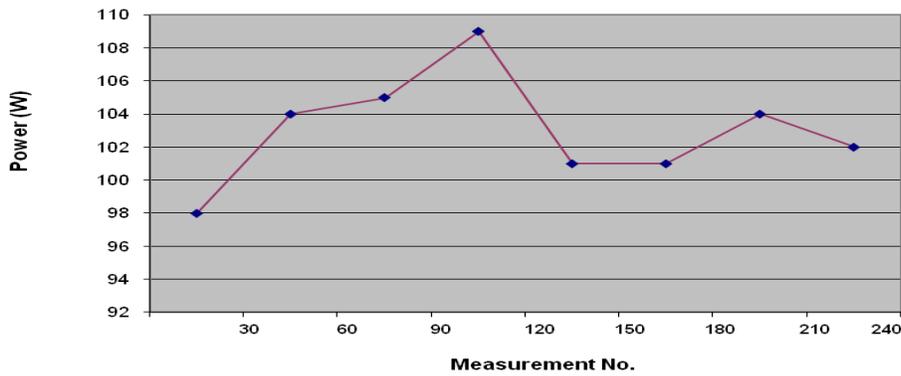


Fig. 5.6: Graph of Variations in Power at every 30th measurement.

Table 5.5: Variation in measured parameters at every 30th measurement (Inductive Load)

Sl. No.	Count No. (Every 30 th Measurement)	Voltage (V)	Current I (A)	Phase Angle ^o (φ)	Power (W)
1	30	232	0.69	76	38
2	60	230	0.7	72	50
3	90	232	0.68	76	38
4	120	232	0.7	72	50
5	150	228	0.71	76	39
6	180	226	0.7	72	49
7	210	232	0.72	72	51
8	240	230	0.68	76	38

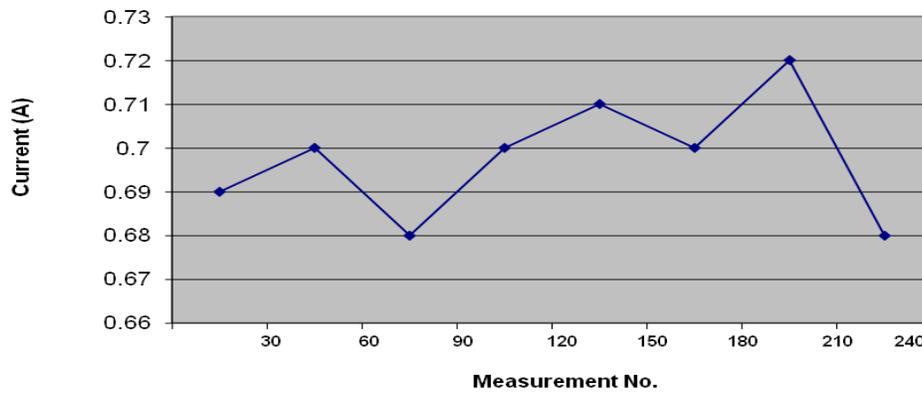


Fig. 5.7: Graph of Variations in Current at every 30th measurement.

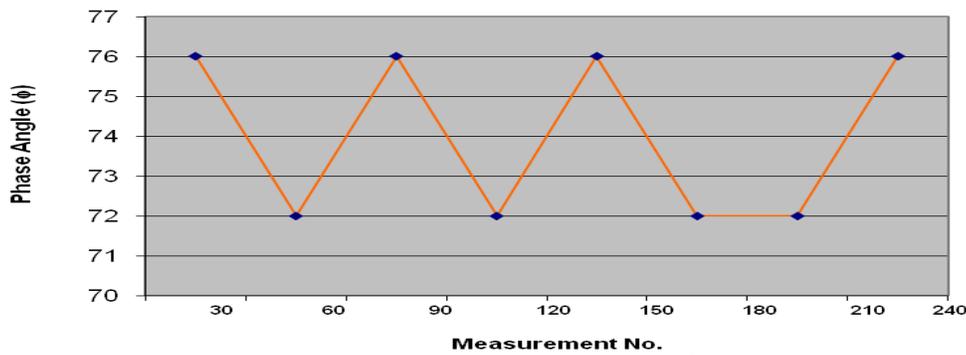


Fig. 5.8: Graph of Variations in phase angle at every 30th measurement.

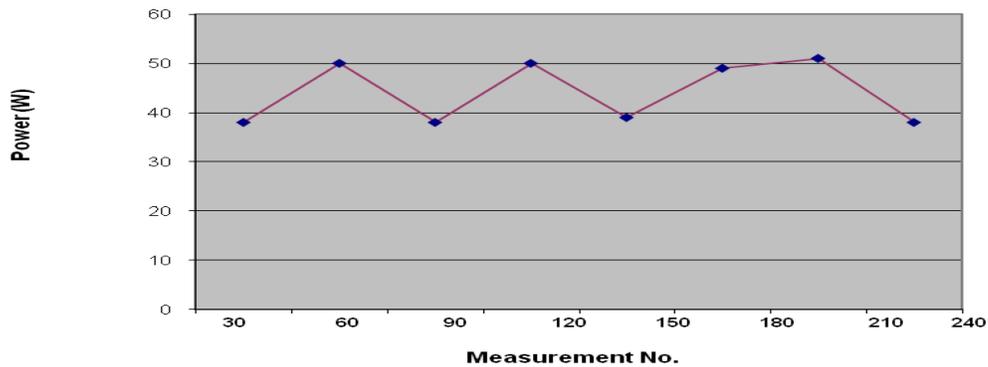


Fig. 5.9: Graph of Variations in Power at every 30th measurement.

Table 5.6: Variation in measured parameters at every 30th measurement (Resistive Inductive Load)

Sl. No.	Count No. (Every 30 th Measurement)	Voltage (V)	Current I (A)	Phase Angle ^o (f)	Power (W)
1	30	236	0.9	22	196
2	60	235	0.92	22	200
3	90	237	0.91	24	197
4	120	238	0.89	25	192
5	150	238	0.86	24	186
6	180	236	0.84	25	179
7	210	232	0.89	22	191
8	240	235	0.9	24	193

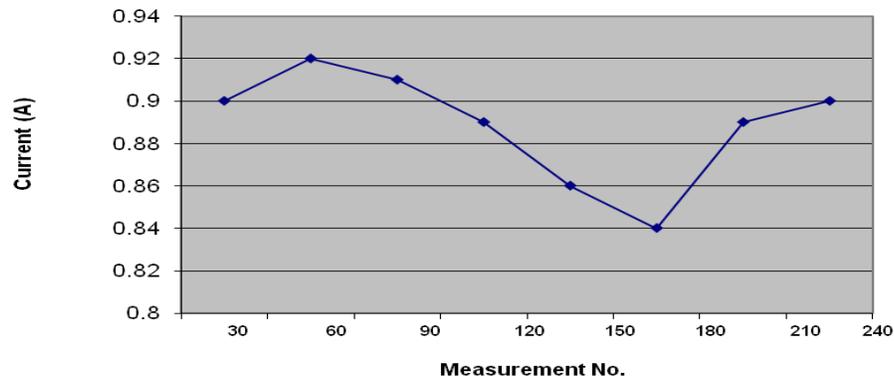


Fig. 5.10: Graph of Variations in current at every 30th measurement.

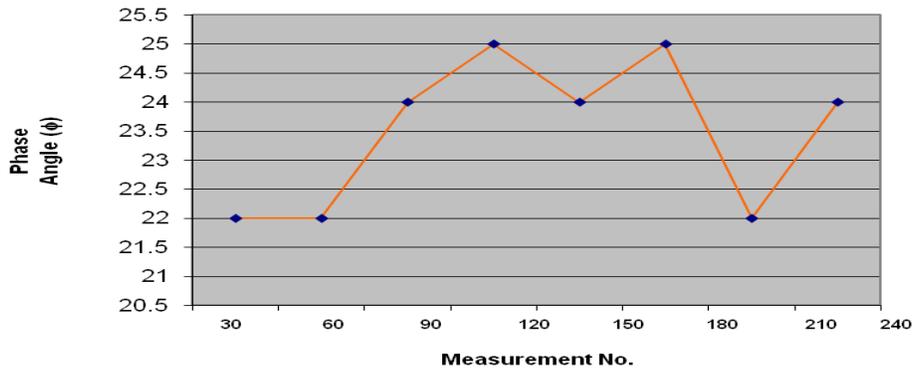


Fig. 5.11: Graph of Variations in phase angle at every 30th measurement.

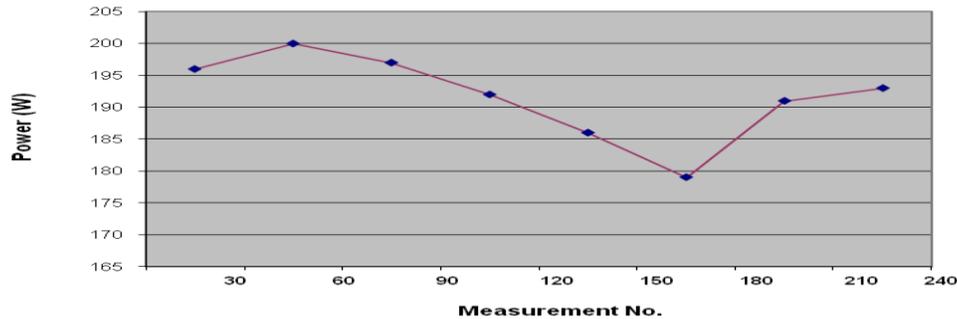


Fig. 5.12: Graph of Variations in Power at every 30th measurement.

. The experimental values and the time of processing data indicates that the present work has improved the overall performance of the system. This result is different as there is no interaction with any conventional energy meter to record the power consumed by the consumer. The experiment has clearly suggested that there is no need to have a separate energy meter for recording energy. It means that the present hardware itself acts as energy meter.

6. Conclusion and Future Scope

6.1 Conclusion

In this work, an SMS based remote metering technique for the energy distribution sector has been proposed. The developed prototype has been tested over the GPRS link. The proposed solution is used to read the voltage and current remotely and the same is verified with those of physically measured values while experimenting. The results computed using remote metering system were accurate, fast, and has avoided human intervention. This technique can well be utilized at various levels of the power generation and distribution sector. The time taken to collect one set of data and compute the power consumed by the client is about 4.5ms. This needs to be added with the network time delay for the completion of the total measurement and computation via SMS. The system can be extended to compute the energy bill for the consumers and the same can be sent through SMS.

Current design and development of the networked digital metering system for AC power distribution can be used in power stations and substations wherein the power generated and its usage can be tracked. Similarly the system can be used in substation where the power consumed and metered can be tracked for high tension and low tension loads.

6.2 Future Scope

Future scope of the work can be as follows:

- The system can be implemented at any stage in the power system right from generation to distribution with slight modifications.
- It is possible to include billing if implemented on the consumer end.
- Details pertaining to the energy consumption can be sent through SMS to the consumer without human intervention.

Acknowledgement

I express my deep respects and profound gratitude to my guide and constant source of inspiration Ms Elavarasi E, Asst. professor, whose support and motivation encouraged me. I thank my guide at KPTCL, Mr. Krishnaiah Paramesha, Executive Engineer, HRD center, Bangalore for all the guidance and support extended during the project.

I take this opportunity to thank the Principal, AMC Engineering College, Bangalore, Dr. T. N. Sreenivasa and the Head of the Department, Electronics & communication Engineering, Dr. Bhuvaneshwari. H. B for providing me an opportunity to undertake the project at KPTCL, Bangalore.

I thank Mrs. Jayashree B, Executive Engineer, KPTCL, HRD Centre, Bangalore.

I thank Mrs. Lakshmi Assistant Executive Engineer, BESCO, Bangalore.

My heartfelt thanks to my family, Rajajinagar, Bangalore-10, for all the support and encouragement.

I am highly indebted to my family for all the constant cooperation.

References:

Books:

1. "The 8051 Microcontroller", Smt. Uma Rao. K and Dr. AndhePallavi, 2009.
2. "The 8051 Microcontroller", Kenneth J. Ayala, Publication PenramInternational ,2007.

Journals:

1. Jaswinder Singh, Ravinder Singh, Abhishek Gaur, YashwantShrimali, Deepak Singh, Vaibhav Gupta. "GIS Based Automatic Meter Reading System for Substation", 2009, NI Lab view meet India
2. ArunSehgalltron Inc., Spokane, Washington, Member IEEE "AMR offers Multiple benefits" published in Pipeline & Gas Technology, April/May 2005 Page nos. 01-05.

About the Author



Mrs. Mahalakshmi N

Mrs. Mahalakshmi N. was born in Bangalore, Karnataka, India on 21 April 1986. She graduated in Electronics & Communication Engineering from R. L. Jalappa Institute of technology, Visveswaraya Technology University, Belguam, Karnataka, India in 2008. At present she is pursuing her M. Tech in Digital Electronics & Communication from AMC college of Engineering, Bangalore. She served as Lecturer, for R. L. Jalappa Institute of technology for two years. And she was a co-ordinator for 'xibit' a national level technical symposium. She has two years of experience in teaching in Electronics and Electrical departments. She has a flair for teaching management subjects.

A study on strength properties of roller compacted concrete with 30% replacement of OPC 53 grade cement by flyash.

Ganapati Naidu. P

M.E (2009-2011)

Department of Civil Engineering
Andhra University, Visakhapatnam

S.Adishes

Associate Professor

Department of Civil Engineering
Andhra University, Visakhapatnam

B. Jagannadh Kumar

M.E (2009-2011)

Department of Civil Engineering
Andhra University, Visakhapatnam

P.V.V.Satayanarayana

Professor

Department of Civil Engineering
Andhra University, Visakhapatnam

Abstract

Roller compacted concrete is a zero-slump concrete, low water content, dense mix consisting of coarse aggregate, sand, cementitious materials, and water. By using conventional vibrators, it is very difficult to compact for larger thicknesses. There is a chance of getting honey combing and inconsistency with respect to laboratory values. To rectify this, roller compacted concrete technique is proposed by preparation of samples. In this an attempt is made to prepare M15 and M20 mixes at their optimum moisture content and tested for compressive and split tensile strengths for various time periods (i.e., 3days, 7days, 14 days and 28 days). From these results, it is observed that higher strengths at early periods are obtained. When cylinders were tested, the compressive strength for M15 and M20 mixes at 28 days are 25Mpa and 31Mpa respectively. Similarly, for M15 and M20 mixes at 28 days, the split tensile strengths are 1.7Mpa and 1.9Mpa respectively. Since high strengths are obtained in compression and tension, it can be used as base course and sub-base course for flexible and rigid pavements.

Keywords: Roller compacted concrete, Compressive strength, Split tensile strength, Grade of concrete and water cement ratio

Introduction

Roller compacted concrete is a zero-slump concrete consisting of dense-graded aggregate and sand, cementations' materials, and water. Because it contains a relatively small amount of water, it cannot be placed by the same methods used for conventional (slump) concrete. Roller compacted concrete has the same basic ingredient as conventional concrete: cement, water, and aggregates. The basic difference is that roller compacted concrete is a much drier mix with practically zero slumps. It is drier, and looks and feels like damp gravel. It does not require any forms, dowels, reinforcing steel & finishing. Also, the method of compaction is different than the conventional compacted concrete and it is compacted by vibratory or pneumatic-tired rollers. The objective of mix design is to produce a roller compacted concrete mix that has sufficient paste volume to coat the aggregates in the mix and to fill in the voids between them. Any of the basic roller compacted concrete proportioning methods like those based on concrete consistency testing, the solid suspension model, the optimal paste volume method, and soil compaction testing may be used for mix design. Roller-Compacted Concrete uses aggregate sizes often found in conventional concrete. However, the blending of aggregates will be different than that done in case of conventional concrete. Crushed aggregates are preferable in roller compacted concrete mixes due to the sharp interlocking edges of the particles, which help to reduce segregation, provide higher strengths, and better aggregate interlock at joints and cracks. The use of roller compacted concrete for pavements at industrial facilities such as port and intermodal container terminals is particularly appropriate because of the ability to construct low-cost concrete pavements over large areas. Roller compacted concrete is also used in Bulk material storage, General cargo storage, Container terminals, Road / rail transfer facilities, Truck parks, Tank roads and parking, Sewage sludge stacking, Composting slabs and Pre-casting yards.

Objective and scope

To carryout soil compaction procedure for M15 and M20 grades for find out optimum moisture content. Casting of cubes and cylinders of M15 and M20 grades with 30% fly ash replacement to know compressive and split tensile strengths.

Materials and Methods

Cement

Ordinary Portland Cement of 53 Grade conforming to Bureau of Indian Standards (I.S. 12269: 1987) was used in the present investigation. The cement was tested for various properties as per I.S. 4031: 1981.

Test name		Result
Consistency		30%
Specific gravity		3.12
Setting time	Initial	45 minutes
	final	375 minutes

Flyash

Flyash used in this experimental work was obtained from National Thermal Power Corporation (NTPC), Visakhapatnam. Flyash is finely grained residue resulting from the combustion of ground or powdered coal. Mean particle size is about 0.1 to 0.2 μm and finer than cement and consist mainly of glassy spherical particles as well as residues of hematite and magnetite. The specific gravity, specific surface area and density of flyash are 2.84, 310 m^2/kg and 1.4 kg/m^3 respectively.

Fine Aggregate

Locally available river sand passing through 4.75 mm sieve and retained on 75 micron sieve were used.

Coarse aggregate

Crushed blue granite of coarse aggregate passing through 24mm sieve and retained on 20mm sieve were used.

Water

Potable water was used for the concrete. The pH value was 7.25. The same water was used for mixing and curing of concrete cubes and cylinder.

Experimental Procedure

Mix design

Mix proportion of 1:2:4 and 1:1.5:3 were assumed on nominal basis for M15 and M20 grades respectively.

Testing methodology

Optimum moisture content by Modified Proctors test for both the grades (M15 and M20) of concrete, 2% of water (i.e. 140ml for M15 and 110ml for M20) is added to its weight of mix proportion and mixed thoroughly in a basin. Empty weight of CBR mould (Height 17.5cm and Diameter 15cm) is taken. Now the mould is placed on proctor test apparatus and filled with the concrete in five layers. After placing of each layer, 56 blows were applied on mix with a free drop of 4.89 kg rammer from a height of 450mm and then the top surface leveled by trowel. Now the weight of the mould with mix is determined. Density of sample is determined by the ratio of compacted mix weight by volume. By this dry density of sample is calculated. This procedure was repeated by increasing the water percentage by 2% up to 10%. A graph is drawn between the moisture content verses dry density. By this optimum moisture content (OMC) and maximum dry density are determined. Using this optimum moisture content cubes and cylinders are prepared.

Slump cone test

The cement and fine aggregate are mixed thoroughly until uniform color on a rigid and non absorbent surface. Then coarse aggregate is added and mixed it thoroughly and then water is added and mixed it until the concrete appeared to be homogenous. Slump cone mould is thoroughly cleaned and oil is applied inside of it. Now the mixed sample is placed in three layers. After placing of every layer in mould, 25 blows were applied on it by using tamping rod taking care to distribute the strokes evenly over the cross section. After complete filling of three layers, the top surface is leveled with trowel. Now slowly lift the slump cone mould vertically without disturbance of placed sample. Height of the sample specimen is noted. The difference between the initial and final height of the sample gives the slump value in mm.

Casting

Cube moulds of 150mm in size were cast in conventional method to know the compressive strengths of grade of concrete. The cylindrical moulds of diameter 150mm X height 300mm were cast to know the split tensile strength test. These moulds were cleaned and applied oil inside surface. Now the moulds were placed on vibration machine and filled with green concrete in five layers by giving vibration to each layer for a period of 2 minutes.

Curing:

The test specimens are stored in moist air for twenty four hours and then specimens are de-mould and submerged in clear fresh water for 28 days.

Procedure (testing):

Specimens are removed from water after curing, wiped out the excess water and the surface is dried. The specimens are placed in CTM and then tested for compressive strength as per IS: 516-1959 and the results were tabulated.

Methodology and Experimental Investigation

M15 Grade concrete cubes of 150mm were used to determine the compressive strength. In total 24 cubes were cast, among which 12 cubes were without flyash and the rest were with flyash as replacement of cement by 30%, and cured in water for a period of 28 days. The compressive strength was carried out after 3, 7, 14 and 28days.

M15 Grade concrete cylinders of diameter 150mm X height 300mm of 48 numbers were cast to study compressive strength and split tensile strength. Out of 48 cylinders 24 numbers were without flyash and the remaining was with flyash. The cylinders of 24 numbers were divided in to 2 sets, as 12 in each set. 1st set is for compressive strength test and the 2nd set is for split tensile strength test. These specimens were tested under CTM as per IS 516: 1959 and strengths were calculated for 3, 7, 14 and 28 days, and the results were tabled.

Similar procedure was carried out for the M20 grade concrete also and the results were tabled.

Mix notation	Shape of specimen	3 day Strength in MPa	7 day Strength in MPa	14 day Strength in MPa	28 day Strength in MPa
M15 C	Cube	15.22	22.44	25.11	30.88
M15 C	Cylinder	10.11	18.22	20.55	24.66
M15 S	Cylinder	1.10	1.20	1.50	1.60
M15 FR30C	Cube	14.55	22.88	25.66	31.44
M15 FR30C	Cylinder	12.33	18.77	21.88	25.56
M15 FR30S	Cylinder	1.10	1.30	1.50	1.70
M20 C	Cube	18.66	24.55	32.00	36.22
M20 C	Cylinder	15.22	20.11	24.99	29.33
M20 S	Cylinder	1.20	1.50	1.60	1.9
M20 FR30C	Cube	18.00	24.22	33.44	37.11
M20 FR30C	Cylinder	15.66	20.66	26.22	31.33
M20 FR30S	Cylinder	1.20	1.50	1.70	1.90

*C: Compressive strength; S: Split tensile strength; FR30: Flyash replacement by 30%

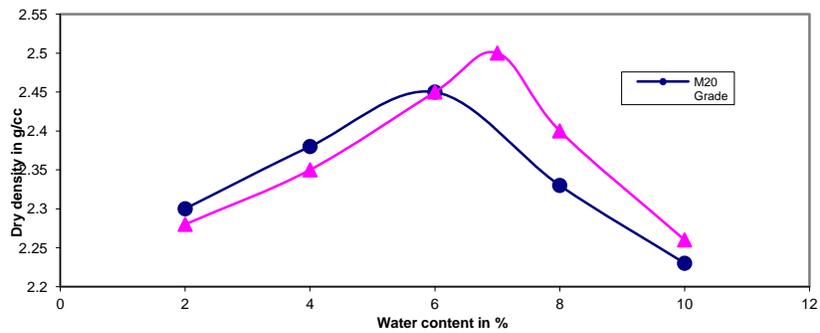
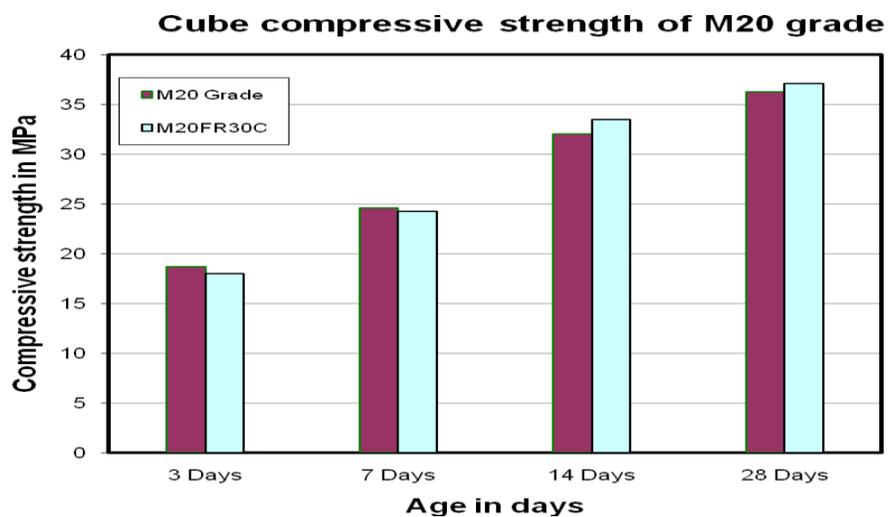
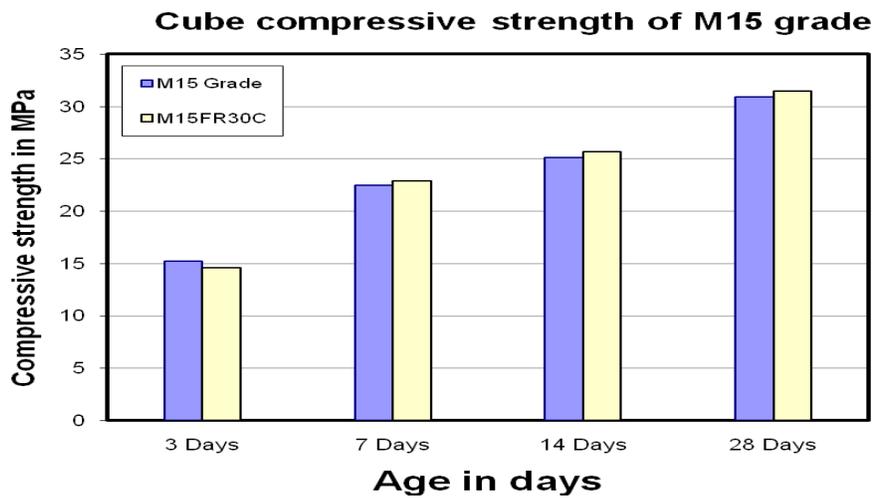
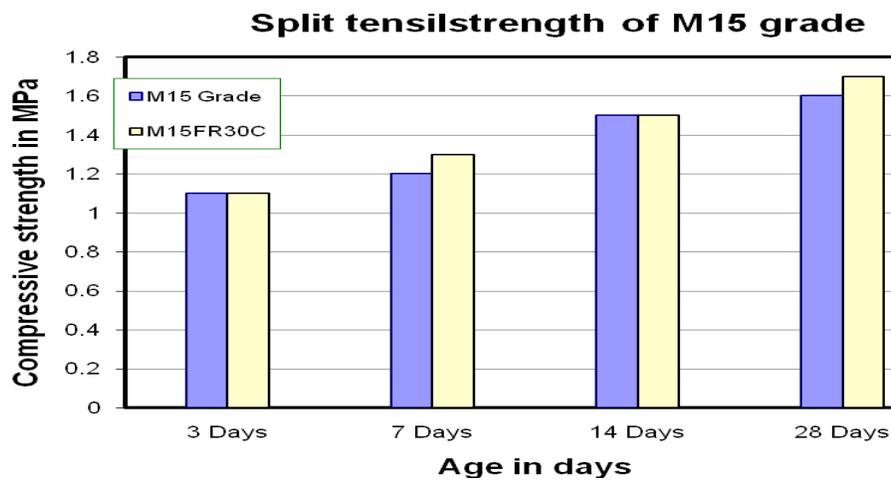
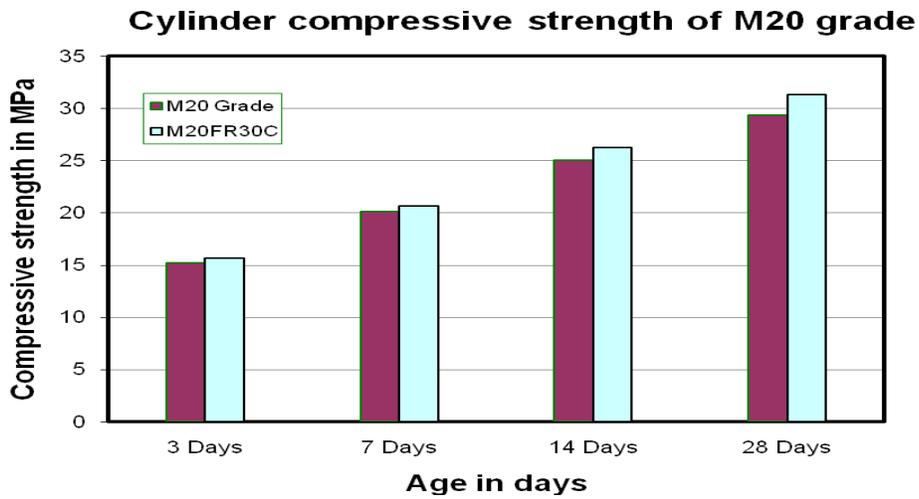
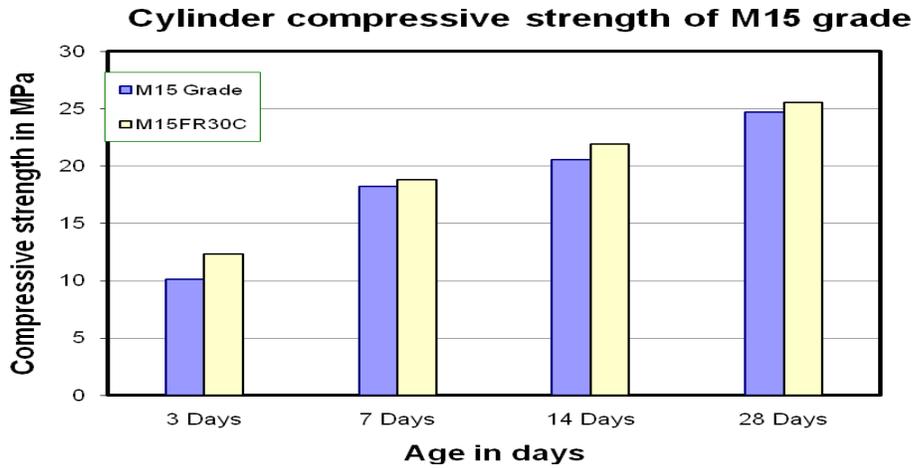
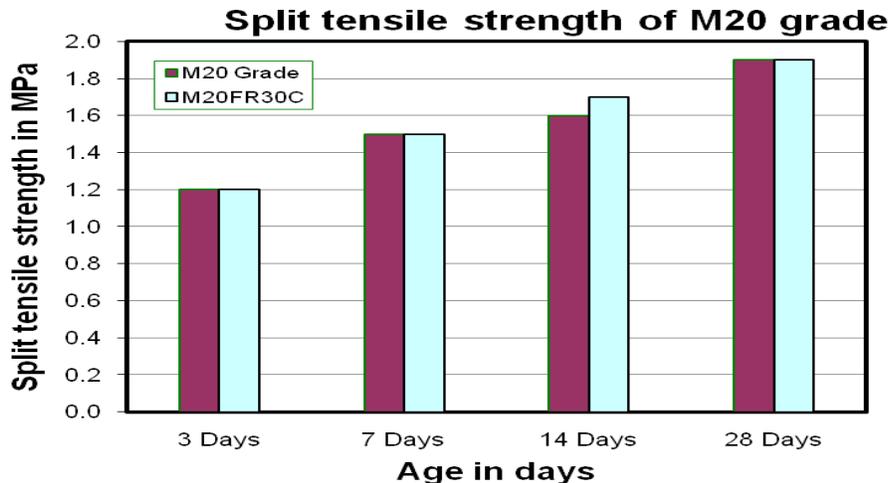


Figure 1 Optimum moisture contents for M15 and M20







Results and Inferences

For M15 and M20 grade mixes, the maximum dry density was observed as 6% and 7% respectively by weight of their proportions and zero slump were noted. The figure 1 represents the same.

The compressive strengths of concrete cubes of M15 and M20 grades of 30% replacement of cement by flyash was performed almost same strengths M15 and M20 grades of concrete mixes without any replacement of cement.

The compressive strengths of roller compacted concrete cylinder specimens of M15 and M20 grades of 30% replacement of cement by flyash was performed almost same strengths M15 and M20 grades of roller compacted concrete mixes without any replacement of cement.

The split tensile strengths of roller compacted concrete cylinders of M15 and M20 grades of 30% replacement of cement by flyash was performed almost equal strengths M15 and M20 grades of roller compacted concrete mixes without any replacement of cement.

From the above investigations it was observed that the compressive strength of cube and cylinder specimens of M15 and M20 grades with and without replacement of 30% flyash was showed same performance with a marginal difference. The same trend was observed in split tensile strength of cylindrical specimens also. This experimental results conclude that roller compacted concrete of M15 and M20 grades with 30% fly ash replacement can be used for both base and sub base course. Because flyash was an industrial waste by product and disposal of it is a great problem.

REFERENCES

1. Myle Nguyen James, Wonchang Choi and Taher, Abu-Lebdeh presented a paper on use of recycled aggregate and fly ash in concrete pavement
2. L.Javier Malvar presented a paper on use of flyash in concrete pavements.
3. M.C.Natarajan and Lenin das discusses about Concrete mix proportioning as per IS 10262:2009 – Comparison with IS 10262:1982 and ACI 211.1-91
4. Synthesis of Current and Projected Concrete Highway Technology, David Whiting, et al, SHRP-C-345, Strategic Highway Research Program, National Research Council.
5. ACI Committee 226. 1987a. Ground granulated blast furnace slag as a cementitious constituent in concrete ACI 226.AR-87. Detroit: American Concrete Institute.
6. Adams, T. H. 1988. Marketing of fly ash concrete. In MSU seminar: Fly ash applications to concrete (January), 1.10, and 5.10. East Lansing: Michigan State University.
7. Admixtures and ground slag for concrete. 1990. Transportation research circular no. 365 (December). Washington: Transportation Research Board, National Research Council.
8. Davis, R. E., R. W. Carlson, J. W. Kelly, and A. G. Davis. 1937. Properties of cements and concretes containing fly ash. Proceedings, American Concrete Institute 33:577-612.
9. Diamond, S. 1981. Effects of two Danish fly ashes on alkali contents of pore solutions of cement fly ash pastes. Cement and Concrete Research 11:383-94.
10. Diamond, S. 1985. Very high strength cement-based materials: A perspective. Materials Research Society Symposia Proceedings 142:223-43.

11. Farbiarz, J., and R. L. Carrasquillo. 1986. Effectiveness of fly ash replacement in the reduction of damage due to alkali-aggregate reaction in concrete. Report no. FHWA/TX-87/15+450-1 (May). Texas State Department of Highways and Public Transportation.
12. Halstead, W. J. 1986. Use of fly ash in concrete. NCHRP 127 (October). Washington: Transportation Research Board, National Research Council.
13. Helmuth, R. 1987. Fly ash in cement and concrete. Skokie, III: Portland cements Association.
14. Hines, D. 1985. Fly ash use in lean concrete base. Colorado Department of Highways final report. Report no. CDOH-SMB-R-85-13 (December).
15. Hobbs, D. W. 1982. Influence of pulverized-fuel ash and granulated blast furnace slag upon expansion caused by the alkali-silica reaction. Magazine of Concrete Research 34:83-93.
16. Idorn, G. M., and K. R. Henrisken. 1984. State of the art for fly ash uses in concrete. Cement and Concrete Research 14 (4):463-70.
17. Joshi, R. C., B. W. Langan, and M. A. Ward. 1987. Strength and durability of concrete with high proportions of fly ash and other mineral admixtures. In Durability of building materials. Vol. 4, 253-70. Amsterdam: Elsevier Science Publishers.
18. Kohubu, M. 1969. Fly ash and fly ash cement. In Proceedings, Fifth international symposium on the chemistry of cement (1968). Part IV, 75-105. Tokyo: Cement Association of Japan.
19. Lane, R. O., and J. F. Best. 1982. Properties of fly ash in Portland cement concrete. Concrete International. Design and Construction 4 (7):81-92. RILEM. 1988. Siliceous by-products for use in concrete. Final report: 73-SBC RILEM Committee. Materials and Structures 21 (121):69-80.
20. Mehta, P. K. 1980. Performance test for sulfate resistance and alkali-silica reactivity of hydraulic cements. In ASTM STP 691: Durability of building materials and components, 336-45.
21. Mehta, P. K. 1983. Pozzolanic and cementitious by-products as mineral admixtures for concrete: A practical review. In ACI special publication SP-79: The use of fly ash, silica fume, slag and other mineral by-products in concrete, ed. V. M. Malhotra, 1-46. Detroit: American Concrete Institute.
22. Olek, J., P. J. Tikalsky, and R. L. Carrasquillo. 1986. Production of concrete containing fly ash for pavement applications. Research report 364-2 (May). Austin: University of Texas Center for Transportation Research
23. Soroushian, P. 1990. Durability characteristics of fly ash concrete. In Recent advances in concrete technology seminar, MSU-CTS 4 (February), 6.1-6.18. East Lansing: Michigan State University.

Motion Detection Method to Compensate Camera Flicker Using an Algorithm

¹Alam Inder Singh, ²Gagandeep Kaur

¹Student, M.tech Electronics and Communication Section, YCOE, Talwandi Sabo, Guru Kashi Punjabi University
Campus, Patiala

²Asst Prof., Electronics and Communication Section, YCOE, Talwandi Sabo, Guru Kashi Punjabi University
Campus, Patiala

Abstract

This paper presents an improved motion detection system that is based on background subtraction method and threshold comparison method. Motion detection is used in many computer vision tasks like human tracking, pose estimation and recognition. It is a basic part for many computer vision tasks. Our purposed method makes background image using 10 previous consecutive frames. Our method detects motion via a standard webcam in real-time YUY2_640x480 resolution. Experimental results showed that the proposed method is more robust in nature as it can avoid the noise in motion detection due to camera flicker and useful to reduce the number of false positive alarms.

Keywords - background subtraction method, consecutive frames, motion detection, threshold comparison **method**.

I. Introduction

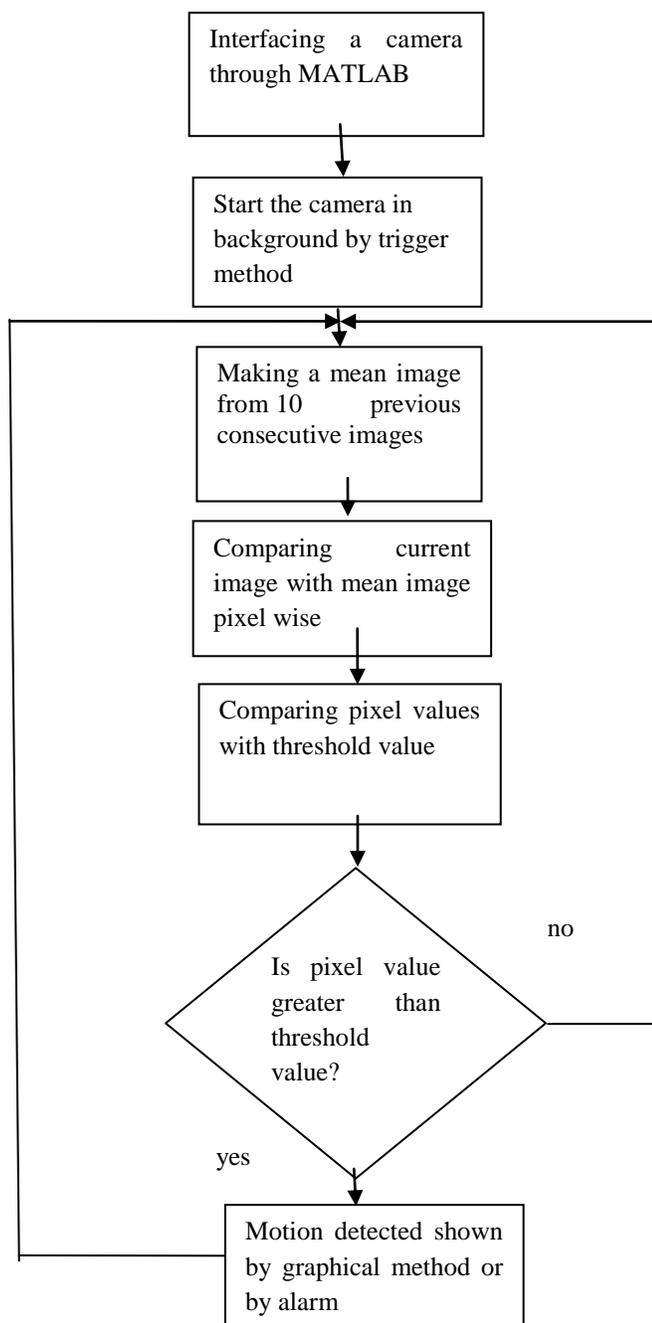
Motion detection is a process of confirming a change in position of an object relative to its surroundings or the change in the surroundings relative to an object. The task of a motion detection system is to detect a “area of motion” present in a “area of environment being monitored”. The area of motion in this case refers to portion of the environment with activity due to the motions of moving objects. Motion detection is usually a software-based monitoring algorithm which, when it detects motions will signal the surveillance camera to begin capturing the event or simply shows the motion detection using graphical method.

Motion detection is used in many computer vision tasks like human tracking^[3], pose estimation^[4] and recognition^[5]. It is a basic part for many computer vision tasks. Many methods are used for motion detection such as optical flow method^[6, 7, 13], background subtraction method^[2, 10] and consecutive frames subtraction method^[1].

Optical motion detection method uses Infrared light or laser technology for motion detection. Devices, such as PIR motion detectors, have a sensor that detects a disturbance in the infrared spectrum, such as a person or an animal. Once detected, an electronic signal can activate an alarm or a camera that can capture an image or video of the motioner. Background motion detection method is a simple method for motion detection by a fixed camera compares the current image with a reference image or background image pixel by pixel and simply counts the number of pixels with change more than the threshold value and thus motion is detected. Consecutive frame subtraction method simply compares the current image with previous image to find the change in value of pixels above threshold to detect the motion.

In this paper, a new approach is purposed which is a combination of background subtraction method and consecutive frame subtraction method. As in this method background image is formed by taking mean of previous ten consecutive frames and then current image is compared pixel wise with the background image to detect motion. This method will brought some robustness in motion detection because previous background subtraction method is very sensitive to the very little motion that can be called noise

II. Flow chart



The focus of this paper is on making a background image from ten previous consecutive frames in real time by trigger method. Then current image is compared pixel wise or subtracted from background image to detect any motion. The image received after subtraction is called difference image. Values of pixels can be positive or negative in difference image. Therefore absolute of difference image is taken and then values of pixels in difference image is compared with threshold value and if the pixel value is more than threshold value then it means there is motion in the area being monitored. This method continuously keep making background image using ten previous frames in real time. To make it practical and useful warning system or graphical method is used. This method will also show the number of objects detected in motion and percentage area of total area in which motion is present. This

background image formation and motion detection process happens in while loop because it continuously detects motion in real time, it did not stop until required it to stop.

III. Image processing

This method analyzes the images captured by camera to detect motion in the area being monitored using MATLAB Software. Using image processing technique to create the mean image or average image of previous ten consecutive frames and then subtracting this mean image from current image to detect the motion. In addition this method also shows number of objects in motion.

IV. Output/warning unit

Two approaches are used for indication of motion detection, one by blowing alarm upon motion detection and second is the graphical method to take a record which shows number of objects in motion and percentage of area in which motion is present.

V. Motion detection

This method is used to detect motion by comparing the pixel value of difference image with threshold value provided by us in the algorithm. As shown in TABLE I, object in motion is said to be detected when pixel value is more than the threshold value. Threshold value is provided according to our requirement depending upon environment. This method can detect number of objects in motion. The motion is said to be caused by object or living being by counting the number of pixels with more than threshold value grouped together. This method can detect more than one object in motion by counting the number of groups of pixels. Detected motion can be considered appropriate depending on the number of pixels present in a group of pixels as we can ignore a group of pixels having less than 5 numbers of pixels in it as it can be due to noise. This method can also show the percentage area in which motion is present using graphical method.

Comparison of threshold value and pixel value	Motion Detection
Pixel value > threshold value	Motion is present
Pixel value < threshold value	Motion is not present

Table1-Comparison of threshold value and motion detection

VI. Conclusion

In this paper, a new method is purposed to detect motion using the algorithm. The proposed method is adjustable to the camera movements which were shown as detected motion in other approaches because of their over sensitivity. Therefore this method is useful to reduce the number of false positive alarms.

In addition this method shows the number of objects in motion detected if there are many and it also shows the percentage of total area in which motion is present. It runs at a 640×480 resolution which is acceptable for real-time scenarios. Experiments showed that the proposed method is more robust in nature as it can avoid the noise in motion detection due to camera flicker. According to the real world experiments motion detection and graphical method used for taking record is taken in the experiment in real time.

References

- [1] Haritaoglu I, Harwood D, David L S., 2000, Real-time surveillance of people and their activities, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 809-830.
- [2] Wren C, Azarbayejani A, Darrell T, 1997, Real-time tracking of the human body, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 780-785.
- [3] G. L. Foresti, C. Micheloni, L. Snidaro, P. Remagnino, and T. Ellis, 2005, Active video-based surveillance system, IEEE Signal Process. Mag., vol. 22, no. 2, pp. 25–37.
- [4] Raju Rangaswami, Zoran Dimitrijević, Kyle Kakligian, Edward Chang, Yuan-fang Wang , 2004, The sphinx video surveillance system, IEEE Conference on Multimedia and Expo.
- [5] Rudy Melli, Andrea Prati, Rita Cucchiara, Lieven de Cock , 2005, Predictive and Probabilistic Tracking to Detect Stopped Vehicles, The 7th IEEE Workshops on Application of Computer Vision, vol. 1, pp. 388-393.
- [6] Barron J, Fleet D, Beauchemin S. Performance of optical flow techniques, International Journal of Computer Vision, 1994,12, pp.43-47.
- [7] Bin Wang, Jianshou Pan, Yanbing Liang, 2004 , Video image motion detect based on Matrox card, Proceeding of Northwest University, pp.38-45.
- [8] Yan Zhao and Jiao-min Liu, 2010, An Improved Method for Human Motion Detection and Application, 3rd International Congress on Image and Signal Processing (CISP2010).
- [9] Haritaoglu I, Harwood D, David L S., 2000 , Real-time surveillance of people and their activities, IEEE Transactions on Pattern Analysis and Machine Intelligence, 22, pp. 809-830.
- [10] A. Elgammal, D. Harwood and L. Davis, 2000, Non-parametric model for background subtraction, Proc. European Conf. Computer Vision.
- [11] Nan Lu, Jihong Wang, Q.H. Wu and Li Yang, An Improved Motion Detection Method for Real-Time Surveillance, IAENG International Journal of Computer Science, 35:1, IJCS_35_1_16, pp.1-10
- [12] Mirko Ristivojević and Janusz Konrad, Multi-frame motion detection for active unstable cameras, XXII Brazilian Symposium on Computer Graphics and Image Processing.
- [13] Jor-El Sy Rivo and Engr. Rhandley Cajote, “Object Motion Detection Using Optical Flow”, Digital Signal Processing Laboratory, Department of Electrical and Electronics Laboratory, University of the Philippines, pp. 1-2.

An Efficient Approach for Mining Frequent Itemsets with Large Windows

K Jothimani¹, S. Antony Selvadoss Thanmani²

¹ Research Scholar, Research Department of Computer Science,
NGM College, 90, Palghat Road, Pollachi - 642 001
Coimbatore District, Tamilnadu, INDIA

² Professor and Head, Research Department of Computer Science,
NGM College, 90, Palghat Road, Pollachi - 642 001
Coimbatore District, Tamilnadu, INDIA

Abstract— The problem of mining frequent itemsets in streaming data has attracted a lot of attention lately. Even though numerous frequent itemsets mining algorithms have been developed over the past decade, new solutions for handling stream data are still required due to the continuous, unbounded, and ordered sequence of data elements generated at a rapid rate in a data stream. The main challenge in data streams will be constrained by limited resources of time, memory, and sample size. Data mining has traditionally been performed over static datasets, where data mining algorithms can afford to read the input data several times. The goal of this article analysing the mining frequent itemsets in theoretical manner in the large windows. By comparing previous algorithms we propose new method using analytical modelling to determine the factors over data streams.

Keywords— Data Mining, Data Streams Frequent Itemset Mining, Large Sliding Window.

I. Introduction

Recently the data stream, which is an unbounded sequence of data elements generated at a rapid rate, provides a dynamic environment for collecting data sources. It is likely that the embedded knowledge in a data stream will change quickly as time goes by. Therefore, catching the recent trend of data is an important issue when mining frequent itemsets from data streams. Although the sliding window model proposed a good solution for this problem, the appearing information of the itemsets within the sliding window has to be maintained completely in the traditional approach. In this paper, for estimating the approximate supports of patterns within the current sliding window, two data structures are proposed to maintain the average time stamps and frequency Changing points of patterns, respectively.

The main challenge is that 'data-intensive' mining is constrained by limited resources of time, memory, and sample size. Data mining has traditionally been performed over static datasets, where data mining algorithms can afford to read the input data several times. When the source of data items is an open-ended data stream, not all data can be loaded into the memory and off-line mining with a fixed size dataset is no longer technically feasible due to the unique features of streaming data [2].

Data from sensors like weather stations is an example of fixed-sized data, whereas again, market basket data are an example of variable size data, because each basket contains a different number of items. By contrast, sensor measurements have a fixed size, as each set of measurements contains a fixed set of dimensions, like temperature, precipitation, etc.

We classify the stream-mining techniques into two categories based on the window model that they adopt in order to provide insights into how and why the techniques are useful. Then, we further analyze the algorithms according to whether they are exact or approximate and, for approximate approaches, whether they are false-positive or false-negative. First, each element in the datastream can be examined only once or twice, making traditional multiple-scan approaches infeasible. Second, the consumption of memory space should be confined in a range, despite that data elements are continuously streaming into the local site. Third, not with standing the data characteristics of incoming stream may be unpredictable; the mining task should proceed normally and offer acceptable quality of results. Fourth, the latest analysis result of the data stream should be available as soon as possible when the user invokes a query

II .RELATED WORK

There are a number of research works which study the problem of data-stream mining in the first decade of 21st century. Many previous studies contributed to the efficient mining of frequent itemsets (FI) in streaming data [4, 5]. According to the stream processing model [20], the research of mining frequent itemsets in data streams can be divided into three categories: *landmark windows* [15, 12, 19, 11, 13], *sliding windows* [5, 6, 14, 16, 17, 18], and *damped windows* [7, 4], as described briefly as follows. In the landmark window model, knowledge discovery is performed based on the values between a specific timestamp called landmark and the present. In the sliding window model, knowledge discovery is performed over a fixed number of recently generated data elements which is the target of data mining.

Two types of sliding window, i.e., *transaction-sensitive sliding window* and *time-sensitive sliding window*, are used in mining data streams. The basic processing unit of window sliding of first type is an expired transaction while the basic unit of window sliding of second one is a time unit, such as a minute or a hour. In the damped window model, recent sliding windows are more important than previous ones.

Besides, there are still some interesting research works [9] [10] [11] on the sliding window model. In [9] a false-negative approach named *LSWIM* was proposed. By employing a progressively increasing function of ms , *LSWIM* greatly reduces the number of potential itemsets and would approximate the set of FIs over a sliding window. According to this definition, each transaction could be represented by the product of the corresponding prime numbers of individual items into the transaction. As the product of the prime numbers unique we can easily check the inclusion of two itemsets by performing a modulo division on itemsets ($Y \text{ MOD } X$). If the remainder is 0 then $X \subseteq Y$, otherwise X is not included in Y .

LSWIM is an approximate approach based on the application of the *Principle of Inclusion and Exclusion* in *Combinatorial Mathematics* [7]. One of the most notable features of *LSWIM* is that it would approximate the count of an arbitrary itemset, through an equation (i.e., Equation (4) in [12]), by only the sum of counts of the first few orders of its subsets over the data stream. There are also two techniques named *counts bounding* and *correction*, respectively, integrated within *LSWIM*. Both techniques are of the purpose to improve the quality of *LSWIM*'s approximation, while they adopt different means to achieve the purpose. By working together with these original techniques, the mining result of DSCA reaches good accuracy. The concept of *Inclusion and Exclusion Principle* [7] is valuable that it may also be applied in mining FIs under different window models other than the landmark window. Based on the theory of *Approximate Inclusion-Exclusion* [8], we devise and propose a novel algorithm, called *SWIM*, to approximate dynamically and discover FIs over the sliding window in a transactional data stream

III. MINING LARGE SLIDING WINDOWS

LSWIM (Large Sliding Window Incremental Miner) algorithm relies on a verifier function and it is an exact and efficient algorithm for mining very large sliding windows over data streams. The performance of *LSWIM* improves when small delays are allowed in reporting new frequent itemsets, however this delay can be set to 0 with a small performance overhead.

A. Problem Statement and Notations

Let D be the dataset to be mined (a window in our case); D contains several transactions (baskets), where each basket contains one or more items. Let $I = i_1, i_2, \dots, i_n$ be the set of all such distinct items in D . Each subset of I is called an itemset, and by k -itemset we mean an itemset containing k different items. The *frequency* of an itemset p is the number of transactions in D that contain itemset p , and is denoted as $\text{Count}(p, D)$. The support of p , $\text{sup}(p, D)$, is defined as its frequency divided by the total number of transactions in D . Therefore, $0 \leq \text{sup}(p, D) \leq 1$ for each itemset p . The goal of frequent itemsets mining² is to find all such itemsets p , whose support is greater than (or equal to) some given minimum support threshold α . The set of frequent itemsets in D is denoted as $\sigma_\alpha(D)$.

Here we consider frequent itemsets mining over a data stream, thus D is defined as a sliding window over the continuous stream. D moves forward by a certain amount³ by adding the new slide ($\delta+$) and dropping the expired one ($\delta-$). Therefore, the successive instances of D are shown as W_1, W_2, \dots . The number of transactions that are added to (and removed from) each window is called its slide size. In this paper, for the purpose of simplicity, we assume that all slides have the same size, and also each window consists of the same number of slides. Thus, $n = \lfloor W/S \rfloor$ is the number of slides (a.k.a. panes) in each window, where W denotes the window size and S denotes the size of the slides.

B. The *LSWIM* Algorithm

Large Sliding Window Incremental Miner (*LSWIM*) always maintains a union of the frequent itemsets of all slides in the current window W , called Pattern Tree (*PT*), which is guaranteed to be a superset of the frequent itemsets over W . Upon arrival of a new slide and expiration of an old one, we update the true count of each pattern in *PT*, by considering its frequency in both the expired slide and the new slide. To assure that *PT* contains all itemsets that are frequent in at least one of the slides of the current window $\cup_i(\sigma_\alpha(S_i))$, we must also mine the new slide and add its frequent itemsets to *PT*. The difficulty is that when a new pattern is added to *PT* for the first time, its true frequency in the whole window is not known, since this pattern wasn't frequent in the previous $n-1$ slides. To address this problem, *LSWIM* uses an auxiliary array, *aux array*, for each new pattern in the new slide.

The *aux array* stores the frequency of a pattern in each window starting at a particular slide in the current window. In other words, the *aux array* stores frequency of a pattern for each window, for which the frequency is not known. The key point is that this counting can either be done eagerly (i.e., immediately) or lazily. Under the laziest approach, we wait until a slide expires and then compute the frequency of such new itemsets over this slide and update the *aux arrays* accordingly. This saves many additional passes through the window. The pseudo code for the *LSWIM* algorithm is given in Figure 1. At the end of each slide, *LSWIM* outputs all itemsets in *PT* whose frequency

at that time is $\geq \alpha \cdot n \cdot |S|$. However we may miss a few itemsets due to lack of knowledge at the time of output, but we will report them as delayed when other slides expire. The following mini-example shows how LSWIM works.

For Each New Slide S

- 1: For each pattern $p \in PT$
 update $p.freq$ over S
- 2: Mine S to compute $\sigma\alpha(S)$
- 3: For each existing pattern $p \in \sigma\alpha(S) \cap PT$
 remember S as the last slide in which p is frequent
- 4: For each new pattern $p \in \sigma\alpha(S) \setminus PT$
 $PT \leftarrow PT \cup \{p\}$
 remember S as the first slide in which p is frequent
 create auxiliary array for p and start monitoring it

For Each Expiring Slide S

- 5: For each pattern $p \in PT$
 update $p.freq$, if S has been counted in
update $p.aux$ array, if applicable
report p as delayed, if frequent but not reported
at query time
delete $p.aux$ array, if p has existed since arrival of S
delete p , if p no longer frequent in any of the current
slides

Fig. 1. LSWIM pseudo code.

Max Delay. The maximum delay allowed by LSWIM is $n - 1$ slides. Indeed, after expiration of $n - 1$ slides, LSWIM will have a complete history of the frequency of all frequent itemsets of W and can report them. Moreover, the case in which a pattern is reported after $(n - 1)$ slides of time, is rare. For this to happen, pattern's support in all previous $n - 1$ slides must be less than α but very close to it, say $\alpha \cdot |S| - 1$, and suddenly its occurrence goes up in the next slide to say β , causing the total frequency over the whole window to be greater than the support threshold. Formally, this requires that $(n - 1) \cdot (\alpha \cdot |S| - 1) + \beta \geq \alpha \cdot n \cdot |S|$, which implies $\beta \geq n + \alpha \cdot |S| - 1$. This is not impossible, but in real-world such events are very rare, especially when n is a large number (i.e., a large window spanning many slides). While LSWIM(Delay=L) represents an efficient incremental mining algorithm, counting frequencies of itemsets over a given dataset ($n - L + 1$ slides in our case) remains a bottleneck. Therefore, faster algorithms are required to compute these counts efficiently.

IV. CONCLUSION

Mining data streams for association rules has proven to be a difficult problem, since techniques developed to mine frequent itemsets on stored data result in excessive costs and time delays. This paper has made two important contributions to the solution of this problem. The first is the introduction of a very fast algorithm to verify the frequency of a given set of itemsets. In fact, our algorithm outperforms the existing state-of-the-art counting algorithms by an order of magnitude. The second contribution is to use our fast verifier to solve the association-rule mining problem under the realistic assumption that we are mostly interested in the new/expiring itemsets.

This delta-maintenance approach effectively mines very large windows with slides, which was not possible before. However, we also explored a second approach that further improves the performance by simply allowing a small reporting delay. Clearly this approach would become desirable in situations where modest delays in reporting new itemsets are acceptable. Such delays are negligible when compared to the time needed for the experts to validate the new rules before they are actually put into use. In summary we have proposed an approach of great efficiency, flexibility, and scalability to solve the frequent pattern mining problem on data streams with very large windows.

REFERENCES

- [1] M.N. Garofalakis, J. Gehrke, & R. Rastogi, Querying and mining data streams: you only get one look (A Tutorial), Proc. 2002 ACM SIGMOD Conf. on Management of Data, Madison, Wisconsin, 2002, p. 635.
- [2] Y. Zhu & D. Shasha, "Stat Stream: statistical monitoring of thousands of data streams in real time", Proc. 28th Conf. on Very Large Data Bases, Hong Kong, China, 2002, pp. 358-369.
- [3] G.S. Manku & R. Motwani, "Approximate frequency counts over data streams", Proc. 28th Conf. on Very Large Data Bases, Hong Kong, China, 2002, pp. 346-357.

- [4] J.H. Chang & W.S. Lee, "A sliding window method for finding recently frequent itemsets over online data streams", Journal of Information science and Engineering, 20(4), 2004, pp. 753-762.
- [5] J. Cheng, Y. Ke, & W. Ng, "Maintaining frequent itemsets over high-speed data streams", Proc. 10th Pacific-Asia Conf. on Knowledge Discovery and Data Mining, Singapore, 2006, pp.462-467.
- [6] C.K.-S. Leung & Q.I. Khan, "DSTree: a tree structure for the mining of frequent sets from data streams," Proc. 6th IEEE Conf. on Data Mining, Hong Kong, China, 2006, pp. 928-932.
- [7] B. Mozafari, H. Thakkar, & C. Zaniolo, "Verifying and mining frequent patterns from large windows over data streams", Proc. 24th Conf. on Data Engineering, Mexico, 2008, pp. 179-188.
- [8] K.-F. Jea & C.-W. Li, "Discovering frequent itemsets over transactional data streams through an efficient and stable approximate approach, Expert Systems with Applications", 36(10), 2009, pp. 12323-12331.
- [9] F. Bodon, "A fast APRIORI implementation", Proc. ICDM Workshop on Frequent Itemset Mining Implementations (FIMI'03), 2003.
- [10] N. Jiang and L. Gruenwald, "Research Issues in Data Stream Association Rule Mining". In SIGMOD Record, Vol. 35, No. 1, Mar. 2006.
- [11] Frequent Itemset Mining Implementations Repository (FIMI). Available: <http://fimi.cs.helsinki.fi/>
- [12] Y. Chi, H. Wang, P.S. Yu, & R.R. Muntz, "Moment: maintaining closed frequent itemsets over a stream sliding window", Proc. 4th IEEE Conf. on Data Mining, Brighton, UK, 2004, pp. 59-66.
- [13] N. Jiang & L. Gruenwald, "CFI-Stream: mining closed frequent Itemsets in data streams", Proc. 12th ACM SIGKDD Conf. on Knowledge Discovery and Data Mining, Philadelphia, PA,USA, 2006, pp. 592-597.
- [14] Quest Data Mining Synthetic Data Generation Code. Available: http://www.almaden.ibm.com/cs/projects/iis/hdb/Projects/data_mining/datasets/syndata.html
- [15] H.F Li, S.Y. Lee, M.K. Shan, "An Efficient Algorithm for Mining Frequent Itemsets over the Entire History of Data Streams", In Proceedings of First International Workshop on Knowledge Discovery in Data Streams 9IWKDD, 2004.
- [16] H.F Li, S.Y. Lee, M.K. Shan, "Online Mining (Recently) Maximal Frequent Itemsets over Data Streams", In Proceedings of the 15th IEEE International Workshop on Research Issues on Data Engineering (RIDE), 2005.
- [17] P. Indyk, D. Woodruff, "Optimal approximations of the frequency moments of data streams", Proceedings of the thirty-seventh annual ACM symposium on Theory of computing, pp.202-208, 2005.



Smt. K Jothimani received her Bachelor degree in Computer Science from Bharathidasan University, Trichy in 2003. She received her Master degree in Computer Applications in 2008. She pursued her Master of Philosophy in Computer Science in the year 2009 from Vinayaka Missions University, Salem. Currently she is a research scholar of the Department of Computer Science, NGM College under Bharathiyar University, Coimbatore. She had three years of experience in the

computer field in technical as well as non technical. She is a life member of Indian Society for Technical Education from the year 2009. Also she is a member of Computer Society of India(CSI). She has published more than ten papers in international and national conferences including standard journals. Her area of Interests Data mining, Knowledge Engineering and Image Processing.



Dr. Antony Selvadoss Thanamani is presently working as Professor and Head, Dept of Computer Science, NGM College, Coimbatore, India (affiliated to Bharathiar University, Coimbatore). He has published more than 100 papers in international/ national journals and conferences. He has authored many books on recent trends in Information Technology. His areas of interest include E-Learning, Knowledge Management, Data Mining, Networking, Parallel and Distributed Computing. He has to his credit 24 years of teaching and research experience. He

is a senior member of International Association of Computer Science and Information Technology, Singapore and Active member of Computer Science Society of India, Computer Science Teachers Association, New York.

Recent Developments in Preparation of Non Conventional Activated Carbons

Prof Shantini Bokil¹ Prof.Dr.R.K.Rai² and Prof.Dr.S.N.Kaul³

¹Asst.Prof.Civil Engg, MIT College of Engineering,Pune

²Lecturer in Civil Engg, Amravati Engineering College, Amravati

³Former Director NEERI Nagpur,

Abstract:

The present paper has been initiated with a view to investigate the adsorption characteristics of available agro based material to be used as an adsorbent for the removal of the contaminants from the secondary treated industry effluents. The main purpose is to characterise the available agro-based waste material as adsorbent and assess the removal efficiency of contaminants by available agro-based waste material. The paper highlights the various evaluation parameters like initial concentration, time, pH, temperature dose of the adsorbent, particle-size, functional group and agitation speed on the removal of contaminants from the secondary treated industry effluents.

Keywords: adsorption, Freundlich, Langmuir isotherms, effluents.

1. Introduction

Water pollution by contaminants from the secondary treated industry effluents like dyes/colours in particular has been a matter of concern for scientific workers since long time. Many of the dyes produce toxic intermediate products during process of removal which are highly toxic. Textile and pulp and paper industries produce large amount of colored effluents. There is no single treatment method which is capable of adequate treatment of the color effluents.

Adsorption is a method which has been widely reported for the removal of dyes from colored aqueous solutions and dye stuff effluents. Activated carbon has been frequently applied for decontamination of colored effluents but its high cost limits its application for large scale treatment of wastewater in developing countries. To reduce the cost of treatment, significant numbers of non-conventional agro-based low-cost adsorbents have been used for the removal of colors from effluents.

Classification Of Pollutants In Waste Water And Methods Of Treatment

The treatment of waste waters from textile industries is carried out in many stages. After pre, primary, secondary treatments waste water still contains dyes and dissolved contaminants which cannot be removed by conventional methods say chemical and biological methods. Such substances can be removed through tertiary polishing treatment processes like adsorption.

Research and development studies have been carried out using activated carbon prepared from number of agricultural wastes and agro based products for the removal of contaminants from the aqueous solutions and are focused on the treatment that are inexpensive and eco-friendly.

Literature review highlights that adsorption through agricultural products such as rice husk, soybean hulls, saw dust, coconut shell, jute, groundnut shell, clay, fly-ash etc., has been demonstrated to be a useful alternative to the conventional treatment systems.

Present work aims at preparation of activated carbon from low cost agro-based waste materials and their subsequent application for the removal of effluent from the secondary treated industrial effluent (live waste water) aqueous solutions.

FUNDAMENTALS OF ADSORPTION:

The surfaces of solids or liquids possess a residual field of force which gives rise to the process of uptake of atoms, molecules, ions from the systems in contact. This phenomenon is known as 'adsorption'.

TYPES OF ADSORPTION

The adsorption is categorized on the basis of interactions involved between the atoms, molecules or ions or adsorbate and adsorbent. According to the forces involved two basic kinds of adsorption processes can be distinguished:

1. Physical adsorption.
2. Chemical adsorption.

PHYSICAL ADSORPTION

Physical adsorption is reversible and a rapid condition of equilibrium is established between adsorbed and unadsorbed adsorbate species. Molecules adsorbed by physical adsorption are held to the adsorbent surface by physical adsorption are held to the adsorbent surface by weak van der Waals forces of attraction.

CHEMICAL ADSORPTION

This type of adsorption takes place due to the action of specific 'chemical force: i.e. those involved in the transfer or sharing of electrons between an adsorbent and adsorbate species. The chemical reaction on the surface of the adsorbent may be either exothermic or endothermic in nature.

ADSORPTION FROM SOLUTIONS

A solid surface in contact with a solution has the tendency to accumulate surface layer of solute molecules due to imbalance of surface forces. Majority of the molecules accumulated at the interface are adsorbed onto the large surface area within the pores of adsorbent and relatively a few are adsorbed on the outside surface of the particle. **FIG.1.** explains the pore structures of the typical adsorbent.



FIG.1. Macropore, mesopore, micropore and submicropore adsorption sites on activated carbon.

FACTORS INFLUENCING ADSORPTION

Molecules of solute are removed from the solution and taken up by the adsorbent during the process of adsorption. The equilibrium distribution of solute between the liquid and solid phases is an important property of adsorption system and helps in defining the capacity of a particular system of equal importance to the engineer are the kinetics of the system which describes the rate at which this equilibrium is reached. The rate of adsorption determines the detention time required for treatment and thus the size of the contacting systems.

The adsorption process takes place in three steps:

1. Macro-transport: movement of adsorbate through water to solid interface by diffusion.
2. Microtransport: Diffusion of organic particles through solid adsorbent.
3. Sorption.

Many factors influence the rate at which adsorption occurs and the extent to which a particular material can be adsorbed. The more important factors are briefly discussed below.

Agitation

The rate of adsorption is controlled by either film diffusion or pore diffusion, depending on the amount of agitation in the system. If relatively little agitation occurs between the adsorbent particle and the fluid, the surface film of liquid around the particle will be thick and film diffusion will be the rate limiting step. If adequate mixing is provided the rate of film diffusion will increase to the point that pore diffusion becomes the rate-limiting step.

CHARACTERISTICS OF ADSORBENTS

Particle size and surface are important properties of adsorbent. The size of adsorbent particle influences the rate at which adsorption occurs; adsorption rates increase as particle sizes decrease. Thus, adsorption rates are faster for powdered carbon than for granular carbon. Adsorption is strongest when the pores are just large enough to permit the molecules to enter.

PH

The pH at which adsorption is carried out has been shown to have a strong influence on the extent of adsorption. Organic acids are more adsorbable at low pH whereas, the adsorption of organic bases is favored at high pH. The optimum pH for any adsorption process must be determined by laboratory testing.

CONCENTRATION

Commonly both equilibrium amounts adsorbed and rate of adsorption increase with concentration and rate constraints increase with concentration approximately in a linear fashion and the degree of linearity may change in other operating conditions. The extent of adsorption in a given adsorbent depends on the adsorbate concentration in solution. The break points have been interpreted as an indication of either a transition in adsorbed phase or as the on-set of the formation of further types. The actual variation of adsorption with concentration at controlled condition is represented by an isotherm. However, best linear normalization can be made by association of rate constants, obtained from plots of concentration, c , at time, t , against \sqrt{t} as a function of initial adsorbate concentration.

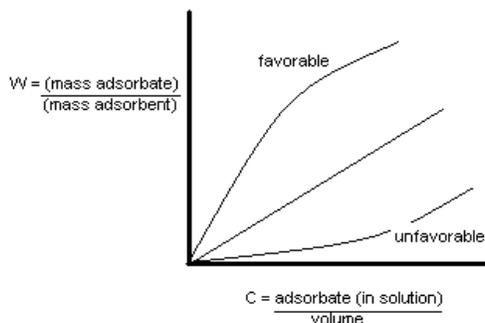
TEMPERATURE

The temperature at which an adsorption process is conducted will affect both the rate and the extent of adsorption. Adsorption rate increases with increase temperature. However, since adsorption is an exothermic process, the adsorption will increase at lower temperature and decrease at higher temperatures.

Generally the amount of material adsorbed is determined as a function of the concentration at a constant temperature, and the resulting function is called an adsorption isotherm.

DEVELOPMENT OF ADSORPTION ISOTHERMS

Adsorption from aqueous solutions involves concentrations of the solute on the solid surface. When solutions contacted with a solid adsorbent, molecules of adsorbate get transferred from the fluid to the solid until the concentration of adsorbate in solution is in equilibrium with that on the surface and no change can be observed in the concentration of the solute on solid surface or in the bulk solution. Equal amounts of solute eventually are being adsorbed and desorbed



simultaneously. This is called adsorption equilibrium.

FIG.2.ADSORPTION ISOTHERMS

Adsorption equilibrium data is typically plotted in the form of an *adsorption isotherm* (i.e. at constant temperature) with the mass adsorbed on the y-axis and the mass in the fluid on the x-axis as in **FIG.3**. The shape of the curve is significant and factors heavily into design. "**Favorable**" isotherms permit higher solid loadings at lower solution concentrations. These tend to start out steep and level out. Isotherms which start out flat are "**unfavorable**", since they only work well at high concentrations of solute. Several fits have been proposed for isotherms.

The shape of the adsorption isotherms gives qualitative information about the adsorption process and the extent of the surface coverage by the adsorbent. All isotherms share a number of common features. These data can then be used for the preliminary evaluation of the proposed system.

Various mathematical relationships have been developed to describe the dynamic equilibrium distribution of adsorbate between the adsorbent and bulk liquid phases. These relationships are applicable when the adsorption tests are conducted at constant temperature and are referred to as adsorption isotherms. Freundlich, Langmuir and BET adsorption isotherms are the most commonly used models for describing the dynamic equilibrium.

FREUNDLICH ADSORPTION ISOTHERM

Freundlich developed an empirical equation or the adsorption isotherm which encompasses the heterogeneity of the surface and the exponential distribution of sites and their energies. This equation can also be explained to the varification of adsorption with concentration over a limited range at constant temperature. The Freundlich adsorption

isotherm can be expressed as: $q_e = K_F C_e^{1/n}$

$\log q_e = \log K_F + 1/n \log C_e$

Plot of $\log q_e$ versus $\log C_e$ gives straight line with intercept $\log K_F$ and slope $1/n$

Where,

q_e = amount of dye adsorbed per unit of adsorbent (mg/g)

C_e = Equilibrium concentration of dye in (mg/L)

K_F = Quantity of dye adsorbed for unit concentration of dye in (mg/g) and

$1/n$ is the measure of adsorption intensity Value of $1/n$ from plot indicates good adsorption potential. Value of $1/n$

obtained is always less than 1. For steep slopes ($1/n$ close to 1) indicate high adsorptive capacity at high concentrations.

For flat slopes ($1/n$ close to 0) indicate less adsorptive capacity at lower concentrations.

LANGMUIR ADSORPTION ISOTHERM

The basic assumptions underlying Langmuir’s model which is also called the ideal localized monolayer model are:

1. The molecules are adsorbed on definite sites on the surface of the adsorbent.
2. Each site can accommodate only on molecule (monolayer). The area of site is fixed, quantity determined solely by the geometry of the surface.
3. The adsorption energy is the same at all types, and
4. The adsorbed molecules cannot migrate across the surface or interact with neighboring molecules.

The Langmuir equation was originally derived from kinetic consideration. The Lagmuir isotherm can be expressed as:

$$C_e/q_e = 1/ab + 1/a (C_e)$$

A plot of C_e/q_e versus C_e gives straight line with intercept $1/ab$ and slope $1/a$.

BET Adsorption Isotherm

Brauer, Emmett and Teller developed the BET adsorption isotherm for the generalization of the ideal localized minolayer treatment (Langmuir Model) to account for multilayer adsorption. The BET model is based on the assumption that each molecule in the first adsorbed layer serves as site for adsorption of a molecule into the second layer serves as a site for adsorption of a molecule into the second layer and so on. The BET equilibrium isotherm can be represented by:

$$q_e = \frac{K_B Q^0 C_e}{(C_s - C_e)[1 + (K_B - 1) C_e/C_s]}$$

OR

$$\frac{C_e}{q_e(C_s - C_e)} = \frac{1}{K_B Q^0} + \frac{K_B - 1}{K_B Q^0} \frac{C_e}{C_s}$$

Some of the other important adsorption isotherms are Redilich-Peterson Isotherm and Henry’s Law Linear Adsorption Isotherm.

Henry’s Law Linear Adsorption Isotherm represents the simplest isotherm in which the amount adsorbed varies directly with the equilibrium concentration of solute. The isotherm is described by:

$$Q_e = K_h C_e$$

The isotherm is obtained under conditions at micro and trace concentrations solute. Adsorption thermodynamic parameters can be obtained from adsorption equilibrium constants with temperatures. Linear regression is frequently used to determine the best-fitting isotherm. The linear least-squares method with linearly transformed isotherms has also been widely applied to confirm experimental data and isotherms using coefficients of determination.

Methodolgy

ADSORPTION PRACTICE

Experimental determination of the isotherm is a standard practice for determining the feasibility of an adsorbing system, selecting the adsorbent for evaluating the dose of the adsorbent. Batch experiments are to be carried out to obtain the equilibrium adsorption data for all the contaminants in this work.

Many investigators have employed the adsorption which is carried out either in a batch wise or in a continuous manner .The review reflects on various types of Batch adsorption system and continuous flow systems.

1. Batch Adsorption Systems

In a batch process the adsorbent and wastewater are mixed together in a suitable reaction vessel until the concentration of adsorbate has been reduced to the desired level. Separation of adsorbent (through filtration, centrifugation or settling leaves an effluent suitable for discharge and the adsorbent itself can be regenerated and reused or discarded).

The number and size of adsorption vessels and the amount of adsorbent needed to achieve the required degree of treatment must be determined for the design of batch adsorption systems. These are influenced by wastewater volume, flow rate, and the rate of the adsorption reaction.

Most batch adsorption systems are operated on a fill and draw basis and are limited to small wastewater volumes, flow rate, and the rate of the adsorption reaction. For small volumes generated over a period of 8-10 hours only one adsorption vessel is used.

Batch absorbers can be operated either as a single or as a multistage units operating in counter-current fashion.

2. Continuous Adsorption Systems

The continuous flow processes are usually operated in fixed bed adsorption columns. These systems are capable of treating large volumes of wastewaters and are widely used for treating domestic and industrial wastewaters. These columns may be operated either as single unit or as multiple units in parallel or series or combined parallel and series. Furthermore, they may be separated either in the up-flow or down-flow modes. Up-flow columns can also be used as extended or fluidized beds. In moving bed absorbers (Pulse bed absorbers), bed of adsorbent particles move downward and removed from the bottom and an equal volume of fresh adsorbent is added to the top.

Down-flow beds also remove suspended solids through filtration. Such dual purpose units have lower capital cost than separate units.

Adsorption Column Design

Continuous column studies for investigating the effects of bed heights, flow rate and initial concentration of contaminants. For designing an unsteady state fixed bed adsorber, model is needed that predicts the solute concentration in the bed as a function of time and position. The rate of adsorption depends upon external mass transfer, internal diffusion and surface adsorption.

Breakthrough Curves

Breakthrough point and breakthrough curve are of significant importance in adsorption studies and designing the operating conditions. The area of the activated carbon bed in which the sorption occurs is called Mass Transfer Zone (MTZ). As top layer of the carbon granules become saturated with organic materials, the MTZ will move further down in the bed until breakthrough occurs.

The curve of the pollutant concentration as a function of time has an S shaped appearance and is commonly called the breakthrough curve as shown in the **FIG 3**. It shows typical breakthrough curve. The time at which the breakthrough curve first begins to rise appreciably is called the break-through point and the time at which the outlet effluent concentration essentially reaches that of the inlet pollutant concentration is called the exhaustion point. The shape of the mass transfer zone depends on the adsorption isotherm, flow rate, and the diffusion characteristics. Usually, the shape is to be determined experimentally.

In designing a column, the length of the adsorption zone represents the minimum bed depth needed to produce a low solute concentration in the effluent. The actual bed depth is usually greater than this minimum to allow a reasonable operating period between regeneration of the adsorbent.

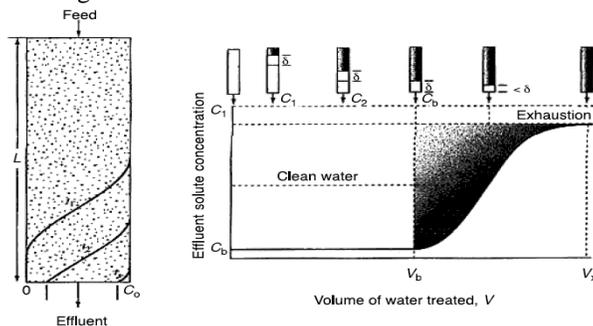
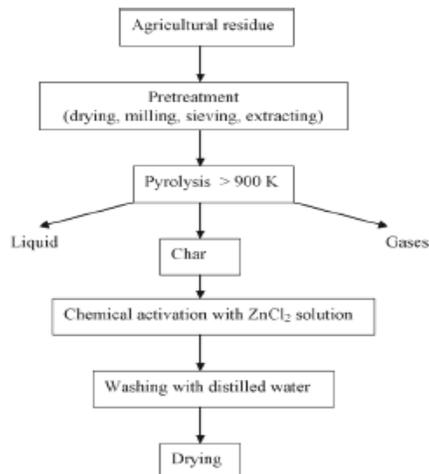


Fig 3.Typical Breakthrough Curve Showing The Movement Of Mass Transfer Zone.

Adsorbents:

Activated carbon is to be prepared by making a char from the organic materials such as coconut shell or walnut hulls or jute. The char is produced by heating the base material to a red heat (less than 700°C) in a retort to drive off the hydrocarbons, but with an insufficient supply of oxygen to sustain combustion as shown in **FIG. 5**.

FIG. 5 Process Flow Diagram Of Preparation Of Activated Carbon From Agricultural Waste



Analytical Procedure

The analysis involves evaluate the effects of various operating parameters and establish the equilibrium constants by Batch adsorption studies. The tests are to be conducted in temperature controlled conditions in various concentrations, maintaining agitation speed for a fixed time up to equilibrium. The supernatant liquid is to be analyzed for dyes spectrophotometrically. Adsorption equilibrium data is to be typically plotted in the form of an adsorption isotherm to define the various constants. The effects of various operating parameters to be studied are:

1.EFFECT OF CONCENTRATION AND TIME

The initial adsorbate concentration and time of between the adsorbate and adsorbant are significant importance as the rapid uptake of pollutants and establishment of equilibrium in a short period signifies the efficiency of the adsorbent for its use in the waste water treatment.

2.EFFECT OF pH

The pH at which adsorption is carried out shows a strong influence on the extent of adsorption. This is partly due to the fact that hydrogen ions themselves are strongly adsorbed partly because the pH influences the ionization, and adsorption of many compounds. Organic acids are more absorbable at low pH whereas, the adsorption of organic bases is favored at high pH. The optimum pH for any adsorption process must be determined by laboratory testing.

3.EFFECT OF AGITATION SPEED

If relatively little agitation occurs between the adsorbent particle and the fluid, the surface film of liquid around the particle will be thick and film diffusion will be the rate limiting step. If adequate mixing is provided the rate of film diffusion will increase to the point that pore diffusion becomes the rate-limiting step.

4.EFFECT OF TEMPERATURE

The temperature at which an adsorption process is conducted will affect both the rate and the extent of adsorption. Adsorption rate increases with increase temperature. However, since adsorption is an exterior process, the adsorption will increase at lower temperature and decrease at higher temperatures.

For evaluating the above parameters effluent solutions of different concentrations were agitated with known amounts of adsorbents till equilibrium is attained.

Adsorption Kinetics

Kinetics of adsorption is quite significant as it decides the residence time of adsorbate at solid solution interface and helps in determining the rate of the adsorption process. This study includes

1. Adsorption Rate study and
2. Intra-particle diffusion study.

1. Adsorption Rate study :

The review of the literature reveals that rate of uptake of adsorbate by adsorbent is of significant importance for the control of efficiency of effluent treatment and for designing the adsorption system. The rate equation for transfer of adsorbate from fluid to solid phase depends upon the nature of interaction between adsorbate and adsorbent process. The rate constants of the adsorption to be determined using first order rate expression proposed by Lagergren:

$$\text{Log}(q_e - q) = \text{log } q_e - K_{ad} / 2.303 \cdot t$$

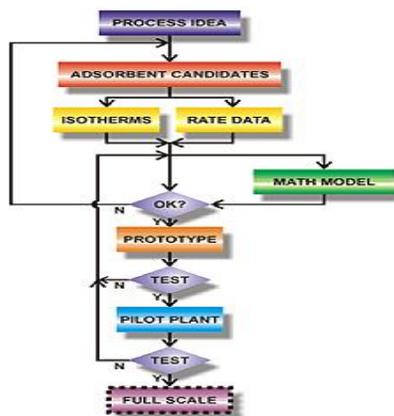
Where, K_{ad} is the rate constant for adsorption

The straight line plots of $\text{Log}(q_e - q)$ versus time t for adsorption of dyes are to be plotted to check for validity and check if it obeys first order rate kinetics. The rate constant can be determined from the slopes of the plots.

2. Intra-particle Diffusion study :

The kinetics of solute transport from solution phase to the surface of adsorbent particles is either by film diffusion, pore diffusion, pore surface diffusion or the combined effect of the above. The analysis involves plots of adsorbate and adsorbed per unit mass of the adsorbent versus square root of time. The plots should be linear and pass through the origin. The profile of the graph indicates whether the process is controlled by film diffusion or intra-particle.

FIG 5. Flow-Chart Presentation Of Adsorption Kinetics



Thermodynamic Parameters

The adsorption reaction at any interface between two phases can be regarded as an equilibrium process. The point of equilibrium is being dictated by the relative energies. These energy values can be defined in terms of thermodynamic parameters like free energy, enthalpy and entropy changes.

The free energy of an adsorption, considering the adsorption equilibrium constant K_a is given by the following equation:

$$\Delta G^\circ = -RT \ln K_a$$

Where,

ΔG° is free energy, standard free energy change (J/mol), R is the universal gas constant (8.314 J/mol K), and T is the absolute temperature (K).

The literature reveals that the negative values of **free energy** indicate the spontaneous nature of adsorption. Thermodynamic parameter, **free energy** could be calculated from adsorption equilibrium constant obtained from langmuir isotherm. The value of **enthalpy and entropy** could be obtained from the slope and intercept of the relationship between **free energy** and reaction temperature.

$$\Delta H^\circ = R(T_2/T_1/T_2 - T_1) - \ln(K_2/K_1)$$

$$\Delta S^\circ = (\Delta H^\circ - \Delta G^\circ) / T$$

If the value of **enthalpy** is positive, indicates that the adsorption reaction is endothermic. Thermodynamic considerations of an adsorption process are necessary to indicate whether the process is spontaneous or not. The free energy change is an indication of spontaneity of a chemical reaction and therefore is an important criterion for assessing the adsorption process. The enthalpy and free energy changes indicate exothermic nature of adsorption. The entropy change suggests the randomness in the adsorption.

Implications

From both environmental and economical considerations, converting these waste materials into more valuable products such as activated carbons for water treatment like dye removal is an attractive option, as it offers a solution to two environmental issues simultaneously. This paper aims to highlight the feasibility of converting these non-conventional waste materials to activated carbon for removing dyes from wastewater.

The main aim is to use live waste water from one of the industries say textile or paper and pulp industry for the removal of contaminants for the utilization, reuse and recycle to meet the requirement of the non potable water as a solution for water scarcity in the industrial sector.

References

1. Metcalf and Eddy waste water treatment disposal and reuse 4th edition, Tata McGrawHill publication.
2. Saini D. R. Feb 2010 Activated carbon adsorbant for pollution control Journal of Indian association for environmental management, vol-37-1, P 51-64.
3. Industrial Waste water Treatment, A.D.Patwardhan, Prentice-Hall, New Delhi.
4. Ayhan Demirbas, Agricultural based activated carbons for the removal of dyes from aqueous solutions: A review Journal of Hazardous Materials 167 (2009) 1–9
5. Ö. Yavuz*, A. H. Aydin, Removal of Direct Dyes from Aqueous Solution Using Various Adsorbents Polish Journal of Environmental Studies Vol. 15, No. 1 (2006), 155-161.
6. S. Saiful Azhar, Dye Removal from Aqueous Solution by using Adsorption on Treated Sugarcane Bagasse American Journal of Applied Sciences 2 (11): 1499-1503, 2005
7. Manfred Clara, Birgit Strenn, Ernis Saracevic, Norbert Kreuzinger Adsorption of bisphenol-A, 17 β -estradiol and 17 α -ethinylestradiol to sewage sludge Chemosphere 56 (2004) 843–851.
8. Y.C.Sharma and S.N.Kaul, Adsorption of Methylene Blue from Aqueous Solutions by an Economically Viable Indigenous Activated Carbon. Environmental Sanitation
9. Shikha Chandra. Gaikwad.S, Decolorization of Pulp-Paper Mill Effluents by Endophytic strains Bacillus spp and Pseudomonas spp of Mimosa pudica, using Batch Reactor.
10. Ageetha and P. Sivkumar Adsorption of Cr (VI) and Pb (II) from Aqueous Solution Using Agricultural Solid waste, Journal of Environ. Sciences and Engg. Vol- 51, No-2, P 151-156, April.
11. Suresh K, Borah M Adsorption of Nickel by Bentonite Clays: A Comparative Study Journal of Environmental Sci. and Engg. Vol 51 – 2, 133-136, April 2009.
12. S.Renganathan, Kinetic studies on sorption of Basic Dye using EICHORNA CRASSIPES, 13.Y.C. SHARMA, UMA S.N. UPADHYAY and F. GODE Adsorptive Removal of Basic Dye from water and wastewater by Activated Carbon, Volume 4, Number 1: 21-28, January-April, 2009, Journal of applied sciences and Journal of Environmental Sci. and Engg.
14. G.H.Sonavane, V.S.Shrivastav, Removal of Basic Dye (Methylene Bl) from Aqueous Solution by Adsorption, Journal of Environmental Sci. and Engg. Vol 51 – 1, 45-52, Jan 2009.
15. Yao-hui Huang, Chan-Li Hsueh, Chun-Ping Huang, Liang-Chih Su, Adsorption thermodynamics and kinetics studies of Pb(II), Separation and Purification Technology. Vol 55 (2007) 23-29.

Credit Card Fraud: Bang in E-Commerce

Khyati Chaudhary¹ Bhawna Mallick²

^{1,2}Galgotias College of Engg. & Technology,
Greater Noida

Abstract

Recent decades have seen a gigantic expansion in the use of credit cards as a true transactional medium. Data mining is rising as one of the chief features of many homeland security\ initiatives. Often, it is used as a means for detecting fraud, assessing risk, as well as product retailing. Data mining is becoming increasingly common in both the private as well as public sectors. Data mining involves the use of data analysis tools to find out formerly unknown, believable patterns and relationships in large data sets. Credit card offers a number of secondary benefits unavailable from cash or checks. Credit cards are safer from theft than is cash. Fraud detection involves monitoring the behavior of populations of users in order to estimate, detect or avoid unwanted behavior. In this paper, we studied various factors required to distinguish transaction and characterizes factors affecting fraud detection with fraud prevention techniques.

Keywords: E-commerce, Credit Card, Fraud detection, Fraud Prevention

1. Introduction

Data mining involves the use of complicated data analysis tools to discover previously unknown, valid patterns and relationships among large data sets. These tools can include mathematical algorithms, statistical models, and machine learning methods (such as Neural Networks or Decision Trees). Consequently, data mining comprises of more than collection and management of data, it also includes analysis and prediction. Data mining can be performed on data represented in textual, quantitative or multimedia forms. Data mining applications can use a range of parameters to observe the data. This includes association, classification, sequence or path analysis, clustering and forecasting. When using normal measure, detection of credit card fraud is a tricky task. Therefore, the development of the credit card fraud detection model has become of much important, whether in the academic, organization or business community recently. Proposed models or existing models are mostly statistics-driven or Artificial Intelligent-based (AI), which have the theoretical advantages in not imposing artificial assumptions on the input variables [1].

Credit Cardholders have many beneficiary schemes as to hold interest-free balances for almost two months with “grace-period”. Suitable information on fraudulent activities is tactical to the banking industry. Banks have huge databases. Extraction of important business information can be done from these data store. Data stores have some patterns into clusters that are natural to the input data. Concept of fraud detection has been laid on data mining techniques which include association rules, clustering and classification. The chief point of research on fraud detection has been focused on pattern matching in which abnormal patterns are identified from the normal ones. Popularity of online shopping is growing day by day on high pace. Nowadays, Credit card is the most popular mode of payment (59 percent), Germany and Great Britain have the largest number of online shoppers. Usually, Retailers like Wal-Mart handle much larger number of credit card transactions online as well as regular purchases [2]. There are numerous options for handing credit card payments on the Internet, as the processing of credit card transactions is generally independent of the type of e-commerce exchange. While a huge portion of e-commerce would consist of credit card purchases, like regular or often. It is more important for businesses and organizations that rely upon on income from e-commerce to know the options available as well as costs linked with credit card transaction processing on the Internet. No one has any hint about transaction being processed are whether a fraudulent transaction or legitimate which has passed the prevention mechanisms. Therefore, the goal of the fraud detection systems is to pre-determine every transaction for the possibility of being fraudulent regardless of the prevention mechanisms and to categorize transaction as fraudulent ones as early as possible after the fraudster has begun to commit a fraudulent transaction. Credit card fraud detection is a tremendous task but also trendy problem to solve. Many fraud detection systems estimate the transactions and generate a suspicion score (commonly a probability between 0 and 1) which shows the chances of that transaction to be fraudulent. Computational procedures of these scores are applicable to the techniques used to build the model(s) in the fraud detection systems. These corresponding scores are used with a predefined threshold value to differentiate between fraudulent transactions from the legitimate ones easily [3].

Introduction of new technologies such as telephone, automated teller machines (ATMs) and credit card systems have enlarged the amount of fraud loss for many banks. Analyzing whether each transaction being processed is legitimate or not is very expensive is another task to determine transaction genuinely. Further, if we check them in all transactions & confirm whether a transaction was done by a client or a fraudster by phoning all card holders is cost

prohibitive. Fraud prevention by automatic fraud detections mechanism can be applied where the well-known classification methods can be identified, where pattern recognition systems has key function. One can learn from fraud happened in the past and categorize new transactions easily. Recently, perhaps the most frequently used technique is Neural Networks in credit card business.

Credit Card Fraud Detection domain presents a number of challenging issues for data mining as well:

- There are millions of credit card transactions processed each day. Mining of such massive amount of data requires highly efficient techniques that scale data efficiently.
- Highly skewed-data,
- Each transaction record has a different dollar amount and there is a chance of variable potential loss.

II. The Fraud Detection Problem

Problem of detecting fraudulent transactions occurs after they have been focused to fraud prevention methods and relevant processes. There is immense literature on wide range of security methods to look after transactions from unauthorized use or exposure of their private/secure information and consequent valuable resources. Still, fraudsters find a mode through which many witty means of circumventing a countless prevention techniques.

On the other side, many transaction media such as ATM, bank cards or debit cards, require the use of pins, passwords, and in some cases “biometrics” to authenticate the legitimate owner. Credit cards create fascinating problems since generally no pin is required for their use; only the name, expiration date and account number is required. Popular means of criminally transacting with credit cards is by stealing someone’s identity & in some cases, creating a new fake identity. Therefore, fraudulent electronic transactions (E-transaction) by credit card are the key problem. Credit cards need not be necessarily physically obtainable to transact and over the internet they can be used to fraudulently transact web better and heavier losses for banks and their customers if caught by fraudsters. The chief idea in fraud detection is that fraud may be detected by noticing significant deviation from the “normal behavior” of a customer’s account. That is why; behavior of an account can thus be used to protect that account. Currently banks have come to realize that a fused, global approach is mandatory to detect fraud, involving the periodic sharing with each other of information about attacks.

iii. Different Type of Fraud Techniques

There are many ways in which fraudsters bring out a credit card fraud. As the technology changes, so does the technology of fraudsters varies and thus the mode in which fraudsters go about carrying out fraudulent activities. Frauds can be broadly categorized into three stages i.e., traditional card related frauds, merchant related frauds and Internet frauds. Different types of methods for committing credit card frauds are:

A. Merchant Connected Frauds (MCF)

Merchant connected frauds are being committed either by owners of the merchant firm or their employees. Different types of frauds initiated by merchants are:

i. Merchant Collusion: When merchant owners or their employees plan to commit fraud using the cardholder accounts or by using the personal information [4].

ii. Triangulation: Triangulation is among the type of fraud which is done and operates from a website. Triangulation includes products or goods that are offered at heavily discounted rates and are being shipped before payment. The phenomenon initiate by the customer while browse the site and if he/she likes the product he/she place the online information such as name, address and valid credit card details to that particular site. However, when the fraudsters get these details, they order goods from a legitimate site using stolen credit card details. Further, after this, fraudster use credit card information for purchasing the products/goods.

B. Internet Associated Frauds (IAF)

Internet is the foundation for the fraudsters to make the frauds in the simple and the easiest technique(s). Currently, fraudsters have begun to operate on a actually transactional level. Internet has become a new worlds market, capturing consumers from countries around the world along with the development of trans-border, economic and political spaces. Some of the most frequently used techniques in Internet fraud are:

i. Site cloning: Site cloning is the process where fraudsters close whole site or simply the pages from which the customer made a purchase. There is no option left with the customers to believe that they are not dealing with the company that they wished to purchase goods or services from because the pages that they are viewing are somehow

matching to those of the real site. Further, cloned site will receive these details and propel the customer a receipt of the transaction through the email just as the real company would do.

ii. False merchant site(s): Actually, some sites offer a contemptible service for the customers. Site(s) requests the customer to fill his/her complete details such as name and address to access the webpage where the customer gets his requisite products. Numerous sites claim to be free but require a valid credit card number to verify an individual's age. In this mode sites collect as many as credit card details. Sites are generally part of a larger criminal network that either uses the details it collects to raise revenues or sells valid credit card details to small fraudster(s).

iii. Credit card generators (CCG): CCG are computer programs that create valid credit card numbers and expiry dates. CCG generates lists of credit card account numbers from a single account number. CCG software works by using the mathematical Luhn algorithm that card issuers use to produce other valid card number combination(s).

iv. Lost/ Stolen Cards: When individual loses his card or a card is stolen by someone or when a legitimate account holder receives a card and loses it or someone else steals the card for criminal purposes. This is the simplest way for the fraudsters where they get the information of the cardholders without investing on the modern technology. It is possibly the hardest form of traditional credit card fraud to embark upon.

v. Account Takeover: Fraud occurs when the valid customer's personal information is taken by the fraudsters. In this fraudster(s) takes control of a legitimate account by providing the customer's account number or the card number. Fraudster then acquaintances the card issuer as the genuine cardholder to ask the mail to redirect to a new address. In some cases, the fraudster reports card lost and asks for a replacement to be sent.

vi. Cardholder-Not-Present (CNP): CNP transactions are performed only on the Internet in such kind of frauds neither the card nor the cardholder is present physically at the point-of-sale. This would have many forms to commit the fraud as there are many types of transactions such as orders made over the phone or Internet, by mail order or fax.

In such transaction(s), retailers are not capable to physically check the card or the identity of the cardholder, which makes the user unknown and able to disguise between their true identities. In this, details of the credit card are usually copied without the cardholder's awareness. Fraudulently obtained card details are usually used with fictitious personal details to make fraudulent CNP purchases. Security Code imprinted on the back of cards can help in prevention of fraud where card details have been obtained but when the card is stolen it won't be helpful. This is the promising method for fraud prevention.

vii. Fake and Counterfeit Cards: Another type of fraud where the formation of counterfeit cards, together with lost or stolen cards poses maximum threat in credit card frauds. Fraudsters are always in search of new and more original ways to create counterfeit cards.

viii. Erasing the magnetic strip: In this type of the fraud, the fraudsters erase the magnetic stripe by using the powerful electro-magnet. Fraudster then tampers with the facts on the card so that they match the details of a valid card which they may have attained. The cashier will then carry on to manually input the card information into the terminal. This type of fraud has high risk because of the cashier would look at the card closely to read the numbers.

ix. Creating a fake card: In present scenario, we have refined machines where we can create a fake card from using the scratch. It is the expected fraud, though fake cards require a lot of effort and skill to produce it. Present cards are having many security features, all designed to make it tricky for fraudsters to make good quality fraudulent. Furthermore, after introducing the Holograms in the credit cards it makes very complicated to forge them effectively.

x. Skimming: Another kind of fraud being committed is skimming which is fast emerging as the most popular form of credit card fraud. Mostly, fraud cases of Counterfeit fraud involve skimming. It is a method where the actual data on a card's magnetic stripe is electronically copied onto another. Fraudster(s) does this even as the customer is waiting for the transaction to be validated through the card terminal. Card holder doesn't know about this activity and it is very difficult for customer(s) to identify. In some of the cases, details obtained by skimming are used to carry out fraudulent card not-present (CNP) transactions by fraudsters.

xi. Phishing: Phishing is a type of fraud planned to steal a person's identity. It is generally committed via spam e-mail or pop-up windows. It is the phenomenon which works by a wicked person sending lots of false e-mails. E-mails received looks like they come from a website or company you trust. Message tells users to provide the company with your personal details including your payment card details. These companies can also claim that the reason for this is a

database crash or the like. For the fraudster might put a link to a website that look exactly like the real one but is in fact a scam site by making the e- mails look even more authentic. These copies are often known as “spoofed websites”.

Iv. Literature Review

4.1 Overview on Credit Card

Credit card frequently used as a necessary mode of payments in today’s society. People used credit card for a range of reason such as obtaining credit facility, cash advance, easy payment, charge card. There are some controversial issues that have been addressed not only in terms of the numbers of credit flooding the nation’s economy, but the amount transactions that end up with payment default and the numbers of credit card fraud as been recorded which endangered the economy should be seriously paying attention [5]. But because of the advances and changing behavior in purchasing activities has considerably contributed to the diffusion of credit card as becoming more significant and applicable in maintaining the purchasing activities. Based on the judgment, it is stated that there is positive connection between usage rate and income. The fact that was frequently stated, most of the card issuers normally allowance a higher credit limit among the higher income group. Lastly, it was stated that higher income clients are the main targets for the credit card issuers. Bulky purchase allows people not to carry cash and is useful in Internet purchases and rental collateral. But the crisis is that it is improper on religious grounds because there will be an interest payments made when the outstanding balance is not repay in full.

In the card issuer’s point of view, numerous problems occurred. Industry is growing and this research would be helpful for the banks offering the credit cards to focus on quite a few factors that pressurize the credit card holders in choosing their preferred credit cards.

Buttafogo began the workshop with a operational definition of credit card fraud as: “Unauthorized account activity by a person for which that account was not planned. Operationally, this is an event for which action can be taken to stop the neglect in progress and incorporate risk management practices to protect against similar actions in the future.” He then described the range of fraudulent activities observed in the industry.

The Internet and the ambiguity associated with card not present (CNP) transactions current unique fraud management challenges. Authentication of the cardholder is a primary requirement in managing fraud on the Internet. There are no commonly accepted solutions. As a result, credit card fraud on the Internet is significantly greater than in the physical, or even, phone environments [6].

Data mining contributed towards fraud detection. Data mining has various categories through which various operations has been performed. Data Mining can be mainly classified into the following categories:

1) **Association rule mining** which uncovers interesting association patterns among a large set of data items by showing attribute- value circumstances that occur together regularly. Market basket analysis is a classic example in which analyzing purchasing habits of customers by finding associations between different items in customers’ “shopping baskets.”

2) **Classification and prediction** is the process of identifying a set of ordinary features and models that to explain and distinguish between classes or concepts. Models are used to guess the class of objects whose class label is unknown. For example, Bank which may classify a loan application as either a fraud or a potential business using models based on uniqueness of the applicant. A huge number of classification models have been developed for predicting future trends of stock market indices and foreign exchange rates (FRI).

3) **Clustering analysis** segments a bulky set of data into subsets or clusters. In this, each cluster is a collection of data objects that are similar to one another within the same cluster but dissimilar to objects in other clusters. Additionally, objects are clustered based on the principle of maximizing the intra-class resemblance while minimizing the inter-class similarity. For example, clustering techniques can be used to recognize stable dependencies for risk management as well as investment management.

4) **Sequential pattern and time-series mining** looks for patterns where one value leads to another later value. Example, after the inflation rate increases, the stock market is likely to go down.

V. Impact of Credit Card Frauds

Fraudulent activity on a card affects every person that is the cardholder, the merchant, the acquirer as well as the issuer. In this section, we analyses the impact that credit card frauds have on all the company involved in transacting business through credit cards.

5.1. Impact of Fraud on Cardholders

Cardholders are the least impacted party due to fraud in credit card transactions as consumer responsibility is limited for credit card transactions by the legislation existing in most countries. This is true for both card-present (CP) as well as card not-present (CNP) scenarios. Banks even have their own principles that limit the consumer's problem to a greater degree. Banks also have a cardholder protection policy that covers losses of the cardholder. The cardholder has to just report doubtful charges to the issuing bank, which in turn investigates the issue with the acquirer and merchant and further processes chargeback for the disputed amount [7].

5.2. Impact of Fraud on Merchants

Merchants are the most exaggerated party in a credit card fraud mainly in the *card-not-present (CNP)* transactions, as they have to accept full accountability for losses due to fraud. Every time a legitimate cardholder disputes a credit card charge, the card-issuing bank will send a chargeback to the merchant, reversing the credit for the transaction. The merchant does not have any physical confirmation available to challenge the cardholder's dispute then it is almost impossible to reverse the chargeback. As a result, the merchant will have to completely absorb the cost of the fraudulent transaction. The cost of a fraudulent transaction consists of:

1. **Cost of goods sold:** Since it is doubtful that the merchandise will be recovered as in a case of fraud, the merchant will have to write off the value of goods involved in a fraudulent transaction. The impact of this kind of losses will be maximum for low-margin merchants.

2. **Shipping cost:** More relevant in a *card-not-present (CNP)* scenario. While the shipping cost is usually bundled in the value of the order, the merchant will also need to take up the cost of shipping for goods sold in a fraudulent transaction. Besides, fraudsters normally request high-priority shipping for their orders to allow fast completion of the fraud, resulting in high shipping costs.

3. **Card association fees:** Visa and MasterCard have put in place quite strict programs that fine merchants generating extreme charge backs. Usually, if a merchant exceeds well-known chargeback rates for any three-month period the merchant could be penalized with a fee for every chargeback.

4. **Merchant bank fees:** In addition to the penalties charged by card associations, the merchant has to pay an extra processing fee to the acquiring bank for every chargeback being transacted.

5. **Administrative cost:** Every transaction that generates a chargeback requires major administrative costs for the merchant. Usually each chargeback requires one to two hours to process. Because for processing a chargeback requires the merchant to receive and research the claim, contact the consumer, and respond to the acquiring bank or issuer with adequate documentation.

6. **Loss of Reputation:** Maintaining reputation and goodwill is very important for merchants.

5.3. Impact of Fraud on Banks (Issuer/Acquirer)

On the basis of the scheme rules defined by both MasterCard and Visa, it is sometimes possible that the Issuer/Acquirer bears the costs of fraud. Still in cases, where the Issuer/Acquirer is not bearing the direct cost of the fraud, there are some indirect costs that will finally be borne by them. As like in the case of charge backs issued to the merchant, there are also administrative and manpower costs that the bank has to sustain. The issuers and acquirers also have to make vast investments in preventing frauds by deploying complicated IT systems for detection of fraudulent transactions.

Vi. Fraud Prevention

Along with all the negative impacts of fraudulent credit card activities such as financial and product losses, fines, loss of reputation and technological advancements in perpetrating fraud, it is simple for merchants to feel offended and helpless. Though scientific advancements, preventing fraud have started to assure for fraud prevention.

Merchants and Acquirers as well as Issuers are creating original solutions to bring down on fraudulent transactions and lower merchant chargeback rates. Key challenge with fraud prevention is the long time pause between the time fraudulent transaction occurs and the time when it gets detected that is the cardholder initiates a chargeback. Analysis results shows that the average lag between the transaction date and the chargeback notification could be as high as 72 days [8]. It means that if no fraud prevention is in place, one or more fraudsters could easily produce significant damage to a business before the affected stakeholders even realize the problem.

6.1. Fraud Prevention Technologies

While fraudsters are using sophisticated methods to achieve access to credit card information and perpetrate fraud, new technologies are being available to assist merchants to detect and prevent fraudulent transactions as well. Fraud detection technologies enable merchants and banks to perform highly automated screenings of incoming transactions and flagging suspicious transactions. Although none of the tools and technologies offered here can by itself eliminate fraud, each technique provides incremental value in terms of detection ability. Various fraud prevention techniques are as follows:

6.1.1. MANUAL REVIEW (MR)

Manual review consists of reviewing every transaction manually for signs of fraudulent activity and involves a exceptionally high level of human involvement. It can prove to be very expensive, as well as time consuming. Furthermore, manual review is incapable to detect some of the more common patterns of fraud such as use of a single credit card multiple times on multiple locations (physical or web sites) in a short distance.

6.1.2. ADDRESS VERIFICATION SYSTEM (AVS)

AVS is applicable in *card-not-present* (CNP) scenarios. AVS matches the first few digits of the street address and the ZIP code information given for delivering the purchase to the corresponding information on record with the card issuers. Code on behalf of the level of match between these addresses is returned to the merchant.

6.1.3. CARD VERIFICATION METHODS (CVM)

The Card Verification Method (CVM) consists of a 3- or 4-digit numeric code printed on the card but is not imprinted on the card and is not accessible in the magnetic stripe. The merchant can request the cardholder to provide this numeric code in case of *card-not present* (CNP) [9] transaction and submit it with authorization. Main idea of CVM is to make sure that the person submitting the transaction is in control of the actual card, since the code cannot be copied from receipts or skimmed from magnetic stripe. CVM provides some protection for the merchant(s) it doesn't defend them from transactions placed on physically stolen cards. Moreover, fraudsters who have provisional possession of a card could read and copy the CVM code.

6.1.4. NEGATIVE AND POSITIVE LISTS (N/P L's)

Negative List (NL) is a database used to recognize high-risk transactions based on specific data fields. For example, negative list would be a file containing all the card numbers that have formed charge backs in the past, used to evade further fraud from repeat offenders. Likewise a merchant can put up negative lists based on billing names, street addresses, emails and internet protocols (IPs) that have resulted in fraud or attempted fraud, effectively blocking any more attempts. An acquirer could form and maintain a list of high-risk countries and decide to review or confine orders originating from those countries. Positive files/List (PL) represents a vital tool to prevent unnecessary delays in processing valid orders.

6.1.5. PAYER AUTHENTICATION (PA)

Payer Authentication is a rising technology that promises to fetch in a new level of security to business-to-consumer (b-c) internet commerce. First implementation of this kind of service is the Verified by Visa (VbV) or Visa Payer Authentication Service (VPAS) program. This program is based on a Personal Identification Number (PIN) linked with the card, comparable to those used with ATM cards and a secure direct authentication channel between the consumer and the issuing bank [10]. PIN is issued by the bank when the cardholder enrolls the card with the program and will be used completely to authorize online transactions. When registered cardholders check out at a participating merchant's site, they will be driven by their issuing bank to provide their password. Likewise, once the password is verified, the merchant may complete the transaction and send the verification information on to their acquirer.

6.1.6. LOCKOUT MECHANISMS (LM)

Automatic card number generators signify one of the new technological tools normally utilized by fraudsters. These programs are easily downloadable from the Web and are able to produce thousands of 'valid' credit card numbers. The qualities of frauds initiated by a card number generator are as follows:

- (i) Multiple transactions with similar card numbers (for example, same Bank Identification Number (BIN))
- (ii) A huge number of declines.

Acquiring merchant/bank sites can put in place prevention mechanisms particularly designed to detect number generator attacks.

6.1.7. FRAUDULENT MERCHANTS (FM)

Both MasterCard and Visa publish a list of merchants who have been known for being occupied in fraudulent transactions in the past. These lists could provide functional information to acquirer right at the time of merchant recruitment preventing potential fraudulent transactions [11].

Vii. Conclusion

Credit card fraud has become more and more widespread in recent years. Building an accurate, efficient and easy-handling credit card risk monitoring system is one of the chief tasks for the merchant banks for improving merchants risk management level in an automatic, scientific and effective way,. In this era of digital world, credit card is of extreme importance to financial organizations, institutions and companies. As credit card becomes the most accepted mode of payment for both online as well as regular purchase, cases of fraud associated with it are also increasing. For the purpose of reducing the bank's risk, various techniques have been employed. In this study, we characterize various fraud commitment and prevention methods as well. However model has been proposed for credit card fraud detection in catching the fraudulent transactions.

References

1. A. Shen, R. Tong, and Y. Deng, "Application of classification models on credit card fraud detection," June 2007
2. Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K. Majumdar, "Credit Card Fraud Detection using Hidden Markov Model," *IEEE Transactions On Dependable And Secure Computing*, vol. 5, Issue no. 1, pp.37-48, January-March 2008.
3. Aihua Shen, Rencheng Tong, Yaochen Deng, Application of Classification Models on Credit Card Fraud Detection, 2007 IEEE.
4. Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster-Shafer theory and Bayesian learning," *Special Issue on Information Fusion in Computer Security*, Vol. 10, Issue no 4, pp.354- 363, October 2009.
5. CLIFTON PHUA1*, VINCENT LEE1, KATE SMITH1 & ROSS GAYLER2A Comprehensive Survey of Data Mining-based Fraud Detection Research
6. Pengyue J. Lin, Behrokh Samadi, Alan Cicolone, Daniel R. Jeske, Development of a Synthetic Data Set Generator for Building and Testing Information Discovery Systems, Proceedings of the Third International Conference on Information Technology: New Generations (ITNG'06)2006 IEEE.
7. Jon T. S. Quah and M. Sriganesh, Real Time Credit Card Fraud Detection using Computational Intelligence, Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 2007.Kou, Y., Lu, C.-T., Sirwongwattana, S., Huang, Y.-P.: Survey of fraud detection techniques. In: Proceedings of the 2004 IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan (2004).
8. Tej Paul Bhatla, Vikram Prabhu & Amit Dua "Understanding Credit Card Frauds," 2003.
9. Y. Sahin, E. Duman "Detecting Credit Card Fraud by ANN and Logistic Regression" 2011.
10. Raghavendra Patidar, Lokesh Sharma, "Credit Card Fraud Detection Using Neural Network", *International Journal of Soft Computing and Engineering (IJSCE)* June 2011.
11. Y. Sahin, E. Duman, Detecting Credit Card Fraud by ANN and Logistic Regression, ©2011 IEEE

Efficient Moving Object Detection Based On Statistical Background Modeling

Kusuma.U¹, S.T Bibin Shalini²

^{1,2}Dept. Electronics and Communication

Address AMC Engineering College, Bannerghatta road, Bangalore

Abstract— Tracking vehicles is an important and challenging problem in video-based Intelligent Transportation Systems, which has been broadly investigated in the past. A robust method for tracking vehicles is implemented in this thesis work.

The proposed algorithm includes three stages: object detection, counting and tracking. Vehicle detection is a key step. The concept of moving object detection is built upon the segmentation method. Background subtraction method is used in this work. According to the segmented object shape, a predict method based on Kalman filter is proposed. By assuming that the vehicle moves with almost a constant acceleration from the current frame to the next, a Kalman filter model is used to tracking and predicting the trace of a vehicle.

The model can be used in the traffic analysis as it is capable of tracking and counting multiple targets in a big area hence forming an effective, efficient, practical vehicle tracking system. The proposed method has been tested on few traffic-image sequences and the experimental results show that the algorithm is robust and can meet the requirement.

1. Introduction

Object tracking is the problem of estimating the positions and other relevant information like trajectory, shape, size and number of moving objects in an image sequence. So, a tracker assigns consistent labels to the tracked objects in different frames of a video. It has several important applications such as security and surveillance, annotation of videos, traffic management, motion-based video compression and interactive games.

Object tracking, in general, is a challenging problem. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of the object and the scene, nonrigid object structures, object-to-object occlusions, and camera motion. The complex task of tracking can be simplified by imposing constraints on the motion and/or appearance of objects. For the given video sequence, we assume that the object motion is smooth with no abrupt changes. The object motion is further constrained to be of constant velocity. Prior knowledge about the number and the size of objects, object appearance and shape, is also made use of to simplify the problem.

We are interested in detecting and tracking moving objects in video at low- to moderate- resolution and frame rate. We have developed a flexible tracking pipeline that allows us to investigate different combinations of foreground extraction, feature extraction and motion correspondence algorithms. In foreground extraction we have explored applications of [background subtraction techniques](#) and [salient region extraction](#). Our background subtraction research includes the investigation of the effectiveness of popular background subtraction techniques and the development of a new technique for background subtraction with foreground validation. We have also examined the application of salient region detection to extracting moving objects.

In feature extraction, we extract features such as the centroid, the size, and the average pixel intensity of each moving object. These are then used in tracking algorithms such as Kalman filters to track the objects from one frame to the next. Additional logic is incorporated for track maintenance to determine constraints such as the number of frames over which an object must be tracked successfully for it to be assigned a track and the number of frames over which no object is assigned to an existing track, making the track disappear.

In this thesis our approach uses frame differencing for background and fore ground separation. Blob detection for counting the no of detected objects. Centroid is the feature extracted and these centroid co-ordinates are used for tracking. Tracking is achieved using Kalman filter.

2. Problem Description

Given a video sequence (recorded), our goal is to detect if there are entities in that sequence that are changing their spatial position. When these entities are detected we should be able to estimate their spatial position (within certain limits) such that if certain events occur (an entity is temporarily disappears shortly from the sequence) we still want to be able to predict its trajectory and behaviour until the event will eventually stop.

We will make any assumption regarding the type of the scene recorded in the video sequence should posses a non complex fixed background. We are in need of a model that is capable in emulating the motion of an object in that particular frame of the video sequence with respect to the background and a model which is robust enough to withstand various influences exerted by the environment. In addition to this, our model will also have to take into account the noise injected by the capturing device.

At the end of the process, our model should provide us with an estimation of the trajectory of the object whose motion it tried to emulate as well as the number of moving entities detected.

3. Proposed Solution

We use background subtraction for segmentation and detection of objects and use Kalman Filter for tracking the objects.

3.1. Background Subtraction

Background subtraction is an approach for detecting changes in the background of the scene. It assumes an initially static scene where there is no motion and builds mean and standard deviation images for the scene during this phase. Since the input video to our system has no such initial frames with no motion, we initialise the mean image by taking the median of the first 100 frames and the standard deviation image as the difference between the frame and the initial median image averaged over the first 100 frames.

$$|I(x, y) - \text{MeanImage}(x, y)| > c \times \text{StdDevImage}(x, y) \quad (1)$$

Once we have our initial background images, we label a pixel as being in foreground if for any channel Eq. (1) holds true. We chose the constant $c = 2.1$ in Eq. (1). Hence we obtain a mask for the foreground pixels in the scene.

Another issue that needs to be addressed here is that the input

video in our case has some camera shake and the camera moves considerably in one or two frames. The sudden camera movement results in some background objects also coming into the foreground. Our approach to solving this problem was trying to absorb static objects into the background. Hence we maintained a counter for each pixel as to if some pixel was continuously in the foreground for a threshold number of frames, we removed it from the foreground mask. We chose this threshold to be 15 frames. Once we removed the pixel from the foreground mask, we updated our background mean and standard deviation images.

This was done as a running average between the previous mean image and the background in the current frame.

The background standard deviation image was updated similarly as a running average between previous standard deviation image and the difference between current frame and mean background image.

3.2. Detection of Objects

Background subtraction gives us candidate foreground pixels. We further need to cluster the foreground pixels into candidate objects. We initially use morphological operations of image opening and closing. Image closing operation fills the holes, if any, in some cluster of foreground pixels. Image opening operation removes potential noise foreground pixels that connect two clusters. After these operations we remove the foreground clusters having area less than 200 pixels and use connected component labelling which gives us potential objects in the scene. We maintain areas, bounding boxes and centroids of each such object in the foreground.

3.3. Tracking of objects - Kalman filter

Once we have detected some objects, we need to find

which of these objects correspond to previously detected objects and those which are detected for the first time. The tracking is done by fitting a discrete time linear dynamical system (LDS) to each object. Each object has an associated state which is a vector $[x, y, v_x, v_y]^T$, where x, y is taken as the center of the bounding box of the object and v_x, v_y are the velocities along x and y directions. We assume that the vehicles move with constant velocities and hence do not take into account the accelerations of the objects. This assumption is valid considering the fact that the velocities of objects in the video do not change considerably between two consecutive frames. The output of the LDS is the actual centers of the bounding boxes of the detected objects.

In principle, the Kalman filter may have applications both to discrete-time systems as well as continuous-time systems but the continuous domain is not to be discussed herein. Our system is treated in discrete-time due to the nature of the digital computing methods used, and the arrival of sensor data in a time-quantized manner. The purpose of the Kalman filter is to estimate the state $x \in R^n$ of a discrete-time process that is perturbed by some noise. The nature of that noise is crucial in the definition of the filter, and in its performance. All sensor and process noise is presumed to be zero-mean-Gaussian (0,). The system being filtered has inherent system properties described by a model. The evolution of the state is governed by the following discrete time stochastic difference equation:

$$X_k = A_k X_{k-1} + B_k U_{k-1} + W_{k-1}$$

Sensor data is used to correct this state, but measurement of this process may not necessarily be performed in the same space as the state itself. It is also assumed to have had some noise introduced in the form of sensor noise. We define the measurement or otherwise called the observation of $Z_k \in R^n$ as:

$$Z_k = H_k X_k + v_k$$

Kalman filter is divided into two distinct stages : Prediction stage and update stage The ability of these two stages to be partitioned forms an important implication. If sensor data is not available for a stretch of time, or if two or more sensors are operating at different frequencies, the filter may continue to operate, giving an estimate of the true state. The prediction stage consists of the dynamics themselves, along with the incorporation of another matrix that represents the degree to which the state can be trusted. Every time the system predicts what the state should be, it also predicts how accurate the state estimation is. On a prediction, the state covariance matrix increases, while every time sensor updates the state, it is reduced. The filtering equations themselves are presented below with annotation as to their significance

Prediction Stage:

Predicted State

$$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_k U_k$$

Predicted estimate covariance

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k$$

Update Stage

- Innovation (error)

$$\tilde{y}_k = z_k - H_k \hat{x}_{k|k-1}$$

Innovation Covariance

$$S_k = H_k P_{k|k-1} H_k^T + R_k$$

Kalman Gain (weighting factor)

$$K_k = P_{k|k-1} H_k^T S_k^{-1}$$

Updated state estimate x

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k$$

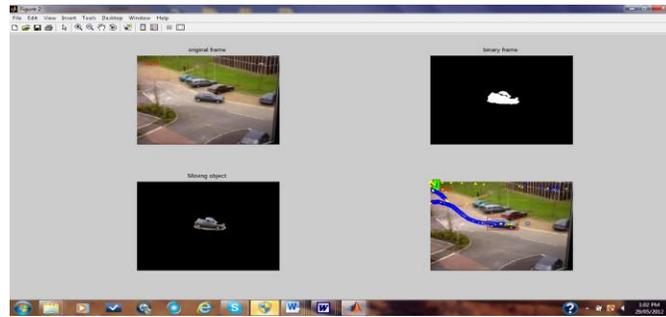
Updated estimate covariance

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}$$

While the derivations themselves will not be presented in this thesis, the implementation of these equations is very straightforward in MATLAB.

Certain initializations are required for the state vector and matrices defined above. If the filter began operating with the wrong assumptions, many time-steps may occur before the filter starts reflecting what the internal state truly is, due to the internal dynamics. Since the tracking example presented assumes the system is initialized while the vehicle is standing still, the state vector is initialized to zero.

4. Results



Conclusions

This work proposes an efficient algorithm for detecting a moving object using background elimination technique. Initially we compute the frame differences (FD) between frames F_i and F_{i+k} . The frame differences obtained are then compared with one another which help in identifying the stationary background image. The moving object is then isolated from the background. In the post processing step, the noise and shadow regions present in the moving object are eliminated using a morphological gradient operation that uses median filter without disturbing the object shape. This could be used in real time applications involving multimedia communication systems.

This paper also discusses an application system of traffic surveillance. Here we develop an algorithm to track and count dynamic objects efficiently. The tracking system is based on a combination of a temporal difference and correlation matching. Tracking system uses Kalman filter to predict the path of the moving object.

The system effectively combines simple domain knowledge about object classes with time domain statistical measures to identify target objects in the presence of partial occlusions and ambiguous poses in which the vehicles are moving. Ground clutter is effectively rejected. The experimental results show that the accuracy of counting vehicles is good although the vehicle detection was computational complexity of our algorithm is linear to the size of a video clip and the number of vehicles tracked. As a future work a combination of higher dimensional features with some additional constraints may be tried so that adverse effects of some features can be compensated by contribution of others.

References

- [1] <http://www.darpa.mil/grandchallenge>.
- [2] <http://www.ieeexplore.org>
- [3] www.google.com
- [4] L.Li I.Y.H Gu, M.K.H. Leung, Q. Tian, "Adaptive background subtraction based on feedback from fuzzy classification"
- [5] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing"
- [6] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing"
- [7] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russel. "Towards robust automatic traffic scene analysis in real-time."
- [8] Chris Stauffer W.E.L Grimson "Adaptive background mixture models for real-time tracking"
- [9] Nir Friedman and Stuart Russell. "Image segmentation in Video sequences: A probabilistic approach,"
- [10] Chris Stauffer W.E.L Grimson, "Adaptive background mixture models for real-time tracking"

Designing Continuous-Time Observers for Linear Hybrid Systems with Application to Three Tank Model

¹Mohammad Amin Zahedi Tajrishi, ²Behrouz Rezaie, ³Reza Ghaderi

^{1,2,3}Department of Electrical and Computer Engineering,
Babol University of Technology, Babol, Iran

Abstract

In this paper, two methods are studied and compared for designing continuous-time observers in linear hybrid systems. The methods are proposed based on Kalman filter and Luenberger observer design methods in linear systems. Three tank system is also considered as a hybrid system with unknown state variables. The linear hybrid model is obtained by linearizing the model of three tank system around its equilibrium point. In addition, it is also assumed that a state feedback controller is utilized for controlling the linearized hybrid system in order to stabilize the system. Simulation results depict the effectiveness and applicability of the proposed methods for state observation in hybrid systems. Comparison of two proposed methods shows that the optimal observer obtained using Kalman filter design method has better performance than Luenberger method.

Keywords: Linear hybrid system, Three tank system, Continuous-time observers, Kalman filter, Luenberger observer.

Introduction

In recent years, the study and control of hybrid systems has been considered by many researchers. Such systems are modeled in a special framework provided for describing processes based on continuous dynamics, discrete dynamics and logic rules. Same as many real systems, the states of these systems may be inaccessible and thereby the control of them is not feasible in this condition. The use of observer is therefore required for controlling hybrid system with unknown state variables such that the estimated values are used instead of state variables which are not reachable. There are well-known methods for designing observers in linear systems such as Kalman filter [1], Luenberger observer [2], alpha-beta [3] and Wiener filter [4]. For hybrid systems, the state variables in continuous or discrete parts or both of them may be unavailable and thus continuous-time or discrete-time observers should be designed for each part of hybrid systems. In other words, hybrid systems have continuous and discrete observers. In such systems, discrete dynamic is usually raised from a logic or switching law which is based on design criteria and can be available. Thus, continuous state estimation associated with the continuous dynamics of hybrid systems is more important for control design purpose.

In [5], the authors revealed how to choose observer parameters so that its estimation error is independent of arbitrary plant disturbances. Contrary to previous researches, in [5], the plant disturbances were not assumed to have any mathematical structure. Also, conditions under which an observer can reject plant disturbances were presented. In [6], Gauthier construct an observer for nonlinear and autonomous systems. In [7], an estimation approach was presented for a class of hybrid discrete-time linear systems using Luenberger observers. The proposed Luenberger observer for such a kind of systems relied on the switching among different gains. It was shown that convergence conditions are to ensure the stability of the error dynamics and the related gains may be selected by solving a set of linear matrix inequalities (LMIs). In [8] a methodology for designing dynamical observers for hybrid plants was proposed. In [9], the observability properties of a class of hybrid systems, whose continuous variables are available for measurement, were considered. They depicted that the discrete variable dynamics can be always extended for observable systems to a lattice in such a way that the extended system has the properties that allow the construction of the LU discrete state estimator. In [10], delay approach to continuous-discrete observer design for Lipschitz nonlinear systems was discussed. In [11], a motion sensorless control was proposed for single-phase permanent magnet brushless dc motor based on an I-f starting sequence and a real-time permanent magnet flux estimation.

A well-known case study of hybrid system is three-tank model introduced in [12]. Three tanks system is composed of three tanks of liquid which are connected to each other through pipes and controllable valves. The control objective in this system is usually to preserve the liquid level of tanks in a desirable value and the control objective can be achieved by controlling the liquid flows provided by pumps or adjusting the interconnecting valves. In [13], observability analysis and observer synthesis were studied for a three-tank water process. Observability of the process was considered under various assumptions on measurements. In observer design, the singularity of nonlinear observers was also considered and water level was controlled by using two pumps by a PI controller. In [14], methodologies for designing limited look ahead supervisory controllers were discussed for a class of embedded systems that can be modeled as switching hybrid system (SHS). The authors discussed the controller design and implementation for a three-tank system test-bed with distributed sensor and actuation units. A set of real-time fault adaptive control experiments demonstrated the effectiveness of the

approach .However , disadvantage of their work was considering the pumps as a discrete input. In [15],an identification method was proposed for a class of hybrid systems which were linear and separable in the discrete variables (that were discrete states and discrete inputs). The method took cognizance of the fact that the separable structure of the hybrid system constrains the evolution of system dynamics. In particular, the proposed method identified models corresponding to a certain number of modes, far fewer than the total possible modes of the system. It then generated the models for the remaining modes without any further requirement for input–output data by exploiting the separable structure of the hybrid system. Because of many new parameters to substitute ,the model has a high complexity and high computational effort.

An important class of theoretical and practical problems in communication and control rises from a statistical nature (Kalman [1]). Such problems are: (i) prediction of random signals,(ii) separation of random signals from random noise and(iii) detection of signals of known form (pulses, sinusoids) in the presence of random noise .In a pioneering work, Wiener [4] revealed that problems (i)and (ii) lead to the so-called Wiener-Hopf integral equation. He also gave a method (spectral factorization) for the solution of this integral equation in the practically important special case of stationary statistics and rational spectra. In [4], the objective is to obtain the specification of a linear dynamic system (Wiener filter) which accomplishes the prediction, separation, or detection of a random signal. In[3](alpha- betha), it was shown that the state vector of a linear system can be reconstructed from observations of the system inputs and outputs. It was shown that the observer, which reconstructs the state vector, is itself a linear system whose complexity decreases as the number of output quantities available increases. The observer may be incorporated in the control of a system which does not have its state vector available for measurement. The observer supplies the state vector ,but at the expense of adding poles to the overall system. In much of modern control theory designs are based on the assumption that the state vector of the system to be controlled is available for measurement. In many practical situations only a few output quantities are available.

Application of theories which assume that the state vector is known is severely limited in these cases. In [2] (Luenberger) shown that the state vector of a linear system can be reconstructed from observations of the system inputs and outputs.

In this paper, the problem of designing a continuous-time states observer to estimate the continuous-time dynamics of hybrid systems is investigated. It is assumed that we know our discrete-time states. In this system, it is assumed that the continuous and discrete dynamics can be linearized. It is also assumed that the system is observable and controllable. Furthermore, a state feedback controller is considered to control the system. The controller operates based on the state estimation obtained from an observer in the continuous-time part of hybrid system. The linear hybrid model is obtained by linearizing the model of three tank system around its equilibrium point. In addition, it is also assumed that a state feedback controller is utilized for controlling the linearized hybrid system in order to stabilize the system. The methods of observer design are based on Kalman filter and Luenberger method .To demonstrate the efficiency of designed observer in estimating the continuous state of model a three tanks system will be considered.

In the next section ,hybrid systems observers will be introduced. In Section 2, three tanks system model is introduced. Section 3presents and compares the proposed methods for designing observers in linearized hybrid systems. Simulation results are presented in Section 4 to show the applicability of proposed methods in three tank model. Finally ,conclusion remarks are drawn in the last section.

1- Observers in hybrid systems

This section is devoted to the introduction of hybrid system formalism and to the formulation of the observability problem for hybrid systems. In this section, we present a general definition of hybrid systems that includes a large class of hybrid systems. The notion of a hybrid system that has been used in the control community is centered around a particular composition of discrete and continuous dynamics. In particular, a hybrid system has continuous evolution and occasional jumps. The jumps correspond to the change of state in an automaton whose transitions are caused by controllable or uncontrollable external events or by the continuous evolution. A continuous evolution is associated to each discrete state, described by differential equations, which may have different structure for each discrete state, and an initial state that is determined whenever a transition into the discrete state of the automaton is taken [16].

Technological innovation pushes towards the consideration of systems of a mixed continuous and discrete nature, which are sometimes called hybrid systems. Hybrid systems arise, for instance, from the combination of an analog continuous-time process and a digital time asynchronous controller. Control theorists and computer scientists (and others) are joining forces to approach the huge challenges in this field. The simplest definition of a hybrid system is a system whose behavior is characterized by several modes of operation. In each mode the evolution of the continuous state of the system is described by its own difference or differential equation. The system switches between the various modes when a particular event occurs. We will typically study hybrid systems in which the continuous input will be designed and the mode transitions are externally induced by state events [12].

Figure1 shows a typical model of hybrid system. In this Figure, σ and Ψ are discrete input and discrete output and u and y are continuous input and continuous output respectively.

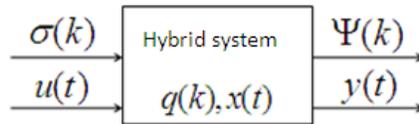


Figure 1. Hybrid system model

State estimation problem has been one of the critical issues in control theory and applications in hybrid systems. A reliable state estimate for a process is indispensable not only for control purpose but also for other applications such as navigation of spacecraft, monitoring, and fault diagnosis in mechanical systems[21]. As is well-known, given the measurements $y =$

$g(x, u)$ for the process $\dot{x} = f(x, u)$, it is possible to reconstruct the full-state by the estimator

$\hat{x} = f(\hat{x}, u) + K(y - \hat{y})$ with innovation process [22]. However, since the mathematical model (f, g) is only an approximation to the physical process and the actual plant is usually affected by external disturbances, there always exists a discrepancy between the real states and estimates. Like the other feedback systems, the uncertainty effect not considered in the plant model could be significantly reduced by increasing the estimator gain K . However, those with innovation process alone certainly have a limit to cover all uncertainties with wide ranges in a frequency spectrum. As a simple example, conventional state estimators such as Luenberger observer or Kalman filter produce biased estimates when the plant is under biased external disturbances. Then Hybrid observers that we use in this paper are Kalman filter and Luenberger method.

2- Three tank system model

The Schematic of three tank system is shown in Figure 2. Tanks 1 and 3 are filled independently by two pumps.

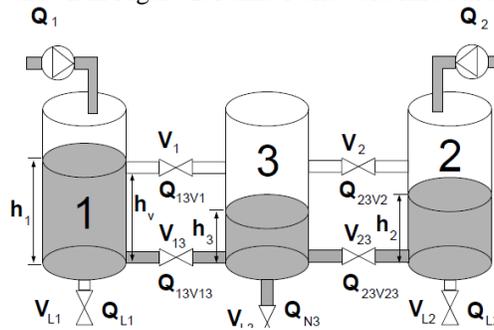


Figure 2. Three tank system model [12]

In this figure, Q_1 and Q_2 are flow of tank 1 and tank 2 respectively changed continuously from 0 to Q_{max} . h_1, h_2 and h_3 are liquid height of each tank, and can be measured by a level sensor. The height of each tank is continuous variable and behavior of each valve is discrete, So, three tank can be considered as a hybrid system. The equations of the three tank can be considered as follows[12]:

$$\dot{h}_1 = \frac{1}{A} (Q_1 - Q_{13V1} - Q_{13V13} - Q_{L1}) \tag{1}$$

$$\dot{h}_2 = \frac{1}{A} (Q_2 - Q_{23V2} - Q_{23V23} - Q_{L2}) \tag{2}$$

$$\dot{h}_3 = \frac{1}{A} (Q_{13V1} + Q_{13V13} + Q_{23V2} + Q_{23V23} - Q_{N3}) \tag{3}$$

Where (definition is same as Figure2) Q with any indices else{1,2} is the flow of liquid else from pumps, $H_i, i = \{1, 2, 3\}$ is liquid height of each tank and for simplicity it is assumed that the surface area of each tank have a same surface area A .

For designing observer, it is assumed that the discrete variables are available, and we consider that valves of V_1 and V_2 and V_{L1} and V_{L2} are closed and V_{13} and V_{23} are opened. Therefore, the equations of three tank will be as follows:

$$\dot{h}_1 = \frac{1}{A}(Q_1 - Q_{13V13}) \quad (4)$$

$$\dot{h}_2 = \frac{1}{A}(Q_2 - Q_{23V23}) \quad (5)$$

$$\dot{h}_3 = \frac{1}{A}(Q_{13V13} + Q_{23V23} - Q_{N3}) \quad (6)$$

The outlet flows in each tank based on Bernoulli's equation in steady states modeled as a simple orifice can be considered as $Q_{i3Vi3} = a_i \sqrt{H_i}$ with indices $i=\{1,2\}$ and $Q_{N3} = a_3 \sqrt{H_3}$ Where, a_1, a_2, a_3 are proportional constants which depend on the coefficients of discharge, the cross sectional area of each orifice and the gravitational constant and H_i $i = \{1, 2, 3\}$ is liquid height of each tank .According to Equations (4)- (6), nonlinear state equations described the system dynamics of the three tanks can be explained as

$$\dot{h}_1 = \frac{1}{A}(Q_1 - a_1 \sqrt{H_1}) \quad (7)$$

$$\dot{h}_2 = \frac{1}{A}(Q_2 - a_2 \sqrt{H_2}) \quad (8)$$

$$\dot{h}_3 = \frac{1}{A}(a_1 \sqrt{H_1} + a_2 \sqrt{H_2} - a_3 \sqrt{H_3}) \quad (9)$$

Suppose that for Q_1 and Q_2 , the fluid level in the tanks are at steady state level H_1, H_2 and H_3 .

If small variation in Q_1 is q_1 and in Q_2 is q_2 , then the resulting perturbation in levels are h_1, h_2 and h_3 respectively. From Equations (7) up to (9), the equation becomes:

$$(H_1 + h_1) = \frac{1}{A}((Q_1 + q_1) - a_1 \sqrt{H_1 + h_1}) \quad (10)$$

$$(H_2 + h_2) = \frac{1}{A}((Q_2 + q_2) - a_2 \sqrt{H_2 + h_2}) \quad (11)$$

$$(H_3 + h_3) = \frac{1}{A}(a_1 \sqrt{H_1 + h_1} + a_2 \sqrt{H_2 + h_2} - a_3 \sqrt{H_3 + h_3}) \quad (12)$$

From (7)-(12), the equations are

$$\dot{h}_1 = \frac{1}{A}(q_1 - a_1 \sqrt{H_1 + h_1} - \sqrt{H_1}) \quad (13)$$

$$\dot{h}_2 = \frac{1}{A}(q_2 - a_2 \sqrt{H_2 + h_2} - \sqrt{H_2}) \quad (14)$$

$$\dot{h}_3 = \frac{1}{A}(a_1 \sqrt{H_1 + h_1} - \sqrt{H_1} + a_2 \sqrt{H_2 + h_2} - \sqrt{H_2} - a_3 \sqrt{H_3 + h_3} - \sqrt{H_3}) \quad (15)$$

$$\sqrt{H_1 + h_1} = \sqrt{H_1} \left(1 + \frac{H_1}{2H_1}\right) \quad \sqrt{H_1 + h_1} - \sqrt{H_1} \approx \frac{h_1}{2\sqrt{H_1}}$$

Then for small perturbations, we have Therefore,

$$\sqrt{H_2 + h_2} - \sqrt{H_2} \approx \frac{h_2}{2\sqrt{H_2}} \quad \text{and} \quad \sqrt{H_3 + h_3} - \sqrt{H_3} \approx \frac{h_3}{2\sqrt{H_3}}$$

Similarly we have the approximations of Equation is

$$h_1 = \frac{1}{A} \left(q_1 - \frac{a_1 h_1}{2\sqrt{H_1}} \right) \tag{16}$$

$$h_2 = \frac{1}{A} \left(q_2 - \frac{a_2 h_2}{2\sqrt{H_2}} \right) \tag{17}$$

$$h_3 = \frac{1}{A} \left(\frac{a_1 h_1}{2\sqrt{H_1}} + \frac{a_2 h_2}{2\sqrt{H_2}} - \frac{a_3 h_3}{2\sqrt{H_3}} \right) \tag{18}$$

$$h_1 = \frac{1}{A} \left(q_1 - q_{01} - \frac{a_1 h_1}{2\sqrt{H_1}} \right) \quad , \quad h_2 = \frac{1}{A} \left(q_2 - q_{02} - \frac{a_2 h_2}{2\sqrt{H_2}} \right) \quad \text{and}$$

The above Equations can also be written as

$$h_3 = \frac{1}{A} \left(\frac{a_1 h_1}{2\sqrt{H_1}} + \frac{a_2 h_2}{2\sqrt{H_2}} - \frac{a_3 h_3}{2\sqrt{H_3}} - q_{03} \right) \quad \text{where } q_{01}, q_{02} \text{ and } q_{03} \text{ represent perturbations in the outflow at the}$$

drain pipes .we haven't any perturbations in output then if we rewritten this equation with laplace transform we have equations as following

$$(T_1 S + 1)h_1(s) = k_1 q_1(s) \tag{19}$$

$$(T_2 S + 1)h_2(s) = k_2 q_2(s) \tag{20}$$

$$(T_3 S + 1)h_3(s) = k_1 h_1(s) + k_2 h_2(s) \tag{21}$$

Where,

$$T_1 = \frac{A}{a_1} \quad T_2 = \frac{A}{a_2} \quad T_3 = \frac{A}{a_3}$$

$$k_1 = \frac{1}{\frac{a_1}{2\sqrt{H_1}}} \quad k_2 = \frac{1}{\frac{a_2}{2\sqrt{H_2}}} \tag{22}$$

Each value of $a_1, a_2, a_3, A, H_1, H_2$ can be obtained from plant and those values are $H_1 = 7, H_2 = 8, H_3 = 10$ and $a_1 = 10.78, a_2 = 11.03, a_3 = 11.03$ and $A = 32$. By solving Equations (22) with those values $a_1, a_2, a_3, A, H_1, H_2$ the value of $T_1, T_2, T_3, k_1,$ and k_2 can be determined as following table.

Table 1: value of parameters of Equation (19)-(21)

$T_1 = 2.2439$
$T_2 = 2.0514$
$T_3 = 1.8349$
$k_1 = 0.0701$
$k_2 = 0.0641$

3- Designing Observers in Hybrid Systems

Schematic of a hybrid system with continuous-time and discrete-time observers for such systems are shown in Figure 3. In

this Figure, \tilde{q} and \tilde{x} are estimation of location observer and continuous observer respectively. It can be seen that both observers can be used for state observation in this system. However, in this paper it is assumed that only continuous-time state variables are unknown.

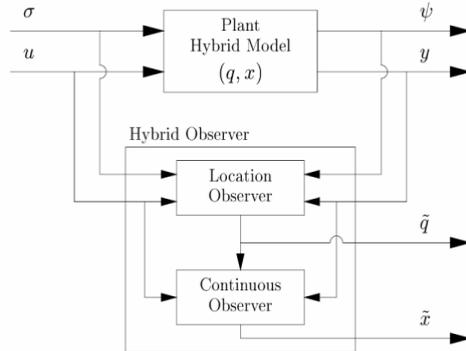


Figure 3. General structure of observers in a hybrid system

Suppose that the discrete states of the three tanks are available. So, only the continuous variable of the three tank system which is the height of liquid in the third tank in the schematic (as shown in Figure 4) should be estimated.

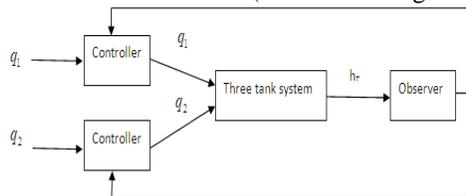


Figure 4. Observer of three tank system

For designing observer, the three tank system equation is considered as defined in Section 2. The structure of Luenberger observer is shown in Figure 5.

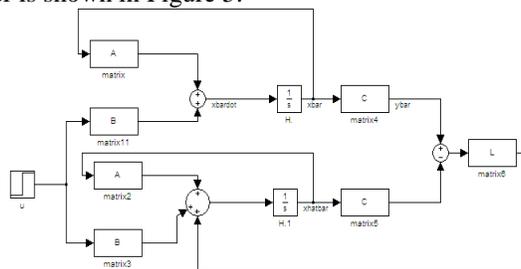


Figure 5. Luenberger Observer schematic

A symmetric positive definite matrix P in Luenberger method as theorem in [7], is the solution of the algebraic Lyapunov inequalities

$$(A_i - L_i C_i)^T P (A_i - L_i C_i) - P < 0 \quad i = 1, 2, \dots, k \quad (23)$$

The observer involves an estimation error which asymptotically convergent to zero. L is the observer gain.

Kalman filter is another method which is used for designing observers. Structure of Kalman observer is shown in Figure 6.

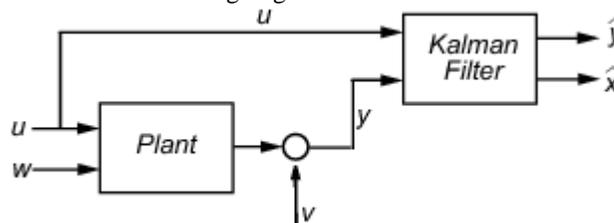


Figure 6. Schematic of Kalman Observer

Where u , w , v and y are input, white process noise, and white measurement noise and measured output, respectively. The state estimates are \hat{y} and \hat{x} . Note that \hat{y} estimates the third tank output.

4- Simulation Results

In this section, the observers are designed using Kalman filter and Luenberger observer design methods and in presence of a state feedback controller. It is supposed that the discrete states of the three tanks are available. So, only the continuous variable of the three tank system i.e. the liquid level of third tank will be estimated, as shown in Figure 4.

It is assumed that both continuous and discrete dynamics of hybrid system are linear. Discrete dynamics of system are known and the system does not have any perturbation.

The matrices A,B,C are known from state-space model of three tank system where

$$A = \begin{bmatrix} -0.9906 & -0.4858 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0 & -1.032 & -0.5313 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}, B = \begin{bmatrix} 0.0625 & 0 \\ 0 & 0 \\ 0 & 0.0625 \\ 0 & 0 \end{bmatrix}, C = [0 \quad 0.03819 \quad 0 \quad 0.03493]$$

and matrices P, L (gain observer of Luenberger method) are selected as follows:

$$P = 10e-003 * \begin{bmatrix} 0.1331 & 0.1175 & -0.0873 \\ 0.1175 & 1.4445 & -0.0371 \\ -0.0873 & -0.0371 & 0.0960 \end{bmatrix},$$

$$L = \begin{bmatrix} 552.4211 \\ -21.2599 \\ 650.3919 \end{bmatrix}$$

In addition, matrix K as in [1] for Kalman filter is designed to be:

$$K = \begin{bmatrix} 0.0009 & 0.0019 & 0.0008 & 0.0017 \\ 0.0008 & 0.0016 & 0.0007 & 0.0014 \end{bmatrix}$$

System output in absence of observer and controller is shown in Figure 7.

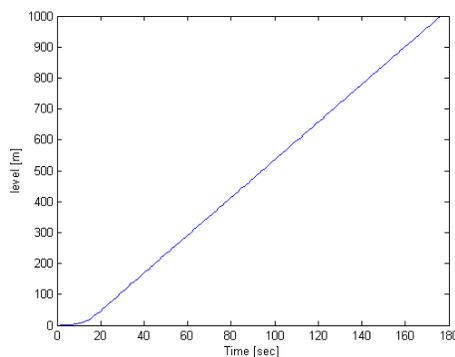


Figure 7. System output without any control and observe

It can be seen that the level of third tank with the swinging form goes to infinity.

When the state feedback control is added to the system, system output (the level of the third tank) is in the form of Figure 8.

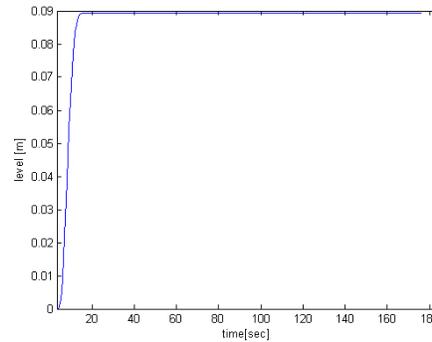


Figure 8. System output when we have state feedback control

When the state feedback control and Luenberger observer are used ,the result could be in the form of Figure 9 and the output could be estimated as 9.97082

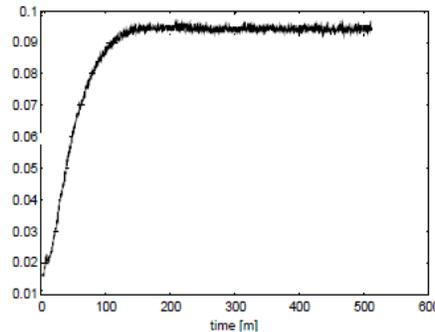


Figure 9. System output when we have state feedback control and Luenberger observer

When the model with the proposed Kalman filter method is used, the output is as the form shown in Figure 10.

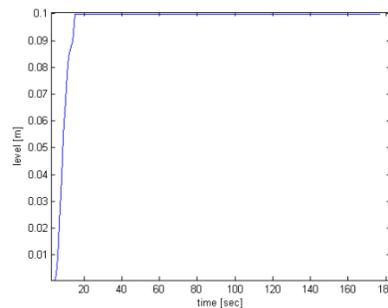


Figure 10. System output with state feedback control and Kalman filter

where the third tank height becomes and keep to 9.99999 cm.

It is obvious that without observer and controller in the system, the output goes to infinity because of the nature of the system. The controller could keep the output to steady state, but it is far from the ideal point. When the Luenberger method is used, the result will come near to ideal but Kalman filter is the better way to force it to ideal point.

5- Conclusion

In this paper, a three tanks system has been considered as a hybrid system and by considering a closed loop control system ,the problem of observers design for this system has been studied .Luenberger and Kalman filter observers have been designed by considering the state feedback controller as the system controller. The system controller uses an estimation of the output ,i.e. the pump inlet ,from observer in every time in order to achieve the desired level of the third tank. This filter can be applied easily to any hybrid systems. Comparing two observers show that Kalman filter can perform better response for hybrid systems specially when there exist noise, disturbance or uncertainty.

References

- [1] R.E Kalman, On a new approach to filtering and prediction problems, ASME Journal of Basic Engineering,1960, vol. 28 ,pp 35-45.
- [2] D.G. Luenberger, Observing the state of linear system, IEEE Trans. On Military Electronics,1964, vol. MIL-8, pp 74-80.
- [3] P. E. Caines and R. Greiner and S. Wang, Dynamical logic observers for finite automata, Proceedings of 27th Conference on Decision and Control, 1988, pp. 226–233.
- [4] N. Wiener, The Extrapolation, Interpolation and Smoothing of Stationary Time Series, John Wiley & Sons, Inc., New York, N.Y,1949.
- [5] K. K. Sundareswaran, P. J. Mclane., and M. Bamyoumi, Observers for Linear Systems with Arbitrary Plant Disturbances, IEEE Transactions On Automatic Control, 1977, Vol. Ac-22, No. 5,pp 870-871.
- [6] J. P. Gauthier, H. Hammouri, and S. Othman, A Simple Observer for Nonlinear Systems, Applications to Bioreactors, IEEE Transactions On Automatic Control,1992, Vol. 37, No. 6 , pp 875-880
- [7] A. Alessandri and P. Coletta, Design of Luenberger Observers for a Class of Hybrid Linear Systems ,HSCC c Springer-Verlag Berlin Heidelberg .2001, pp. 7-18.
- [8] A. Balluchi, L. Benvenuti, M.D. Di Benedetto, and A.L. Sangiovanni-Vincentelli, Design of Observers for Hybrid Systems , In Lecture Notes in Computer Science 2289, C.J. Tomlin and M.R. GreensreetEds, 2002, pp. 76-89.
- [9] D. Del Vecchio and R. M. Murray, Existence of Discrete State Estimators for Hybrid Systems on a Lattice, 43rd IEEE Conference on Decision and Control, 2004,vol. 1 pp 1-6.
- [10] Salim Ibrir, Hybrid observer design for a class of nonlinear systems,Proceedings of the 7th Asian Control Conference,Hong Kong, China,2009,pp142-147.
- [11] LiviuIoanIepure, Hybrid I-f Starting and Observer-Based Sensorless Control of Single-Phase BLDC-PM Motor Drives,IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 2012,VOL. 59, NO. 9,pp 3436-3444.
- [12] D. Mignone,Control and Estimation of Hybrid Systems with Mathematical ptimization,A dissertation submitted to the Swiss Federal Institute of Technology , Diss. ETH No. 14520,2002.
- [13] M. Hou, Y. S. Xiong, and R. J. Patton, Observing a Three-Tank System, IEEE Transactions On Control Systems Technology, 2007, Vol. 13, No. 3, pp478-484.
- [14] Jian Wu, A Hybrid Control System Design and Implementation for a Three Tank Testbed , IEEE Conference, 2005.
- [15] N. Nandola, S. Bhartiya, Hybrid system identification using a structural approach and its model based control: An experimental validation,journalNonlinear Analysis: Hybrid Systems ,2009, vol. 3, no. 2, pp. 87-100.
- [16] E. De Santis, M. Di Benedetto,S. Di Gennaro, G.Pola, Hybrid Observer Design Methodology , Distributed Control and Stochastic Analysis of Hybrid Systems SupportingSafety Critical Real-Time Systems Design (HYBRIDGE),2003.
- [17] C. M.Ozveren and A. S. Willsky, Observability of discrete event dynamic systems, IEEE Trans. on Automatic Control, 1990, pp 797-806.
- [18] P. Aixian, G. Yun, Design and Application of Rough Controller in Three-tank System , IEEE ,conference on computing ,control and Industrial engineering , 2010,pp144-148.
- [19] F. Hamdi, N. Messai, N. Manamanni, Observer Based State Feedback control Design for Switched Linear Systems: A Differential Petri Net Approach , IEEE Conference on Control and Fault Tolerant Systems., 2010,pp 167-172.
- [20] C. Liu, J. Zuo, J. Wu, D. Liu, Observability analysis of AC-DC hybrid power system , IEEE Conference on Power and Energy Society , 2010,pp 1-7.
- [21] S. Kwon ,Robust Tracking Control and State Estimation of Mechanical Systems , Division of Mechanical and Industrial Engineering Pohang University of Science and Technology,2002.
- [22] Friedland, B., Advanced Control System Design, Prectice Hall, 1996.

Single Precision Floating Point Divider Design

¹Serene Jose, ²Sonali Agrawal

^{1,2}Department of Electronics and Communication, Amrita School of Engineering
Kasavanahalli, Bangalore- 560035, India

Abstract— Growth in Floating Point applications and mainly its usage in reconfigurable hardware have made it critical to optimize floating point units. Divider is of particular interest because the design space is large and divider usage in applications varies widely. The design presented in this paper covers a range of performance, area and throughput constraints. Floating point numbers can be represented by single and double precision respectively. A design for single precision floating point divider was done in Verilog and was synthesized using Xilinx and Synopsys tool. The path delay, device utilization was also determined successfully.

Keywords— Divider, Floating Point Unit (FPU), and Single precision

I. INTRODUCTION

Reconfigurable hardware mainly FPGAs and DSP processors use floating point operations in configuring the real time data. This paper presents the implementation of a 32 bit single precision IEEE 754 standard floating point divider which is designed with the power of handling rounding and exceptional cases of division like Not A Number along with special cases of floating point numbers like de-normal, real zero, infinite positive and negative values.

In recent FPUs, importance has been given in developing fast adders, subtractors and multipliers with very less emphasis on dividers. But to deal with the upcoming new technologies, an efficient implementation of division is inevitable. As such many algorithms were developed for dividers including subtractive methods, functional iterations and high radix algorithm for faster computation which further required multipliers and thus consumed large area and power. To account for this drawback, digit recurrence algorithm[1],[2],[3],[4] which uses subtraction along with non-restoring method for computation as it consumes much less area and power.

In the beginning, floating-point divider dealt with obtaining the value of denominator inverse from the LUT which is then multiplied with numerator[5],[6]. But this method was favourable only for small dividers. Following this, divider design with newton-Raphson, Taylor series iterations was developed which had the drawback of high latency[7],[8].To speed up the dividing process, an algorithm based on higher radix SRT [9],[10]was introduced which is complex in nature and is commonly employed for microprocessors.

This paper gives architecture for floating point divider where the divider core employs a fixed point division based on binary digit recurrence algorithm .The area, power and delay parameters are also found out using Synopsys design compiler and IC compiler and was also synthesised in Xilinx.

II. FLOATING POINT REPRESENTATION

Floating point is a method of representing a wide range of real values. A floating point number can be represented in the form: $\pm m \times b^e$, where e the exponent, m the mantissa and b the base or radix. The IEEE Standard 754 for binary floating point number representation which includes single precision with 32 bit wide and double precision with 64 bit wide format.

This section briefly examines the single precision floating point representation. The MSB starts from the left which is one bit width sign, following eight bit width exponent and 23 bit width mantissa (significand) or fractional part.

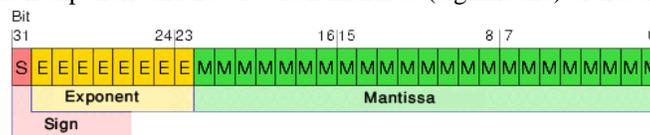


Fig 1.1 Single Precision Format

The value can be represented as:

Normalized value: $(-1)^s \times 2^e \times (1.f)$

Denormal or Subnormal value: $(-1)^s \times 2^{-126} \times (0.f)$ where,

$s = 0$ for positive and 1 for negative

$f = (b_{23}^{-1} + b_{22}^{-2} + b_{21}^{-3} + b_i^{-n} + \dots + b_0^{-23})$; $b_i^{-n} = 0$ or 1

When the leading bit of mantissa is 1, commonly referred to as hidden bit followed by the radix point and the 23 bit fractional part, then the representation is said to be normalized.

$e =$ unbiased exponent

The exponent can hold both positive and negative value. The value range of unbiased signed integer is [-128,127] or unsigned integer [0,255]. A bias of 127 is added to actual exponent e , to make negative exponents possible without using a sign bit.

For the biased exponent E, $E_{max}=255$ and $E_{min}=0$ are reserved for special quantities: The number zero which can be positive or negative is represented with $E=0$ and $f=0$. The hidden bit is always 0; Denormalized numbers which can also be positive or negative are represented with $E=0$ and $f \neq 0$, which implies the value of $e = -128$ and mantissa $M=0.f$. The hidden bit is 0; Positive or negative Infinity have $E=255$ and a fraction $f=0$ including the hidden bit; Finally, an exponent $E=255$ and a fraction $f \neq 0$ represents the symbolic unsigned entity NaN (Not a Number), which can be QNaN (Quiet NaN) produced by operations like $0/0, \infty/\infty$ and SNaN (Signalling NaN) is signalled if operations like division by 0 or negative square roots are encountered. The exceptional cases of underflow and overflow property should be also encountered. i.e. whenever the floating point division result exceeds the minimum or maximum value of the restricted exponent range,

To increase the precision of the floating point division result, rounding is necessary. For this, three bits known as guard, round and sticky bits were temporarily added internally to the actual fraction. While guard and round bits are the normal storage holders, the sticky bit is turned '1', whenever a 1 is shifted out of range. Here, the standard default rounding employed is 'Round to nearest even mode'. The value is rounded up or down to the nearest infinitely precise result. If the value is exactly halfway between two infinitely precise results, then it should be rounded up to the nearest infinitely precise even.

III. SINGLE PRECISION FLOATING POINT DIVISION ARCHITECTURE

This division architecture comprises of sign, exponent and mantissa computation sectors. Initially, the divider receives two 32 bit floating point numbers as its operands. The most significant steps in the calculation of the quotient Q of two normalized numbers A, the dividend and B, the divisor are as follows:

The primary aim in the computation process is to unpack these numbers by separating the floating point numbers into the corresponding sign bits, exponent bits and mantissa bits. The sign logic is a simple XOR operation. The exponent is calculated by adding the bias to the subtracted result of exponents. i.e. $E_Q = E_A - E_B + 127$. Mantissa division block performs fixed point division using digit recurrence algorithm, which includes the following steps: First, the remainder of the division is set to the value of the dividend followed by subtracting the divisor value from the remainder. If the result is positive or zero, the MSB of the quotient is 1 or else 0. The remainder is shifted left each time after the subtraction process and the whole process continues for n times (equal to mantissa bit length) to obtain the quotient value. After this, the output of the mantissa has to be normalized i.e. mantissa part has to be left shifted so as to make the MSB 1 and correspondingly, changes has to be made to the exponent part also. After normalization, the output has to undergo the rounding control process, where the three bits namely 'guard', 'round' and 'sticky' bits are stacked to the LSB of the 24 bit significant part and the round to the nearest even is performed.

Along with this, corresponding changes are to be made through exponent adjustment block.

Finally, the output from sign block, exponent adjustment block and the rounding block are concatenated in the packing block to produce the final quotient.

To deal with the denormal inputs, a special block with a priority encoder and left shifter is included. The priority encoder functions to detect the position of leading one in the inputs and the left shifter shifts the mantissa part to so much value, so that the mantissa part is presently in normalized form and can be presented to the divider core.

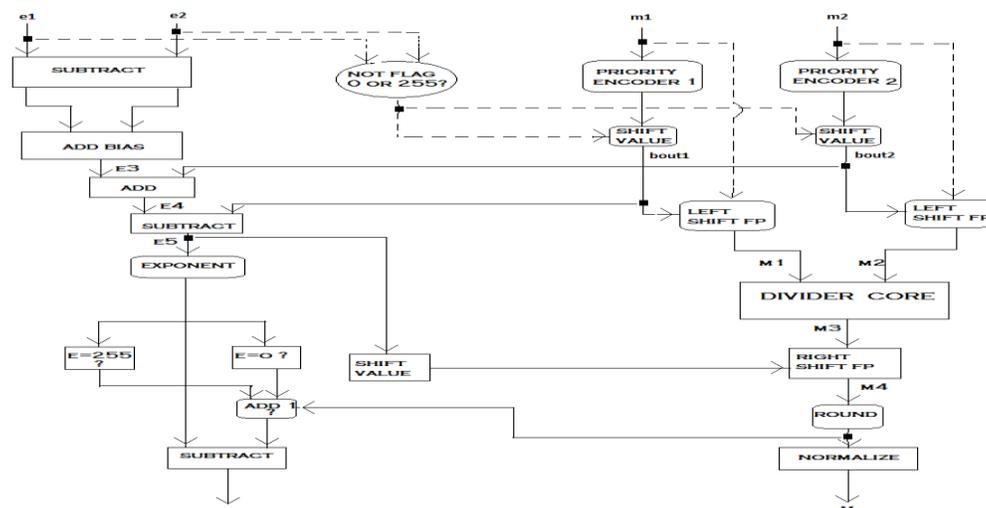


Fig1.2 Architecture of Single Precision floating point divider

IV. RESULTS

A. Simulation results

The simulations were carried out in Model Sim 5.7g. The implementation of the single precision floating point divider was done using Verilog and result was verified with different values.

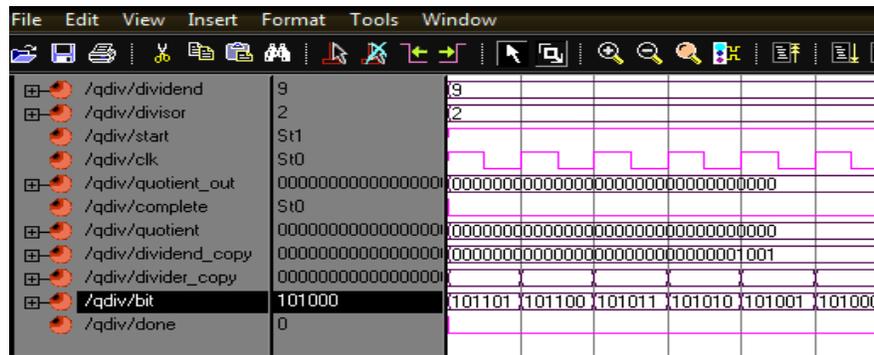


Fig 1.3.Simulation result of fixed point divider using digit recurrence algorithm

B. Device Utilization

The hardware amount utilization was obtained from Xilinx. This report gives the information that the divider design is FPGA synthesizable, with the efficient utilization of 1% of the device.

Fig 1.4.Device Utilization Report

LOGIC UTILIZATION	USED	AVAILABLE
Number of slices	166	14752
Number of slice Flip Flops	131	29504
Number of 4 Input LUTs	256	29504
Number of Bonded IOBs	125	376

C. Delay Calculation: Divider compilation gave 10.305 ns of maximum combinational path delay.

D. Area and Power

COMBINATIONAL AREA USAGE	4756.377nm ²
TOTAL CELL AREA	8439.099nm ²
TOTAL DYNAMIC POWER	45.2305μW
CELL LEAKAGE POWER	27.1779μW

Fig 1.5. Area and Power Characteristics

V. CONCLUSION

This paper presents a vibrant design for single precision floating point divider with exceptional cases handling capability. The characteristic features of the divider such as area, power, latency etc. were determined using Xilinx and Synopsys tools.

V1.References

- [1] Baesler M,Voigt S,Teufel T, “FPGA implementation of Radix-10 Digit Recurrence Fixed and Floating-point dividers”,IEEE international conference on Reconfigurable computing and FPGAs,December 2011,pp:13-19
- [2] Rust I,NollnT.G,“A radix-4 single-precision floating-point divider based on digit set interleaving”, Circuits and systems,proceedings of IEEE international symposium ,june 2010,pp:709-712.
- [3] Thakkar A J,Einioui A, “Design and implementation of double precision floating point division and square root on FPGAs”, IEEE, Aerospace conference on july 2006,pp:7
- [4] Stavros Paschalakis ,Peter Lee, “Double Precision Floating-point arithmetic on FPGAs”, In proc.IEEE symposium on FPGA for custom computing machines,2004,pp:155-167.
- [5] Xiaojun Wang, Braganza S, Leaser M, “Advanced components in the variable precision Floating-point library”, Field-Programmable custom computing machines, IEEE symposium on Dec 2006, pp.: 249-268
- [6] J.Dido ,N.Geraudie, “A Flexible floating point format for optimizing data paths and operators in FPGA based DSPs”, in International Symposium on Field Programmable Gate Arrays, Feb. 2002, pp.50-55.
- [7] Suhap Sahin, Adnan Kavak, Yasar Becerikli,H.Engin Demiray, “Implementation of floating point arithmetic using an FPGA”, IEEE,2003,pp:12-17
- [8] E.Roesler and B.E Nelson “Novel optimizations for hardware floating point units in a modern FPGA architecture”, in 12th International Conference on Field-Programmable Logic, pp.637-646,2002.
- [9] X.Wang and B.E. Nelson, “Tradeoffs of designing Floating point division and square root on Virtex FPGAs”, in 11th IEEE Symposium on Field Programmable Custom Computing Machines, April 2003,pp.195-203.
- [10] Mohamed Anane,Hamid Bessalah,Mohamed Issad,Nadjia Anane,Hassen Salhi,“Higher Radix and Redundancy Factor for floating-point SRT Division”,IEEE transactions on VLSI,June 2008,vol 16,pp:774-779.

About The Authors



Serene Jose received her Bachelor of Engineering in Electrical And Electronics Engineering from Mahatma Gandhi University, Kerala in the year 2010 and currently pursuing Master of Technology in VLSI design in Amrita University, Bangalore. Her main research interests include Low Power VLSI Design, Digital Design.



Sonali Agrawal graduated from Pandit Ravishankar Shukla University in 2002 .She completed her post-graduation from NIT, Rourkela in 2008 .Presently ,she is working as an Asst. Professor at Amrita Viswa Vidyapeetham and has been engaged with research in Low Power VLSI Design, Digital Design, Hardware Description languages,DSP.

Improvement in Enhanced Privacy Protection in LBS

Rashmi Samele

¹Department of Computer Science
& Engineering, GCET-Gr. Noida

Vivek Singh Senger

²Department of Computer Science & Engineering, GCET-Gr. Noida

Abstract

In order to access location-based services, mobile users have to disclose their exact location information to service providers. However, adversaries could collect the location information for purposes against mobile users' privacy, such as tracking and stalking. The most popular solutions for privacy protection are utilizing the K-anonymity model to blur user's exact location information. System architectures applying the K-anonymity principle to construct the cloaking region for the privacy preserving spatial query. Our technique, progressively increase the privacy level of the users.

Keywords - Cloaking Algorithm, Privacy Protection, road networks, Location-Based Services.

1. Introduction

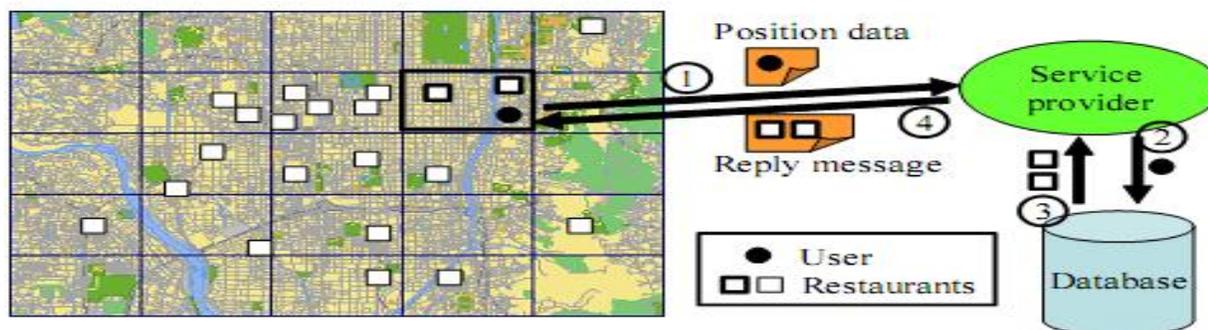
As the result of recent advances in wireless technologies, more and more personal mobile devices (e.g., cell phones, PDAs, etc.) possess the ability to access the Internet ubiquitously. In addition, Global Positioning System (GPS) receiver modules are becoming a standard component in the recent generation hand-held devices. Consequently, novel location based services (LBS) allow users to launch location-dependent queries ubiquitously. Sample queries of such location based services include find me the nearest ATM and show me the gas station with the lowest price within one mile. In order to fulfill these queries, mobile users have to reveal their current location information to service providers. However, service providers may disclose the trajectory of a certain user to malicious users which can decrease the dependability of LBS. For protecting mobile users privacy, recent research in [6] proposed a framework for location based services without compromising location privacy by leveraging the K-anonymity concept [9].

1.1 Location Based Services

Definition 1: A location-based service (LBS) is an information and entertainment service, accessible with mobile devices through the wireless network and utilizing the ability to make use of the geographical position of the mobile device [11].

Definition 2: A wireless-IP service that uses geographic information to serve a mobile user. Any application service that exploits the position of a mobile terminal. There are several different examples of Location Based services:

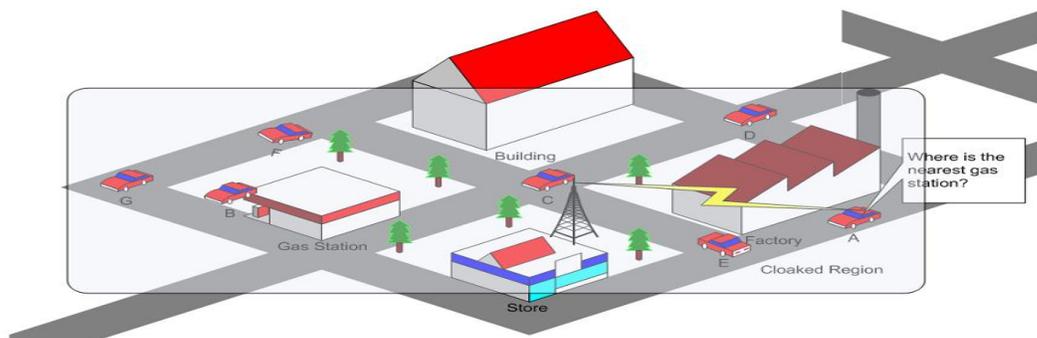
- *Resource tracking with dynamic distribution* -Taxis, service people, rental equipment, doctors, fleet scheduling.
- *Resource tracking* - Objects without privacy controls, using passive sensors or RF tags, such as packages and train boxcars.
- *Finding someone or something* - Person by skill (doctor), business directory, navigation, weather, traffic, room schedules, stolen phone, emergency calls.
- *Proximity-based notification (push or pull)* - Targeted advertising, buddy list, common profile matching (dating), automatic airport check-in.
- *Proximity-based actuation (push or pull)* - Payment based upon proximity (EZ pass, toll watch).



EXAMPLE OF L.B.S.

1.2 Privacy Concerns in Location Based Services - Information that users used in the location based services is considered as very personal and critical information for many reasons. Users' information included in the location based services consisted with four major components:

- Users' identity information.
- Users' location information - the location coordinates.
- The context of the location based services.
- Timestamp.



EXAMPLE

2. Related Work

privacy protection is very necessary and urgent in the location based services. And a lot of related research and projects have been done. However, the conventional privacy protection mechanisms highly rely on encryption and pseudonym techniques to safeguard users communication and hide users' identities. However, the queries launched by users may contain other sensitive information, which potentially

Could hurt users' privacy. Recently, a user can request for a cloaked area which cover the location of at least $K - 1$ closest peers to blur its actual location.

2.1 Peer to Peer Spatial Cloaking- the peer to peer spatial cloaking system doesn't have the trusted third part to construct the cloaked region. Because the trusted third party, the location anonymizer, is not only a performance bottleneck but also a oblivious attacking target for the malicious user, the peer to peer spatial cloaking triesto construct the cloaked region by peer themselves.

2.1.1 Peer to Peer System Architecture-the peer to peer system doesn't have the location anonymizer as the middle-ware between the base station and the service provider. The cloaked region is constructed by users' collaboration and is sent to the service provider directly. Consequently, the peer to peer spatial cloaking system contains two main components: mobile clients and location -based database server

2.1.2 Peer to Peer Cloaked Region Construction-The entire area that the peer to peer spatial cloaking system covers is divided into grid. Users in the peer to peer system communicate with each other to discover other $k-1$ peers to meet the k -anonymity requirement described in their privacy profile. In that way, user can hid his exact location into a cloaked spatial region that is the minimum grid area covering at least the $k-1$ peers and itself, and satisfies the minimal area requirement as well.

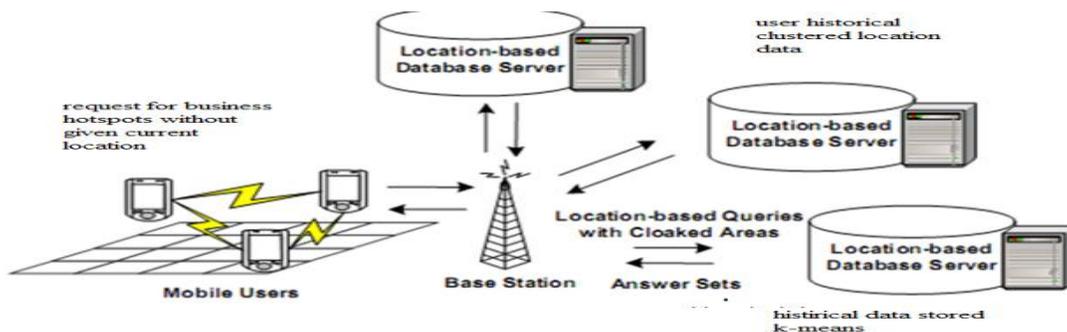


Fig.1 Peer to Peer Spatial Cloaking System Architecture

3. Roadnetworks based spatial cloaking-

3.1 System Overview - In the most attacking scenarios, the attackers are always able to obtain additional information about the targeted area. Combining this area information, the attacker can easily exclude the users who are very unlikely to issue the query(which indicates the query issuer) in the cloaked region. And in that way,K-anonymity principal will be compromised, and the user's location privacy will be sueffred. K-anonymity in the location privacy is quite similar with it is the data privacy. We try to hide the actual query requester within k-1 other users. In order to achieve the K-anonymity there are two criteria: first, there must be k-1 other users to satisfy the privacy in terms of anonymity quantity. Secondly, it is the quality of the anonymity, which is the indistinguishability of these k-1 users. In the original formal definition, these k entries (including the one we want to protect) should not be able to tell the difference and identify.

3.2 System Architecture – Assumptions

Clients have the capability to communicate with the service provider.

Clients also can talk with each other via the wireless connections.

Clients a GPS module to get its location.

Clients follow the protocol and are willing to work corporately.

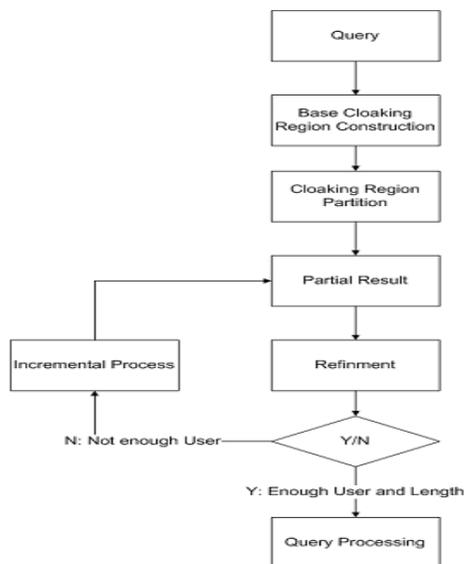
^Clients have a limited storage on board that it cannot hold all the map information. But it has the ability to store the local map in their devices

The cluster head will maintain the information about the total number of the client nodes it manage and their exact locations which will be used in the refine phrase of the cloaking region construction algorithm.

3.3 Query Process- Lemma: If the result set includes all the edge points' K nearest neighbor and the POIs inside the cloaking region, every inside points' K nearest neighbor will be included in the result set.

Proof: Assume we have point P inside the cloaking region and its Kth nearest neighbor K is not included in the result set. And K is outside of the cloaking region. So the shortest path between P and K, DPK, must come through an edge point E of the cloaking region. However, K is not included in E's K nearest neighbor set. So the distance between K and E must be greater than the upper bound of E's K nearest neighbors. So there must be a better candidate inside E's K nearest neighbor set or there is another shortcut between P and K which is not via E. Both of them are both against the assumptions.

Fig -2 Query Process



The basic idea for this search mechanism is that for each edge points in the cloaking region we will do a K Nearest Neighbor search to guarantee that the result set will include all points' K Nearest Neighbor.

A priority queue will be constructed for each edge points based on the distance between it to the searching point. And the search will be executed like the round robin scheduling mechanism in Operating Systems. Every edge point will execute a K nearest neighbor query partially and turn by turn until the Search Threshold for each round is met. When the current searching distance is equal or more than the search distance for the current round which is the distance of the first item in priority queue,the search in this priority will be stopped and the process will be continued to next priority queue. by peer-to-peer based cloaking on road networks.

ALGORITHM

- ⊙ Step1:start
- ⊙ Step2: :Read a first record first user recent location database
- ⊙ Step3: : If direction of current record = direction of user recent location information
- ⊙ **Step4:** Then calculate distance of current record and user recent location information, otherwise Loop: next record in the user recent location information and go to step2.
- ⊙ Step5 : After calculating distance compare that current distance < minimum distance($distance = ((x1-x2)^2+(y1-y2)^2)^{1/2}$)
- ⊙ Step6 :If this condition holds then minimum distance = current distance, otherwise Loop: next record in the user recent information and go to step2.
- ⊙ Step7 :transfer data in historical database
- ⊙ Step8 : add new record in priority base
- ⊙ Step9 :estimated location = location of current record
- ⊙ Step10 :result= nearest POI(hide user location time difference)

4 Simulator Implementation-

We implemented the technologies and algorithms mentioned above which protect location privacy by peer-to-peer based cloaking on road networks. We name the prototype system as PROS. With PROS, a mobile user forms a cloaked road segment set by collaborating with her peers when she needs to retrieve information from location-based service providers.

4.1 Cloaking Phrase-

- PROS employs road segment sets to represent the cloaked region to achieve advanced location privacy protection effect . A cloaked road segment set will be expanded iteratively until the privacy profile is satisfied.

4.2 Searching Phrase-

As the result of peer-to-peer cloaking, the cloaked road segment set is submitted with the LBS query to the location-based service provider for evaluation. In PROS, we adopt the privacy protected spatial network query algorithms, PSNN and PSRQ, as proposed in [2], for answering K nearest neighbor queries and range queries on road networks. Notice that since the cloaked regions discussed in [2] are based on grid cells rather than road segment sets, we extend both PSNN and PSRQ to retrieve inclusive query result sets based on the input cloaked road segment set and the underlying road network. Finally, the query initiator filters out the exact query results from the inclusive query result set returned by the Location-based Service Provider.

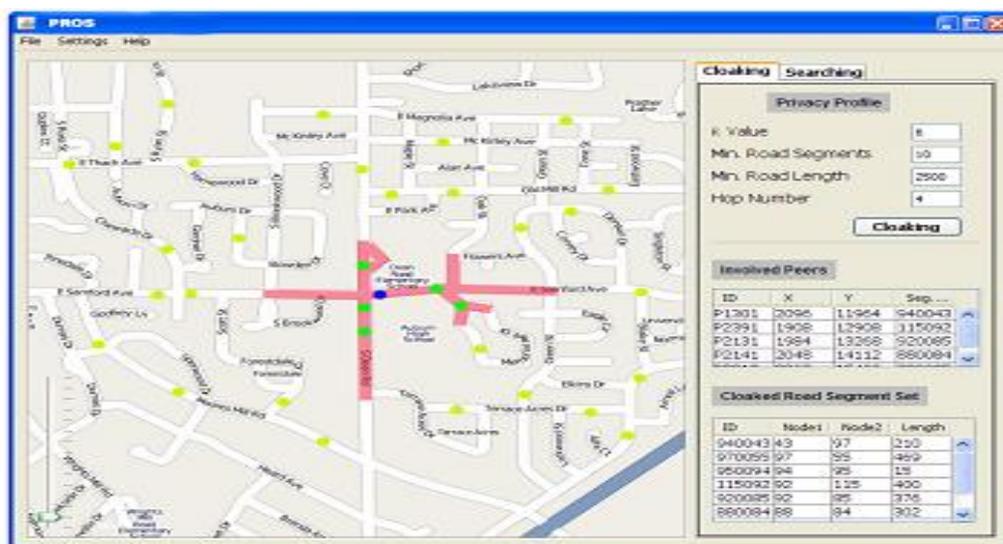


Fig -3 cloaked road segment set

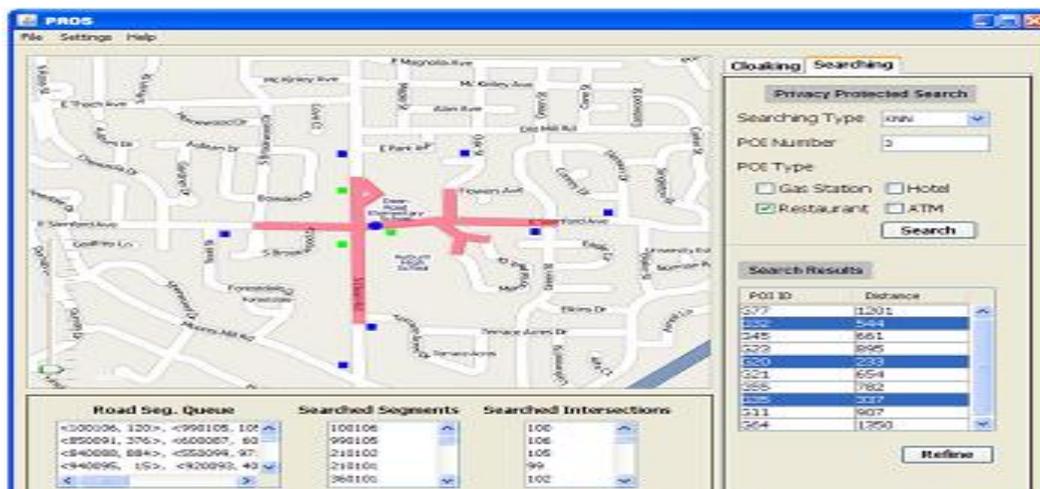


Fig 4 – Exact query result

5. Conclusion and future work –

We introduced the basic ideas of location based services, while raised the potential privacy issues in such kind of the services. And we also surveyed some existing and popular privacy protection solutions in location based services. Moreover, we proposed a road network based decentralized spatial cloaking method as well as the corresponding incremental spatial query processing algorithms. And we also first raised the issue of indistinguishability for the K-anonymity idea. By using an incremental cloaking region construction process, we are able to provide a higher privacy protection for the query issuer. And we extend the framework to adapt with the road networks to further ,for example, the road network indexing method and the peer management schemas. We hope to address these issues in the future researches and propose some more efficient solutions.

REFERENCES

- [1] Claudio Bettini, Xiaoyang Sean Wang, Sushil Jajodia. Protecting Privacy Against Location-Based Personal Identification. In Security Data Management, pages 185-199, 2005.
- [2] Alon Efrat and Arnon Amir. Buddy Tracking - Efficient Proximity Detection Among Mobile Friends. In IEEE International Conference on Computer Communications (INFOCOM), 2004.
- [3] Bugra Gedik and Ling Liu. Location Privacy in Mobile Systems: A Personalized Anonymization Model. In Proceedings of the 25th International Conference on Distributed Computing Systems, pages 620-629, 2005.
- [4] Marco Gruteser and Dirk Grunwald. Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In Proceedings of the First International Conference on Mobile Systems, Applications, and Services (MobiSys), 2003.
- [5] Haibo Hu, Jianliang Xu, Wing Sing Wong, Baihua Zheng, Dik Lun Lee, Wang-Chien Lee. Proactive Caching for Spatial Queries in Mobile Environments. In Proceedings of the 21st International Conference on Data Engineering (ICDE), pages 403-414, 2005.
- [6] Mohamed F. Mokbel, Chi-Yin Chow, and Walid G. Aref. The New Casper: Query Processing for Location Services without Compromising Privacy. In Proceedings of the 32nd International Conference on Very Large Data Bases (VLDB), pages 763-774, 2006.
- [7] Kyriakos Mouratidis, Panos Kalnis, Gabriel Ghinita and Dimitris Papadias. Preventing Location-Based Identity Inference in Anonymous Spatial Queries. IEEE Trans. Knowl.Data Eng., 2007.
- [8] Sarah Spiekermann. General Aspects of Location Based Services. Location-Based Services, pages 9-26. 2004.
- [9] Latanya Sweeney. k-Anonymity: A Model for Protecting Privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(5):557-570, 2002.
- [10] Yu Chen, Jie Bao , Wei-Shinn Ku, and Jiun-Long Huang, Cache Management Techniques for Privacy Protected Location-based Services, Proceedings of the 2nd International Workshop on Privacy-Aware Location-based Mobile Services (PALMS), in conjunction with the 9th International Conference on Mobile Data Management (MDM 2008), 2008.
- [11] Location Based Services, http://en.wikipedia.org/wiki/Location-based_services.
- [12] Gruteser, M., Grunwald, D., Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking, MobiSys pp. 31-42, 2003.
- [13] Du, J., Xu, J., Tang X. and Hu, H., iPDA: Support- ing Privacy-Preserving Location- Based Mobile Services, MDM pp. 212-214, 2007.

Reliability Analysis: The Mathematical Expression

Pooja Parnami, Ruchi Dave, Neha Singhal, Ankur Dutt Sharma

“Department of Computer Science, Information Technology & Mechanical Engineering”

Abstract - : The reliability engineering discipline has undergone evolutionary development and breakthroughs during the last six decades. The need for reliable products was first sensed in both commercial and military sectors in early 1950s. Since then enormous progress has been made in the area of reliability engineering. Before 1950s, the focus was either on quality control or on machine maintenance problems. Literature suggests that before World War II reliability was intuitive in nature and the basic concept of reliability was born during this time period.

In this paper with the help of Mathematical simulation, the reliability analysis will be proved for particular products. The concept of reliability analysis will be more useful to check any product or physical element’s durability and confidence of quality. To decide the standardization of any product, this reliability analysis will be important and key of success. The need of satisfaction related to physical products can be checked by these analysis also. Some of Mathematical approaches like Boolean algebra, Logarithm equations some of the formulas and mathematical expressions will be introduced. The condition between failure and durability will be considered to check the quality & reliability.

Some of the different types of distributions like Exponential, Weibull, and Gamma & Lognormal can be expressed in PDF or CDF curves for reliability analysis standardizations and proof.

Finally to check the reliability of substances these analyses will be useful and easy to prove product quality and other features.

I. Introduction

Recently, due to the increased competition, complex product design and development, the use of increasingly sophisticated manufacturing processes, particularly in the area of defense and space technology, and increasing focus on customer satisfaction, the question of reliability has become a matter of great interest.

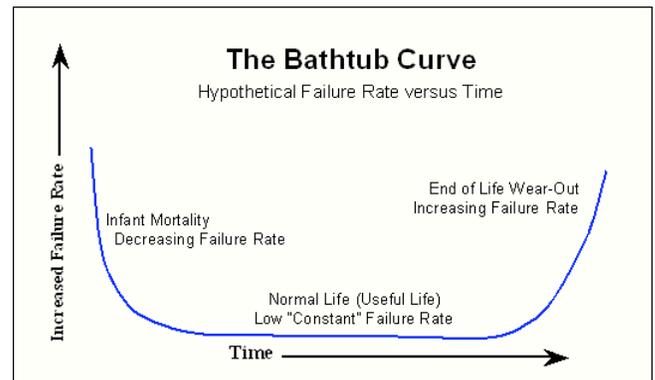
Reliability is defined as the probability that a component, device, system, or process will perform its intended function without failure for a given time when operated correctly in a specified environment. Reliability deals with reducing the frequency of failures over a time interval and is a measure of the probability for failure-free operation during a given interval, i.e., it is a measure of success for a failure free operation. It is often expressed as

$$R(t) = \exp(-t/MTBF) = \exp(-\lambda t) \dots\dots\dots(1)$$

where λ is constant failure rate and MTBF is mean time between failure. MTBF measures the time between system failures and is easier to understand than a probability number. For exponentially distributed failure modes, MTBF is a basic figure-of-merit for reliability (failure rate λ , is the reciprocal of MTBF). Also reliability may be the product of many different reliability terms such as-

$$R = R_{utilities} * R_{feed-plant} * R_{processing} * R_{packaging} * R_{shipping}$$

The life of a population of units can be divided into three distinct periods. Figure 1.1 shows the reliability “bathtub curve” which models the cradle to grave instantaneous failure rates vs. time. This way wear out should never occur during the useful life of a module.



1.1 Mean Time between Failures (MTBF)

Reliability is quantified as MTBF (Mean Time Between Failures) for repairable product and MTTF (Mean Time To Failure) for non-repairable product. A correct understanding of MTBF is important.

A power supply with an MTBF of 40,000 hours does not mean that the power supply should last for an average of 40,000 hours. According to the theory behind the statistics of confidence intervals, the statistical average becomes the true average as the number of samples increase. An MTBF of 40,000 hours, or 1 year for 1 module, becomes 40,000/2 for two modules and 40,000/4 for four modules.. The formula for calculating the MTBF is-

$$\theta = T/R \dots\dots\dots(2)$$

θ = MTBF (Mean time between failures)
 T = Total time
 R = Number of failures
 MTBF calculations do not consider suspensions whereas MTTF does. MTTF is the number of total hours of service of all devices divided by the number of devices. It is only when all the parts fail with the same failure mode that MTBF converges to MTTF. Then the formula for calculating the MTTF is-

$$\gamma = T/N \dots\dots\dots (3)$$

γ = MTTF (Mean time to failure)
 T = Total time
 N = Number of units under tes

2 Diagram based model for reliability analysis.

In system reliability analysis, it is important to model the relationship between various items as well as the reliability of the individual items in order to determine the reliability of the system as a whole. Diagram-based models provide a visual representation of the system and permit a better understanding of the target system. The visual (or physical) representation of an item that belongs to a system is often used to model system reliability. Diagram-based models involve Reliability Block Diagrams (RBDs), Failure Modes and Effects Analysis (FMEA), Fault Tree Analysis (FTA), Event Tree Analysis (ETA), Decision Tree Approach (DTA) and Root Cause Analysis (RCA) that are frequently used for reliability analysis.

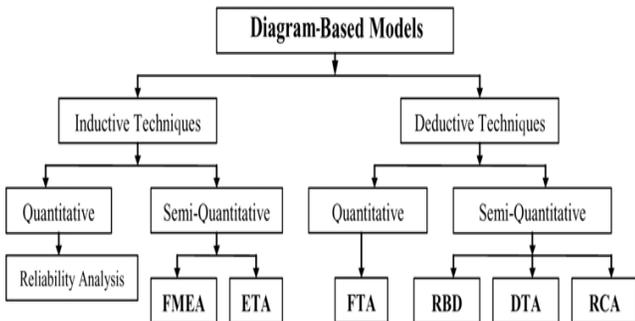


Fig. 2.1 Reliability block diagrams

System modeling tools can help you calculate the reliability, availability, and cost metrics of systems with complex interdependent component relationships. Relx OpSim provides the block diagram and phase diagram tools for system modeling combined with powerful optimization and simulation analysis techniques. Easy, intuitive graphical tools and a wide selection of maintenance-related calculations let you model complex system configurations including those with parallel, load-sharing, and standby redundancy types. Determine optimal values for spares,

preventive maintenance intervals, and inspection intervals while accounting for multiple facets of component maintenance, such as spares availability, personnel requirements, and degradation factors.

2.1 Failure modes and effects analysis (FMEA)

A failure modes and effects analysis (FMEA) pronounced fah-me-ah, is a procedure in operations management for analysis of potential failure modes within a system for classification by severity or determination of the effect of failures on the system. It is widely used in manufacturing industries in various phases of the product life cycle and is now increasingly finding use in the service industry. Failure modes are any errors or defects in a process, design, or item, especially those that affect the customer, and can be potential or actual. Effects analysis refers to studying the consequences of those failures.

2.2 Types of FMEA

- Process: analysis of manufacturing and assembly processes
- Design: analysis of products prior to production
- Concept: analysis of systems or subsystems in the early design concept stages
- Equipment: analysis of machinery and equipment design before purchase
- Service: analysis of service industry processes before they are released to impact the customer
- System: analysis of the global system functions

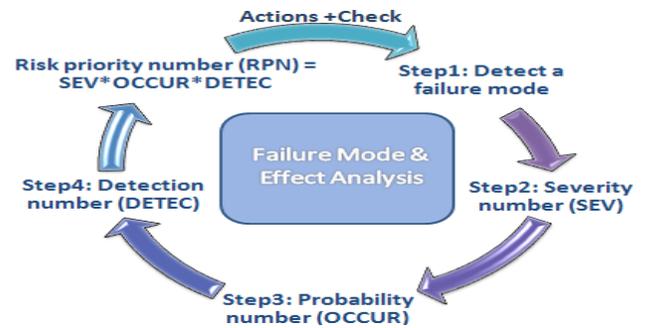


Fig. 3.2 FMEA

2.3 Element of FMEA

1. Failure mode: "The manner by which a failure is observed; it generally describes the way the failure occurs."
2. Failure effect: Immediate consequences of a failure on operation, function or functionality, or status of some item
3. Indenture levels: An identifier for item complexity. Complexity increases as levels are closer to one.
4. Local effect: The Failure effect as it applies to the item under analysis. Next higher level effect: The Failure effect as it applies at the next higher indenture level.
5. End effect: The failure effect at the highest indenture level or total system.
6. Failure cause: Defects in design, process, quality, or part application, which are the underlying cause

of the failure or which initiate a process which leads failure.

7. Severity: "The consequences of a failure mode. Severity considers the worst potential consequence of a failure, determined by the degree of injury, property damage, or system damage that could ultimately occur."

2.4 Reliability Analysis Procedures

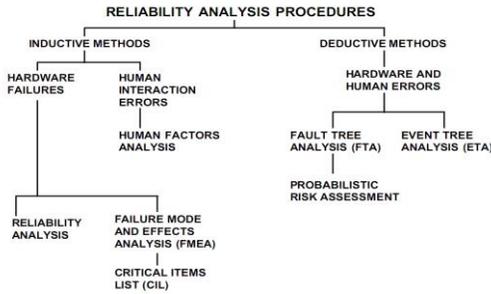


Fig.3.3 Reliability Analysis Procedures

Step 2.4.1: Severity---Determine all failure modes based on the functional requirements and their effects. Examples of failure modes are: Electrical short-circuiting, corrosion or deformation. It is important to note that a failure mode in one component can lead to a failure mode in another component. Therefore each failure mode should be listed in technical terms and for function. A failure effect is defined as the result of a failure mode on the function of the system as perceived by the user. In this way it is convenient to write these effects down in terms of what the user might see or experience. Examples of failure effects are: degraded performance, noise or even injury to a user.

Step2.4.2:Occurrence---- In this step it is necessary to look at the cause of a failure and how many times it occurs. This can be done by looking at similar products or processes and the failures that have been documented for them. A failure cause is looked upon as a design weakness. All the potential causes for a failure mode should be identified and documented.

Step 2.4.3: Detection--When appropriate actions are determined, it is necessary to test their efficiency. Also a design verification is needed. The proper inspection methods need to be chosen. First, an engineer should look at the current controls of the system, that prevent failure modes from occurring or which detect the failure before it reaches the customer.

2.5 Uses of FMEA

Development of system requirements that minimize the likelihood of failures. Development of methods to design and test systems to ensure that the failures have been

eliminated. Evaluation of the requirements of the customer to ensure that those do not give rise to potential failures. Identification of certain design characteristics that contribute to failures, and minimize or eliminate those effects. Tracking and managing potential risks in the design. This helps avoid the same failures in future projects.

Example-

Table 3.5.1 FMEA Table For Ball-Pen

Part	Function	Potential Failure Mode	Potential Effect of Failure	Severity	Potential Cause of Failure	Occurrence	Failure Detection	Detection	RPN	ACTIONS
Outer Tube	Provide Grip for writer	Hole gets Blocked	Vacuum on ink Supply Stops flow	7	Debris Increases Into hole	3	Check Clearance Of hole	5	105	Make hole larger
Ink	Provide Writing medium	Incorrect Viscosity (Low)	High Flow	4	Too Much Solvent	2	QC on ink supply	4	32	Introduce More Rigid QC
Ink	Providing Writing medium	Incorrect Viscosity (High)	Low Flow	4	Too Little Solvent	2	QC on ink supply	3	24	Introduce More Rigid QC

3 Benefits

Improve the quality, reliability and safety of a product/process, Improve company image and competitiveness, Increase user satisfaction ,Reduce system development timing and cost .

Collect information to reduce future failures, capture engineering knowledge ,Reduce the potential for warranty concerns ,Early identification and elimination of potential failure modes ,Emphasize problem prevention ,Minimize late changes and associated cost ,Catalyst for teamwork and

idea exchange between functions ,Reduce the possibility of same kind of failure in future ,Provides a wide range of system metrics, including Failure Rate, MTBF, MTTF, Reliability, and Availability, among others.,Model your complete operational profile using phase modeling ,Specify a calculation goal: minimize costs, maximize reliability, or maximize capacity ,Assign costs to various capacity levels ,Simulation log visually represents failures over time as each simulation is performed.

4 Limitations

Since FMEA is effectively dependent on the members of the committee which examines product failures, it is limited by their experience of previous failures. If a failure mode cannot be identified, then external help is needed from consultants who are aware of the many different types of product failure. FMEA is thus part of a larger system of quality control, where documentation is vital to implementation. General texts and detailed publications are available in forensic engineering and failure analysis. It is a general requirement of many specific national and international standards that FMEA is used in evaluating product integrity. If used as a top-down tool, FMEA may only identify major failure modes in a system. Fault tree analysis (FTA) is better suited for "top-down" analysis. When used as a "bottom-up" tool FMEA can complement FTA and identify many more causes and failure modes resulting in top-level symptoms.

5 Conclusions

Reliability engineering has come a long way over the last six decades and will go further to meet increasing global competition and customer expectations. The future thrust areas which need further research work by the researchers in the area of reliability engineering can be uncertainty quantification, failure analysis of complex systems and life testing of equipments.

5.1 Uncertainty quantification

The greatest risk (uncertainty) associated with reliability predictions is the variability or non-deterministic nature of the distribution parameters. This variability is further enhanced by three types of deviations in product design, parameters and/or characteristics. The first type of deviation occurs from permanent changes in product design or features.

5.2 Failure analysis of complex systems

The increasing efforts to improve time-to-market and enhance product functionality are throwing different challenges to reliability community. Traditional failure analysis tools (including FMEA) provide a valuable way to the incorporation of latest technological advancement into new designs resulted in more complex product designs. This increased product complexity leads to an emergence of unpredictable failure behavior. These pressures and ever increasing competition in global market are challenging

reliability community to devise more efficient and effective failure analysis methods.

5.3 Reliability testing

Testing for reliability is about exercising an application so that failures are discovered and removed before the system is deployed. Because the different combinations of alternate pathways through an application are high, it is unlikely that you can find all potential failures in a complex application. However, you can test the most likely scenarios under normal usage conditions and validate that the application provides the expected service. As time permits, you can apply more complicated tests to reveal subtler defects.

6 References

1. Abdelaziz, A.R. (1999) 'A fuzzy-based power system reliability evaluation', *Electric Power System Research*, Vol. 50, pp.1-5.
2. Aggrawal, K.K. and Gupta, J.S. (1975) 'On minimising the cost of reliable systems', *IEEE Transactions on Reliability*, Vol. R-24, No. 3, p.205
3. Barringer, H. Paul and David P. Weber (1995), Where Is My Data For Making Reliability Improvements?, Fourth International Conference on Process Plant Reliability sponsored by Hydrocarbons Processing and Gulf Publishing Company, Houston, TX.
4. Barringer, H. Paul and David P. Weber (1996a), Life Cycle Cost Tutorial, Fifth International Conference on Process Plant Reliability sponsored by Hydrocarbons Processing and Gulf Publishing Company, Houston, TX.
5. Barringer, H. Paul (1996c), An Overview Of Reliability Engineering Principles, Energy.
6. Barringer, H. Paul (1997), Life Cycle Costs & Reliability For Process Equipment, Energy.
7. Blanchard, B. S., Dinesh Verma, Elmer L. Peterson 1995, Maintainability: A Key to Effective Serviceability and Maintenance Management, Prentice-Hall, Englewood Cliffs, NJ.
8. Clements, Richard Barrett (1991), Handbook of Statistical Methods in Manufacturing, Prentice Hall, Englewood Cliffs, New Jersey.
9. Dugan, J.B. (1999) 'Fault-tree analysis of computer-based systems', *1999 Tutorial notes*, Reliability and Maintainability Symposium, Washington, DC.
10. Kececioglu, Dimitri 1995, Maintainability, Availability, & Operational Readiness Engineering, Prentice Hall PTR, Upper Saddle River, NJ.
11. Mitra, Amitava, "Fundamentals of quality control and Improvement" Pearson education.
Newton, David "Practical Reliability engineering", John wiley and sons Ltd.
12. Pecht, Michael 1995, Product Reliability, Maintainability, and Supportability Handbook, CRC Press, New York.

Simulation of Beamforming Solution of Interference Reduction for High Altitude Systems

¹Bharati.L.Rathod, ²Mr.Hemanthkumar.P, ³Mr.Aaquib.Nawaz.S

¹IV Semeste, M.Tech (DEC)

²Asst.Prof.ECEDept

AMCEC, Bangalore,

Abstract-

Interference reduction is vital for being able to effectively communicate with mobile users in rugged terrain and mountainous regions .It is proposed to outfit a high flying airborne node with a Code Division Multiple Access (CDMA) base station in order to provide line of sight communications and continual coverage to the remote users in a highly congested environment. The results will One approach to increasing capacity and coverage zones for the servicing wireless station is to use smart antennas.

This paper simulates many Beamforming algorithms namely SMI, LMS, VSSLMS, Griffiths, VSSG, EDNSS and ENSS. The algorithms provides different ways by which we can calculate the phase shifts and apply to individual antenna elements so that main beam is formed in the look direction and nulls or reduced radiation is formed in the jammer directions.

The two algorithms namely EDNSS & ENSS from adaptive filtering and applied to smart antenna concepts of beamforming. The ENSS algorithms achieves less error and high convergence as compared to existing beamforming other algorithms.

Keywords—Smart Antenna, Beam forming Algorithms, least mean square (LMS),variable step size LMS, Griffith's variable step size Griffith's, variable error data normalized step size(EDNSS) error normalized step size(ENSS).

I. Introduction

In recent years a substantial increase in the development of broadband wireless access technologies for evolving wireless internet services and improved cellular systems has been observed. It is widely foreseen that in the future an enormous rise in traffic mobile and personal communication systems. This is due to an increase in number of users and introduction of new high bit rate data services. This trend is observed for second-generation systems as well and it will most certainly continue for third-generation systems.

The rise in traffic put a demand on both manufacturers and operators to provide sufficient Capacity in the network, this becomes a major challenging problem for the service providers to solve.

There are certain negative factors like co-channel interference and fading in the radiation environment contributing to the limit in the capacity. Smart antenna is one such development in this direction to full fill the feature requirements of mobile networks.

Smart Antenna with signal processing algorithm used to identify spatial signature such as DOA of signal and to locate desired beam .

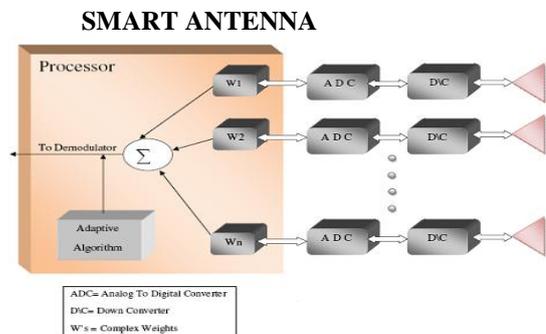


Fig-1: Block diagram of smart antenna system

It locates signal of interest (SOI) using Direction of Arrival (DOA) algorithm. The two basic functions of any smart antenna are (DOA) Estimation and Adaptive antnnas) arranged in linear fashion (uniform linear array) to extract the spatial information from the waves.The signal received at sensor is sent for computation of weights.

II. Implementation

Let $w(n)$ denote unit sample response of the FIR wiener filter that produces minimum mean square.A beamforming is a set of sensors arranged in a linear fashion (uniform linear array).It is a signal processing technique used in sensor array for directional signal transmission or reception.

The signal received at sensor is sent for computation of weights.

$$W(n+1)=w(n)+\Delta w(n)$$

Where $W(n)$ is beamforming array weights, $\Delta w(n)$ is correction applied to new weights.

1. Least Mean Square (Lms)

The LMS algorithm is the most widely used adaptive beamforming algorithm, being employed in several communication applications. It has gained popularity

due to its low computational complexity and proven robustness. The LMS algorithm changes the weight vector $w(n)$ along the direction of the estimated gradient based on the steepest descent method. In employing the LMS algorithm, it is assumed that sufficient knowledge of the reference signal is present.

The weight vector update equation assumes a particular simple form as given by

$$w(n+1) = w(n) + \mu e(n) x^*(n) \quad (1.1)$$

Where, μ is the step size which can be in the range given by

$$0 \leq \mu \leq \frac{2}{3tr(R_{xx})} \quad (1.2)$$

2. Variable Step Size Lms (Vss-Lms)

Variable Step-Size LMS (VSS LMS) algorithms are used, with the intention of decreasing misadjustment and to maximize convergence rate. Step size is larger when the estimate is far from the optimum value and a smaller step-size as it approaches the optimum value. The performance of this method is promising especially in non stationary environment.

The step size is calculated, during each iteration by using equation

$$\mu(n+1) = \alpha \mu(n) + \gamma |e(n)|^2 \quad \text{Where,} \quad (2.1)$$

' α ' indicates the correlation of the present step size to the previous step size, it is in the range $0 < \alpha < 1$ and γ is used to control convergence characteristics of VSS-LMS algorithm, $\gamma = 0.5$ and error signal $e(n)$ is given by

$$e(n) = d(n) - y(n) \quad (2.2)$$

During each iteration the step size is changed as

$$\begin{aligned} \mu(n+1) &= \mu_{upper} && \text{if } \mu(n+1) > \mu_{upper} \\ &= 0 && \text{if } \mu(n+1) < 0 \\ &= \mu(n+1) && \text{otherwise} \end{aligned} \quad (2.4)$$

3. Griffiths' Algorithm

Griffiths' algorithm [16] utilizes certain a priori knowledge (when available) to create an effective real time adaptation process. The algorithm can be used for a number of applications including Noise Control, Adaptive Beamforming and Acoustic Signal Processing. The algorithm is ideal when the cross-correlation between the desired response and the input signal vector (i.e. the inputs to the weights) are known a priori. In this case, the algorithm can be executed without the need for a real time response input $s(n)$.

The weight update equation for Griffith's algorithm can be derived by using LMS algorithm, which is given by

$$\begin{aligned} w(n+1) &= w(n) + 2\mu e(n)x(n) \\ &= w(n) + 2\mu (d(n) - y(n))x(n) \\ &= w(n) + 2\mu r_{sx} - 2\mu y(n)x(n) \end{aligned} \quad (3.2)$$

Where, μ is the step size, r_{sx} is the cross-correlation and $y(n)$ is the array output.

4. Variable Step Size Griffiths' (Vssg)

The Variable Step Size Griffiths' algorithm [17] is a combination of the Variable Step Size LMS algorithm and the Griffiths' algorithm. The algorithm is expected to combine the merits of the Variable Step Size LMS algorithm and Griffiths' algorithm.

The motivation behind this algorithm is, the use of an algorithm which would achieve faster convergence through the use of a variable step size LMS algorithm and a smoother gradient through the use of the Griffiths' algorithm..

The weight update equation in case of VSSG algorithm is given by

$$W(n+1) = w(n) + \frac{\mu l}{\|x\|} \quad (4.1)$$

The value of l is computed by taking the difference between cross correlation of reference signal with induced signal and cross correlation between array output and induced signal given by

$$l = r - y^* x(n) \quad (4.2)$$

The upper bound for step size in VSSG is given by

$$\mu_{upper} = 0.07 \quad (4.3)$$

5. Variable Error Data Normalized Step Size Algorithm (Ved-Nss)

VED-NSS algorithm [8] has greater noise reduction performance and performs convergence analysis of Error Data Normalized Step-Size (EDNSS) algorithm. Adopting VED-NSS provides fast convergence at early stages of adaptation, while ensuring small final misadjustment.

In VED-NSS algorithm step size is varied between two bounds viz; μ_{upper} and μ_{lower} , which will provide fast convergence. The upper and lower bound are given by [22]

$$\mu_{upper} = \frac{2}{3tr(\text{COV}(x))} \quad (5.1)$$

$$\mu_{lower} = \frac{2}{tr(R_x)} \quad (5.2)$$

Where $tr(\text{COV}(x))$ is the trace of covariance matrix. $tr(R_x)$ is trace of auto correlation matrix. The weight update equation in VED-NSS algorithm is given by

$$w(n+1) = w(n) + \frac{\mu(n+1) e(n) x(n)}{(\alpha_1 e_s + ((1 - \alpha_1))} \quad (5.3)$$

Where, α_1 is constant chosen for fast convergence, $\mu(n+1)$ is defined as in equation (3.4) and e_s is the normalized error given by

$$e_s = e^*(n) e(n) \quad (5.4)$$

The step size is varied between two bounds during each iteration by using

$$\begin{aligned} \mu(n+1) &= \mu_{upper} && \text{if } \mu(n+1) > \mu_{upper} \\ &= \mu_{lower} && \text{if } \mu(n+1) < \mu_{lower} \\ &= \mu(n+1) && \text{otherwise} \end{aligned}$$

6. Error Normalized Step Size (Enss)

ENSS algorithm is same as VED-NSS [8] except that the weight update consists of normalization of error vector. The weight updation in ENSS algorithm is given by

$$w(n+1) = w(n) + \left[\frac{\mu(n+1)}{(1 + \mu(n+1)e_s(n))} \right] e(n)x(n) \quad (6.1)$$

III. Simulation Results:

Case1: Less Antenna Elements and Less Jammer with one desired angle Desired Angle=45 No.of Antenna=8.

Case2: more Antenna Elements and Single Jammer. Desired Angle=30 No.of Antenna Elements=100.

Case3: Less Antenna Elements and More Jammer. Desired Angle=45 No.of Antenna=8 Jammer Angle=5, 30, 60.

Case4: more Antenna Elements and MoreJammer. Desired Angle=30 No.of Antenna Elements=100 Jammer Angle=5, 10, 60.

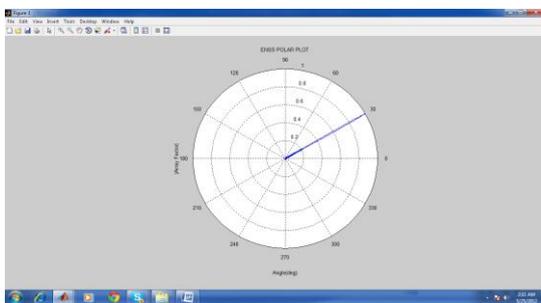


Fig 2: Polar plot of ENSS which forms main beam at 30,while reducing jammers in non signal of interest (NSOI).

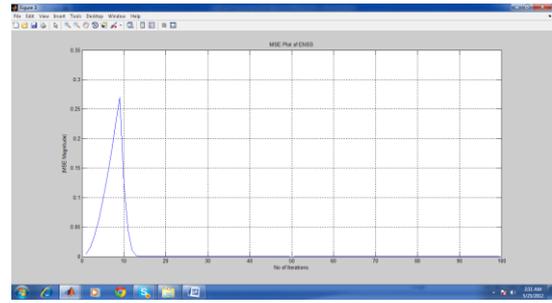


Fig3: MSE plot of ENSS

Converges at about 15iteration with less error compare to all other algorithms.

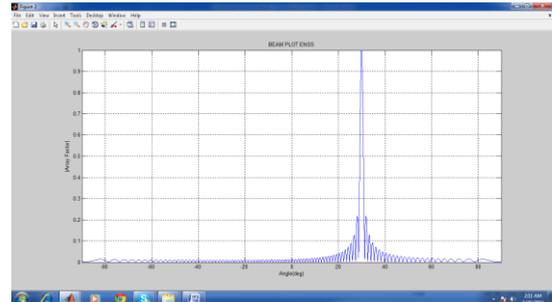


Fig4:Beam Plot with Desired Angle=30

Converges around 15 iterations whereas other algorithms converges about 100 iteration

Table: comparison of MSE Plots(100iterations)

	Case1	Case2	Case3	Case4
LMS converges	More than 100	More than 100	More than 100	More than 100
VSSLM S	Better than LMS	Better than LMS	Better than LMS	Better than LMS
Griffiths	90	90	98	95
VSSG	Less error	Less error	Less error	Less error
EDNSS	better	better	better	Better
ENSS	40	15	40	15

I. Conclusion

In the phase the Beamforming algorithms namely; VSS,VSS-LMS,Griffiths,VSSG,VED were simulated and compared. These algorithms were able to produce main beam towards desired direction and direct nulls towards interference directions. The VSSLMS provides faster convergence as compared to LMS. The weight calculation is performed by varying the step size with an upper limit on the step size so that the algorithm does not diverge from the optimum value. The Griffiths' algorithm will use a constant step size and it requires knowledge of both autocorrelation and cross correlation. The VSSG algorithm combines the advantages of both VSSLMS and Griffiths' in order to

improve convergence. Simulation result of beamforming algorithms showed that the convergence of VSSG is fast followed by Griffiths and VSS-LMS.

References

- [1] George V Tsoulos , “Smart antennas for mobile communication systems: Benefits and Challenges” , *IEEE proceedings of international conference on information technology*, Vol-11, Issue-2 ,pp 84-94.
- [2] D. G. Manolakis, Vinay K Ingle and Stephen M Kogon, “Statistical and Adaptive Signal processing spectral estimation, signal modeling, Adaptive filtering and array processing,” Mc Graw Hill 2000.
- [3] Jack H. Winters , “Smart Antennas for Wireless Systems” , *proceedings of IEEE international conference on signal processing*, June 2005, pp 107-109.
- [4] Okello, Y.Fukui, I.Nakanishi, Kobayashi, M ”A New Modified Variable Step Size Algorithm for the VSSLMS algorithm”, *IEEE Proceedings of the International Symposium on Circuits and Systems*, June 1998 , Vol 5,pp 170 – 173.
- [5] L. C. Godara, “Application of antenna arrays to mobile communication. II. Beam-forming and direction-of-arrival considerations”, *Proceedings of IEEE*, Vol 85, No 8, 1997, pp 1195-1245.
- [6] Chen sun, Nemai Chandra Karmakar, “Direction of Arrival Estimation based on a Single-Port Smart Antenna using MUSIC Algorithm with periodic signals”, *International Journal of Signal Processing*, 2004,Vol 1, No 2, pp153-162.
- [7] Jeffrey Foutz, Andreas Spanias, and Mahesh K. Banavar, “Narrowband Direction of Arrival Estimation for Antenna Arrays”, *Morgan & Claypool Publishers series*, 2008.
- [8] Alexander D.Poularikas, Zayed M.Ramadan, “Adaptive filtering Primer with MATLAB”, *CRC Press*, 2006.
- [9] Lal Chand Godara, “ smart antennas”, *CRC press*, 2004.
- [10] Frank B. Gross, “ Smart Antennas for Wireless Communications”, *McGraw-Hill*, 2005.
- [11] Ahmed El Zooghby, “Smart Antenna Engineering”, *Artech House mobile communications series*, 2005.
- [12] Dimitris G. Manolakis, Vinay K. Ingle, Stephen M. Kogon, “Statistical and Adaptive Signal Processing”, *ARTECH HOUSE publication, INC*, 2005.
- [13] Nasser Kehtarnavaz, “Real-Time Digital Signal Processing Based on the TMS320C6000”, *Newnes publishers*, 2005
- [14] Monson H. Hayes, “Statistical Digital Signal Processing and Modeling”, *John Wiley and Sons*, 1966.
- [15] Constantine A. Balanis, Panayiotis I. Ioannides, “Introduction to smart antenna”, *Morgan & Claypool publishers*, 2007
- [16] L.J Griffiths, “A simple Adaptive algorithm for real time processing in Antenna arrays”, *proceedings of IEEE*, Oct -1969, Vol-57.
- [17] Kumaswamy .V,S.Ravishankar,”Implementation of Variable Step Size Griffiths algorithm for beamforming “,*Proceedings of IEEE International symposium on microwaves*,2008,Vol-8,pp:31-35.

Website

- [1] www.ti.com
- [2] www.spectrumsignalprocessing.com

Floating Point Unit Implementation on FPGA

Deepa Saini¹, Bijender M'dia²

^{1,2}(Electronics & Communication, M.D.University, INDIA

ABSTRACT

As densities of FPGA are increasing day by day, the feasibility of doing floating point calculations on FPGAs has improved. Moreover, recent works in FPGA architecture have changed the design tradeoff space by providing new fixed circuit functions which may be employed in floating-point computations. By using high density multiplier blocks and shift registers efficient computational unit can be developed. This paper evaluates the use of such blocks for the design of floating-point units including adder, subtractor, multiplier and divider.

1. Introduction:

Floating-point unit is a part of a computer system specially designed to carry out operations on floating point numbers. Addition, subtraction, multiplication, division are the typical operation of floating point unit. Floating-point operations are often pipelined. In superscalar architectures without general out-of-order execution, floating-point operations were sometimes pipelined separately from integer operations.

Over the past few years the use of FPGAs in compute-intensive applications has been growing. The vast majority of applications have employed fixed-point arithmetic due to its smaller size. The key advantage of floating-point over fixed-point is its ability to automatically scale to accommodate a wide range of values using its exponent. Floating-point is thus preferred by programmers for non-integer computations when it is available on CPUs due to its ease of use. However, this scaling behavior comes at the cost of reduced accuracy. A 64-bit fixed point Representation can have more accuracy (but less range) than a 64-bit floating-point representation.

2. Floating Point Unit Organization:

The FPU chip performs all floating-point functions for microprocessor chip set. The FPU has two fully- pipelined execution units, allowing two floating-point mathematical operations and two floating-point memory operations every cycle. The FPU register file contains 32, 64-bit entries and has eight read ports and four write ports. Load and store data queues provide a pipelined interface between the IU and the FPU, streamlining data flow and minimizing unused cycles. The FPU offers peak performance of 300 double-precision MFLOPS with a clock frequency of 75 MHz. The IU places floating-point instructions in the floating-point instruction queue in the IU. Each entry in the floating-point instruction queue is arranged as a quad word. The dispatch mechanism is similar to that of the IU: from

Zero to four floating-point instructions can be dispatched by the IU to the FPU each cycle. Instructions are dispatched only when the FPU has adequate resources available to execute the instructions. The floating-point instruction queue provides temporary storage for floating-point instructions while any dependencies that might prevent the floating-point instructions from being executed (such as waiting for dependent loads to complete, etc.) are cleaned up.

Floating-point instruction dispatches and floating-point loads and stores to the data streaming cache are controlled by the IU. The IU is responsible for generating all address and control signals for floating-point loads and stores to the data streaming cache. Accesses that miss in the data streaming cache and require interfacing to main memory are handled by the off-chip cache controller. During these miss cycles, the FPU continues to execute floating point instructions already in the queue.

Once floating-point data is retrieved from the data streaming cache, it is placed in the load data queue. For store operations, the result is placed in the store data queue. As soon as the corresponding address information from the tag RAM is available, the data is written out to the data streaming cache. Figure 1 diagrams the floating-point data path.

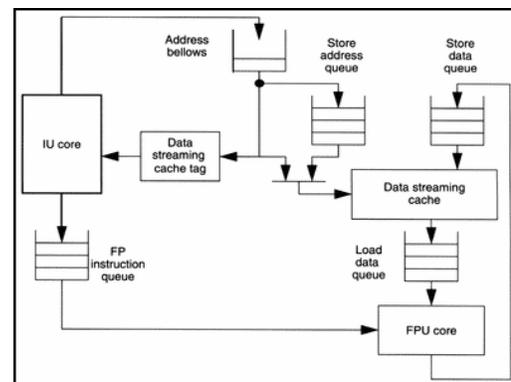


Figure1: Floating Point data path

3. Conceptual Overview:

The input operands are separated into their mantissa and exponent components. The comparison of the operands to determine which is larger only compares the exponents of the two operand, so in fact, if the exponents are equal then both the input numbers are treated equally to populate the registers.

This is not an issue because the reason the operands are compared is to find the operand with the larger exponent, so that the mantissa of the operand with the smaller exponent can be right shifted before performing the addition. If the exponents are equal, the mantissas are added without shifting.

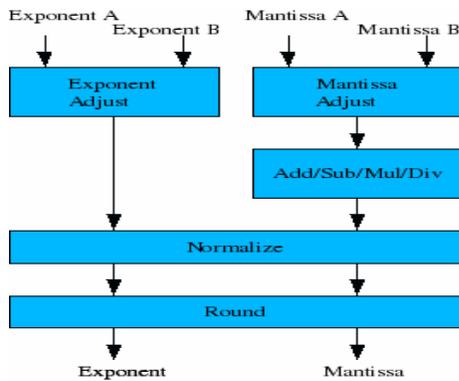


Figure 2: Conceptual overview of FPU

4. Floating Point Numbers

There are several ways to represent real numbers on computers. Floating-point representation - the most common solution - basically represents real in scientific notation. Scientific notation represents numbers as a base number and an exponent. For example, 123.456 could be represented as 1.23456×10^2 . In hexadecimal, the number 123.abc might be represented as $1.23abc \times 16^2$.

Floating-point solves a number of representation problems. Fixed-point has a fixed window of representation, which limits it from representing very large or very small numbers. Also, fixed-point is prone to a loss of precision when two large numbers are divided. Floating point on the other hand, employs a sort of "sliding window" of precision appropriate to the scale of the number. This allows it to represent numbers from 1,000,000,000,000 to 0.0000000000000001 with ease.

4.1 STORAGE LAYOUT

IEEE floating point numbers have three basic components: the sign, the exponent, and the mantissa. The mantissa is composed of the fraction and an implicit leading digit (explained below). The exponent base (2) is implicit and need not be stored.

The following figure shows the layout for single (32-bit) and doubles (64-bit) precision floating-point values. The number of bits for each field are shown (bit ranges are in square brackets):

	Sign	Exponent	Fraction	Bias
Single Precision	1 [31]	8 [30-23]	23 [22-00]	127
Double Precision	1 [63]	11 [62-52]	52 [51-00]	1023

Figure 3: Layout for single and double bit precision floating point values

4.1.1 THE SIGN BIT

The sign bit is as simple as it gets. 0 denotes a positive number; 1 denotes a negative number. Flipping the value of this bit flips the sign of the number.

4.1.2 THE EXPONENT

The exponent field needs to represent both positive and negative exponents. To do this, a bias is added to the actual exponent in order to get the stored exponent. For IEEE single-precision floats, this value is 127. Thus, an exponent of zero means that 127 is stored in the exponent field. A stored value of 200 indicates an exponent of (200-127), or 73. The exponents of -127 (all 0s) and +128 (all 1s) are reserved for special numbers.

For double precision, the exponent field is 11 bits, and has a bias of 1023.

4.1.3 THE MANTISSA

The mantissa, also known as the significand, represents the precision bits of the number. It is composed of an implicit leading bit and the fraction bits.

In order to maximize the quantity of representable numbers, floating-point numbers are typically stored in normalized form. This basically puts the radix point after the first non-zero digit. In normalized form, five is represented as 5.0×10^0 .

A nice little optimization is available to us in base two, since the only possible non-zero digit is 1. Thus, we can just assume a leading digit of 1, and don't need to represent it explicitly. As a result, the mantissa has effectively 24 bits of resolution, by way of 23 fraction bits.

4.2 Double Precision Floating Point Number

The IEEE 754 standard defines how double precision floating point number are represented. 64 bits are used to represent a double precision floating point number.

Sign	Exponent	Mantissa
63	62.....52	51.....0

Figure 4: Double precision floating point number bit format

5 Floating Point Unit Ip Core

The floating point IP core is separated into 9 source files:

1. fpu_double.vhd (top level)
2. fpu_add.vhd
3. fpu_sub.vhd
4. fpu_mul.vhd
5. fpu_div.vhd
6. fpu_round.vhd
7. fpu_exceptions.vhd
8. fpupack.vhd
9. comppack.vhd

5.1 HIERARCHY:

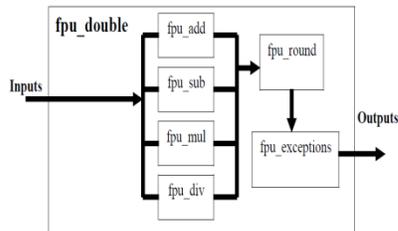


Figure 7: Hierarchy of various source file

5.2 TOP LEVEL

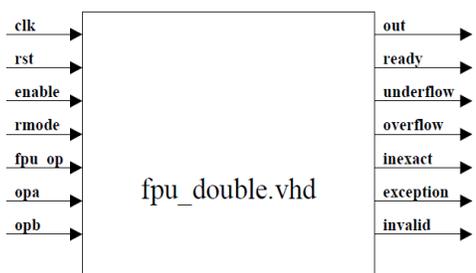


Figure 8: Top level module of Floating point unit

The input signals to the top level module are the following:

1. clk (global)
2. rst (global)
3. enable (set high to start operation)
4. rmode (rounding mode, 2 bits, 00 = nearest, 01 = zero, 10 = pos inf, 11 = neg inf)
5. fpu_op (operation code, 3 bits, 000 = add, 001 = subtract, 010 = multiply, 011 = divide, others are not used)
6. opa, opb (input operands, 64 bits)

The output signals from the module are the following:

7. out_fp (output from operation, 64 bits)
8. ready (goes high when output is available)
9. underflow
10. overflow
11. inexact
12. exception
13. invalid

The top level, fpu_double, starts a counter (count_ready) one clock cycle after enable goes high. The counter (count_ready) counts up to the number of clock cycles required for the specific operation that is being performed. For addition, it counts to 20, for subtraction 21, for multiplication 24, and for division 71. Once count_ready reaches the specified final count, the ready signal goes high, and the output will be valid for the operation being performed. fpu_double contains the instantiations of the other 6 modules,

which are 6 separate source files of the 4 operations (add, subtract, multiply, divide) and the rounding module and exceptions module. If the fpu operation is addition, and one operand is positive and the other is negative, the fpu_double module will route the operation to the subtraction module. Likewise, if the operation called for is subtraction, and the A operand is positive and the B operand is negative, or if the A operand is negative and the B operand is positive, the fpu_double module will route the operation to the addition module. The sign will also be adjusted to the correct value depending on the specific case.

6 Simulation Result

The generic and Spartan3-E optimized designs were similar for the add/sub and multiplier units, with the optimized designs simply using shift registers and the 18x18 built-in multipliers. The divider units were fundamentally different from one another. The significand divider for the generic unit was a restoring array divider, while the optimized design used the 18x18 built-in multipliers. The word sizes tested which show the best performance for the optimizations presented include 16-bits (9-bit significand), 23-bits (16-bit significand), and 41-bits (32-bit significand).

This is due to there being a good match between the significand size and the width of the available multiplier blocks in Spartan-3E. In addition, a standard IEEE 32-bit format was run (23-bit significand) which shows less benefit due to not as good a match between significand size and multiplier block.

In a configurable computing environment there may be no special significance to using the standard IEEE word sizes other than they match what is used on CPUs, simplifying validation. In many cases, however, non-standard word sizes may be profitably employed Three different versions of each module are represented — the generic module, an optimized module which uses both built-in multipliers and shift registers. Of those area savings, the multiplier and divider the majority of the area savings was due to the use of the multiplier blocks Results of synthesize and simulation are shown in figure 9 ,10 and 11.g

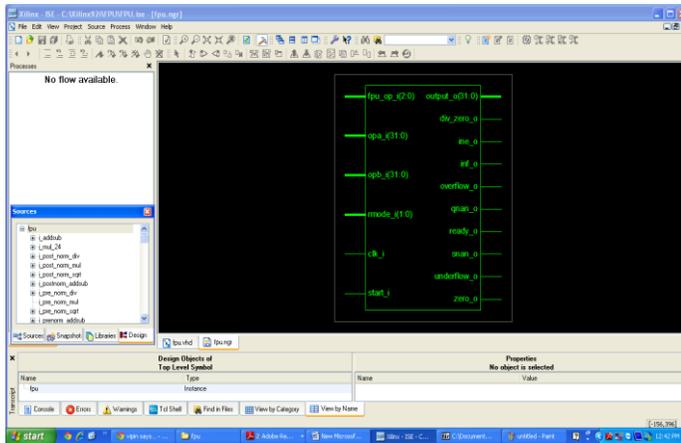


Figure 9: Synthesize Result of Floating Point unit

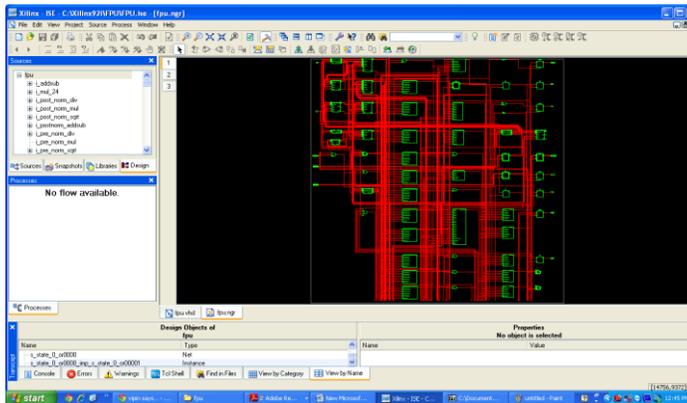


Figure 10: Detailed view of Floating point unit after synthesize

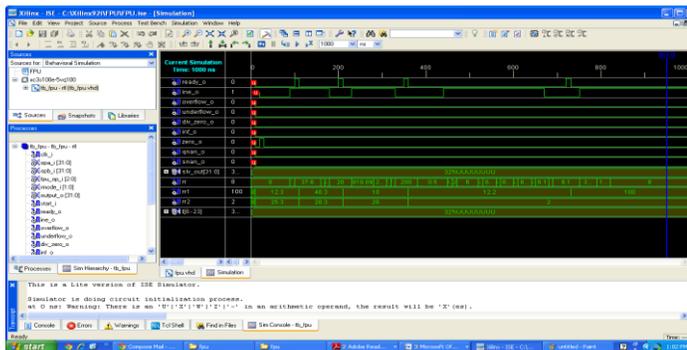


Figure 11: Simulation Result of Floating Point unit

7, Conclusions

This paper see the effects of new FPGA features like multiplier blocks and shift registers on the designing of floating point unit that performs addition, subtraction, multiplication and division. And research shows that area required by Floating point unit for doing multiplication and division by using newly added block is much less in comparison to floating point unit using LUT's FF's only.

References

- [1] N. Shirazi, A.Walters, and P. Athanas, "Quantitative analysis of floating point arithmetic on FPGA-based custom computing machines," in *Proceedings of IEEE Workshop on FPGAs for Custom Computing Machines*, D. A. Buell and K. L. Pocek, Eds., Napa, CA, Apr. 1995, pp. 155–163.
- [2] Behrooz Parhami, *Computer Arithmetic*, Oxford Press, 2000.
- [3] Joseph J. F. Cavanagh, *Digital Computer Arithmetic*, McGraw-Hill, 1984.
- [4] *What Every Computer Scientist Should Know About Floating-Point Arithmetic*, by David Goldberg, published in the March, 1991 issue of *Computing Surveys*. Copyright 1991, Association for Computing Machinery, Inc., reprinted by permission.
- [5] LOW COST FLOATING-POINT UNIT DESIGN FOR AUDIO APPLICATIONS by *Sung-Won Lee and In-Cheol Park* Division of Electrical Engineering, Department of EECS, KAIST 373-1 Gusong-dong Yusong-gu, Taejon, 305-701, KOREA