International Conference on Computing and Communication '16

(ICCC'16)

January 28 ~ 29, 2016, Kerala, India

Mar Athanasius College of Engineering
Kothamangalam, Kerala, India

Volume Editors

AbyAbhahai T,
Computer Science and Engineering Department
Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
Email ID :abytom@gmail.com

Sidharth Shelly,
Electronics and Communication Department
Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
Email ID: sidhushelly@mace.ac.in

Arun K L,
Electronics and Communication Department
Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
Email ID: arunkl@mace.ac.in

ISSN: 2231 - 5403

ISBN:978-1-921987-47-2

DOI :10.5121/ijci.2016.5201 - 10.5121/ijci.2016.5244

Preface

The International conference on Computing and Communication (ICCC'16) was held in Mar Athanasius College of Engineering, Kothamangalam in Kerala during January 28 - 29, 2016 in association with ACM, Cochin chapter. Kothamangalam is a place better known as "The gateway to western ghats" is a land known for fables and churches. The college has constantly produced renowned personalities in the field of engineering for the past 54 years.

The objective of ICCC16 is to provide a platform for researchers, engineers, academicians as well as industrial professionals from all over the world to present their research results and development activities in Computing and Communication. There has been a rapid development in technology related to these areas during the past few years and has resulted in change in techniques adopted in various applications. The scope for further development always exist and the conference would provide the opportunity to discuss the state of the art and to explore the avenues for future work. All the submitted papers were reviewed by a panel of national and international experts. In addition we had workshops & tutorials relevant to the industry & academia, and well known speakers for the plenary and invited sessions during the conference.

We would like to thank the Chief patron, Patron, General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research.
It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

AbyAbhahai T,

Sidharth Shelly,

Arun K L

Organization

Chief Patron

Winny Varghese	Secretary, M.A College Association

Patron

Soosan George T.	Principal, MACE Kothamangalam

Advisors

Madhavan Ganesh	Director of Informatics , RPGEH, Kaiser, Permanente, Division of Research, Ooklan, CA, USA
BintoGeroge	Professor, Western Illinois University, USA
Shiby Thomas	Vice President, Enterprise Data Warehouse & Analytics, Lahey Health Systems, USA
S. Joseph Antony	Associate Professor, University of Leeds, LS29JT, UK
SathyaPeri	Assistant professor ,IIT Hyderabad,India
Ing. EtiennaNtagwirumugara	Professor, University of Rwanda, Kigali-Rwanda
A. B. Chattopadhyay	Professor, BITS- PILANI, Dubai, U.A.E
G. SenthilKumaran	Dean, University of Rwanda, Kigali-Rwanda
AnubarataDey	Asst. Professor , IIT Roorkee, India
S. Vadivel	Professor, BITS- PILANI, Dubai, U.A.E
Siva Kumar K	Asst. Professor, IIT, Hyderabad, Andhra Pradesh, India
R Vijayakumar	Professor, School of Computer Science, Kottayam, India
Rajasree M S	Director & Prof., IIITM-K,India

Technical Programme Committee

John Jose	Assistant Professor, IIT, Guwahati, India
Thomas T.G	Dean & Professor BITS- PILANI, Dubai, U.A.E
SreekrishnaBhat	Research Scientist, Ricoh Innovation Private Limited, Bengaluru, India
S. D Madhukumar	Associate Professor, Department of Computer Science, NIT Calicut, India
Sheena Mathew	Professor, School of CUSAT, India
Shiny Gopinath	Associate Professor, College of Engineering, Trivandrum, India
Ms. Sindu Thomas	Product Development Manager, Freddie Mac, Virginia, USA
Rajesh Cherian Joy	VJCET, Vazhakkulam, India
Samson Thomas	MIIM, Kuttikkanam, India
Lethakumari B	UCE, Thodupuzha, India
K. Gunavati	PSG, Coimbatore, India
Neelakantan P.C	ASIET, Kalady, India
P.T.Vanathi	Professor, PSG, Coimbatore, India
Ms.MiniUlanat	CUSAT,Cochin India
Mr.Anil P Y	CUSAT, Cochin, India
P.S. Subin	Dean & Professor, VISAT, Ernakulam, India
Bos Mathew Jose	Associate Professor,EEE, MACE
Vinodkumar Jacob	Professor, ECE, MACE
Mathew.K	Asst. Professor, ECE,MACE
JinsaKuruvilla	Assistant Professor, MACE,Kothamangalam, India

General Chair

Surekha Mariam Varghese	Professor and Head, Dept. of Computer Science &Engg, MACE

Program Chair

Sunny Joseph	Professor and Head, Dept. of Electronics & Communication Engg, MACE
Manu John	Head, Dept. of Computer Applications, MACE

Organizing Chair

Mary Joseph	Assoc. Professor, Dept. of Electronics & Communication Engineering, MACE
Beena Jacob	Asst. Professor, Dept. of Computer Applications, MACE
Linda Sara Mathew	Asst.Professor, Dept. of Computer Science and Engg, MACE

Finance Chair

Thomas George	Professor,Dept. of Electronics & Communication Engineering, MACE
Jisha P Abraham	Assoc.Professor,Dept. of Computer Science & Engineering, MACE
NishaMarkose	Asst.Professor, Dept. of Computer Applications, MACE

Website Chair

Eldo P Elias	Asst.Professor, Dept. of Computer Science & Engineering, MACE

Publicity Chair

Babu P Kuriakose	Assoc.Professor, Dept. of Electronics & Communication Engineering, MACE

Proceedings

Vinod Kumar Jacob	Professor, Dept. of Electronics & Communication Engineering, MACE
Joby George	Assoc.Professor, Dept. of Computer Science & Engineering, MACE

Technically Sponsored by

ACM Cochin

Organized by

Mar Athanasius College of Engineering
Kothamangalam, Kerala, India

Proceedings by

International Journal on Cybernetics & Informatics (IJCI)

AIRCC Publishing Corporation

TABLE OF CONTENTS

INTERNATIONAL CONFERENCE ON COMPUTING AND COMMUNICATION (ICCC 2016)

INVIVO PATTERN RECOGNITION AND DIGITAL IMAGE ANALYSIS OF SHEAR STRESS DISTRIBUTION IN HUMAN EYE

S. Joseph Antony
School of Chemical and Process Engineering
Faculty of Engineering, University of Leeds, UK

ABSTRACT

Human eye is made of a number of structural components to deliver vision, and cornea is the front window of the eye. Human cornea is made of soft biological materials. It is unfriendly for exposures to the common energy levels of X-ray scans for repeated probing of its structural architecture. Here we study an alternative imaging methodology by using a normal white-light source. By exploiting the natural birefringent property of cornea, the shear stress distribution pattern and its directional characteristics on the surface of cornea is recognized in vivo. Digital image processing of corneal retardation helps us to locate the stress concentration zones on its surface and to study their features along preferential directions. Such digital image outputs could be used in future to bench mark the health standard of cornea as well as a potential identity signature of people's eyes.

KEYWORDS

Digital imaging, pattern recognition, cornea stress, eye strain

1. INTRODUCTION

Medical imaging tools such as X-ray scanners, Magnetic resonance imaging (MRI), positron emission tracking (PET) and ultrasound scanners help us to detect and treat anomalies in human bodies [1]. A better eye sight is perhaps the most important factor influencing the quality of human life. Human eye comprises different transparent substances including cornea, the aqueous humour, the crystalline lens and the vitreous humour [2] Cornea resides at the front of the eye [2] as illustrated in Fig.1.

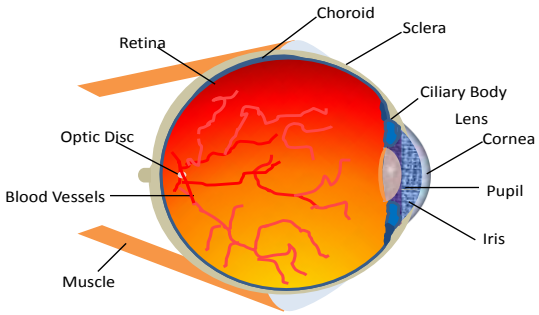


Figure 1: Schematic diagram of a human eye [3]

DOI: 10.5121/ijci.2016.5201

Previous studies on birefringent measurements in human eyes are mainlyfocused on the medical treatments Stress distribution in cornea could result due to the inherent molecular architecture of the corneal tissues [4,5] as well as contributed by the level of IOP in the eye (i.e, intra-ocular pressure measured conventionally as the average, stress microscopically is scarce in the literature. In a recent study [3], the author reported in vivo sensing of maximum shear stress distribution on human cornea using digital photo stress analysis tomography (PSAT). PSAT exploits the birefringent property of cornea tissues (Fig. 2). When a circularly polarised light from a normal white light source falls on the cornea, the out coming light is elliptically polarised [3]. Using an optical analyser, the light vectors of the elliptically polarised light at any point on the surface of the cornea can be characterised digitally (Fig.2). Using this information, the retardation of the light between the major and minor principal axis can be determined and related to the maximum shear stress using the stress-optic law [3].

Previous studies on the birefringence in human eyes focused on the medical treatments perspectives for example, to characterise retinal structures [7], their age effects [7] and in vitro analysis of keratoconus [8]. Using X-ray diffraction, some studies have shown diamond-like architecture of the collagen fibrils [9] which provides mechanical stability to the cornea [8]. However these studies are mostly in vitro as well as perhaps not suitable for repeated examinations of eye tissues. The shear stress patterns of cornea using PSAT had also correlated with the diamond-like architecture of the stress bearing fibril elements of the cornea [3].

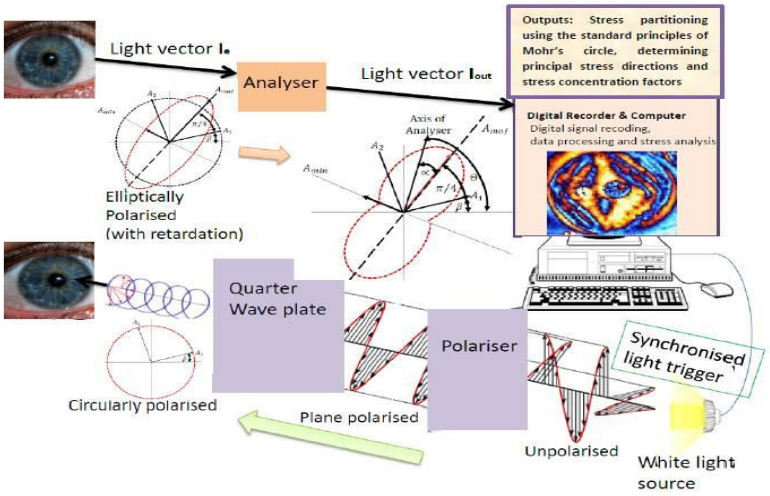


Figure 2: Schematic diagram of digital PSAT to recognise stress patterns in the eye [3]

2. STRESS DISTRIBUTION CHARACTERITICS OF CORNEA AS A BIOMARKER

The current study is the extension of our previous study [3] in which the shear stress concentration factors of a healthy human cornea in vivo were reported. For comprehensive details on the PSAT methodology and the partitioning of the stress components on cornea, the readers could refer to our previous work [3]. Here, we present some key features of the shear stress distribution characteristics (Fig.3, [10]) of a healthy cornea, and then digitally analyse for the direction of the major principal stress on its outer surface. We hope that such an analysis could potentially serve as a route map for bio marking the cornea in the individuals in future. Furthermore, bi-polar digital imaging scheme was applied to enhance the edge detection [11,12] of the stress bearing fibrils and this enhanced the visibility of diamond-like structure of the cornea fibrils.

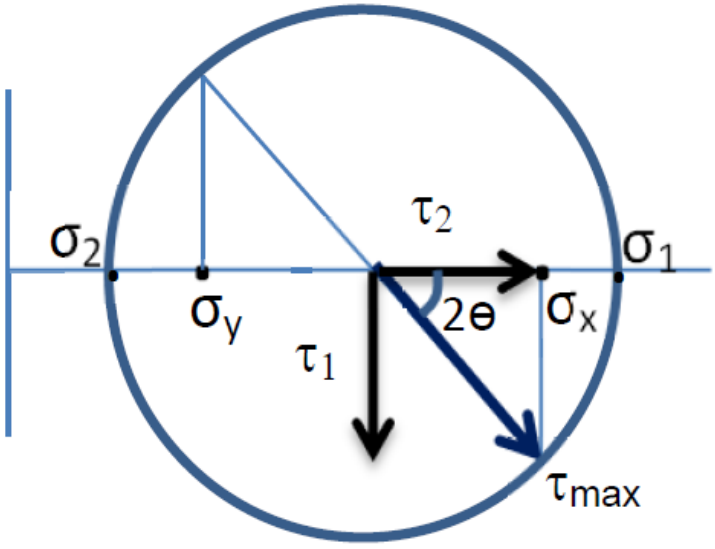


Figure 3: Partitioning of maximum shear τ_{max} at different planes using the principles of Mohr's circle [10]: τ_1 represents shear stress acting in horizontal and vertical planes whereas τ_2 represents shear stress acting at 45^{o} to the horizontal and vertical planes. τ_{max} acts at 45^{o} to the principal stress direction.

3. RESULTS AND DISCUSSION

3.1 Diamond-like stress-bearing structure of cornea fibrils

Figure 4 shows the edge detection of the fibrils of cornea distributing the maximum shear stress. Here the output is presented in terms of the retardation of the light (which can be also converted to the Pa unit [3]).

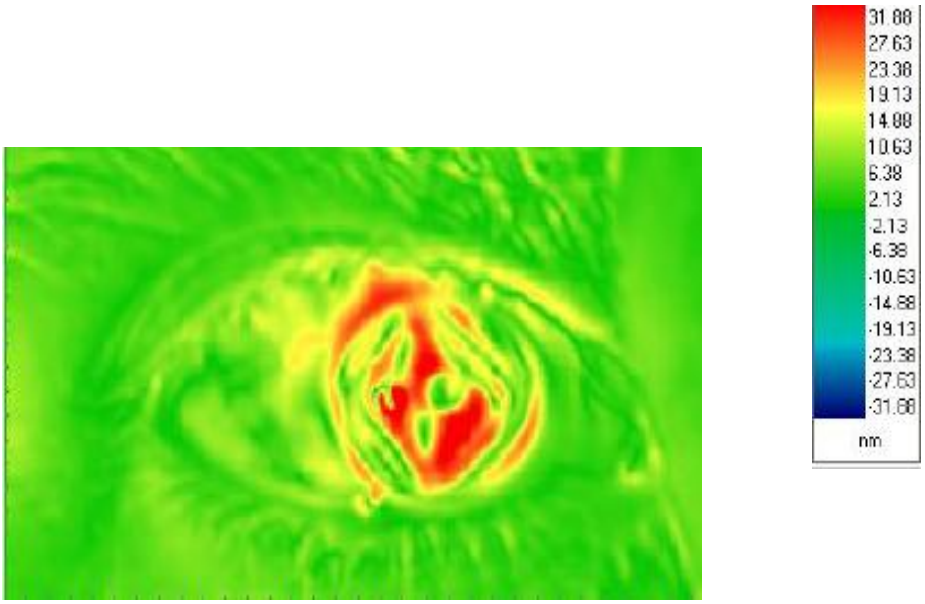


Figure 4: Diamond-like distribution of τ_{max}along the horizontal (+ve) and vertical (-ve) planes on cornea.

The digital stress map presents a diamond-like structure similar to the structure displayed by X-ray scans of cornea in vitro [9]. However, it is worth noting that the current approach does not involve the use of any X-ray source. Hence the current approach could be more suitable for any repeated measurements of stresses in the cornea in future applications.

3.2 Identification Of The Feature Sof Maximum Principal Stress Along The Major Diagonal Axis Of The Corneaal Diamond

In Figure 5, the axes of the diamond structure are illustrated for the purpose of digital analysis of its mechanical characteristics in the following sections.

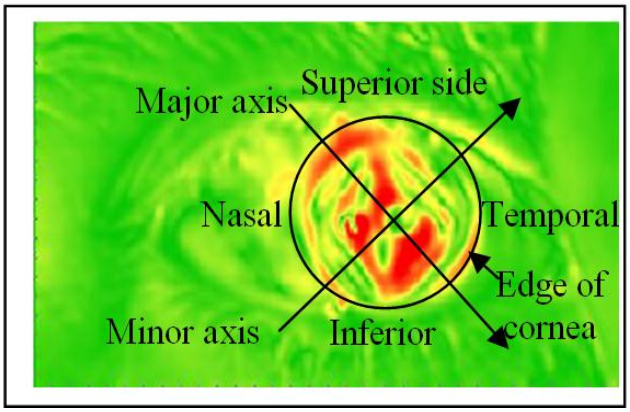


Figure 5: Illustration of the major and minor elliptical axes of the cornea diamond-like structure shown in Fig.4 and the anatomical terminologies.

From the digital information of the light r etardation discussed above, further analysis is performed to get the signature of the cornea by plotting the direction of major principal stress along the major axis of the 'corneal diamond'. This result is presented in Fig.6a, exhibiting equal numbers of three troughs (+ve degree) and crests (-ve degree) with well-defined spacing.

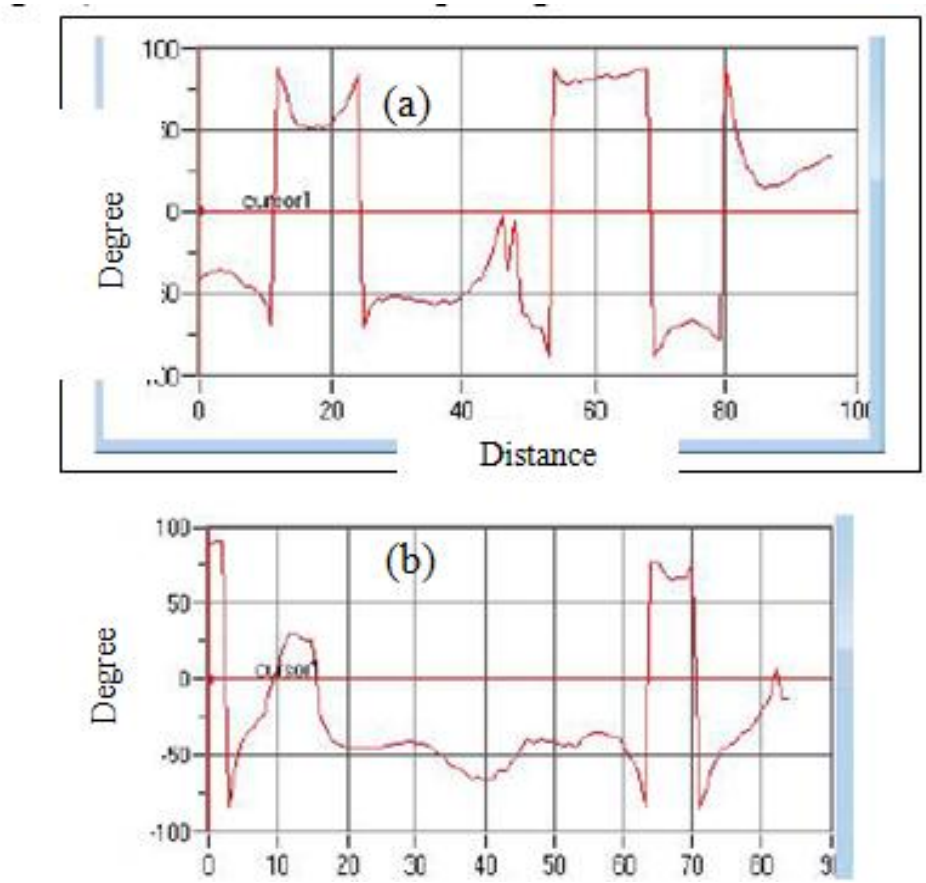


Figure 6: Variation of the direction of major principal stress along (a) major and (b) minor axes of the corneal diamond within the edges of the cornea

This information along the minor axis (Fig.6b) results unequal lengths of troughs and crests, and more dominantly with longer troughs along the middle region of the cornea. Hence along the minor axis, the direction of the major principal stress acts mostly along the vertical plane whereas along the major axis, the maximum shear stress is distributed along the horizontal and vertical planes periodically. Hence the distinctive features of this signature along the major axis can be studied in future, for example the spacing and width of the troughs and crests in relation to the angle of major principal stress as a signature of cornea in different individuals. For completeness, the maps of the direction of the major and minor principal stresses on the surface of cornea are presented in Fig.7.

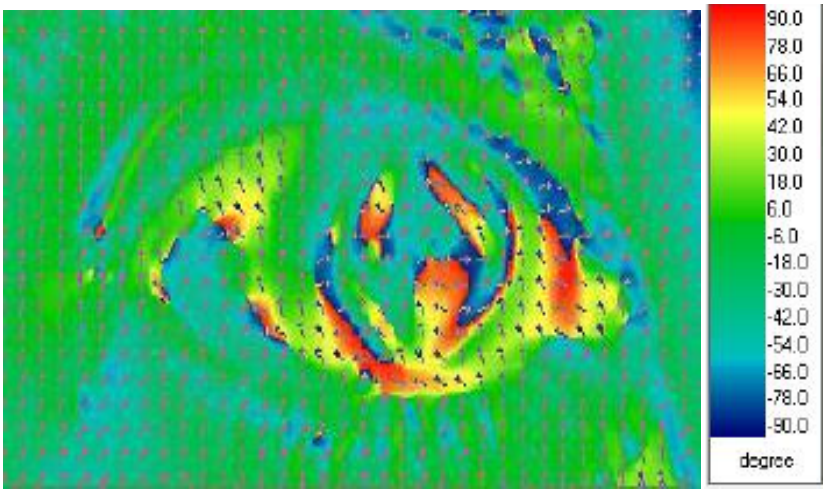
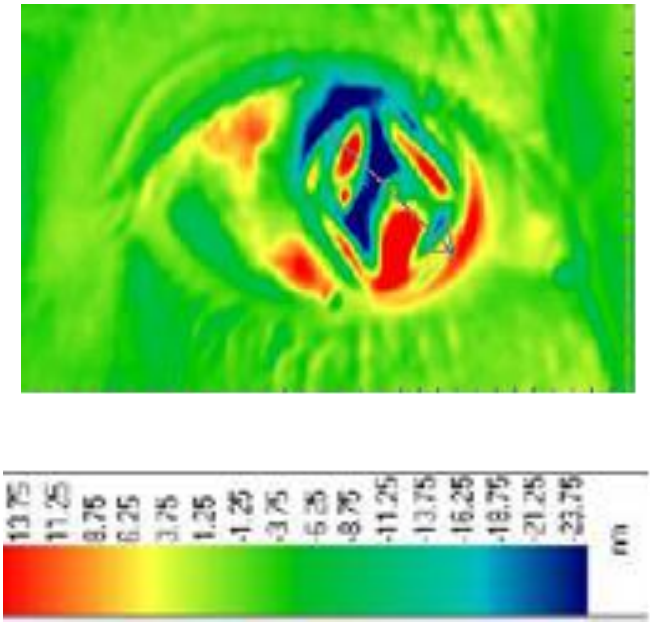


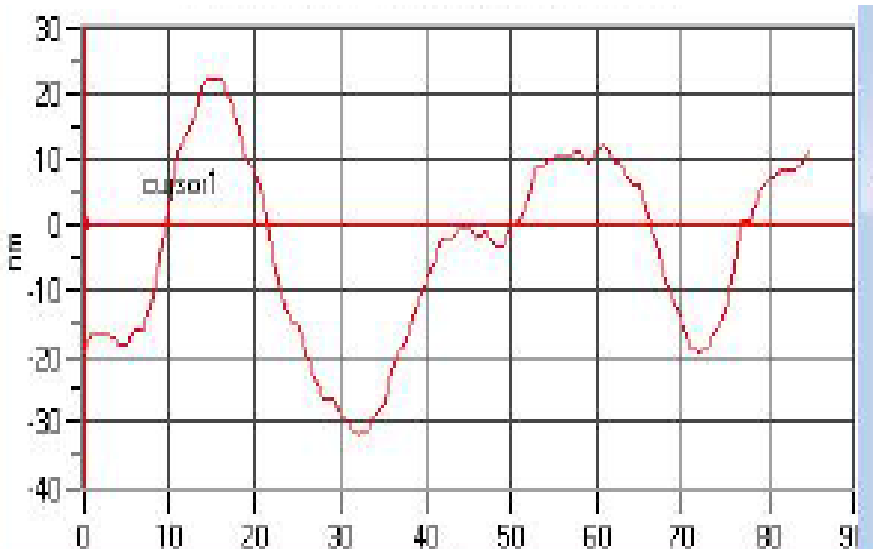
Figure 7: Major principal stress distribution is colour coded and the arrows show the direction of the minor principal stress on the cornea

3.3 Distribution of shear stress acting along the horizontal and vertical planes of cornea

Similar to Fig.6, the distribution of shear stress along the horizontal (+ve) and vertical (-ve) planes of the cornea are presented in Fig.8. In this plot, the digital information of these along the major principal axes of the diamond-like part of the cornea are also extracted from the retardation measures and quantified.



(a)



(b)

Figure 8: (a) Distrubtion of shear stress acting along the horizontal and vertical planes on cornea and (b) their quantification along the major axis of the corneal diamond as marked by the arrow in (a).

4. CONCLUSION

The stress distribution patterns on a healthy cornea of human eye are reported here using PSAT. The method senses the retardation of the light components between the major and minor optical axes at any point of interests on the surface of cornea. This information is digitally stored and a subsequently stress analysis is performed in detail. The bi-polar image enhancement clearly reveals a diamond-like stress distribution pattern on the surface of cornea. Previous structural studies of cornea using X-ray scans in vitro have shown such diamond-like structures. Based on the features of stress distribution, the current non-X ray based sensing provided an alternative way of bio-marking the cornea using ordinary white light source.

ACKNOWLEDGEMENT

The author acknowledges Mr G Calvert for his suggestions and support to this work.

REFERENCES

[1] P. Suetens, Fundamentals of medical imaging, Cambridge press, London, 2009.

[2] M. Millodot, Dictionary of optometry and visual science, Butterworth-Heinemann, London, 2009.

[3] S. J. Antony, "Imaging shear stress distribution and evaluating the stress concentration factor of the human eye", Scientific Reports,Nature Publishing Group, vol. 5, 8899, 2015,DOI:10.1038/srep08899

[4] J. Last, S. Thomasy, C. Croasdale, P. Russell and C. Murphy, "Compliance profile of the human cornea as measured by atomic force microscopy", Micron, vol. 43, 2012, pp.1293–1298

[5] K.M. Meek and N.J. Fullwood, "Corneal and scleral collagens-a microscopist's perspective", Micron 32, 2001, 261–272

[6] J. Jorge, J. Gonza´lez-Me´ijome, A. Queiro´s, P. Fernandes and M. Parafita, "Correlations between corneal biomechanical properties measured with the ocular response analyzer and ICare rebound tonometry", Jl. Glaucom., vol. 17, 2008, pp. 442-448.

[7] D. VanNasdale, A. Elsner, T. Hobbs and S. Burns, "Foveal phase retardation changes associated with normal aging", Vis. Res., vol. 51,2011,2263–2272

[8] E. Go¨tzinger et al, "Imaging of birefringent properties of keratoconus corneas by polarization-sensitive optical Coherence Tomography", Invest. Ophthalmol. Vis. Sci., vol. 48, 2007, 3551-3558.

[9] Boote, C., Dennis, S., Huang, Y., Quantock, A., and Meek, K.M. Lamellar orientation in human cornea in relation to mechanical properties. Jl. Struc. Biol., vol. 149, 2005, pp. 1–6.

[10] S.P. Timoshenko and J.N. Goodier, Theory of elasticity, McGraw-Hill, Singapore,1982.

[11] W. T. Rhodes, "Bipolar point spread function synthesis by phaseswitching," Appl. Opt.vol. 16, 1977, pp 265–267.

[12] B. Therese and S. Sundravadevelu, "Bipolar ioncoherent image processing for edge detection of medical images", Int.Jl. Recent Trends in Eng., Vol.2, 2009, pp. 229-232.

EFFECTIVE BANDWIDTH ANALYSIS OF MIMO BASED MOBILE CLOUD COMPUTING

[1]Suremya Varghese and [2] Ganesan Subramanian

[1] PG Student, School of Engineering & IT, Manipal University, Dubai
[2]Assistant Professor, School of Engineering & IT, Manipal University, Dubai

ABSTRACT

Digital Disruption is all around us. Mobile is overtaking desktop, Social Media is beating search, Messaging Application are challenging e-mails and everything around us is becoming connected. Mobile devices especially the smart phones are fueling the culture of "Anytime, Anywhere, And Anything''. Smartphone is not only ubiquitous but also the primary computing device for many .These paradigm shifts are fueled by the explosive growth of smart phones which has touched a volume of 1.6 billion units globally. Smartphone growth has also triggered the explosive growth of mobile applications and cloud computing .Together, Mobile cloud computing is now a potential technology for mobile services .MCC overcomes obstacles related to battery life, storage capacity and low bandwidth. Current smart phones uses 2x2 MIMO which gives a speed 300Mbps, by using massive MIMO technology speed can be enhanced up to 1Gbps. This paper gives a BER (Bit Error Ratio) analysis to prove that by increasing number of transmitting and receiving antennas the performance can be enhanced.

KEYWORDS

BER, Cloud computing, mobile, MCC, MIMO

1. INTRODUCTION

Mobile devices and apps are becoming an essential part of human life. Mobile devices are most convenient way of communication which are not bound by time and place. Mobile users enjoy various services from mobile applications, which run on the device or/and on remote servers via wireless networks. However, mobility is facing many problems such as resource scarceness, finite energy and low connectivity. Mobile cloud computing can address these problems by executing mobile applications on resource providers external to the mobile device, on the cloud.

1.1. Cloud Computing

"Cloud computing refers to both the applications delivered as services over the internet and the hardware and systems software in the data centres that provide those services"[3].In the simplest terms, services and solutions that are delivered and consumed in real time over internet are cloud services. For example when you store your photos online or using a social media site or e-mails, you are using a "cloud computing service". Just as a "cloud" in the sky is diffuse and capable of hiding things, a "cloud network" is a diffuse network of computers connected in a hidden fashion [2]. Cloud computing allows users to use infrastructure (server, network and storage), platforms (middleware services and OS) and software's (application programs).

DOI: 10.5121/ijci.2016.5202

Clouds can be Public clouds—providers shares resources over internet to public. Private clouds – dedicated and secured clouds or Hybrid clouds –and integration of public and private clouds.

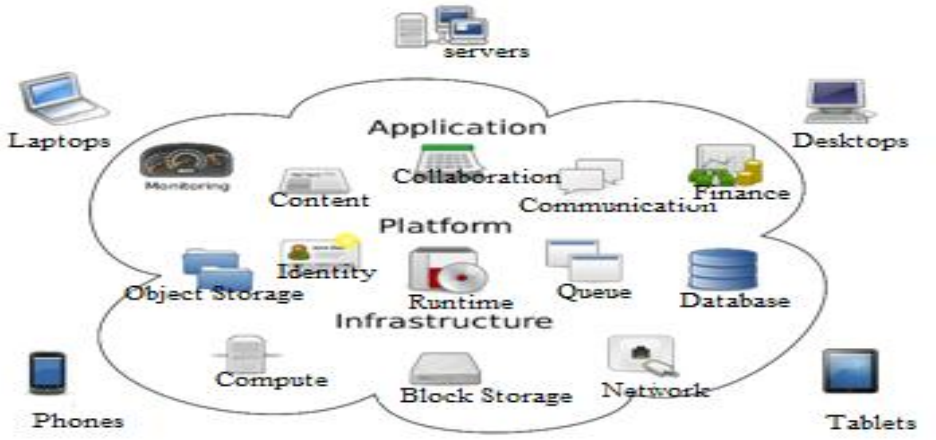


Figure 1 Cloud Architecture.

1.2. Mobile Cloud Computing

Themobile cloud computing is a development of mobile computing, and an extension to cloud computing. In mobile cloud computing, mobile device based intensive computing, data storage, mass information processing have been transferred to cloud, which are then accessed over the wireless connection based on a thin client. This reduces the computing capability and storage requirement of mobile devices. This brings mobile cloud computing not just smart phone users but also to simple mobile phone users. Mobile clouds are changing the work culture from brick and mortar office to just an internet connected phone.

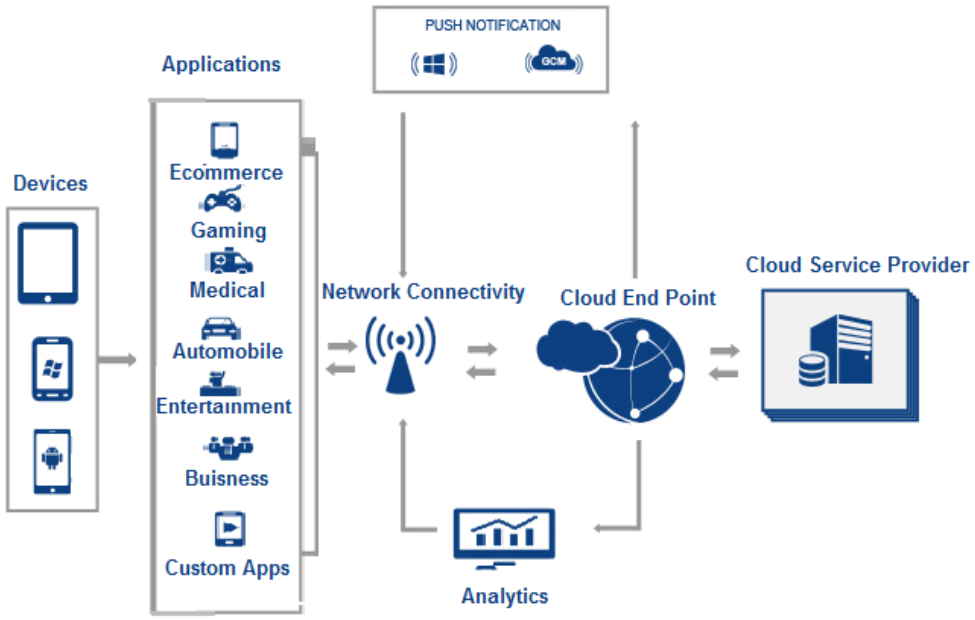


Figure 2---MCC architecture.

1.2.1. Advantages:

1. Improved battery life – As the dependency on mobile devices are increasing the battery life become critical. By offloading the power consuming computations to the cloud, computation time and power can be saved. For example, offloading a compiler optimization for image processing can reduce 41% for energy consumption of a mobile device [1].
2. High data storage capacity – Users can store and access large data on cloud using wireless which overcomes the limited storage of mobile devices.
3. Increases reliability and availability – By storing data and applications on cloud reduces the chances of data lost from device. And with an internet connection users can access it from anywhere.
4. Improved processing power – By sharing computational work with resourceful cloud improves processing time and power.

1.2.2. ISSUES

The mobile cloud computing services is still in early stages of development and facing several issues and challenges like,

1. Security and privacy--Protecting user privacy and data/application is of concern. Mobile devices are exposed to numerous security threats like malicious codes, hacking and viruses. Some GPS services gives out information about user's location [20].
2. Computing issues— Even though computational offloading is one the main features of MCC it faces some issues like selection of data/application to be offloaded and

dynamically updating of that files. Sometimes it may not save energy and time as expected [19].

3. Network connectivity-- since MCC is based on internet, the quality of network connectivity is very important. But it faces issues like longer latency of wireless connections, frequent disconnections, traffic congestion, weather conditions and mobility of users.

Due to limited radio sources the bandwidth is low and the high speed internet is not available everywhere. There should be a seamless connection handover between different access schemes (3G, 4G, GPRS, and WLAN Etc.).

4. Mobile application issues: Mobiles devices uses different operating system the cloud should be able to handle data across multiple O.S. To get the advantage of cloud computing data should be distributed properly.

1.3. MIMO (MULTIPLE INPUT MULTIPLE OUTPUT)

Compared to normal wired channels, the wireless communication is suffered by the multipath radio channel between the transmitter and receiver and the scarcity of available spectrum. Antenna diversity can take advantage of multiple signals to improve the transmission. The various possible configurations are shown in Figure 1.3, are referred as Single Input Single Output (SISO), Single Input Multiple Output (SIMO), Multiple Input Single Output (MISO) and Multiple Input Multiple Output(MIMO). SIMO is a form of receive diversity and MISO uses transmit diversity. MIMO combines transmit and receive diversity.

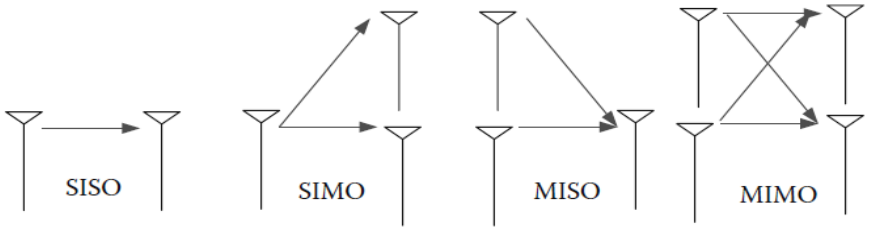


Figure 3—Multiple Antenna Configurations.

Each antenna element on a MIMO system operates on the same frequency and therefore does not require extra bandwidth. A MIMO system consumes no extra power due to its multiple antennas as the total power through all antenna is less than or equal to that of a single antenna system [4].

In current 3g and 4g communications traditional MIMO (2x2) is used. When we increases the number of antennas the capacity of the system increases linearly. But after 8 antennas, due to size and position issues of antennas the capacity do not increase. Latest studies show that if number of antennas is increased in the range of 100 the capacity improved well –known as massive MIMO [5].

2. RELATED WORK

Advances in computing and communication has made Mobile cloud computing an attractive area of research over the past few years. There have been a considerable number of researches to provide MCC with better coverage and quality, power and bandwidth efficient in all diverse environment.

Article [10] describes how computers configured to use in single location changed to portable devices with mobile computing and wireless connection. It also highlights the challenges of mobile computing like low bandwidth, address migration, low power and risk to data.

Article [8] explains the need for mobile cloud computing, advantages of MCC, current approaches, its issues and suggested some improvement methods. Some researchers [11],[7] ,[14] describes the architecture, applications and challenges of mobile cloud computing.

Yi Xu And Shiwen Mao reviewed mobile cloud computing, with focus on the technical challenges of MCC for multimedia applications [9].

The article [5] presents an overview of potential network architecture and some of the potential technologies to employed in future 5^{th} Generation Systems which includes
Non orthogonal multiple access (NOMA), Full duplex, Device to Device communication, Cognitive radio, millimeter Wave communication and Massive MIMO.

With the help of a simple 2 branch transmit diversity scheme, article [6] explains antenna diversity is most practical and effective technique to reduce multipath fading.

In [4], the author provided an overview of MIMO systems and MIMO channel modelling techniques. With the help of MATLAB simulations showed that equal diversity gain does not imply equal performance, but limited their analysis till 4X4 MIMO.

Some researchers analysed performance of 4X4 MIMO- OFDM hybrid technology using different modulation schemes like QPSK and 16 QAM [16], [17]. In [18], the performance of 4X4 MIMO is analysed practically in laboratory.

3. RESEARCH OBJECTIVES

This objective of this research is to identify and present recommendation for improving transmission efficiency in Mobile Cloud Computing. This is done analysing the BER performance mathematically and simulating in MATLAB SIMULINK. This research is done in three phases

Phase 1: Analysis the BER performance of different modulation techniques
Phase 2: Analysis the BER performance of different fading channels.
Phase 3: Analysis the BER performance multiple antenna configurations.

4. ANALYSIS

BER (Bit Error Ratio) testing is a powerful methodology for testing digital transmission system. In simple terms BER is the ratio of number of total errors tothe total number of transferred bits during the time interval. BER is a unit less performance measure, often expressed as a percentage. BER depend on transmitter, receiver and the medium between them.

4.1. BER ANALYSIS OF DIFFERENT MODULATION TECHNIQUES

The most fundamental digital modulation techniques are:

- PSK (phase-shift keying)
- FSK (frequency-shift keying)
- ASK (amplitude-shift keying)
- QAM (quadrature amplitude modulation).

Simulation have been done using MATLAB, for BPSK, BDPSK, BFSK and 16 QAM. For the purpose of simulation different SNR (Eb/No) is introduced (-5db to 15db) in the AWGN channel. Figure 4 shows the performance of different modulation in AWGN CHANNEL. In the graph as Eb/No increases the BER decreases for all modulation.

Graph and the table values shows that BPSK has better performance. QPSK performance is very similar to BPSK. All carrier signals other than BPSK and QPSK have some level of correlation between their signals which make it to harder to identify the signal at receiver end. This makes BPSK and QPSK to have lowest bit error.

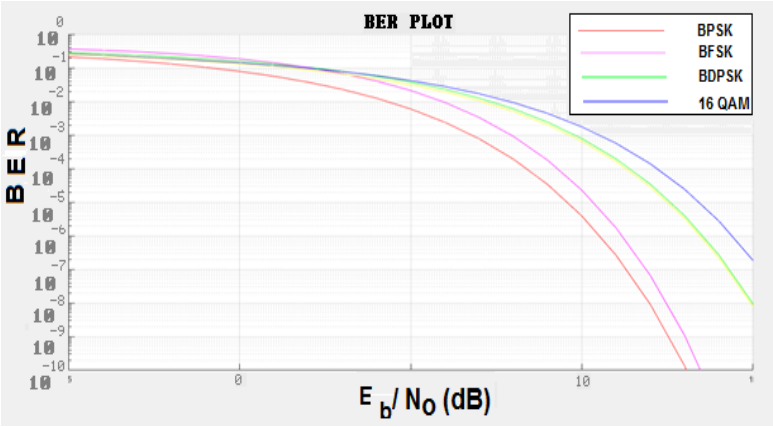


Figure 4. BER Analysis for Different modulation

Table 1. Theoretical BER Calculation

MODULATION TECHNIQUE	FORMULE	BER WHEN $(E_b\|N_o) = 5$
BPSK	$P_b = Q\left(\sqrt{\frac{2E_b}{N_o}}\right)$	0.0008
BFSK	$P_b = Q\left(\sqrt{\frac{E_b}{N_o}}\right)$	0.0129
BDPSK	$P_b = \frac{1}{2}exp\left(-\frac{E_b}{N_o}\right)$	0.0036
16 QAM	$P_b \cong \frac{3}{4}Q\left(\sqrt{\frac{4E_b}{5N_o}}\right)$	0.0171

4.1.BER ANALYSIS OF DIFFERENT CHANNELS.

The above simulation is done considering only AWGN channel but signals undergo fading due to multipath propagation, attenuation and scattering. Fading severally affects the performance of the system. It can be clearly understand from the graph, the BER is increased considerably in Rician and Rayleigh fading channels than AWGN channel. A Simulink model is used to estimate BPSK BER performance in AWGN, Rayleigh and Rician channel is shown in the graph with Doppler shift = 0.01Hz jakes model.

Table 2. Theoretical BER Calculation

CHANNEL	FORMULE	BER WHEN$(E_b\|N_o) = 5$
AWGN	$P_b = Q\left(\sqrt{\frac{2E_b}{N_o}}\right)$	0.0008
RAYLEIGH	$P_b = \frac{1}{2}\left[1 - \sqrt{\frac{\overline{\gamma_b}}{1+\overline{\gamma_b}}}\right]$	0.0436
RICIAN (k= 2)	$P_b = \frac{1}{2}erfc\left[\sqrt{\frac{k\overline{\gamma_b}}{k+\overline{\gamma_b}}}\right]$	0.0455

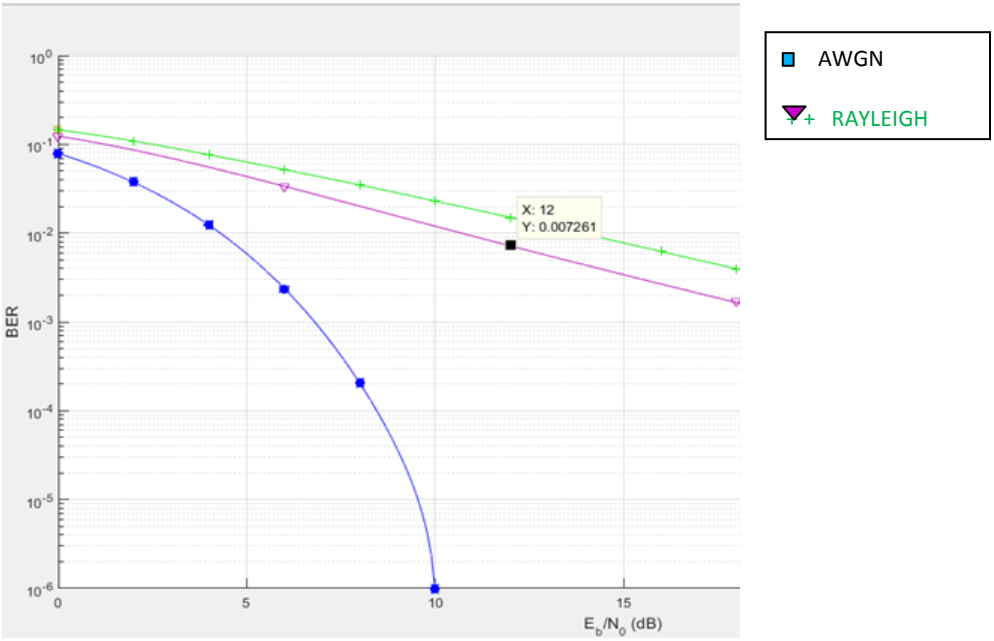


Figure 5 BER Analysis for fading channel

4.3 MULTIPLE ANTENNA CONFIGURATIONS

To order improve the data rate and reliability of communication over multipath fading channel, spatial diversity is used. Data is encoded by channel coding method Space Time Block Codes (STBC). The encoded data is split into N_t streams that are simultaneously transmitted by Nt antenna. The received signal at each receiver linear super position of the N_t transmitted signal plus noise. Performance is determined by the matrices constructed from pairs of distinct set code sequences. The channel matrix can be written as

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

The received signal at antenna j,

$$r_t^j = \sum_{i=1}^{N_t} h_{j,i}\, x_t^i + n_t^j$$

The received signal vector

$$r_t = \left(r_t^1, r_t^2, \ldots, r_t^{N_r}\right)$$
$$r_t = H_t x_t + n_t$$

MIMO system with N_t transmit antenna and N_r receiving antenna has diversity gain equal to N_tN_r. Simulations are done with BPSK modulation in Rayleigh channel.

4.3.1 SIMO- SINGLE INPUT MULTIPLE OUTPUT

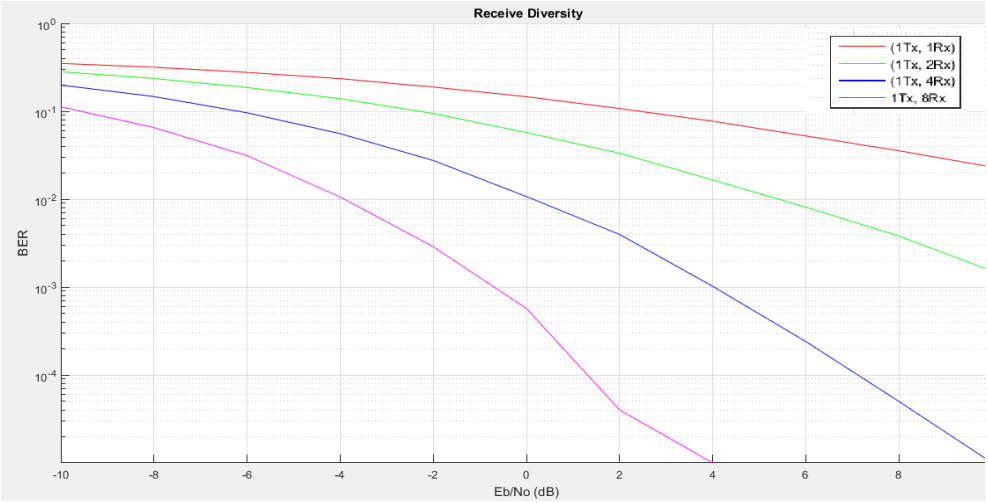


Figure 6.Receiver Diversity

Table 3. BER comparison of SIMO

Eb/No	1 X 1	1 X 2	1 X 4	1 X 8
0	0.1483	0.0578	0.03097	0.00032
2	0.1082	0.0332	0.0033	06E-05
5	0.02415	0.00152	1E-05	0

4.3.2 MISO- MULTIPLE INPUT SINGLE OUTPUT

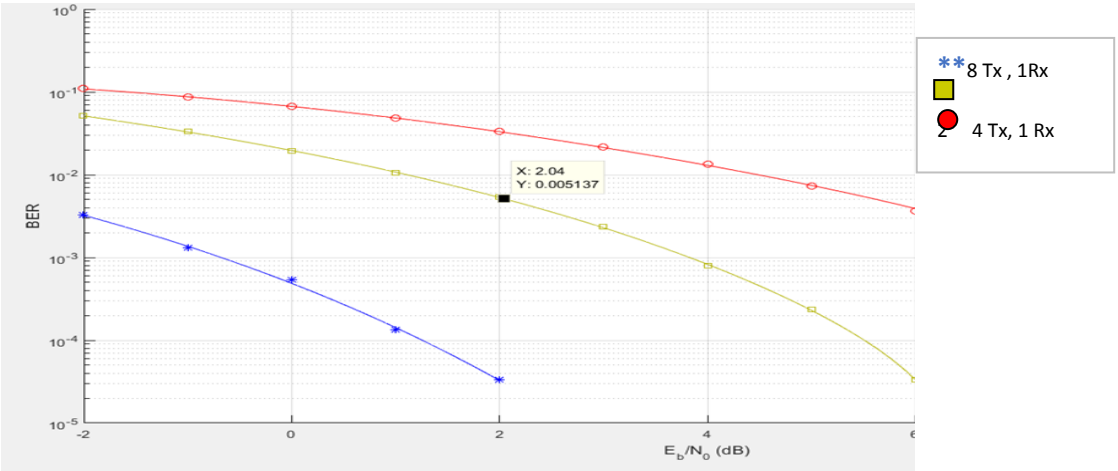


Figure 7. Transmitter Diversity

Table 4. BER comparison of MISO

Eb/No	2 X 1	4 X 1	8 X 1
0	0.0671	0.033	0.0032
2	0.0310	0.00513	5.3E-4
5	0.0042	9.2E-05	1.3E-05

4.3.3 MIMO- MULTIPLE INPUT MULTIPLE OUTPUT

Table 5. BER comparison of MIMO

Eb/No	2 X 2	4 X 4	8 X 8
0	0.0380	0.002	0.001
2	0.0152	4.33E-04	3.3E-05
3	0.0082	1.33E-04	0

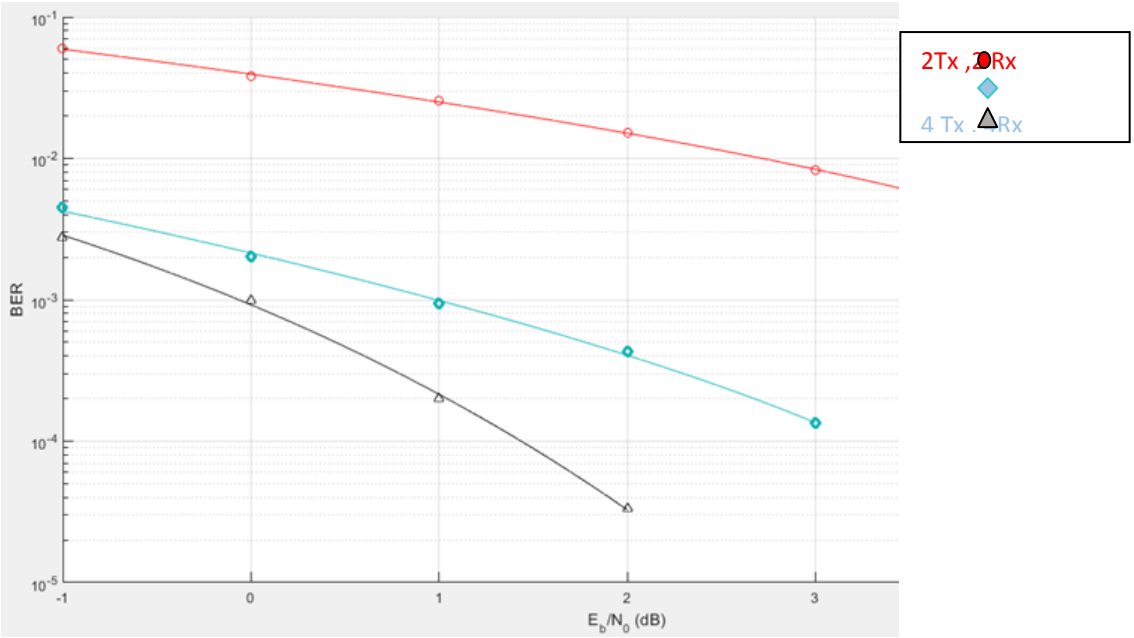


Figure 8. BER Analysis of MIMO

5. CONCLUSION

This paper provided basic overview of mobile cloud computing and its need and challenges. We briefly discussed Space Time Block Coding with different antenna diversity. Its proved that as number of antenna increases the bit error rate also decreases. Eventhough diversity is same for SIMO 1X4 , MISO 4X1 and MIMO 2X2 the BER is different. MIMO system will have higher capacity. Having more number of transmitter end is more feasble than at the receiver end as its easy to add atennas at base station depending upon requirment rather adding at each handset.We can propose MIMO 8X8 is better pereference for 5G integrated MCC.

REFERENCES

[1] S.C Hsueh, J.Y Lin, M.Y Lin, "Secure cloud storage for conventional data archive of smart phones," in: Proc. 15th IEEE Int. Symposium on Consumer Electronics, ISCE '11, Singapore, June 2011.

[2] B.Gabriel, "NASA Turns to Online Giant Amzone for Cloud Computing Services for Mars Rover Curiosity," 12 August 2012 . [Online]

[3] Miss Priyanka J. Pursani, Prof. P.L Ramteke, "Mobile Cloud Computing," : International Journal of Advanced Research in Computer Engineering And Technology(IJARCET), Volume 2, Issue 4, April 2013.

[4] Luis Miguel Cortes-Pena, "MIMO Space time block coding (STBC): Simulation and Results", Design Project, Personal and Mobile Communications, Georgia Tech, April-2009.

[5] M.A Zheng, ZHENG Zheng Quan, DING ZhiGuo, FAN PingZhi and LI HengCHao," Key Techniques for 5G wireless communications, Network Architecture , physical layer and MAC layer Perspective" , Science China , Vol 58, April 2015.

[6] S. M. Alamouti, , "A simple transmitter diversity scheme for wireless communication" , IEEE Journal of Select. Areas Communications. , vol. 16 ,Number 8, pp.1451 -1458 , 1998.

[7] Hoang T. Dinh, Chonho Lee, Dusit Niyato*, Ping Wang, A Survey Of Mobile Cloud Computing: Architecture, Applications, And Approaches, Wireless Communications and Mobile Computing, Volume 13, Issue 18, December 2013, pp:1587-1611.

[8] Niroshinie Fernando' ', Seng W. Loke', Wenny Rahayu , Mobile cloud computing: A survey,Future Generation Computer Systems,Volume 29, Issue 1, January 2013, Pages 84–106.

[9] Yi Xu , Shiwen Mao, A Survey Of Mobile Cloud Computing For Rich Media Applications, IEEE Wireless Communications, June 2013, pp 46-53.

[10] George H. Forman and John Zahorjan, "The challenges of Mobile Computing", University of Washington. April 1994.

[11] Han Qi and Abdullah Gani, "Research on mobile cloud computing: Review, Trend and Perspectives", University of Malaya, Malaysia.

[12] J. V. Vishniakova , A. I. Luchaninov "application of antenna theory with nonlinear elements for mimo analysis" International Conference on Antenna Theory and Techniques, 2013, Odessa, Ukraine

[13] Swati Chowdhuri, Sayan Chakraborty, Nilanjan Dey, Ahmad Taher Azar,. Mohammed Abdel-Megeed M. Salem,, Sheli Sinha Chaudhury, Pranab Banerjee, Recent Research on Multi Input Multi Output (MIMO) based Mobile ad hoc Network: A Review, International Journal of Service Science, Management, Engineering, and Tech., 5(3), 54-65, July-Sept. 2014.

[14] Dipayan Dev and Krishna Lal Baishnab , "A Review and Research towards Mobile Cloud Computing" , 2014 2nd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering,2014.

[15] Bernard Sklar, Digital Communications: Fundamentals andApplications - Prentice Hall Communications Engineering and Emerging Technology, Jan 2001.

[16] Deepak Sharma , Praveen Srivastava, OFDM Simulator Using MATLAB , International Journal of Emerging Technology and Advanced Engineering, Volume 3, Issue 9, September 2013 .

[17] Vibha Rao , T. Malavika, Performance analysis of MIMO-OFDM for multiple Antennas,International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 3, Issue 5, May 2014

[18] Oomke Weikert and Udo Zolzer, A flexible laboratory MIMO system using four transmit four receive antennas, Proeedings of the 10th International OFDM-Workshop, August 2005, pp. 298-302 .

[19] Macro v.Barbera, Sokol Kosta , Alessandro Mei and Julinda Stefa, "To offload or Not to Offload? The Bandwidth and Energy Costs of Mobile Cloud Computing", Sapienza University of Rome, Italy.

[20] Abdul Nazir Khan, M.L Mat Kiah, Samee U. Khan and Sajjad A.Madani, "Towards secure mobile cloud computing: A survey", Elsevier Future Generation Computer Systems 29(2013) 1278-1299.

WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput.
2013; 13:1587–161

AUTHORS

Suremya Varghese

She has obtained her Bachelors in Electronics and communication engineering in 1994 from Mahatma Gandhi University and currently doing her Masters in Digital electronics and advanced communication at, Manipal University Dubai Campus. She got 1 year teaching and 5 years industrial experience.

Ganesan Subramanian:

Prof. GANESAN .S obtained his Bachelors in Electronics and communication Engineering in 2000 from Madurai Kamaraj University, and Master of Engineering in Digital Communication & Networking Engineering in 2004 from Anna University, Chennai. He has more than 15 years of teaching experience and published 20 research papers in International Journals. Currently he is working as Assistant Professor in the Engineering Department, Manipal University Dubai Campus. Area of Interest in Teaching: Cloud computing, Mobile Cloud, Information Theory and Coding Techniques, Wireless Communication, Wireless Sensor Networks

STANDARDISATION AND CLASSIFICATION OF ALERTS GENERATED BY INTRUSION DETECTION SYSTEMS

Athira A B[1] and Vinod Pathari[2]

[1]Department of Computer Engineering ,National Institute Of Technology Calicut, India
[2]Department of Computer Engineering ,National Institute Of Technology Calicut, India

ABSTRACT

Intrusion detection systems are most popular de-fence mechanisms used to provide security to IT infrastructures. Organisation need best performance, so it uses multiple IDSs from different vendors. Different vendors are using different formats and protocols. Difficulty imposed by this is the generation of several false alarms. Major part of this work concentrates on the collection of alerts from different intrusion detection systems to represent them in IDMEF(Intrusion Detection Message Exchange Format) format. Alerts were collected from intrusion detection systems like snort, ossec, suricata etc. Later classification is attempted using machine learning technique, which helps to mitigate generation of false positives.

KEYWORDS

Intrusion Detection Systems, IDMEF, Snort, Suricata, ossec& WEKA

1. INTRODUCTION

Due to the widespread use of Internet, providing security against attacks on network is a challenging job today. Most of the organisations use intrusion detection systems (IDS) for providing security. Need for IDS can be summed up as simple principle of security: Defence in Depth. It is a layered approach involving multiple overlapping controls in preventing, detecting and responding to suspected intrusions.

1.1. INTRUSION DETECTION SYSTEM

Intrusion detection systems are most popular defence mechanisms used to provide security to IT infrastructures. Intrusion is a sequence of related actions performed by a suspicious adversary, which result in the form of compromise of a target system [7]. These kinds of actions violates certain security policy of the system. The process of identifying and responding to suspicious activities of target system is called Intrusion Detection [7].

DOI: 10.5121/ijci.2016.5203

1.2. MOTIVATION

Organisations frequently use several IDSs from different vendors since each has its relative strengths. One may be strong at host-based intrusion detection while another may be strong at network based intrusion detection. Organisations need best performance, do not prefer to take a chance with security and hence use multiple IDSs from different vendors. Different IDSs will be using different protocols and generate alert events in different formats. If we fail to integrate the outputs from all these properly, the volume of data generated will be high and accordingly more false positives occur. Large volume of IDS false alarms is unacceptable to security administrators as it hinders smooth functioning of any organization. To reduce the cost of operation and increase the reliability of a security system, it is required to tackle the excess of false alarms.

2. PROBLEM STATEMENT

To develop an approach to collect alerts from different sensors and standardize them into IDMEF. Later these alerts will be classified into false alarms and attacks attempted using machine learning technique.

3. RELATED WORKS

KleberStoreh et al. [9] proposed an approach for correlat-ing security events using machine learning technique. Layered approach is followed here. Apart from normal methods they analyse alerts generated from different sensors, which are normalised, fused into meta-alerts and are then used for classification into alerts or false alarms. ChampaDey [7] proposed a similar approach for reducing false alarms using incremental clustering algorithm. Only data from snort IDS is used for analysing purpose. The alert data is then processed using incremental clustering algorithm and classified into alerts or false alarms.

4. PROPOSED METHOD

In the proposed system, format difference in alert from different sensors is overcome by representing them into IDMEF (Intrusion Detection Message Exchange Format) format. Later classification of parsed IDMEF alerts into false alarms and attacks is achieved using machine learning technique. In this work, we collect alerts from different intrusion detection systems and proceed as follows:

- Convert collected alerts into a common format (ID-MEF is identified as common format).
- Labelling of alerts.
- Classification of alerts into false alarm or attack using machine learning technique.

Detailed work flow for the proposed system is shown in Figure 1.

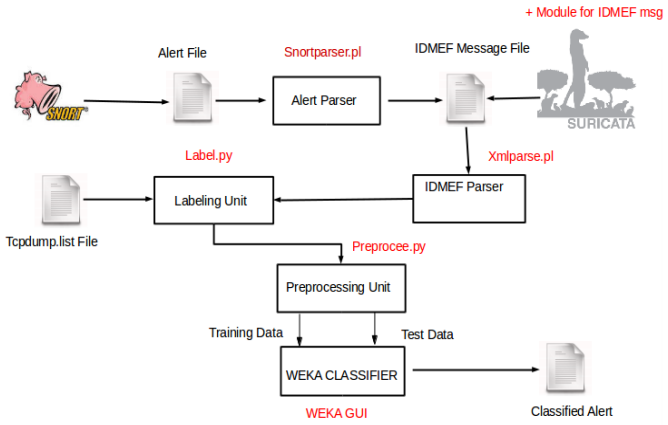


Figure 1. Detailed Work Flow

5. STANDARDISATION OF ALERTS

5.1. IDMEF

IDMEF(Intrusion Detection Message Exchange Format) is an object oriented representation of alert data generated by intrusion detection systems. The goal of IDMEF is a standard representation of alert data in an unambiguous manner. IDMEF data model can be summarised as Figure 3.1 [11] .Two types of implementation for IDMEF was proposed by Intrusion Detection Working Group (IDWG) [11]. One method is using Structure of Management Information (SMI) [11] and the other is using XML. During second phase of our work, we need to process the IDMEF messages. Software tools for processing XML documents are widely available, in both commercial and open source forms [11]. Hence we chose to implement IDMEF in XML format.

5.2. IDMEF GENERATION

DARPA (Defence Advanced Research Project Agency) [1] data sets are used for testing. DARPA simulate American air force based local network being attacked in different ways. Attack information are provided in the form of log files. DARPA data set is replayed using different IDSs.We considered Snort and Suricata IDSs. Alerts were gathered from them and IDMEF messages were generated. IDMEF message generation details are explained in the following sections.

5.2.1. Snort

Snort is a widely used open source signature based network intrusion detection system, configured to operate on Network IDS mode. In Network IDS mode, snort will perform actual analysis to determine malicious traffic and alerts are generated. To conduct testing DARPA 1998 data sets were downloaded from MIT Lincoln Labs website [1]. This dataset contains simulated network traffic embedded with marked attacks. snort was configured in network intrusion detection system to use this data set. We wrote a perl script to attain the task of standardisation phase in work flow diagram. Alert file serves as input to this program. Required alert attributes

are obtained through parsing and IDMEF message is obtained with the help of XML::IDMEF library. The IDMEF messages obtained from snort alert file is shown in Figure 2.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<idmef:IDMEF_Message xmlns:idmef="http://iana.org/idmef version =1.0">
 <idmef:Alert messageid="1">
  <idmef:Analyzer analyzerid="527">
   <idmef:Node category="8" />
  </idmef:Analyzer>
 </idmef:Alert>
 <idmef:CreateTime ntpstamp="06/16-11:22:28.515219" />
 <idmef:Source>
  <idmef:Node>
   <idmef:Address category="ipv4-addr">
    <idmef:address>0.0.0.0</idmef:address>
   </idmef:Address>
  </idmef:Node>
  <idmef:Service>
   <idmef:priority> 2</idmef:priority>
   <idmef:protocol>IGMP</idmef:protocol>
  </idmef:Service>
 </idmef:Source>
 <idmef:Target>
  <idmef:Node>
   <idmef:Address category="ipv4-addr">
    <idmef:address>224.0.0.22</idmef:address>
   </idmef:Address>
  </idmef:Node>
  <idmef:Service>
   <idmef:protocol>IGMP</idmef:protocol>
  </idmef:Service>
 </idmef:Target>
 <idmef:Classification text=" Potentially Bad Traffic" />
</idmef:IDMEF_Message>
```

Figure 2. IDMEF Message Generated by Snort

5.2.2. Suricata

Suricata, a rule-based IDS, take advantage of the externally developed rule sets to monitor sniffed network traffic and provide alerts when suspicious events take place. Suricata uses the Yaml format for configuration. suricata.yaml file included in source code is the example configuration file of Suricata. After packet analysis Suricata generates alert outputs. Output section in suricata.yaml controls the output structure for alerts generated. Default log directory is /var/log/suricata. There are several types of output structures like fast.log, http.log, stats.log etc. To generate IDMEF messages an output structure as mentioned above was developed. For this we have developed a C program, which will write data into buffer in IDMEF format. Program files were appended to source code. Re-installation of Suricata was performed. Suricata was configured to use DARPA data set. Alerts were generated from the suricata. IDMEF messages are generated at default directory /var/log/suricata/fast.log as shown in Figure 3.

```
<?xml version="1.0" encoding="UTF-8"?>
<idmef:IDMEF-Message version="1.0"xmlns:idmef= http://iana.org/idmef/>
<idmef:Alert messageid= 2200075 >
<idmef:Analyzer analyzerid=1>
</idmef:Analyzer>
<idmef:CreateTimentpstamp=03/12/2015-15:36:31.734831 >
<idmef:Source>
<idmef:Node>
<idmef:Addresscategory=ipv4-addr>
<idmef:address>192.168.4.71</idmef:address>
</idmef:Address>
</idmef:Node>
<idmef:Service>
<idmef:priority>3</idmef:priority>
<idmef:protocol>UDP</idmef:protocol><idmef:port>65419</idmef:port>
</idmef:service>
</idmef:source>
<idmef:Target
<idmef:Node>
<idmef:Addresscategory=ipv4-addr>
<idmef:address>192.168.254.2</idmef:address>
</idmef:Address>
</idmef:Node>
</idmef:Target>
<idmef:Service>
<idmef:protocol>UDP</idmef:protocol>
<idmef:port>53</idmef:port>
</idmef:service>
</idmef:Target>
<idmef:Classificationtext=(null)>
</idmef:Classification>
</idmef:Alert>
</idmef:IDMEF-Message>
```

Figure 3. IDMEF Message Generated by Suricata

6. FALSE POSITIVE REDUCTION

6.1. ALERT CLASSIFICATION

As we discussed earlier the main objective of intrusion detection system is to distinguish between attacks and normal events. Most of intrusion detection systems face a common problem which is the generation of high false alarms. An IDS is efficient when it contains less number of false positives and false negatives. One way to tackle this problem is using machine learning technique. Machine learning techniques can be used to distinguish between attacks and false alarms.

a. MACHINE LEARNING

DARPA data set provide tcpdump.list files. For each online traffic, information about attacks in each connection will be included in tcpdump.list files. Connection is a sequence of TCP packets starting and ending at some well defined time interval. Between this connections data flow from one source IP address to target IP address under the control of a protocol. Input to labelling unit are two files, alertlog file and tcpdump.list file. tcpdump.list file contain information about start date, duration, service, source port, destination port, source IP, destination IP, attack score and

attack type. Attack score is a binary valued attribute. Presence of an attack is indicated by an attack score 1 while 0 indicates the absence of an attack. Attacks are mainly divided into five classes DOS, Probe, R2L, U2R, DATA . Algorithm for Labelling Alerts The algorithm is implemented in python. The labelled alert file is used for classification. Classification is attempted using machine learning algorithm. We use WEKA tool for this approach.

Input: Tcpdump.list File, Alertlog (parsed IDMEF file) File
Output: Labelled Alerts

1. For each row in tcpdump.list files
   If row is a labelled attack then add the row to the new file AttackList
2. For each row in alertlog file
   Create key with three attributes timestamp, srcip, destip
   IF
   The key exists in the AttackList file, Identify the attack class for the type of attack found. Label the selected row with the type of attack class.
   Else
   Label the selected row as normal
3. Return the AlertList file

Algorithm 1: Algorithm for Labelling Alerts

6.3. WEKA

Weka(Waikato Environment for Knowledge Analysis) [5] is a free and open source tool used for data mining tasks. Weka has many applications like Explorer, Experimenter, Knowledge Flow and Simple CLI. We attempt classification using Weka Knowledge Explorer.

6.4. WEKA EXPLORER

The classifier panel in Weka Explorer allows us to configure and execute any weka classifier on the current data set. We take data set with known output values and use this to build a data model. Whenever we have new data points with unknown output values, we put it through model and produce our expected output. This model requires one extra step, shown as pre-processing unit in Detailed work flow diagram in Figure 1. Entire training set will be taken and divided into two parts. We will take about 60-80 % data and put into our training set, which will be used to create the data model. Then take the remaining data and use it as test set, which will be used for testing the accuracy of our model after creating it. A Naive Bayesian Learner (bayes.Naive Bayes) algorithm will be used for classification.

6.5. RESULTS AFTER CLASSIFICATION

One way to evaluate IDS is by its prediction ability to give a correct classification of events to be attacks or normal behaviour. According to real nature of an event the prediction from an IDS has four possible outcome which is called confusion matrix.

Table1. Confusion Matrix

	Normal	**Attack**
Normal	TN	FP
Attack	FN	TP

- TN :-True Negatives are actually normal events and successfully labelled normal.
- TP:- True Positives are attack events and successfully labelled as attacks.
- FP:- False Positives are normal events being classified as attacks.
- FN:- False Negatives include attack events incorrectly classified as normal events.

True negatives and True positives corresponds to the correct operation of the IDS.

False Positive Rate(F P R) = FP/FP+TN (1)

Also known as false alarm rate. Rate at which normal data will be falsely detected as attacks. High FPR will degrade the performance of IDS.

False Negative Rate(F NR) = FN/TP+FN (2)

If FNR is high system is vulnerable to attacks.

True Positive Rate(T P R) = TP/TP + FN (3)

True Negative Rate(T N R) = TN/TN + FP (4)

Also known as detection rate or sensitivity. It is the ratio of detected attacks among all attack events.

Accuracy =TP + TN/TP + TN + FP + FN (5)

It is the ratio of events classified as accurate type in total events.

Precision = TP/TP+FP (6)

Figure 4 shows result after classification.

```
Time taken to build model: 0.04 seconds

=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances        1399               80.2179 %
Incorrectly Classified Instances       345               19.7821 %
Kappa statistic                          0.0484
Mean absolute error                      0.133
Root mean squared error                  0.2555
Relative absolute error                106.4934 %
Root relative squared error            102.2921 %
Total Number of Instances             1744

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0         0.005      0         0         0          0.294    dos
                 0.975     0.911      0.83      0.975     0.896      0.765    normal
                 0.031     0.019      0.143     0.031     0.051      0.245    probe
                 0         0.012      0         0         0          0.002    R2L
                 0         0          0         0         0          0.304    U2R
Weighted Avg.    0.802     0.749      0.694     0.802     0.74       0.675

=== Confusion Matrix ===

    a    b    c    d    e   <-- classified as
    0  130   16    2    0 |    a = dos
    4 1394   14   18    0 |    b = normal
    4  151    5    1    0 |    c = probe
    0    1    0    0    0 |    d = R2L
    0    4    0    0    0 |    e = U2R
```

Figure 4. Result After Classification

3. CONCLUSIONS

Organisations frequently use several IDS from different vendors since each have relative strengths and weaknesses. The use of diverse IDS solution leads to generation of too many false positives. If we fail to tackle the problem it will effect the performance of organisations. In the proposed system, format difference in alert from different IDSs are overcome by representing them into IDMEF format. Alert data can be handled efficiently by representing alerts into IDMEF message. Later classification of parsed IDMEF alerts into false alarms and attacks is achieved using machine learning technique. Parameters obtained by parsing IDMEF were not optimised in our approach. This will further improve the performance of alert classification.

ACKNOWLEDGEMENTS

We would like to show our gratitude to everyone for sharing their pearls of wisdom with us during the course of this research, and who provided insight and expertise that greatly assisted the research.

REFERENCES

[1] DARPAdataset, http://www.ll.mit.edu/mission/communications/cyber/CSTcorpora/ideval/data/. Accessed on 03-December-2014.

[2] Ossec, http://www.ossec.net//. Accessed on 03-December-2014.

[3] Snort, https://www.snort.org/. Accessed on 03-December-2014.

[4] Suricata, https://redmine.openinfosecfoundation.org/projects/suricata/ wiki/Suricatayaml. Accessed on 2-February-2015.

[5] HadiBahrbegi Mir Kamal Mirnia Mehdi BahrbegiElnazSafarzadeh Amir AzimiAlastiAhrabi, Ahmad HabibizadNavin and Ali Ebrahimi, "A New System for Clustering and Classification of Intrusion Detection System Alerts Using Self-Organizing Maps", International Journal of Computer Science and Security, 4, 2004.

[6] Neethu B, "Classification of Intrusion Detection Dataset using machine learning Approaches", International Journal of Electronics and Com-puter Science Engineering, 1956.

[7] ChampaDey, "Reducing ids false positives using Incremental Stream Clustering (isc) Algorithm", Dept of Computer and Systems Sci-ences, Royal Institute of Technology, Sweden, page March, JULY-SEPTEMBER 2009.

[8] Debar H and Wespi A, "Aggregration and Correlation of Intrusion-Detection Alerts", In Proceedings of the 4th International Symposium on Recent Advances in Intrusion detection (RAID), Springer Verlang, California, USA, pages 85–103, 2001.

[9] KleberStroeh, Edmundo Roberto Mauro Madeira, and Siome Klein Goldenstein, "An approach to the correlation of security events based on machine learning techniques", Journal of Internet Services and Applications, 2013.

[10] SebastiaanTesink, "Improving intrusion detection systems through machine learning", ILK Research Group,Technical Report Series no. 07-02, Tilburg University, page March, JULY-SEPTEMBER 2007.

[11] FredrikValeur, Giovanni Vigna, and Christopher Krue, "Modeling In-trusion Alerts using idmef", IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, 1(3), JULY-SEPTEMBER 2004.

Authors

Athira A B- She received the B.Tech. Degree in computer science and engineering from University of Calicut, Kerala, India, in 2012, and M.Tech.in computer science and engineering (Information Security) from the National Institute of Technology (NIT) Calicut, Kerala, India in 2015.

VinodPathari- He is working as a full time faculty in the Computer Science and Engineering Department of NIT Calicut, Kerala, India. In addition to information security related topics he is also interested in teaching functional programming and software engineering.

INTENTIONAL BLANK

A NOVEL APPROACH TO ERROR DETECTION AND CORRECTION OF C PROGRAMS USING MACHINE LEARNING AND DATA MINING

Prof. KhushaliDeulkar[1] , Jai Kapoor[2] , Priya Gaud[3] , Harshal Gala[4]
Department Of Computer Engineering D.J Sanghvi College Of Engineering ,Mumbai, India

ABSTRACT

There has always been a struggle for programmers to identify the errors while executing a program- be it syntactical or logical error. This struggle has led to a research in identification of syntactical and logical errors. This paper makes an attempt to survey those research works which can be used to identify errors as well as proposes a new model based on machine learning and data mining which can detect logical and syntactical errors by correcting them or providing suggestions. The proposed work is based on use of hashtags to identify each correct program uniquely and this in turn can be compared with the logically incorrect program in order to identify errors.

KEYWORDS-COMPONENTS

Machine Learning Device(MLD), Data Mining Device(DMD), Databases, Hash-tag.

1. INTRODUCTION

The conventional text-book based learning has gradually been replaced by a more convenient and affordable computer-assisted learning considering its availability and accessibility. [1]With the Introduction to Programming course being made compulsory in almost all academic institutions students with no background in coding struggle to get hold of the syntax and logics of the programs.

Correcting such errors can be a very tedious job when you can't get the meaning of the compile time error messages that you receive. The evaluation of such a huge amount of programs, given the vast amount of pupils the institution has to take, can be very tasking and takes a lot of time.

Similarly, at an organization which involves coding from its employee, sitting and correcting all the syntactical errors will waste a lot of energy and resources. To deal with such issues, we have proposed an automatic error detection and correction system for programs in C. There have been various studies and tools used previously to assist novice programmers to form an error free program and get a correct output.[7] Our system mainly deals with compile time errors i.e. syntactical errors since syntactical errors are significant for the novice programmers. This paper proposes a system which focuses on integration of machine learning, data mining and system programming to detect the errors in the program.

Data mining can be defined as process of analysing data from different and summarizing it into useful information. It is a process of finding correlation or patterns in the information available. [8]The five major steps that will constitute the data mining aspect of the system are: collecting data to mine, determining the table to assist, pre-processing the data, extracting relevant data from raw, data cleaning and formatting it, adapting a mining algorithm, and at last applying mining results.

The four major categories of mining algorithms are[9]:

1) Pattern matching: Finding the instance to data for given pattern.
2) Clustering: Assembling the data in clusters.
3) Classification: Classifying data on the basis of the already classified data.
4) Frequent pattern mining: Locating the frequently occurring pattern.

Now, the correct programs will be stored in the database and organized using data mining. Each of the correct programs will be assigned a hash-tag with help of system programming identification. Whenever an incorrect program with that particular logic will be entered, the system will make a hash-tag based comparison with correct program and using data mining it will detect those errors.

With the help of machine learning, the system will analyse the data, recognize the pattern, learn and then display the errors in the program and suggestions to correct those errors. This will allow the programmers to save a lot of energy and thus get the output faster.

In section II, we reanalyse the previous works on the methodologies and techniques used. In section III we discuss our proposed solution for the issue. In section IV, we have written advantages and disadvantages of the proposed solution. In section V we conclude our paper while describing the scope of the project and the future work possibilities.

2. RELATED WORK

Our research aims to design error detection and correction methods in C programs using data mining and machine learning. Several other researchers have previously worked on the similar domain.

K .K Sharma and Kunal Banerjee [1] have concentrated on the problem of the precedence of the if-else statements and the incorrect ordering of conditions leading to a logical error which standard compilers fail to determine. This is later resolved by tabulating the if-else statements using a set of systematic steps. Firstly, the precedence of the if-else conditions is identified. Secondly, after ordering according to the precedence the innermost conditions are executed and they are compared with the rule table. Comparison is done by converting these statements in normalized form. After normalization of each condition, we now check the ordering of conditions in an else-if construct. Thirdly, a complex analysis is done to compute the time complexity and make it more efficient. The time complexity is done taking two things in consideration: the complexity of comparing two normalized (conditional) expressions and the number of times such comparisons have to be done. This can be used in our project to tackle the if-else condition

problem and will reduce the possibility of logical errors related to the if-else conditions to occur.

YuriyBrun and Michael D. Ernst [2] propose a technique for identifying program properties that indicate errors. The technique generates machine learning models of program properties known to result from errors, and applies these models to program properties of user-written code in order to classify and rank properties that may lead the user to errors. Given a set of properties produced by the program analysis, the technique selects a subset of properties that are most likely to reveal an error. An implementation, the Fault Invariant Classifier, demonstrates the efficacy of the technique. The implementation uses dynamic invariant detection to generate program properties. It uses learning techniques like support vector machine and decision tree to classify these properties. It is done by technique in which it has two steps: training i.e. pre-processing step that extracts properties of programs containing known errors and classification i.e. the tool applies the model to properties of new code and selects the fault-revealing properties. This is followed by creation of models and detection of fault-revealing properties. This technique may be most useful when important errors in a program are hard to find. It is applicable even when a developer is already aware of (low-priority) errors.

Tatiana Vert, Tatiana Krikun and Mikhail Glukhikh [3] in their paper "Detection of Incorrect Pointer Dereferences for C/C++ Programs using Static Code Analysis and Logical Inference" have done static code analysis precision using classic code algorithm with dependencies. The key characteristics of error detection methods are based on soundness, precision, and performance. To achieve all the characteristics is contradictory as one of them is to be compromised to increase the efficiency of the other two. This is solved by a logical interface tool by constructing a source code model for the program. The most convenient model which can be used for code analysis is a control flow graph (CFG). In predicate analysis, information about exact values of variables and relations between them is extracted during analysis of program statements. These values and relations can be represented as logical predicates. Later, in pointer analysis, the rules are: pointer correctness, deference of a pointer to a simple variable, deference of a pointer to an element of complex object and summation of pointer and integer constants. Future considerations of the parameters of the rule can be defect detection rule are incorrect pointer deference , buffer overflow and array out of bounds.

The paper by George Stergiopoulos, PanagiotisKatsaros and DimitrisGritzalis [4] is based on automated detection of logical errors based on profiling the intended behaviour behind the source code. The territory of logical errors has yet been untouched and this paper is an attempt to put a light on this subject based on a profiling method that is combined with analysis of information and cross checks dynamic values with its natural symbolic execution and use of fuzzy logic. Errors are classified using fuzzy logic membership ie severity- values from a scale quantifying the impact of a logical error, with respect to how it affects the AUT's execution flow and vulnerability- with values from a scale quantifying the likelihood of a logical error and how dangerous it is. Profiling is done based on- intended program functionality as Rules (Dynamic Invariants), Program states and their variables, Source code profiling for logical error detection and severity(critical source code points).

The paper by PrakashMurali, AtulSandur and Abhay Ashok Patil [5] is about a logical error correction system in C using genetic algorithm techniques for error correction along with statistical control flow techniques. It is done using expression mining by considering a reduced

subset of the C language to probe the challenges of error correction. The logical error occur in the expression is hypothesized of the input program. The first task is to employ data mining on the input program and extract the expressions in the program. The stepwise analysis is shown in the paper. This algorithm can include logical errors in non-mathematical expressions by suitably modifying the code.

M. I. Glukhikh, V. M. Itsykson, and V. A. Tsesko [6], analyse the development of dependency analysis methods in order to improve static code analysis precision "Using Dependencies to Improve Precision of Code Analysis". They explain the reasons for precision loss when detecting defects in program source code using abstract interpretation methods. Dependency interpretation based on logic inference using logic and arithmetic rules is proposed by them. Defect detection based on abstract interpretation that provides both soundness and high precision is characterized by high computational complexity. The main cause is the need to analyze all the possible execution traces and to store the values of all the reachable objects in these traces. This complete dependency analysis is shown in this paper. The presented research demonstrates that extraction and interpretation of data dependencies is one of the most important aspects of code analysis. There are several directions to improve the suggested approach: development of more precise dependency merging rules, and use of automated theorem proving methods when interpreting dependencies.

3. PROPOSED WORK

Logical Errors has been a rather intangible area to be touched on. Logical error correction can help in saving loads of time along with the desired efficiency and better time complexity. The proposed approach is to compare the two programs and find out errors by detecting the missing invariants using machine learning and data mining. An invariant can be a missing element, statement or number of iterations in a program. There will be a predefined database which will have the profiling of missing invariants. Profiling is basically used for defining the basic use of each invariant using a predefined database .The use of missing invariants will be recognised and organised using data mining and this use will allow the machine learning device to learn the use of each invariant with respect to a particular program. Functions of each problem have similar set of statement unless it has been solved by a different algorithm. With this realisation, each program should be written as a function. So, the two functions can be compared to detect errors. The suggestions provided by machine learning device based on profiling and previous knowledge are displayed to the user wherein the user has a choice of modifying the code or embedding the correct code that is only stored in the database . Embedding the code can be done by replacing the whole code.

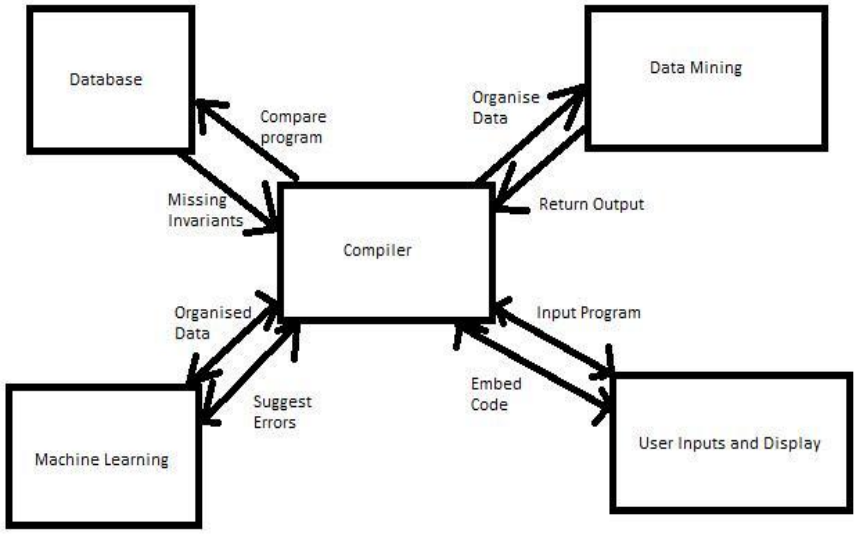


Figure 1:Block Diagram

Step 1: Compiler Construction

The compiler will consist of 4 parts-Machine Learner Device (MLD),Data Mining Device(DMD),C compiler, Database 1 and Database 2.Database 1 will be used to store the correctly executed programs with a hash-tag attached to it(shown in *Figure 2*). A new algorithm for the same problem can be stored separately. Database 2 will be the database used by MLD and DMD for profiling of each variant possible in a program. The data mining device is used to uncover the use of each invariant. The machine learning device will identify the use of each invariant in the user's program.

The compiler should be capable of comparing two programs efficiently and should also be able to find out the missing invariants in the program. The comparison should not be time consuming and the database and the software should work in tandem with that compiler.

Step 2: Programming Construction

The program should be divided into modules or functions. The function provides a base for comparison of the program. The function should provide the same logic as that of the correct program even if variables are defined differently unless a different .Modularity is a must. The construct willdepend upon-indentation, function use and parameters. It would be feasible if parameters in the function are same as it will provide better comparison structure and better efficiency.

Step 3: Comparing the programs

The compiler will compile two programs-the correct program and the incorrect program. The correct program will be the one stored in the database. One thing needs to be taken care of is that the logic should be same. If logic is different but a correct one, then that program will be stored in database. If the logic is different or incorrect one, then correct program will replace the incorrect one without showing the errors. If logic is same, it will evaluate the program and find out the missing code. The comparison should not take much time and should be done efficiently. This missing code will be transferred on to DMD for base profiling.

Step 4: Deducing the errors

Errors will be deduced by predefining the use of each operator, variables and functions etc. The code which was missing will be designated and organised using data mining. The data will be organised according to program and then the use of each element will be checked by DMD. Here, the use of machine learning will play an important part. The machine will learn the use of each element and will implement it according to that particular program. The efficiency and accuracy of the logical error correction should be a prime concern. The machine learning will deduce the errors based on the program as well as the use of profiling.

Step 5: Classifying the errors

The errors here will be classified into logical, syntactical and runtime errors. The syntactical errors can be found out using pre defined profiling. The logical errors require the description of each element along with machine learning of the program. Each of the missing elements have a logic and they need to be identified and classified accordingly. The syntactical and the run time errors faced can be stored in the database as the MLD can learn from these errors for providing future suggestions. They can be used as a reference while executing a similar program. The logical errors need to be individually processed for each program as each program can possibly have multiple logics. The new logic needs to be stored in the database with the same subject name/hash-tag. A hash-tag is basically a reference for storing and retrieving a particular program having a particular logic. The rules for assigning a hash-tag are as follows:

1) Each Hash-tag should be unique.

2) Each Hash-tag will contain a program of unique logic.

3) Hash-tags are case-sensitive. This needs to be taken care of during accessing of program through a hash-tag.

Step 6: Recommending and giving the right solution

The right solution will be based on the type of error detected. If the logic is different and the output is also incorrect, the correct code from the database will directly replace the incorrect code, else it will suggest the solution so that the user is aware of the mistakes. The solution if implemented, works, then that solution can be learnt by the machine so that in the future use, the machine realises the different implementations of logic.

Step 7: Embedding the correct solution in a program

The correct solution can be embedded in the form of macros or new functions as per the user requirement. System programming will play a part here if macros are required else functions will ser etter rea ieve his ime co

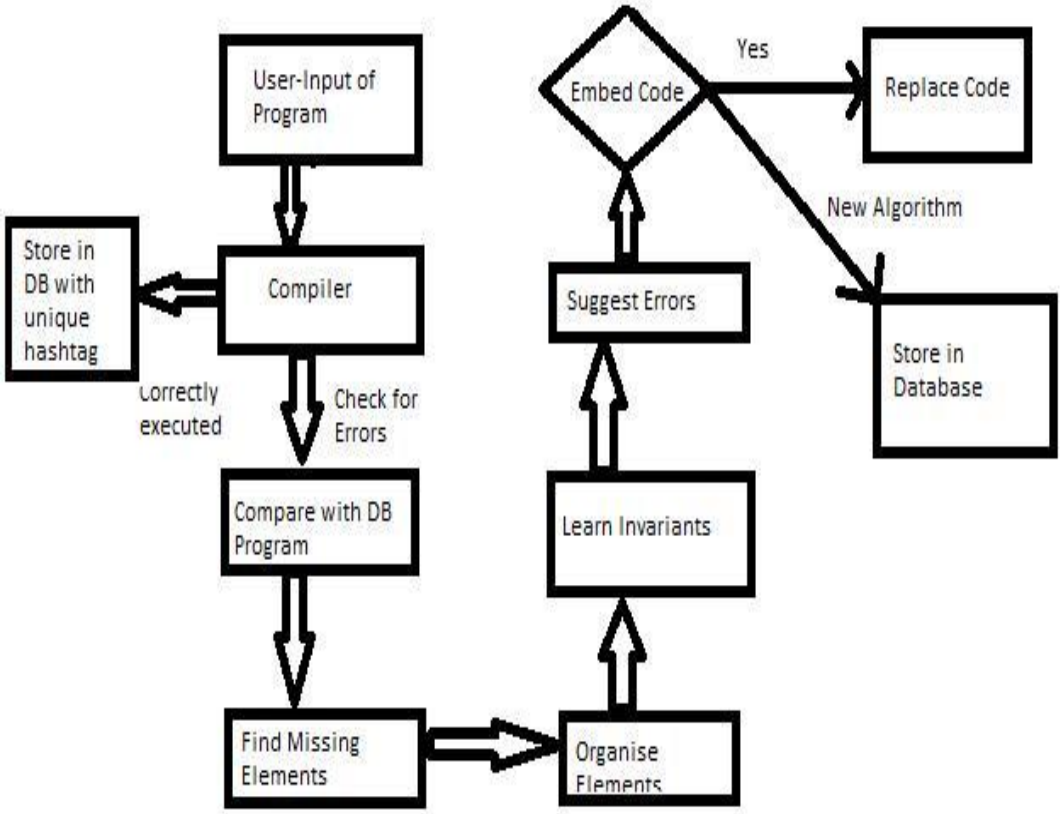


Figure 2*:* Flow Chart

4. ADVANTAGES AND DISADVANTAGES OF PROPOSED WORK

The advantages of our proposed system are as follows:

- Reverse Engineering will become simpler and efficient as understanding of the programming
- logic will become much easier.
- Time taken for debugging will reduce significantly as logical errors will be suggested which will make user realize of the mistake in the program
- Use of cloud computing will allow sharing of code across the world which in turn will

provide more test cases for machine learning.

- Time complexity and space complexity can also be compared of each correct program, thereby saving loads of time and space in a program.
- It can be implemented across all languages and all platforms.
- The program will only be stored if it is executed correctly. So, the comparison with the program in the database will be syntactically correct.

The disadvantages of our proposed system are as follows:

- Logical errors cannot be detected if the algorithm used to a problem is different. This can however be corrected by adding a new approach (correct program) in the database so that machine can learn the new approach.
- The correct program which is stored in database for reference should be logically correct itself. For example, if the program for addition is stored in the hashtag for subtraction, the comparison of the program will be done incorrectly.
- The logic and operators which is not defined in the database will not be compared which will reduce the efficiency to detect the errors.
- 100 percent efficiency will never be achieved.
- The system is susceptible to various security issues.

5. CONCLUSION

Developing the program is impossible unless the code gets compiled correctly. Therefore it is a very important part of the error correction process. A methodology has been proposed which will assist the novice programmers in realizing various types of syntactical errors and how they can be dealt with. The scope of this project deals with inclusion of more complex run-time and logical errors and correcting programs written in different programming languages like C++ and java.

REFERENCES

[1] K K Sharma, Kunal Banerjee, IndraVikas, ChittaranjanMandal, "Automated Checking of the Violation of Precedence of Conditions in else-if Constructs in Student's Programs", IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), 2014

[2] YuriyBrun, Michael D. Ernst, "Finding latent code errors via machine learning over program executions", Proceedings of the 26th International Conference on Software Engineering (ICSE).,2004

[3] Tatiana Vert, Tatiana Krikun, Mikhail Glukhikh, "Detection of Incorrect Pointer Dereferences for C/C++ Programs using Static Code Analysis and Logical Inference", Tools& Methods of Program Analysis, 2013.

[4] George Stergiopoulos, PanagiotisKatsaros, DimitrisGritzalis, "Automated detection of logical errors in programs", Springer-Verlag Berlin Heidelberg 2014.

[5] PrakashMurali, AtulSandur, Abhay Ashok Patil, "Correction of Logical Errors in C programs using Genetic Algorithm Techniques", International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009.

[6] M. I. Glukhikh, V. M. Itsykson, and V. A. Tsesko, "Using Dependencies to Improve Precision of Code Analysis", Automatic Control and Computer Sciences, 2012.

[7] V. Neelima, Annapurna. N, V. Alekhya, Dr. B. M. Vidyavathi, "Bug Detection through Text Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering, May 2013.
[8] Data Mining, available at: https://www.wikipedia.org/
[9] DataMining,availableat: http://www.anderson.ucla.edu/faculty/jason.frand/teacher/palace/datamining.html

INTENTIONAL BLANK

ARTIFICIAL NEURAL NETWORK FOR DIAGNOSIS OF PANCREATIC CANCER

Sanoob M.U [1], Anand Madhu[2], Ajesh K.R[3] and Surekha Mariam Varghese[4]
Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala

ABSTRACT

Cancer is malignant growth or tumour which forms due to an uncontrolled division of cells in a part of body which may even lead to death. These are of different types depending upon the part of body affected. If it is Pancreas then the disease is termed as Pancreatic Cancer. This paper presents an Artificial Neural Network model to diagnose pancreatic cancer based on a set of symptoms. An ANN model is created after analysing the actual procedure of disease diagnosis by the doctor. An approach to detect various stages of cancer affected in pancreas is presented in the paper. Results of the study suggest the advantage of using ANN model instead of manual disease diagnosis.

KEYWORDS

Neural Network, Diagnosis, Fuzzy Logic, Cancer, Pancreatic Cancer

1. INTRODUCTION

Artificial Neural Network (ANN) is a relatively raw model based on the brain's neural structure. In various clinical situations which are considered difficult, ANN has been used successfully as a non-linear pattern recognition technique in making diagnostic and prognostic decisions [1]. Now a days, many medical diagnosis problems are being solved using NN techniques. Artificial Neural Networks are applied to medicine mainly for the task which is based on the measured features to assign the patient to one of a small set of classes [2][3]. A number of researches are going on worldwide on the applicability of neural networks in medical diagnosis [4][5]. Accuracy as well as the objectivity of medical diagnosis has been increased using neural networks. In ANN, the processing element is called as neurons. An artificial Neural Network is a network of such interconnected neurons operating in parallel. Biological nervous systems are the main inspiration behind the concept of artificial neurons. Functioning of a network greatly depends on the connection between the elements in the network. These processing elements are subdivided into several subgroups called layers in the network. The first and the last layers are called the input and output layers respectively. There may be some additional layers, called hidden layers, in between the input and output layers. A neural network can be trained to perform a particular functionality. This is done by adjusting the values of the connections between the elements, called weights. The number of cancer related deaths worldwide is increasing day by day and researches shows that pancreatic cancer is the eighth most common cause of cancer-related deaths worldwide and fourth worldwide.

Malignant type neoplasm pancreatic cancer is originated from transformed cells which arise in tissues form the pancreas. Pancreas is a spongy organ that is around 6-inch long. This is located

in the back of the abdomen behind the stomach. Pancreatic juices, insulin, and hormones are created by exocrine and endocrine glands. These glands are contained within the pancreas. The exocrine glands make the enzymes or pancreatic juices, which are then released to the intestines through a series of ducts. This will help carbohydrates, proteins and fat to digest. Islets of Langerhans are the small clusters of these endocrine cells. The glucagon and insulin are released into the bloodstream by these islets of Langerhans. The levels of sugar in blood are managed by two of these hormones. Improper working of these hormones will often result in diabetes. Tumours are formed by the abnormal pancreas tissues continue splitting and create masses or lumps of tissues. The major functions of pancreas are then interfered by these tumours. Benign is the situation when a tumour stays at one location and shows limited growth.

This paper tried to show that if the particular condition of a patient is given, then the NN can be used to make an accurate prognosis of each individual [6] [7]. How human intelligence can be applied in health sector [8] [9] is the major concern behind this paper. A self-learning intelligent system can be developed using NN which can overcome the uncertainties in the diagnosis of pancreatic cancer. Some symptoms are taken from the patient's previous medical records as well as from the doctor, and by using these data the neural network model is trained to detect the presence/absence of pancreatic cancer in that patient. To diagnose the pancreatic cancer properly using this intelligent model fuzzified symptoms values are applied.

In prediction, NN models are widely being used, especially in medical diagnosis. Studies show that for the diagnosis of different medical diseases, ANN have been used very successfully. In 2004, a NN based model is proposed by Kamruzzaman et al. for the diagnosis of heart diseases [10]. In 2008, a Genetic Algorithm (GA) based technique for classifying tumour mass in breast and to identify breast cancer has been introduced [11]. A new method for Predicting Blood Cancer and Disorder is then developed by Payandeh [12] et al. later. One of the latest works in this is Artificial Neural Network for predicting headache which is done by Bahar et al. In 2011 also, a lot of NN based disease diagnosis has been done. Artificial Neural Network to pre-diagnosis of Hypertension [13], using Back-Propagation training algorithm, Artificial Neural Network model to diagnose skin diseases by Backpo [14] et al. etc are some of them. Similarly ANN models are also developed for breast cancer detection [15], Kidney stone diseases [16] etc.
In the following sections the paper will be dealing with the details of implementation of the ANN model for detecting the pancreatic cancer. In section II, the Methodology used to in the paper is discussed. The results of the experiment and Discussions are included in Section III. Then by Section IV, the paper is concluded.

2. METHODOLOGY

The initial step towards performing the process of medical diagnosis was initiated by examining a number of patients by a group of medical experts and identifying the symptoms. The next step was to propose a neural network which could be used in diagnosing PC diseases. The proposed model contains three layers namely input layer, hidden layer, and output layer. A single hidden layer consisting of 20 hidden layer neurons was created and trained. The input samples and output or target samples were divided for training, validation and test sets automatically.

For training the network, the training set was made use of. Training was continued until there was network stops improving the validation set. The test set was completely independent of the

measure of network accuracy. During the training phase, the patterns in data were learned by hidden neurons and the relationship between input and output pairs are mapped. In the hidden layer, a transfer function was used by each neuron for processing the data it receives from input layer and the processed information to the output neurons was transferred, for further processing using a transfer function in each neuron.

2.1. DATASET

Dataset used for the diagnosis of the pancreatic cancer is shown in Table 1. The data set consist of 11 possible symptoms and 3 outcomes possible. The outcome is purely dependent on the significance of symptoms for a particular patient. The entire dataset consists of measured features of 120 patients in which 90 samples were used for training the network and the remaining for testing purpose.

The set of symptoms represented as S = {Jaundice (J), Loss of Appetite (LA), Weight Loss (WL), Pain in Upper Abdomen (PUA), Irritability (I), Gall Bladder Enlargement (GBE), Swelling Lymph (SL), Diabetes Mellitus (DM), Deep Venous Thrombosis (DVT), Acholic Stool &Steatorrhea (AS&S) and Fatty Tissue Abnormalities (FTA) }.

The set of possible outcomes of diagnosis represented as D = {Disease Detected, Disease might be Detected and Disease not Detected}.

Disease outcomes and the corresponding label assigned for each of them is explained in Table 2. Next, on the basis of fuzzy set, the paper describes each symptom by its membership value. The basic block diagram that explains about the operational procedure is shown in Figure. 1.

Table 1. Symptoms Significance and Result

J	LA	WL	PUA	I	GBE	SL	DM	DVT	AS&S	FTA	Remarks
0.45	0.26	0.60	0.80	0.30	0.10	0.35	0.15	0.55	0.72	0.18	1
0.18	0.00	0.62	0.59	0.17	0.78	0.82	0.50	0.14	0.36	0.47	2
0.73	0.69	0.32	0.33	0.25	0.55	0.13	0.20	0.61	0.49	0.86	2
0.24	0.63	0.28	0.08	0.39	0.36	0.70	0.55	0.23	0.17	0.63	3
0.59	0.68	0.43	0.75	0.73	0.29	0.37	0.13	0.70	0.55	0.43	1
0.44	0.53	0.56	0.69	0.57	0.63	0.19	0.41	0.34	0.72	0.38	1
0.52	0.63	0.72	0.30	0.60	0.19	0.40	0.21	0.42	0.39	0.52	3
0.13	0.12	0.63	0.24	0.11	0.47	0.23	0.51	0.69	0.10	0.63	2
0.63	0.38	0.33	0.21	0.49	0.72	0.62	0.24	0.77	0.43	0.40	1
0.24	0.55	0.78	0.30	0.54	0.41	0.78	0.43	0.64	0.18	0.23	1

Table 2. Assigned Labels and Outcome

Label	Outcome
1	Detected
2	Might be Detected
3	Not Detected

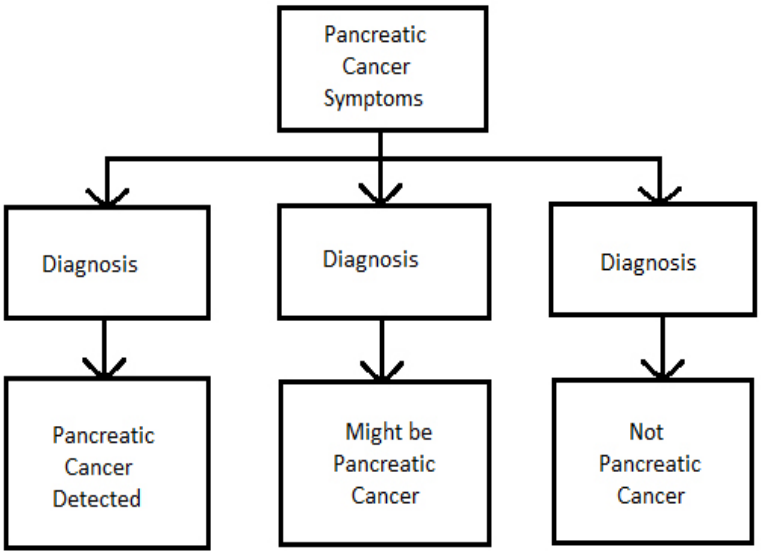


Figure 1. Operational Procedure

2.2. TRAINING OF PARAMETERS

The network will become ready to be trained, only once that network is organized and structured for the targeted application. The initial weight has to be chosen at random for starting this process. Training will be started after that. The network is trained by using existing set of data which is directly obtained from various patients and whose output is well known. The neurons in hidden layer will learn the data pattern while training and map relation between input pairs and output pairs. Each hidden layer neuron made use of a transfer function for processing data that accepts from input layer and transfers the information which is processed to the output neurons for continuing the processing in each neuron using a transfer function.

3. RESULTS AND DISCUSSION

Matlab R2011a's toolbox for Neural network is used for performance evaluation for the new networks which consists of a 11 number three layer feed forward network of inputs and sigmoid hidden neurons and linear output neurons is suggested. The new approach used Levenberg-Marquardt algorithm for back propagation for training the network where training stops automatically when generalization stops improving, as indicated by an increase in the MSE (Mean Square Error) of the samples used for validation. The proposed neural network is shown in Figure. 2.

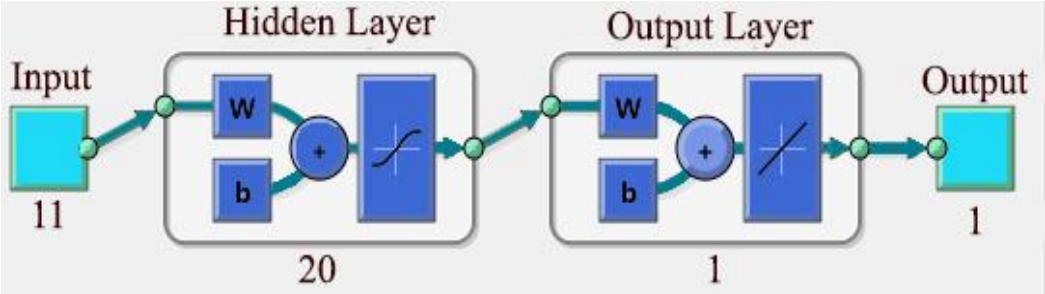


Figure 2. Proposed Neural Network

A membership based fuzzification scheme is adopted here for converting our dataset to to a fuzzified set of symptoms. After an interview with physicians, a linear membership function was again selected for each symptom. Three to five linguistic variables were assigned to each symptom normally, and then repeated the classification tests.

The experimental results of implementing the new ANN methodology to distinguish between Pancreatic Cancer affected and non-affected patients based upon specified symptoms represents good capabilities of the network to learn the training patterns corresponding to symptoms of the patients. The experimental setup is shown in Figure. 3

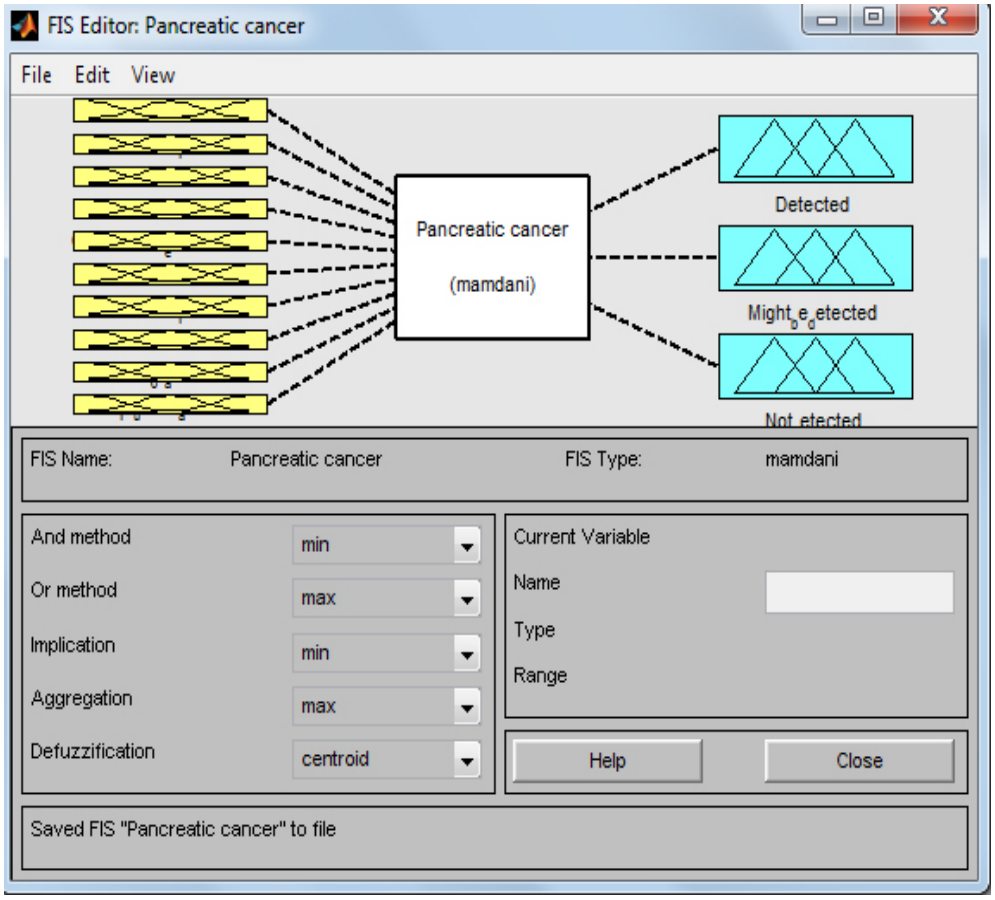


Figure 3. FIS Editor

A triangular membership function is used for fuzzifying the inputs. Using Matlab, a membership function editor is used. This Membership function editor is shown in the Figure. 4.

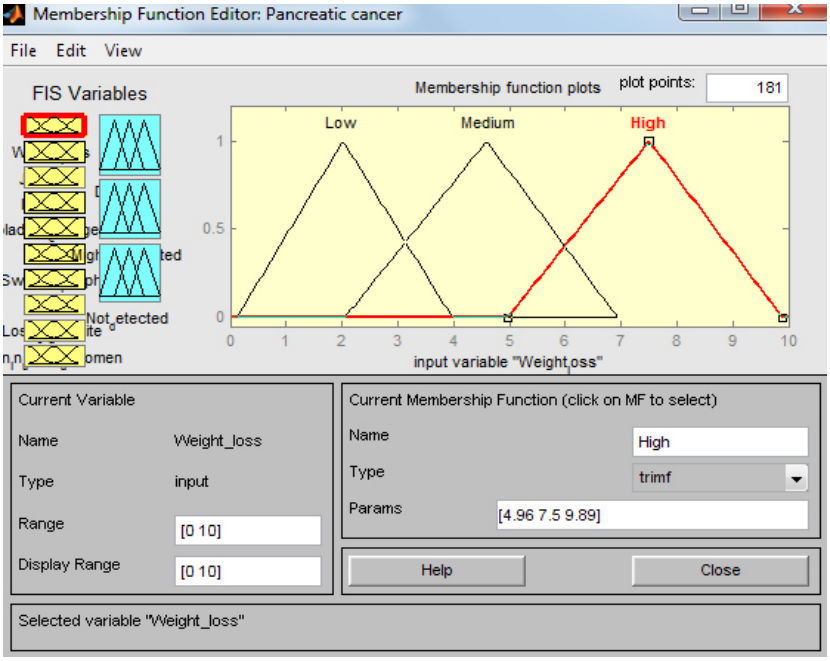


Figure 4. Membership Function Editor

The performance graph plotted based on the results obtained is shown in Figure. 5.

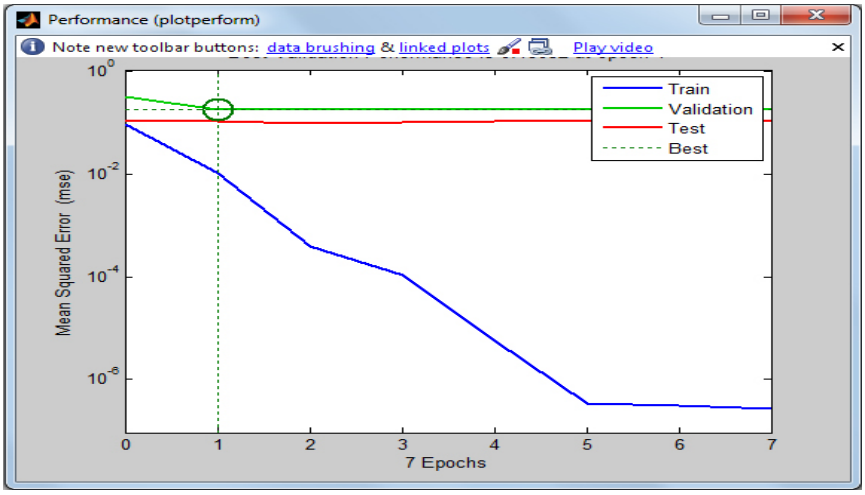


Figure 5. Performance Plot

4. CONCLUSION

An approach for the diagnosis of out-of-controlled cell growth in pancreas based on Artificial Neural Network is explained in this paper. Detection of cancer in the pancreatic cells at its early stage is necessary for its better treatment and cure. Hence it is important to detect pancreatic cancer automatically to mitigate the real-world medical problems. Here we have presented how effectively we can made use of neural networks in the detection and diagnosis of cancer in pancreatic cells. The construction of a diagnostic system which is highly accurate based on neural network model is done here. Also an investigation on the performance of neural network structure is done in this paper. This model is designed in such a way that it made use of fuzzy values instead of normal values for the incorporation of the neural network context. The model has provision to accept symptoms in a patient and it will inform the present condition of that patient, based on the evaluation made on the input symptoms. So this model is an interactive model. So the early detection can be done and hence it will help the doctor to plan and do the better medication for the patient. The symptoms are fuzzified and network is implemented based on around 20 neurons, for getting the better performance and accurate diagnosis. Outcome of the experiments indicates that the noval approach is able to valuate data in most efficient way compared to other normal approaches. Future works can be extended for other similar disease detection comprising complex and related datasets with similar or better accuracy..

REFERENCES

[1] N.Salim, Medical Diagnosis Using Neural Networks, 2004.

[2] W. David Aha and Dennis Kibler, Instance-based prediction of heart disease presence with the Cleveland database

[3] J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C., Johannes, R. S.,Using the ADAP learning algorithm to forecast the onset of diabetes mellitus, Proc. Symp. on Computer Applications and Medical Care, pp. 2615, 1988.

[4] SuvarnaMahavirPatil and R.R. Mudholkar, An Osteoarthritis classifier using back-propagation neural network,International Journal of Advances in Engineering & Technology, Sept 2012, ISSN: 2231-1963.

[5] J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C., Johannes, R.S.,Using the ADAP learning algorithm to forecast the onset of diabetes mellitus, Proc. Symp. on Computer Applications and Medical Care, pp. 2615, 1988.

[6] Imianvan Anthony Agboizebeta., and Obi Jonathan Chukwuyeni, Application of Neuro-Fuzzy Expert System for the Probe and Prognosis of Thyroid Disorder, International Journal of Fuzzy Logic Systems(IJFLS) Vol.2, No.2, April 2012.

[7] Obi J.C. Imianvan A.A, Interactive Neuro-Fuzzy Expert system for diagnosis of Luukemia,, Global Journal of Computer Science and Technology, Volume 11 Issue 12 Version 1.0 July 2011.

[8] W. David Aha and Dennis Kibler, Instance-based prediction of heart disease presence with the Cleveland database, 3rd ed. Harlow, England: Addison-Wesley, 1999.

[9] Prof. A. Maithili, Dr. R. VasanthaKumari Mr. S. Rajamanickam, Neural Networks towards medical, ,International Journal of Modern EngineeringResearch (IJMER), Vol.1, Issue1, pp-57-64 ISSN: 2249-6645.

[10] S. M. Kamruzzaman, Ahmed RyadhHasan, Abu BakarSiddiquee and Md. EhsanulHoqueMazumder, Medical diagnosis using neural network,, ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh

[11] Arpita Das and Mahua Bhattacharya, GA based Neuro Fuzzy Techniques for breast cancer Identification,, 3rd ed. IEEE, 978-7695-3332-2/08, 2008. DOI: 10.1109/IMVIP.2008.19
[12] Payandeh M, MehrnoushAeinfar, VahidAeinfar, Mohsen Hayati, A New Method for Diagnosis and Predicting Blood Disorder and Cancer Using Artificial Intelligence, IJHOSCR, Vol. 3, No.4; 2009.
[13] B. Sumathi,Dr. A. Santhakumaran, Pre-Diagnosis of Hypertension Using Artificial Neural Network, Global Journal of Computer Science and Technology,Global Journal of Computer Science andTechnology Volume 11 Issue 2 Version 1.0 February 2011
[14] Bakpo, F. S. and Kabari, L. G, Diagnosing Skin Diseases Using an Artificial Neural Network DOI:10.5772/16232.
[15] BipulPandey, Tarun Jain, Vishal Kothari and Tarush Grover, Evolutionary Modular Neural Network Approach for Breast Cancer Diagnosis, IJCSI International Journal of Computer Science Issues. Vol.9, Issue 1, No 2, January 2012.
[16] Koushal Kumar, Abhishek, Artificial Neural Networks for diagnosis of kidney stones disease, I.J. Information Technology and Computer Science, 2012, 7, 20-25
[17] W. David Aha and Dennis Kibler, Instance-based prediction of heart disease presence with the Cleveland database, 3rd ed. Harlow, England: Addison-Wesley, 1999.

AUTHORS

Sanoob M.U. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. He completed his B.Tech from AdiShankara Institute of Engineering and Technology, Kalady. His areas of research are Machine Learning and Databases.

Ajesh K.R. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. He completed his B.Tech from AdiShankara Institute of Engineering and Technology, Kalady. His areas of research are Machine Learning and Image Processing.

AnandMadhu is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. He completed his B.Tech from University College of Engineering, Thodupuzha. His areas of research are Machine Learning and Data Mining.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems, Machine Learning and Distributed Computing. She has published 17 papers in international journals and international conference proceedings. She has been in the chair and reviewer for many international conferences and journals.

INTENTIONAL BLANK

A HYBRID K-HARMONIC MEANS WITH ABCCLUSTERING ALGORITHM USING AN OPTIMAL K VALUE FOR HIGH PERFORMANCE CLUSTERING

Sithara E.P and K.A Abdul Nazeer

Department of Computer Science and Engineering, National Institute of Technology, Calicut, Kerala, INDIA

ABSTRACT

Large quantities of data are emerging every year and an accurate clustering algorithm is needed to derive information from these data. K-means clustering algorithm is popular and simple, but has many limitations like its sensitivity to initialization, provides local optimum solutions. K-harmonic means clustering is an improved variant of K-means which is insensitive to the initialization of centroids, but still in some cases it ends up with local optimum solutions. Clustering using Artificial Bee Colony (ABC) algorithm always gives global optimum solutions. In this paper a new hybrid clustering algorithm (KHM-ABC) is presented by combining both K-harmonic means and ABC algorithm to perform accurate clustering. Experimental results indicate that the performance of the proposed algorithm is superior to the available algorithms in terms of the quality of clusters.

KEYWORDS

Data Mining, Clustering, K-means Clustering, K-Harmonic means Clustering, Artificial Bee Colony Algorithm

1.INTRODUCTION

Cluster analysis is one of the important data analysis method which is used in the areas like data mining, vector quantization, image analysis and compression. Clustering is a process which sequentially takes data as inputs and outputs clusters as results. The aim of clustering is to assemble a set of similar objects into a group that in some sense belong together because of related characteristics [1][2].

Among the various clustering methods available, K-means is very simple clustering method to cluster data sets, but this method highly depends on the initial selection of centroids and usually converges to the local optimum solutions [1][3][4][5]. Similarly, K-harmonic Means clustering is a centroid based clustering algorithm in which the harmonic mean of the distances between the centroids and each data point is used as the main component of the performance function [4][6].

K-harmonic means provide better clustering results than K-means [2][3]. In K-means the bond between data points and the nearest centroid is very strong, so that data point is strongly attached to a cluster centroid and it is moving to another cluster only when it is too close to another centroid. This powerful bond stops the centroids from shifting out of the surrounding locality of data. In K-harmonic means, the harmonic means function is used to establish the link between data points and centroids. This association is distributed and make the algorithmnsensitive to

DOI: 10.5121/ijci.2016.5206

initialization. In [7], Bin Zhang provided a performance chart for sensitivity to initialization, given in Figure 1. In this paper KHM_p is the K-harmonic means clustering algorithm in which the p^{th} power of distance function is used, usually p = 2. The K-harmonic algorithm always converge faster than K-means, but sometimes it provides less accuracy clustering results. If the problem contain many local minima then the algorithm will fall into a local optimum solution [8].

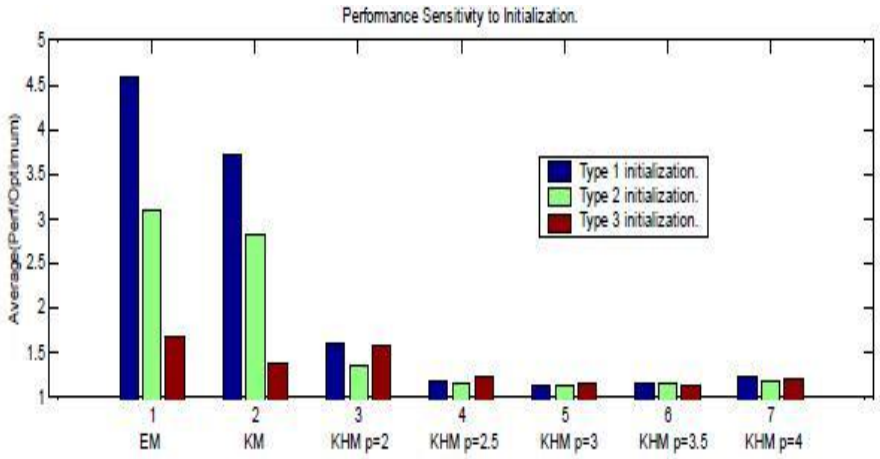


Figure 1. Performance sensitivity to initialization [7]

A better mode of clustering is by the use of Artificial bee colony algorithm (ABC). ABC was introduced by Dervis Karaboga in 2005 [9][10] which was established on the foraging characteristics of honey bees, but can be used for solving numerical optimization problems. It is a population based optimization algorithm. Here, the primary idea is to use ABC to generate best solution by providing non-local moves for the cluster cores

2. RELATED WORK

Several researchers have highlighted their work in the field of clustering. The continuing work on this area has brought about novel and enhanced methods for clustering.

2.1. K-HARMONIC MEANS CLUSTERING

K-harmonic means algorithm is a centre based clustering algorithm, in which the harmonic means between centroids and data values are taken as the main constituent of the performance function [6][7]. The performance function is given in equation 1, where x_1 to x_n are data points, c_1 to c_k are centroids and k is the required number of clusters.

$$perf_{KHM} = \sum_{i=1}^{n} \frac{k}{\sum_{j=1}^{k} \frac{1}{||x_i - c_j||^2}} \qquad (1)$$

It was demonstrated that K-harmonic means is essentially insensitive to the centroid initialization. In K-means, arithmetic mean is used and the value is always near to the higher of two values. But harmonic mean value is near to the minimum of two values. This property enables K-harmonic means clustering to suppress the effect of outliers. The performance of K-harmonic means in improving the quality of clusters, was better than K-means.

The hybrid clustering algorithm introduced by Ravindra Jain [9] is based on applying K-means and K-harmonic means in tandem to find cluster mean until termination condition. It shows that K-harmonic means yields better accuracy in the clustering than K-means algorithm.

Fangyan Nie, Tianyi Tu, et al. proposed a combined Particle swarm optimization and K-harmonic means clustering (PSOKHM) algorithm in [8]. It was a hybrid algorithm which combines the benefits of both algorithms. K-harmonic means clustering sometimes fall into local optimum solutions, so a Particle swarm optimization (PSO) was used to evolve non-local centroid movements and leads to better solutions. In this hybrid algorithm, two methods are used to find the cluster means, thus the algorithm produces better results because it combines the advantages of both the techniques.

2.2. ARTIFICIAL BEE COLONY ALGORITHM

Artificial bee colony (ABC) algorithm was introduced as a swarm-based algorithm [10]. The algorithm is explained by categorising the bees into three groups: the employees, onlookers and scouts. The number of employed bees and the number of onlooker bees are same as that of the number of solutions. In other words, the number of solutions are equal to the number of food sources around the hive.

Employed bees are searching for the food sources. After identifying a food source it returns back to the hive and dance in the dancing area of the hive. Onlooker bees observe this activity to get an idea about the nectar quantity of food sources and choose a food source with a good amount of nectar. The food source with least amount of nectar is considered as abandoned and that bee is acting as a scout and starts to search for a new food source. A possible solution to the problem is represented by the position of food source and the quality (fitness) of a solution is represented by nectar amount of the food source.

In [10] Changsheng Zhang, et al. discussed how ABC can be used for clustering and they proved that the ABC algorithm can work effectively, by analysing the computation time of the ABC algorithm and other well-known techniques [10][11]. A performance comparison table is given in Table 1 taken from [10]. It shows that the ABC algorithm achieves better results. In [12] Bahriye Akay and Dervis Karaboga compared ABC algorithm with other algorithms like genetic algorithm, evolutionary algorithm, particle swam optimization, etc. They proved that ABC algorithm has better performance compared to the mentioned algorithms.

Table 1. Performance comparison table [10]

The average fitness computation numbers and computation time.

Data set		GA	ACO	K–NM–PSO	ABC
Iris	Time (s)	105.53	33.72	48.13	**29.68**
	Numbers	38128	10998	**4556**	8658
Thyroid	Time (s)	153.24	102.15	118.46	**85.26**
	Numbers	45003	25626	**7245**	24136
Wine	Time (s)	226.68	68.29	589.40	**48.85**
	Numbers	33551	9306	46459	17554

Giuliano Armano and Mohammad Reza Farmani [3] proposed a hybrid algorithm as a

combination of artificial bee colony algorithm and K-means algorithm. K-means algorithm is highly dependent on the initialization of centroids and usually gets stuck in local optima. The ABC algorithm performs a global search in the entire solution space and it can generate good and global results. The authors propose a new combination algorithm which makes use of the combined benefits of K-means and ABC algorithms for solving clustering problems.

3. PROPOSED APPROACH

Even though K-harmonic means is insensitive to initial-ization of centroids, the cluster quality needs to be improved by finding global optimum solutions. Thus a well performed optimization algorithm, ABC algorithm is utilized for non-local movement of centroids and to obtain more promising results.

In this hybrid approach k value should be fixed before executing the algorithm. Gap statistics method and Average silhouette width method are used to identify the optimal k value and the value thus obtained is used to fix the number of initial food sources in the proposed algorithm.

3.1. IDENTIFYING OPTIMAL K VALUE

At the pre-processing stage optimal k value is estimated using Gap Statistics method [13]. The obtained value is verified using Average silhouette width method [14].

3.1.1. GAP STATISTICS METHOD

Run a K-means algorithm on the given set of data to find number of clusters, and sum the distance of all points from their cluster mean, this is the dispersion. Generate some number of sample data sets of original and find the mean dispersion of these sample data sets. Each gap is defined as the logarithmic difference between the mean dispersion of reference data sets and dispersion of the original data set. Take the minimum value of k for which the gap is maximized.

3.1.2. AVERAGE SILHOUETTE WIDTH METHOD

Run a PAM (Partition Around Medoids) algorithm on the original data set for values of k, in the range 2 to 10. The average silhouette width of clusters formed is calculated for each iteration. Observe the highest value. The k value corresponding to the highest average silhouette width is taken as the optimal k value.

3.2. PROPOSED ALGORITHM

In the proposed approach, ABC algorithm helps the K-harmonic means clustering algorithm to set the global optimum solutions. The K-harmonic means performance function is used to calculate the fitness of each solution and the characteristics of ABC algorithm leads to global optimum solutions rather than local optimum solutions. These properties of both the algorithms provide better performance in the quality of clusters. The method is formulated in algorithm 1

Algorithm 1 Pseudo-code of the proposed algorithm

Input : Number of data values indicated as $x_i..x_n$

: Values for control parameters SN (number of food sources which is same as k), limit and MCN (maximum cycle number).

Output: SN number of clusters.

1) Begin

2) Initialize trial counter array with values zero
3) Load data set values.
4) Initialize food sources(centroids) c_i where i = 1…SN

5) Evaluate the fitness values (fit_i) of the food sources using k-harmonic means performance function.

6) Set cycle to 1
7) Repeat until the termination criteria met (cycle = MCN)

8) For each employed bee

   a) Produce new food source (v_{ij}).
   b) Use k-harmonic means to evaluate the new fitness values.

   c) Compare them with the original one, if the fitness value does not improve increment the corresponding trial counter value.

   d) Better food source will be memorized and delivered to onlooker bee.

9) Evaluate probability values (p_i) of food sources.

10) For each onlooker bee

   a) Select a food source depending on probability.
   b) Produce new food source.
   c) Apply k-harmonic means to find new fitness values

   d) Compare them with the original one, if the fitness value does not improve increment corresponding trial counter value and memorize the best food source

11) Check if trial counter value > = limit value then the food source is abandoned by the bee and that employee bee become scout. Abandoned solution is replaced by the scout with new randomly produced food source.

12) Memorize the best solutions achieved so far
13) cycle=cycle+1
14) End

Fitness calculations are given in equations 2 and 3.

$$f_i = \sum_{i=1}^{n} \frac{k}{\sum_{j=1}^{k} \frac{1}{||x_i - c_j||^2}} \qquad (2)$$

$$fit_i = \frac{1}{1 + f_i} \qquad (3)$$

Probability equation and equation to produce a new random solution are given by equations 4 and 5 respectively ϕ_{ij} is a random number between -1 and 1.

$$p_i = \frac{fit_i}{\sum_{m=1}^{k} fit_m} \quad (4)$$

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) \quad (5)$$

4. EXPERIMENTAL RESULTS

This section presents the outcome of the experiments carried out for evaluating the performance of the suggested algorithm in improving the cluster quality.

4.1. ALGORITHM IMPLEMENTATION

The standard K-harmonic function and the proposed algorithm (KHM-ABC) were coded in R programing. The data sets used are iris, wine, yeast and spam base downloaded from UCI learning repository.

4.2. RESULTS

4.2.1. OPTIMAL K VALUE IDENTIFICATION

Gap statistics method is used to identify the optimal k value. A plot for wine data set is shown in Figure 2. From the figure, it clearly shows that the minimum value of k with maximum gap is 3. Thus the optimal value of k is 3.

4.2.2. VERIFYING THE OPTIMAL K VALUE

The obtained k value is verified using Average silhouette width method and the plot is given in Figure 3 for wine data set. From the given plot we can infer that the highest average silhouette width value is found at k = 3. Thus the optimal k value can be taken as 3.

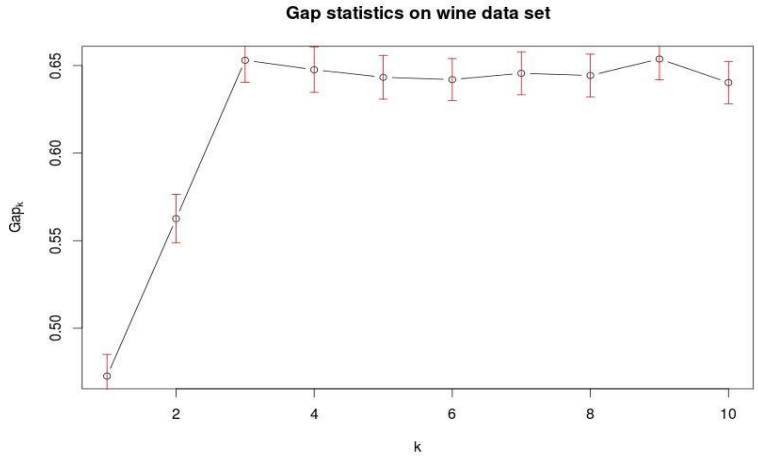


Figure 2. Gap statistics for wine data

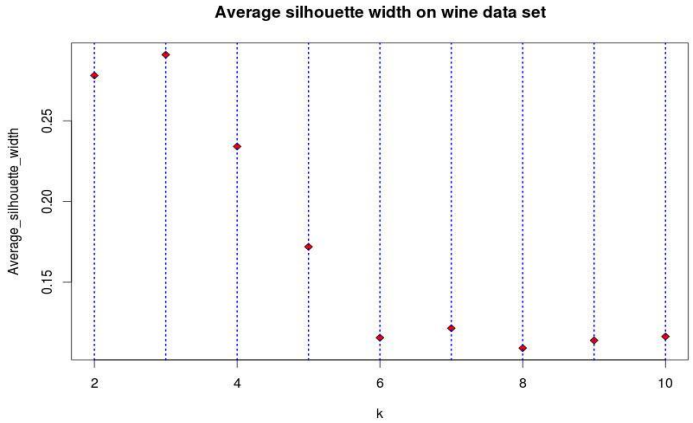


Figure 3. Average silhouette width for wine data

4.2.3. PERFORMANCE SCORES

Silhouette index scores are used to evaluate the performance. Silhouette index scores for clustering algorithms K-means, K-harmonic means, PAM, ABC and KHM-ABC are calculated on different data sets iris, wine, yeast and spam base. The results of the experiments are tabulated and given in Table 2.

The performance comparison graph is given in Figure 4, which shows the improvement of KHM-ABC algorithm in terms of accuracy.

5. CONCLUSIONS AND FUTURE SCOPE

K-harmonic means algorithm overcomes many of the limitations of K-means, but still it may get trapped into local optimum solutions. The proposed method (KHM-ABC) used artificial bee colony algorithm to optimize K-harmonic means clustering algorithm to improve the clustering quality. ABC algorithm always provides global optimum solutions. This feature helps K-harmonic algorithm to fix a good set of initial centroids. The proposed method guarantees the cluster quality. Cluster quality was checked using silhouette index scores. Silhouette index scores are calculated for KHM-ABC and other related popular algorithms ABC, K-means K-harmonic means and PAM. The results showed that the performance of KHM-ABC was better compared to

the other algorithms.

One of the main constraints of the proposed algorithm is that the value of k is not self-learned. In the pre-processing stage the k value was fixed using gap statistics method. The k value thus obtained is verified using silhouette width method. Some statistical method with a systematic approach is worth investigating for determining the value of k at run time.

Table 2. Performance comparison table

Sl.no.	Method	Data set	Optimal k	Silhouette index score
1	K-means	Iris	3	0.55
		Wine	3	0.57
		Yeast	6	0.16
		Spambase	3	0.68
2	KHM	Iris	3	0.55
		Wine	3	0.57
		Yeast	6	0.15
		Spambase	3	0.67
3	PAM	Iris	3	0.55
		Wine	3	0.57
		Yeast	6	0.15
		Spambase	3	0.68
4	ABC	Iris	3	0.53
		Wine	3	0.56
		Yeast	6	0.16
		Spambase	3	0.66
5	KHM-ABC	Iris	3	0.55
		Wine	3	0.57
		Yeast	6	0.2
		Spambase	3	0.71

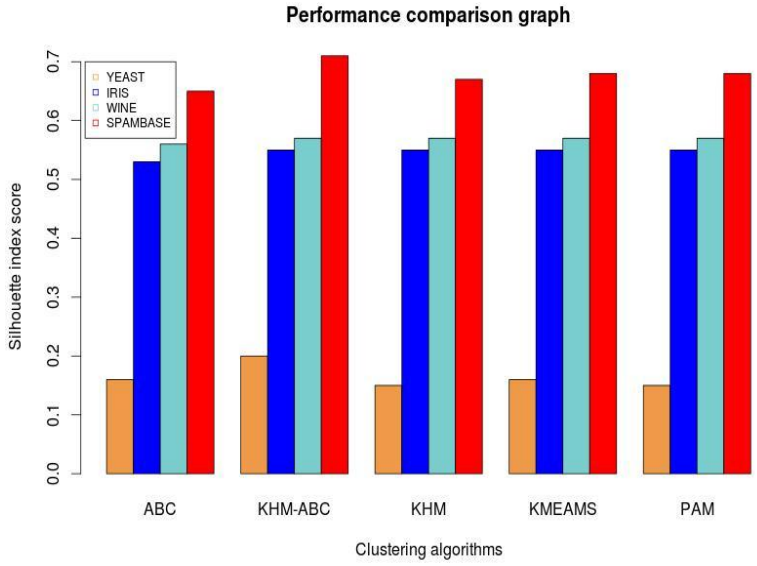


Figure 4. Performance comparison graph

REFERENCES

[1] K. A. A. Nazeer and M. P. Sebastian, "Improving the Accuracy and Efficiency of K–means clustering algorithm", Proceedings of the World Congress on Engineering 2009, vol. Vol I, 2009.

[2] K.Thangavel and N. Visalakshi, "Ensemble based Distributed K-Harmonic Means Clustering", International Journal of Recent Trends in Engineering, vol. Vol 2, NO.1, pp 125–129, 2009.

[3] G. Armano and M. R. Farmani, "Clustering Analysis with Combination of Artificial Bee Colony Algorithm and k-means technique", International Journal of Computer Theory and Engineering, Vol 6, Part 2, pp 141-145, 2014

[4] S. Reyya, M. Pushpa, and et al, "Increasing Comparison Performance using K-Harmonic Mean", International Journal of Management, Information Technology and Engineering, vol. Vol 2, Issue 3, pp 11–18, 2014.

[5] N. Alldrin, A. Smith, and D. Turnbull, "Clustering with EM and K-Means", Department of Computer Science, University of California, San Diego

[6] B. Zhang, M. Hsu, and U. Dayal, "K-Harmonic Means - A Data Clustering Algorithm", Software Technology Laboratory,HP Laboratories Palo Alto, HPL– 1999-124, 1999.

[7] B. Zhang, "Generalized K-Harmonic Means – Dynamic Weighting of Data in Unsupervised Learning", Hewlett-Packard Laboratories, 1999.

[8] F. Nie, T. Tu, and et al, "K-Harmonic Means Data Clustering with PSO Algorithm", in Advances in Electrical Engineering and Automation, AISC, Springer Verlag, 2012, pp. 67–73.

[9] R. Jain, "A Hybrid Clustering Algorithm for Data Mining", School of Computer Science IT, Indore, India,

[10] C. Zhang, D. Ouyang, and J. Ning, "An Artificial Bee Colony approach for Clustering", Expert Systems with Applications, Elsevier, pp. 4761–4767, 2010.

[11] D. Karaboga and C. Ozturk, "A novel clustering approach: Artificial Bee Colony (ABC) algorithm", Ap-plied Soft Computing, Elsevier, pp. 652–657, 2011.

[12] B. A. Dervis Karaboga, "A comparative study of Artificial Bee Colony algorithm", Applied Mathematics and Computation, Elsevier, pp. 108–132, 2009.

[13] R. Tibshirani, G. Walther, and T. Hastie, "Estimating the Number of Clusters in a Dataset via Gap Statistic", J.R.Statist.Soc.B, Sanford University USA, vol. Vol 63, Part 2, pp. 411–423, 2001.

[14] J. Rahnenfuhrer and F. Markowetz, "Exploratory Data Analysis — Clustering Gene Expression Data", Practical DNA Microarray Analysis, Saarbrucken, 2005.

AUTHORS

Sithara E. P obtained herB.Tech (Computer Science and Engineering) from College of Engineering Vadakara, and M.Tech (Computer Science and Engineering) from NIT Calicut. She is currently working as Assistant Professor in Computer Science and Engineering Department, College of Engineering Vadakara. Her areas of interest include Bioinformatics, Data Mining, Data structures and algorithm analysis.

K. A Abdul Nazeer obtained his B.Tech (Computer Science and Engineering) from TKM College of Engineering, Kollam, University of Kerala, M.Tech (Computer Science and Engineering) from IIT Madras and Ph. D from NIT Calicut (Thesis Title: Improved Clustering Algorithms for Bioinformatics Data Analysis). He is currently Associate Professor and Head of Computer Science and Engineering Department, NIT Calicut. His areas of Interest includes

INTENTIONAL BLANK

FUZZY FINGERPRINT METHOD FOR DETECTION OF SENSITIVE DATA EXPOSURE

Staicy Ulahannanl[1] and Roshni Jose[2]

[1] Student, Department of Computer Science Engineering, MBITS Nellimattom
[2]Assistant Professor, Department of Computer Science, MBITS Nellimattom

ABSTRACT

Protecting confidential information is a major concern for organizations and individuals alike, who stand to suffer huge losses if private data falls into the wrong hands. Network-based information leaks pose a serious threat to confidentiality. This paper describes network-based data-leak detection (DLD) technique, the main feature of which is that the detection does not require the data owner to reveal the content of the sensitive data. Instead, only a small amount of specialized digests are needed. The technique referred to as the fuzzy fingerprint – can be used to detect accidental data leaks due to human errors or application flaws. The privacy-preserving feature of algorithms minimizes the exposure of sensitive data and enables the data owner to safely delegate the detection to others.

KEYWORDS

Network Security, Privacy, Data Leak, Detection, Collection Intersection

1. INTRODUCTION

Information leaks are a major problem of computer systems. The leak of confidential data either be it accidental or intentional, may cause huge losses to the data owner. Though there are number of systems designed for the data security by using different encryption algorithms, there is a big issue of the integrity of the users of those systems. It is very hard for any system administrator to trace out the data leaker among the system users. It creates a lot many ethical issues in the working environment.

Typical approaches to preventing data leak are under two categories – host-based solutions and network-based solutions. Host-based approaches may include encrypting data when not used and enforcing policies to restrict the transfer of sensitive data. Most of the host-based solutions require the use of virtualization or special hardware to ensure the system integrity of the detector. This paper present a novel network-based data-leak detection (DLD) solution that is both efficient and privacy-preserving In comparison to host-based approaches, network-based data-leak detection focuses on analyzing the (unencrypted) content of outbound network packets for sensitive information.

Another motivation for the privacy-preserving DLD work is cloud computing, which provides a natural platform for conducting data-leak detection by cloud providers as an add on service. In cloud computing environments, an organization (data owner) may have already outsourced its

DOI: 10.5121/ijci.2016.5207

services to a cloud provider, such as the email service for its own employees. The cloud provider may offer additional services such as inspecting email traffic for inadvertent data leak and serves as a DLD provider. This add-on DLD service requires minimal changes to the cloud provider's infrastructure and makes the cloud service more attractive. However, privacy is a major roadblock for realizing outsourced data-leak detection. Conventional solutions require the data owner to reveal its sensitive data to the DLD provider.

However, the DLD provider is always modeled as an honest-but-curious (aka semi-honest) adversary who is trusted to perform the inspection, but may attempt to learn about the data. Existing work on cryptography-based multiparty computation is not efficient enough for practical data leak inspection in this setting.

This paper design, implement, and evaluate a new privacy preserving data-leak detection system that enables the data owner to safely deploy locally, or to delegate the traffic inspection task to DLD providers without exposing the sensitive data. In this model, the data owner computes a special set of digests or fingerprints from the sensitive data, and then discloses only a small amount of digest information to the DLD provider [3]. These fingerprints have important properties, which prevent the provider from gaining knowledge of the sensitive data, while enable accurate comparison and detection. The DLD provider performs deep-packet inspection to identify whether these fingerprint patterns exist in the outbound traffic of the organization or not, according to a quantitative metric. To prevent the DLD provider from gathering exact knowledge about the sensitive data, the collection of potential leaks is composed of real leaks and noises. It is the data owner, who post-processes the potential leaks sent back by the DLD provider and determines whether there is any real data leak. Data leak is intentional or unintentional release of secure information to an untrusted environment.

These technical contributions are summarized as follows.

1. This paper describes a novel fuzzy fingerprint method for detecting inadvertent data leak in network traffic. Network security consists of the policies adopted to prevent and monitor unauthorized access, misuse, modification, or denial of computer network and computer accessible resource. Its main feature is that the detection can be performed based on special digests without the sensitive data in plaintext, which minimizes the exposure of sensitive data during the detection. This strong privacy guarantee yields a powerful application of fuzzy fingerprint method in the cloud computing environment, where the cloud provider can perform data-leak detection as an add-on service to its clients. This paper describes the quantitative privacy model, algorithms, and analysis in fuzzy fingerprint. The privacy model is useful beyond the specific fuzzy fingerprint problem studied. The detection is based on the fast set-intersection operation between the set of fingerprints generated from the payload of intercepted traffic (done by the DLD provider) and the set of fingerprints generated from the sensitive data (done by the data owner).
2. This paper implement detection system and perform extensive experimental evaluation on 2.6 GB Enron dataset, Internet surfing traffic of 20 users, and also 5 simulated real-world data-leak scenarios to measure the privacy guarantee, detection rate, and efficiency of proposed technique. The results indicate high accuracy performed by underlying scheme with very low false positive rate. It also shows that the detection

accuracy does not degrade when partial sensitive-data digests are used. In addition, these partial fingerprints fairly represent the fully set of data without any bias.

There are two technical challenges associated with network-based DLD detection. First, the DLD provider gains knowledge about the sensitive data when the traffic contains a leak. The challenge is how to restrict the degree of information that can be learned by the DLD provider in case of data leaks – the DLD provider has the access to the plaintext packet payload. The second challenge is how to make the detection noise-tolerant, for example, the intercepted packet payload may contain unrelated bytes or the sensitive data is truncated.

2. RELATED WORKS

Rabin fingerprint based on shingles was used previously for identifying similar spam messages in a collaborative setting, as well as collaborative worm containment, virus scan, Web template detection, and fragment detection.

This work fundamentally differs from the shingle based studies. Consider the new problem of data-leak detection in a unique outsourced setting where the DLD provider is not fully trusted. Such privacy requirement does not exist in the virus-scan paradigm, for the virus signatures are non-sensitive. In comparison, data-leak detection is more challenging because of the additional privacy requirement, which limits the amount of data that can be used during the detection and the amount of sensitive information gained by the DLD provider. In the meantime, the provider's detection accuracy cannot be compromised with partial digests based on the sensitive data. Fuzzy fingerprint method is new, and this work describes the first systematic solution to privacy preserving data-leak detection with convincing results.

Information leak through outbound web traffic was studied by Borders and Prakash [1]. Both works detect suspicious data flow on unencrypted network traffic. Their approach is based on the key observation that network traffic has high regularities and that information (e.g., header data) may be repeated. They proposed an elegant solution that detects any substantial increase in the amount of new information in the traffic.[10] Their anomaly-detection method detects deviations from normal data-flow scenarios, which are captured in rules. In comparison, this work inspects traffic for signatures of sensitive-data and does not require any assumption on the patterns of normal header fields or payload. Furthermore, solution provides privacy protection of the sensitive data against semi-honest DLD providers. This paper also gives performance evidences indicating the efficiency of the solution in practice.

The method of deep packet inspection is also widely used in network intrusion detection system. They focus on designing and implementing efficient string matching algorithms to handle short and flexible patterns in network traffic. However, NIDS is not designed for various kinds of sensitive data (e.g. long non-duplicated data), it may cause problems (e.g. large amount of states in an automata) in data leak detection scenarios. On the contrary, solution is not limited to very special types of sensitive data, and provides a unique privacy-preserving feature for service outsourcing. An alternative to this approach for privacy-preserving computation is to use cryptographic mechanisms.

Another category of approaches for data-leak detection is tracing and enforcing the sensitive data flows. The approaches include data flow and taint analysis [6], legal flow marking, and file-descriptor sharing enforcement [8]. These approaches are different from this paper because they

do not aim to provide an remote service. However, pure network-based solution cannot handle maliciously encrypted traffic [3], and these methods are complementary to our approach in detecting different forms (e.g., encrypted) of data leaks.

3. PROBLEM DEFINITION

According to current Statistics from various security organization research firms and government institutes suggest that there has been a rapid growth of data leak in past 8 years. There are various reasons for data leaks amongst which human errors is most endorsing of all. There are many ways in which this paper can have an audit trail verifying for data leak in networks, but still not all consider human errors as an important check factor, which makes them more prone to fail. Some of the common solutions include keeping a copy of sensitive data at the providers end and maintaining an audit trial for checking whether there is any data leakage in networks, which in turn notifies the organization about data leakage. All these methods in turn are prone to data attacks, as the keep copy of data while auditing. Also methods like deep packet analysis which searches for any relating data patterns. In this method, payloads of TCP/IP packets isanalyzed for any alter in data or any data pattern matching which may form sensitive data collection in network. This data is then compared with the threshold value, and if it exceeds the threshold value, the detection system alerts or notifies the organization.

From the detection perspective, a straightforward method is for the DLD provider to raise an alert if any sensitive fingerprint matches the fingerprints from the traffic. However, this approach has a privacy issue. If there is a data leak, there is a match between two fingerprints from sensitive data and network traffic. Then, the DLD provider learns the corresponding shingle, as it knows the content of the packet. Therefore, the central challenge is *to* prevent the DLD provider from learning the sensitive values even in data-leak scenarios, while allowing the provider to carry out the traffic inspection. This propose an efficient technique to address this problem. The main idea is to relax the comparison criteria by strategically introducing matching instances on the DLD provider's side without increasing false alarms for the data owner.

4. FUZZY FINGERPRINT METHOD

This paper describe the technical details of fuzzy fingerprint mechanism for privacy-preserving data-leak detection, by first introducing shingle and Rabin fingerprint, and then presenting randomization method for detection. The Rabin fingerprint scheme is a method for implementing fingerprints using polynomials over a finite field.

There are two players in this model: the organization (i.e. data owner) and the data-leak detection (DLD) provider.

- Organization owns the sensitive data and authorizes the DLD provider to inspect the network traffic from the organizational networks for anomalies, namely inadvertent data leak. However, the organization does not want to directly
  reveal the sensitive data to the provider.
- DLD provider inspects the network traffic for potential data leaks. The inspection can be performed offline without causing any real-time delay in routing the packets. However, the provider may attempt to gain knowledge about the sensitive data. This paper model

the DLD provider as a honest-but-curious adversary (aka semi-honest), who follows this protocols to carry out the operations, but may attempt to gain knowledge about the sensitive data.

The workflow in a network-based data-leak detection framework is as follows: DATA PRE-PROCESSING by the data owner, TRAFFIC PRE-PROCESSING AND DETECTION by the DLD provider, and ANALYSIS by the data owner. Data pre-processing is where the data owner takes the sensitive dataset and computes the corresponding set of digests. Traffic pre-processing and detection is where the DLD provider gathers network packets and inspects the content for data leaks. Analysis is where the data owner efficiently examines the alerts generated by the DLD provider, identifies and investigates the true leak instances and ignore false positives.

The privacy goal in our fuzzy fingerprint mechanism is to prevent the DLD provider from inferring the exact knowledge of the sensitive data; the DLD provider is given the fingerprints of sensitive data and the content of network traffic which may or may not contain data leak. In our model, this paper aim to hide the sensitive values among other nonsensitive values, so that the DLD provider is unable to pinpoint sensitive data among them even under data-leak scenarios. This paper define our privacy goal as follows, following the K-anonymity privacy definition in the relational databases

Our privacy goal is defined as follows. The DLD provider is given digests of sensitive data from the data owner and the content of network traffic to be examined. The DLD provider should not find out the exact value of a piece of sensitive data with more than 1 K probability, where K is an integer representing the number of all possible sensitive-data candidates that can be inferred by the DLD provider. This describe a novel fuzzy fingerprinting mechanism in the next section to improve the data protection against semi honest DLD provider, by utilizing simple and effective randomization technique in fingerprint generation. The privacy guarantee is much higher than 1 K when there is no leak in traffic, because the adversary's inference can only be done through brute-force guesses.This paper will propose the algorithmic steps with the help of below data flow diagram:

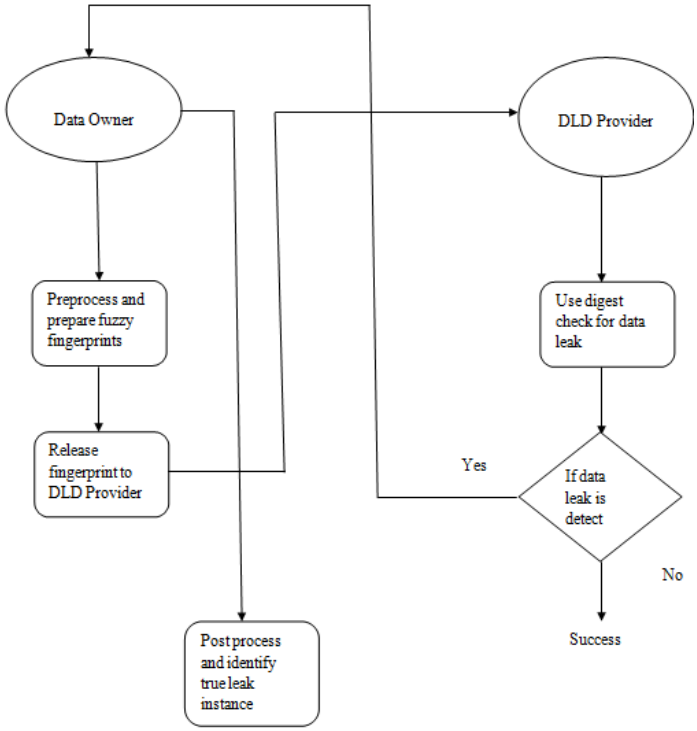


Figure 1. Data Flow Diagram

So, the above flow diagram can be explained as follows:

1. First the data owner will mark his set of sensitive data.
2. Secondly he will pre-compute all data and create a set of fuzzy fingerprint along with set of data digest.
3. He will add noise to the exposed digest, in order to assure that the semi honest provider does not gain complete knowledge about the sensitive data.
4. Then he will release the digest to the semi honest provider, to keep a track of any data leak detection in the network.
5. The DLD provider on receiving the digest, will start to check for any data leak using the digest.
6. If the provider finds any data leak in the current traffic network, he will notify it to the data owner.
7. The data owner on receiving the notification will post compute the data digest neglecting the noise he added, to check whether there was any data leakage in real time.

So, trying to give the data owner control rights i.e. he can decide what part of data to be revealed to the semi honest provider and which not to reveal. In this survey paper, propose a data leak detection framework which can be used as a semi honest provider in the network itself or can also be outsourced. In this system implementing a fuzzy finger print technique that is an additional security check parameter for data leakage method [2].

4.1 SHINGLES AND FINGERPRINTS

To achieve the privacy goal, the data owner generates a special type of digests, which call fuzzy fingerprints. Intuitively, the purpose of fuzzy fingerprints is to hide the true sensitive data in a crowd. It prevents the DLD provider from learning its exact value. The DLD provider obtains digests of sensitive data from the data owner. The data owner uses a sliding window and Rabin fingerprint algorithm to generate short and hard to-reverse (i.e., one-way) digests through the fast polynomial modulus operation. The sliding window generates small fragments of the processed data (sensitive data or network traffic), which preserves the local features of the data and provides the noise tolerance property. Rabin fingerprints [9] are computed as polynomial modulus operations, and can be implemented with fast XOR, shift, and table look-up operations. The Rabin fingerprint algorithm has a unique min-wise independence property, which supports fast random fingerprints selection (in uniform distribution) for partial fingerprints disclosure.

The shingle-and-fingerprint process is defined as follows. A sliding window is used to generate *q*-grams on an input binary string first. The fingerprints of *q*-grams are then computed.

A shingle (*q*-gram) is a fixed-size sequence of contiguous bytes. For example, the 3-gram shingle set of string abcdefgh consists of six elements {abc, bcd, cde, def, efg, fgh}. Local feature preservation is accomplished through the use of shingles. Therefore, this approach can tolerate sensitive data modification to some extent, e.g., inserted tags, small amount of character substitution, and lightly reformatted data. The use of shingles for finding duplicate web documents first appeared in [13] and [14].

This method proves to be faster than any another method and is based on one way computation of exposure of sensitive data. It gives the data owner rights of integrating data specific content securely to the DLD without actually exposing the sensitive data. So, this ensures that the semi honest provider has a very less amount of knowledge of the actual sensitive data and given provisions wherein individual can themselves mark their sensitive data and ask the admin of their local repository to check for any data leak. In the solution procedure, compute a method where the owner of the data contains a set of fingerprints or information digests of his own from the marked data, and can expose a small amount of part of the sensitive digest to the semi honest provider. The provider will then check for any data leak detection in that part of digest, where the digest is composed of real leaks and noise.

Using the min-wise independent property of Rabin fingerprint, the data owner can quickly disclose partial fuzzy fingerprints to the DLD provider. The purpose of partial disclosure is two-fold: *i)* to increase the scalability of the comparison in the DETECT operation, and *ii)* to reduce the exposure of data to the DLD provider for privacy. The method of partial release of sensitive data fingerprints is similar to the suppression technique in database anonymization.[11][12]

5. CONCLUSION

Preventing sensitive data from being compromised is an important and practical research problem. This paper proposed a fuzzy fingerprint framework and algorithms to realize privacy-preserving data-leak detection. Using special digests, the exposure of the sensitive data is kept to

a minimum during the detection. This paper described its application in the cloud computing environments, where the cloud provider naturally serves as the DLD provider. This paper defined privacy goal by quantifying and restricting the probability that the DLD provider identifies the exact value of the sensitive data. The extensive experiments validate the accuracy, privacy, and efficiency of the solutions.

REFERENCES

[1] X. Shu and D. Yao, "Data leak detection as a service," in Proc. 8th Int. Conf. Secur. Privacy Commun.Netw., 2012, pp. 222–240.

[2] Risk Based Security. (Feb. 2014). Data Breach Quick-View: An Executive's Guide to 2013 Data Breach Trends.[Online].Available:https://www.riskbasedsecurity.com/reports/2013-DataBreachQuickView.pdf, accessed Oct. 2014.

[3] Ponemon Institute. (May 2013). 2013 Cost of Data Breach Study: Global Analysis. [Online]. Available: https://www4.symantec.com/mktginfo/whitepaper/053013_GL_NA_WP_Ponemon-2013-Cost-of-a-Data-Breach-Report_daiNA_cta72382.pdf, accessed Oct. 2014.

[4] Identity Finder. Discover Sensitive Data Prevent Breaches DLP Data Loss Prevention. [Online]. Available: http://www.identityfinder.com/, accessed Oct. 2014.

[5] K. Borders and A. Prakash, "Quantifying information leaks in outbound web traffic," in Proc. 30th IEEE Symp. Secur. Privacy, May 2009, pp. 129–140.

[6] H. Yin, D. Song, M. Egele, C. Kruegel, and E. Kirda, "Panorama: Capturing system-wide information flow for malware detection and analysis," in Proc. 14th ACM Conf. Comput. Commun.Secur., 2007, pp. 116–127.

[7] K. Borders, E. V. Weele, B. Lau, and A. Prakash, "Protecting confidential data on personal computers with storage capsules," in Proc. 18th USENIX Secur. Symp., 2009, pp. 367–382.

[8] J. Kleinberg, C. H. Papadimitriou, and P. Raghavan, "On the value of private information," in Proc. 8th Conf. Theoretical Aspects Rationality Knowl., 2001, pp. 249–257.

[9] M. O. Rabin, "Fingerprinting by random polynomials," Dept. Math., Hebrew Univ. Jerusalem, Jerusalem, Israel, Tech. Rep. TR-15-81, 1981.

[10] S. Xu, "Collaborative attack vs. collaborative defense," in Collaborative Computing: Networking, Applications and Worksharing(Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering), vol. 10. Berlin, Germany: Springer- Verlag, 2009, pp. 217–228.

[11] G. Aggarwalet al., "Anonymizing tables," in Proc. 10th Int. Conf. Database Theory, 2005, pp. 246–258.

[12] R. Chen, B. C. M. Fung, N. Mohammed, B. C. Desai, and K. Wang,"Privacy-preserving trajectory data publishing by local suppression," Inf.Sci., vol. 231, pp. 83–97, May 2013.

[13] A. Z. Broder, "Some applications of Rabin's fingerprinting method," in Sequences II. New York, NY, USA: Springer-Verlag, 1993, pp. 143–152.

[14] A. Z. Broder, "Identifying and filtering near-duplicate documents," in Proc. 11th Annu. Symp.Combinat. Pattern Matching, 2000, pp. 1–10.

[15] GTB Technologies Inc. SaaS Content Control in the Cloud. [Online]. Available: http://www.gtbtechnologies.com/en/solutions/dlp_as a_service, accessed Oct. 2014.

[16] S. Geravand and M. Ahmadi, "Bloom filter applications in network security: A state-of-the-art survey," Comput. Netw., vol. 57, no. 18, pp. 4047–4064, Dec. 2013.

[17] B. Wang, S. Yu, W. Lou, and Y. T. Hou, "Privacy-preserving multikeyword fuzzy search over encrypted data in the cloud," in Proc. 33th IEEE Conf. Comput. Commun., Apr./May 2014, pp. 2112–2120.

AUTHORS

StaicyUlahannanis currently pursuing M.Techin Cyber Security in MBITS, Nellimattom. She completed her B. Tech. in Computer Science and engineering from MBITS, Nellimattom. Her areas of research are Network Security and Information Forensics

Roshni Jose is currently working as Assistant Professor in Department of Computer Science and Engineering in MBITS, Nellimattom. She received her B-Tech Degree in Computer Science and Engineering from College of Engineering, thodupuzha and M.Tech in Computer Science and engineering from Sathyabama Institute of Science & Technology. Her areas of research are Network Security and Computer Organisation.

INTENTIONAL BLANK

OVERALL PERFORMANCE EVALUATION OF ENGINEERING STUDENTS USING FUZZY LOGIC

Arya A Surya, Merin k kurian and Surekha Mariam Varghese

Department of Computer Science and Engineering, M.A College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

In this paper we use Fuzzy logic instead of the classical methods of performance evaluation of the students. In classical methods mathematical calculations are being used. This performance evaluation is done for the engineering students mainly. The overall evaluation cannot be just based on the total marks he/she obtained in various subjects. A complete engineer is the one who is skilled in lab experiments , theory papers as well as in projects. So Through this paper we put forward a fuzzy method for the same. Even though this method requires additional software this is very helpful for teachers to evaluate a student. This method is flexible as they can change the membership function and also its value.

KEYWORDS

Fuzzy Logic,Fuzzy Expert Systems, Membership Function and Performance Evaluation ,Project, Theory exam ,Lab exam.

1.INTRODUCTION

Evaluation of performance of the students is usually expressed numerically, based on examination results. Classical evaluation therefore consists of a judgment based on the total percentage or the score he/she obtained in various subjects, laboratory exams, projects etc. Measurement and evaluating are inspirable and important parts of the educational process. Using this method the success or failure is based on a threshold of the total marks. This is not actually a efficient method to evaluate skill of a student. For example an engineering student who is scoring 80 in theories,89 in project whereas just 40 in lab exams will have a score of 70 in classical method. If threshold is 65 we consider him as successful but this is not correct as he is not having skill to perform in projects.

The fuzzy logic tool was introduced in 1965 by LotfiZadh, which is a mathematical tool for dealing. It offers a soft computing partnership which is the important concept of computing with words. It provides a technique to deal with imprecision and information granularity. The fuzzy theory provides a mechanism for representing linguistic constructs such as many, low, medium, often few. In general, the fuzzy logic provides an inference structure that enables appropriate human reasoning capabilities. Fuzzy logic theory emerged during the twentieth century and, by the beginning of the twenty-first century, was predicted to be applied extensively in many fields (Altrock, 1995). One of the applications of the fuzzy logic theory is the measurement and evaluation in education. In this context, the aim of this paper is to define the "impact of the fuzzy logic theory on the measurement of student's performance" (Semerci, 2004). The selection of students is based on their scores the their performance should be evaluated based on their overall skill.

2. METHODOLOGY

2.1 THE AIM OF THE STUDY

The aim of the study is to determine engineering students' performance using a fuzzy logic model instead of classical evaluation methods.

The study aimed to

- Find the performance using classical method
- Find the performance using fuzzy logic method
- Compare both methods.

2.2 FUZZY LOGIC

The idea of fuzzy logic was first advanced by Dr.LotfiZadeh of the University of California at Berkeley in the 1960s. Dr.Zadeh was working on the problem of computer understanding of natural language. Natural language (like most other activities in life and indeed the universe) is not easily translated into the absolute terms of 0 and 1. (Whether everything is ultimately describable in binary terms is a philosophical question worth pursuing, but in practice much data we might want to feed a computer is in some state in between and so, frequently, are the results of computing.)uzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which the modern computer is based[1].

Fuzzy Logic incorporates a simple, rule-based IF X AND Y THEN Z approach to a solving control problem rather than attempting to model a system mathematically. The Fuzzy Logic model is empirically-based, relying on an operator's experience rather than their technical understanding of the system. For example, rather than dealing with temperature control in terms such as "SP =500F", "T <1000F", or "210C <TEMP <220C", terms like "IF (process is too cool) AND (process is getting colder) THEN (add heat to the process)" or "IF (process is too hot) AND (process is heating rapidly) THEN (cool the process quickly)" are used. These terms are imprecise and yet very descriptive o what
Must actually happen [2].

One of the famous applicationx of fuzzy logic and fuzzy set theory is Fuzzy inference system (FIS) (Guillaume, 2001). FIS are knowledge-based or rule-based systems that contain descriptive if-then rules created from human knowledge and experience (Kharola and Gupta, 2014)[4]. A basic fuzzy architecture consists of three components fuzzifier, FIS and defuzzifier. Fuzzifier maps crisp numbers into fuzzy sets whereas the defuzzifier maps output sets into crisp numbers. The FIS represents the core of fuzzy logic controllers (FLC's). It is built of rule-base and data-base, which constitute the knowledge base and inference engine. A view of basic architecture of fuzzy system is shown in figure1

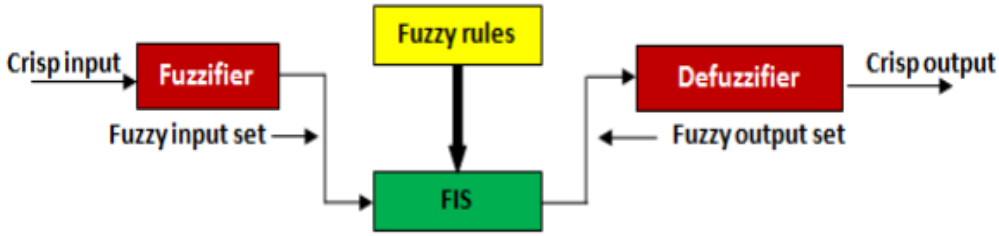


Figure 1: Basic architecture of fuzzy system

2.3 STEPS IN EVALUATION

Performance evaluation involves following steps.

- Fuzzification of input lab exam, theory exam, project results and output performance value.
- Determination of application rules and inference method.
- Defuzzification of performance value

Students appear for three types of exams, so there are three input variables. The output variable is the performance value, which is determind by fuzzy logic (Figure 1).

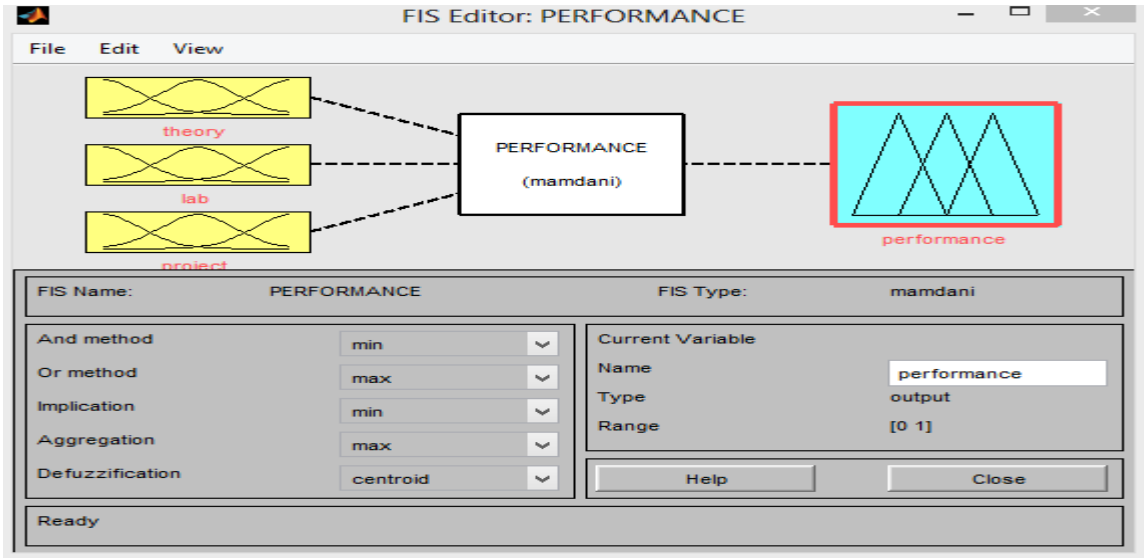


Figure 2: FIS editor

2.4. INPUT VARIABLES

Fuzzification of exam results was carried out using input variables and their membership functions of fuzzy sets. Each student has three exam results, each of which form input variables of the fuzzy logic system. Lab and theory input variable has five triangle membership functions and project input variable has three membership functions[2]. The fuzzy set of input variables lab, theory and project is shown Table 1 and Table 2.

Table1:Fuzzy set of input variables lab and theory

Linguistic Expression	Symbol	Interval
Very Low	VL	(0, 0, 25)
Low	L	(0, 25, 50)
Average	A	(25, 50, 75)
High	H	(50, 75, 100)

Table2:Fuzzy set of input variables Project

Linguistic Expression	Symbol	Interval
Poor	P	(0, 0, 40)
Good	G	(30, 55, 80)
Excellent	E	(70, 100,100)

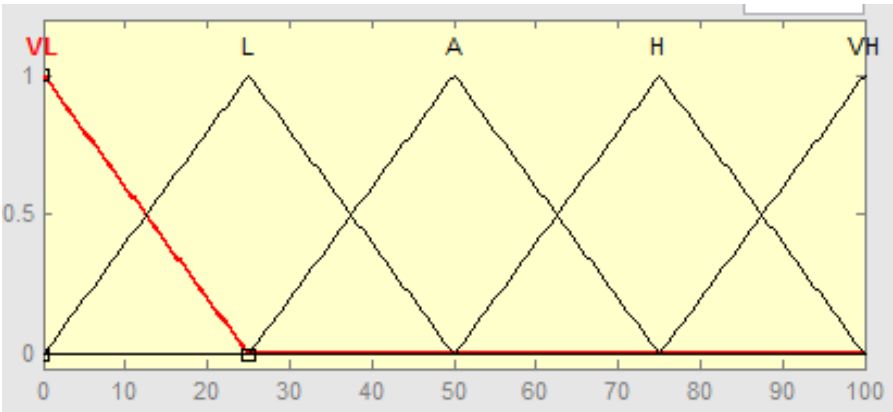


Figure 3: Membership function of input variables lab and theory

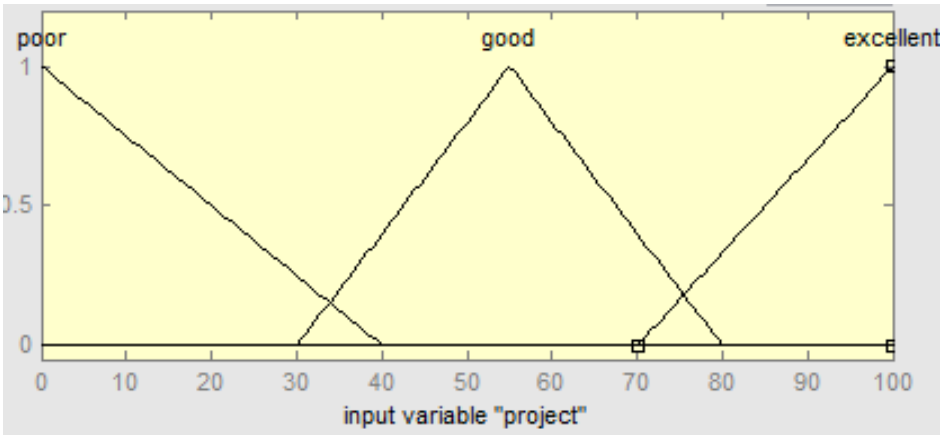


Figure4: Membership function of input variables project

2.5. OUTPUT VARIABLES

The output variable, which is the performance value, is entitled "Performance" and has five membership functions. For reasons of convenience within the application, a value range between 0 and 1 was chosen (Table 3 and Figure 4).

Table3. Fuzzy set of output variable

Linguistic Expression	Symbol	Interval
Very Unsuccessful	VU	(0, 0, 0.25)
Unsuccessful	U	(0 0 25 0 5)
Average	A	(0.25, 0.5, 0.75)
Successful	S	(0.5, 0.75, 1)
Very Successful	VS	(0.75, 1, 1)

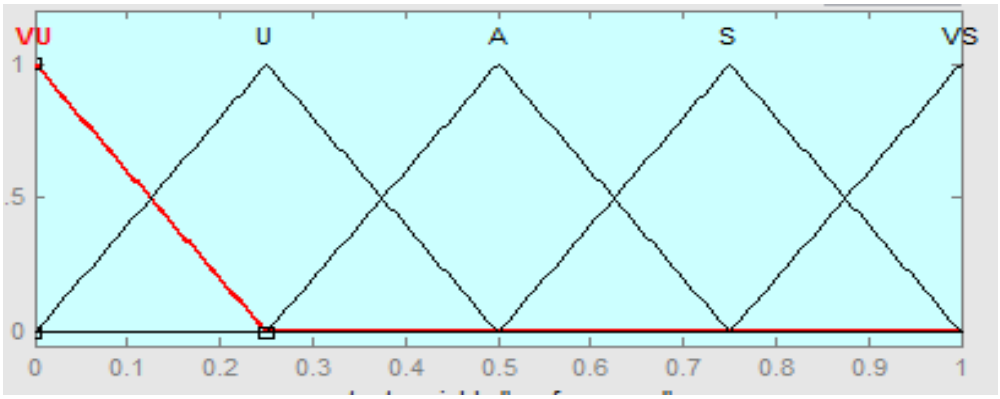


Figure5:Membership function of output variable

2.6. RULES AND INFERENCE

The rules determine input and output membership functions that will be used in inference process. These rules are linguistic and also are entitled "If-Then" rules (Altrock, 1995; Semerci, 2004).

1. If (theory is VL) and (lab is VL) and (project is poor) then (performance is VU)
2. If (theory is VL) and (lab is VL) and (project is good) then (performance is VU)
3. If (theory is VL) and (lab is VL) and (project is excellent) then (performance is U)
4. If (theory is VL) and (lab is L) and (project is poor) then (performance is VU)
5. If (theory is VL) and (lab is L) and (project is good) then (performance is U)
6. If (theory is VL) and (lab is L) and (project is excellent) then (performance is U)
7. If (theory is VL) and (lab is A) and (project is poor) then (performance is U)
8. If (theory is L) and (lab is A) and (project is excellent) then (performance is A)
9. If (theory is H) and (lab is VL) and (project is poor) then (performance is U)
10. If (theory is H) and (lab is VL) and (project is good) then (performance is U)
11. If (theory is H) and (lab is VL) and (project is excellent) then (performance is A)

12. If (theory is H) and (lab is L) and (project is poor) then (performance is U)
13. If (theory is H) and (lab is L) and (project is good) then (performance is A)
14. If (theory is H) and (lab is L) and (project is excellent) then (performance is A)
15. If (theory is H) and (lab is A) and (project is poor) then (performance is U)
16. If (theory is H) and (lab is A) and (project is good) then (performance is A)
17. If (theory is H) and (lab is A) and (project is excellent) then (performance is S)
18. If (theory is VH) and (lab is L) and (project is poor) then (performance is U)
19. If (theory is VH) and (lab is VH) and (project is poor) then (performance is U)
20. If (theory is VH) and (lab is VH) and (project is good) then (performance is VS)

In case of several rules are active for the same output membership function, it is necessary that only one membership value is chosen. This process is entitled "fuzzy decision" or "fuzzy inference". Several authors, including Mamdami, Takagi-Surgeno and Zadeh have developed a range of techniques for fuzzy decision-making and fuzzy inference. The present study uses the method proposed by Mamdami, shown in Equation (1) (Semerci, 2004; Zadeh, 1965; Rutkowski, 2004).

$$\mu_c(y) = \max(\min(\mu_A(\text{input}(i)), \mu_B(\text{input}(j)))) \quad \text{.....} \quad (1)$$

This expression determines an output membership function value for each active rule. When one rule is active, an AND operation is applied between inputs. The smaller input value is chosen and its membership value is determined as membership value of the output for that rule. This method is repeated, so that output membership functions are determined for each rule. To sum up, graphically AND (min) operations are applied between inputs and OR (max) operations are between outputs[7].

3. RESULT

On the application of fuzzy logic we get following results. The Figure5 shows Active rules and performance value for corresponding inputs.

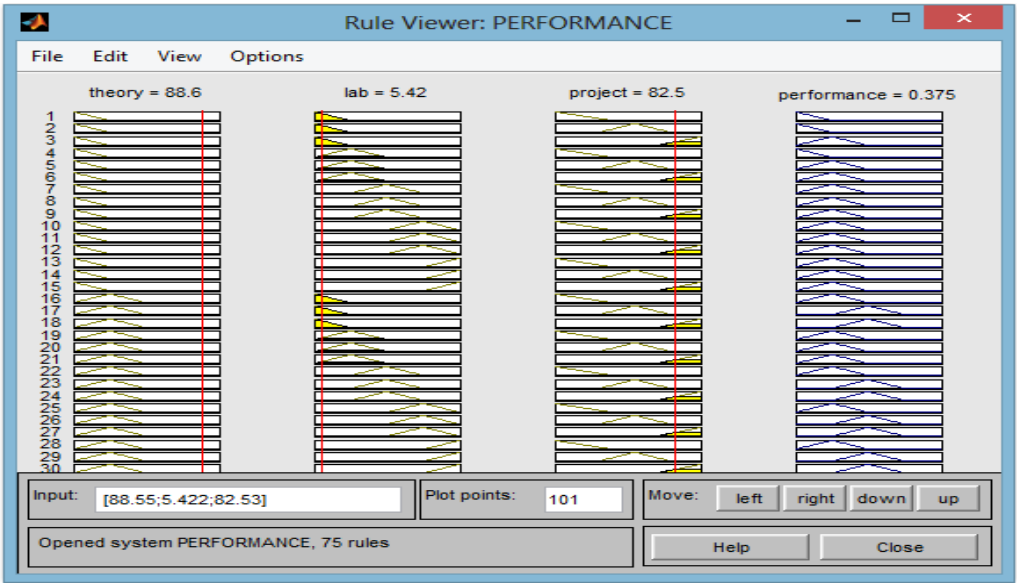


Figure6: Active rules and performance value

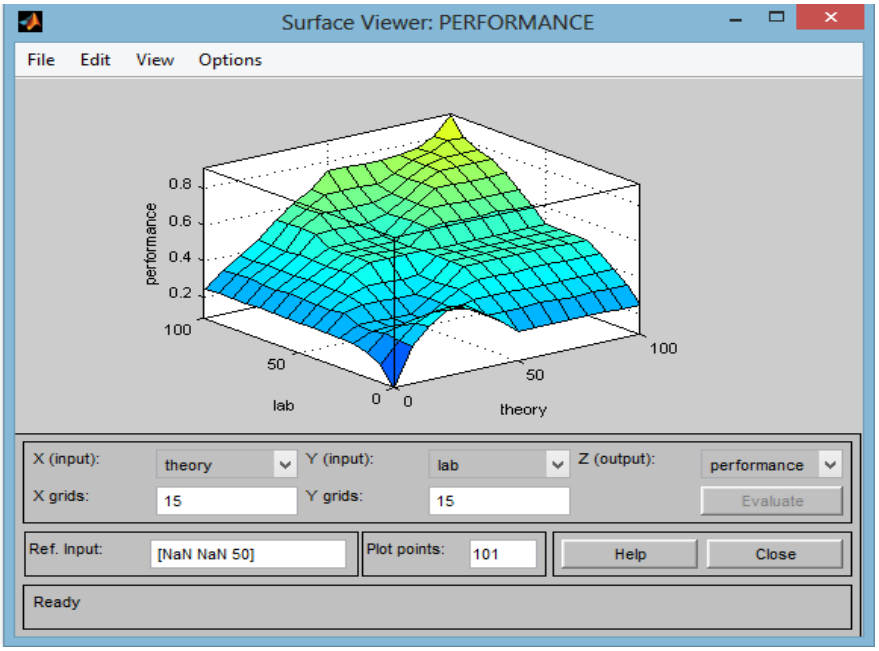


Figure7: Surface View

Table4: Comparison of Performance Value obtained in Classical and Fuzzy Logic Method.

THEORY	LAB	PROJECT	PERFORMANCE Using Classical method	PERFORMANCE Using Fuzzy method
44	88.6	78.9	0.705	0.625
15.1	23.5	46.4	0.283	0.278
68.1	52.4	71.7	0.641	0.554
93.4	97	89.8	0.934	0.855
93.4	59.6	52.4	0.685	0.602
47.6	25.9	92.2	0.552	0.466
47.6	58.4	57.2	0.544	0.5
88.6	100	92.2	0.936	0.907
88.6	5.42	82.5	0.590	0.375

4. CONCLUSION

In this paper, we have proposed Dynamic Fuzzy Expert system for modelling students' academic performance evaluation based Fuzzy logic. When the results are evaluated, a difference in outcomes is seen between the classical method and the proposed fuzzy logic method. While the classical method adheres to a constant mathematical rule, evaluation with fuzzy logic has great flexibility. At the application stage, course-conveners can edit rules and membership functions to obtain various performance values but it is important that the same rules and membership functions are used for all students taking the same lesson. It is also important for the students to understand the assessment criteria before taking exams. For this reason, members of the

educational board should communicate with each other and come to an agreement on rules, membership functions and any other criteria.

5. REFERENCES

[1] K. Mankad, P.S. Sajja and R. Akerkar, "Evolving Rules Using Genetic Fuzzy Approach: An educational case study", International Journal on Soft Computing. 2(1), pp. 35-46, 2011.

[2] R. Biswas, "An Application of fuzzy sets in Students' Evaluation", Fuzzy sets and System, ELSEVIER, pp. 187-194, 1995.

[3] L.A. Zadeh, "Fuzzy sets. Information and Control", 8, pp. 338-354, 1965.

[4] M.S. Upadhyay, "Fuzzy Logic Based of Performance of Students in College", Journal of Computer Applications (JCA), 5(1), pp. 6-9, 2012.

[5] H. White, "Learning in Artificial Neural Networks: A Statistical Perspective", Neural Computation, 1, pp. 425-464, 1989.

[6] J.C. Giarratano and G. Riley, "Expert System: Principles and Programming", Fourth ed., PWS Publishing Com. Boston, MA, USA, 2005.

[7] M. Schneider, G. Langholz, A. Kandel and G. Chew, "Fuzzy Expert System Tools", Jhon Willy and Sons, USA, 1996.

[8] GokhanGokmen et al. / Procedia Social and Behavioral Sciences 2 (2010) 902–909

AUTHORS

Arya A Surya is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering, Kothamangalam. She completed her B.Tech from AdiShankara institute of engineering and technology, Kalady. Her areas of research are DataMining and Machine Learning.

Merin K Kurian is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering, Kothamangalam. She completed her B.Tech from AdiShankara Institute of engineerring and technology, Kalady. Her areas of research are DataMining and Machine Learning.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B -Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems, Machine Learning and Distributed Computing. She has published 17 papers in international journals and international conference proceedings. She has been in the chair and reviewer for many international conferences and journals.

SOFTWARE TOOL FOR TRANSLATING PSEUDOCODE TO A PROGRAMMING LANGUAGE

Amal M R , Jamsheedh C V and Linda Sara Mathew

Department of Computer Science and Engineering, M.A College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

Pseudocode is an artificial and informal language that helps programmers to develop algorithms. In this paper a software tool is described, for translating the pseudocode into a particular programming language. This tool takes the pseudocode as input, compiles it and translates it to a concrete programming language. The scope of the tool is very much wide as we can extend it to a universal programming tool which produces any of the specified programming language from a given pseudocode. Here we present the solution for translating the pseudocode to a programming language by implementing the stages of a compiler.

KEYWORDS

Compiler, Pseudocode to Source code, Pseudocode Compiler, c, c++

1. INTRODUCTION

Generally a compiler is treated as a single unit that maps a source code into a semantically equivalent target program [1]. If we are analysing a little, we see that there are mainly two phases in this mapping: analysis and synthesis. The analysis phase breaks up the source code into constituent parts and imposes a grammatical structure on them. It then uses this structure to create an intermediate representation of the source code. If the analysis phase detects that the source code is either syntactically weak or semantically unsound, then it must provide informative messages. The analysis phase also collects information about the source code and stores it in a data structure called a symbol table, which is passed along with the intermediate representation to the synthesis phase. The synthesis phase constructs the target program from the intermediate representation and the information in the symbol table [2], [3]. The analysis phase is often called the front end of the compiler; the synthesis phase is the back end.

Compilation process operates as a sequence of phases, each of which transforms one representation of the source program to another. Compilers have a machine-independent optimization phase between the front end and the back end. The purpose of this optimization phase is to perform transformations on the intermediate representation; so that the backend can produce a better target program than it would have otherwise produced from an un-optimized intermediate representation.

In this paper, it is intended to produce a user specified programming language from pseudocode. This tool requires a single pseudocode and it can produce the programming language that is specified by the user. Its significance is that it can be extended to a universal programming tool that can produce any specified programming language from pseudocode.

2. COMPILING PSEUDOCODE

The process of compiling the pseudocode consists of certain analysis and operations that has to be performed on it.

2.1. LEXICAL ANALYSER

The Lexical Analyser module analyses the pseudocode submitted by the user by using the transition analysis. The keywords, Identifiers and tokens are identified from the given input.

2.2. SYNTAX ANALYSER

The Syntax Analyser module creates the Context Free Grammar from the pseudocode submitted .The resultant grammar is then used for creation of the parse tree. Then pre order traversal is done to obtain the meaning of the syntax.

2.3. SEMANTIC ANALYSER

The semantic Analyser uses the syntax tree and the information in the symbol table to check the source program for semantic consistency with the language definition. It also gathers type information.

2.4. INTERMEDIATE CODE GENERATOR

In this module, an intermediate code is generated from the input pseudocode. The generated intermediate code is that code which is used for the conversion to any other languages.

2.5. INTERMEDIATE CODE OPTIMIZER

In this module, code optimization is applied on the intermediate code generated. The generated optimized intermediate code is that code which is used for mapping to the concrete languages.

2.6. CODE GENERATOR

The optimized intermediate code is converted into the required programming language in this module. The result might be obtained in the language selected by the user.

2.7. LIBRARY FILE MANAGER

In this module the administrator manages the library files of the target language and also manipulates files in the library package.

3. PROBLEM STATEMENT

3.1. INTERPRET THE PSEUDOCODE

The main task is to identify and interpret the pseudocode given by the user. Each user has his own style of presentation, variation in using keywords or terms (E.g.: - Sum, Add, etc. to find sum of

two numbers), structure of sentence, etc. So initial task is make the tool capable of develop the tool is to interpret and identify the right meaning of each line of the pseudocode.

3.2. TRANSLATE THE PSEUDOCODE INTO PROGRAMMING LANGUAGE

The second step involves the translation of the interpreted pseudocode into programming language. User can specify output in any of the available programming languages. So the tool must be able to support all the available programming language features (in this paper we are concentrating on C and C++ only). That is it must support the object oriented concepts, forms and so on.

4. METHODOLOGY

4.1. LEXICAL ANALYSIS

The first phase of the software tool is called lexical analysis or scanning. The lexical analyser reads the stream of characters making up the pseudocode and groups the characters into meaningful sequences called lexemes. For each lexeme, the lexical analyser produces as output a token of the form {token- name, attribute-value} that it passes on to the subsequent phase, syntax analysis. In the token, the first component token- name is an abstract symbol that is used during syntax analysis, and the second component attribute-value points to an entry in the symbol table for this token. Information from the symbol-table entry 'is needed for semantic analysis and code generation. For example, suppose a source program contains the declare statement[1],[2].

Declare an integer variable called sum# (1.1)

The characters in this assignment could be grouped into the following lexemes and mapped into the following tokens passed on to the syntax analyser:

1. *Declare a***,** is a lexeme that would be mapped into a token (Declare, 59), where Declare is a keyword and 59 points to the symbol table entry for position.
2. *Integer,* is a lexeme that would be mapped into a token (Integer, 112), where Integer is a keyword and 112 points to the symbol table entry for position.
3. *Variable,* is a lexeme that would be mapped into a token (Variable, 179), where Variable is a keyword and 179 points to the symbol table entry for position.
4. *Called,* is a lexeme that would be mapped into a token (Called, 340), where Called is a keyword and 340 points to the symbol table entry for position.
5. *Sum,*is a lexeme that would be mapped into a token (sum, 740), where sum is an identifier and 740 points to the symbol table entry for position.
   (Blanks separating the lexemes would be discarded by the lexical analyser.)

4.2. SYNTAX ANALYSIS

The second phase of the compiler is syntax analysis or parsing. The parser uses the first components of the tokens produced by the lexical analyser to create a tree-like intermediate representation that depicts the grammatical structure of the token stream. A typical representation is a syntax tree in which each interior node represents an operation and the children of the node represent the arguments of the operation [12], [13]. The syntax of programming language constructs can be specified by context-free grammars or BNF (Backus-Naur Form) notation; Grammars offer significant benefits for both language designers and compiler writers.A grammar gives a precise, yet easy-to-understand, syntactic specification of a programming language. From

certain classes of grammars, we can construct automatically an efficient parser that determines the syntactic structure of a source program. As a side benefit, the parser-construction process can reveal syntactic ambiguity and trouble spots that might have slipped through the initial design phase of a language. The structure imparted to a language by a properly designed grammar is useful for translating source programs into correct object code and for detecting errors. A grammar allows a language to be evolved or developed iteratively, by adding new constructs to perform new tasks. These new constructs can be integrated more easily into an implementation that follows the grammatical structure of the language.

4.3. CONTEXT-FREE GRAMMARS

Grammars were introduced to systematically describe the syntax of programming language constructs like expressions and statements. Using a syntactic variable 'Stmt' to denote statements and variable 'expr' to denote expressions, In particular, the notion of derivations is very helpful for discussing the order in which productions are applied during parsing. The Formal Definition of a Context-Free Grammar (grammar for short) consists of terminals, non-terminals, a start symbol, and productions.

1. Terminals are the basic symbols from which strings are formed. The term "token name" is a synonym for "terminal" and frequently we will use the word "token" for terminal when it is clear that we are talking about just the token name. We assume that the terminals are the first components of the tokens output by the lexical analyser.
2. Non terminals are syntactic variables that denote sets of strings. The set of string denoted by non-terminals helps to define the language generated by the grammar. Non terminals impose a hierarchical structure on the language that is the key to syntax analysis and translation.
3. In a grammar, one nonterminal is distinguished as the start symbol, and the set of strings it denotes is the language generated by the grammar. Conventionally, the productions for the start symbol are listed first.
4. The productions of a grammar specify the manner in which the terminals and non-terminals can be combined to form strings. Each production consists of:

a) A nonterminal called the head or left side of the production; this production defines some of the strings denoted by the head.
b) The symbol --+. Sometimes:: = has been used in place of the arrow.
c) A body or right side consisting of zero or more terminals and non-terminals.

The Context Free Grammar generated from 1.1 by the Software tool is

Stmt -> declare_an <DataType> variable Called <Identifier> (1.2)
DataType->integer
Identifier->sum

4.4. SEMANTIC ANALYSIS

The semantic analyser uses the syntax tree and the information in the symbol table to check the source program for semantic consistency with the language definition. It also gathers type information and saves it in either the syntax tree or the symbol table, for subsequent use during intermediate-code generation An important part of semantic analysis is type checking, where the compiler checks that each operator has matching operands. For example, many programming language definitions require an array index to be an integer; the compiler must report an error if a floating-point number is used to index an array .The language specification may permit some type

conversions called coercions. For example, a binary arithmetic operator may be applied to either a pair of integers or to a pair of floating-point numbers. If the operator is applied to a floating-point number and an integer, the compiler may convert or coerce the integer into a floating-point number. The Parse Tree generated from 1.2 by the Software tool is by array representation as follows,

The pre order traversal:

Stmt-> declare_an DataType integer variable Called Identifier sum (1.3)

4.5. INTERMEDIATE CODE GENERATION

In the process of translating a source program into target code, a compiler may construct one or more intermediate representations, which can have a variety of forms. Syntax trees are a form of intermediate representation; they are commonly used during syntax and semantic analysis. After syntax and semantic analysis of the source program, many compilers generate an explicit low-level or machine-like intermediate representation, which we can think of as a program for an abstract machine[10],[11]. This intermediate representation should have two important properties: it should be easy to produce and it should be easy to translate into the target machine. In our software tool intermediate code is generated to convert the code to various languages from single pseudocode.

The intermediate code for 1.3 is as follows:*149 i780 300o (1.4)*

4.6. CODE GENERATION

The code generator takes as input an intermediate representation of the source program and maps it into the target language. If the target language is machine Code, registers or memory locations are selected for each of the variables used by the program. Then, the intermediate instructions are translated into sequences of machine instructions that perform the same task.

The resultant program code for 1.4 is as follows:*int sum; (1.5)*

5. SCHEMA DESCRIPTION

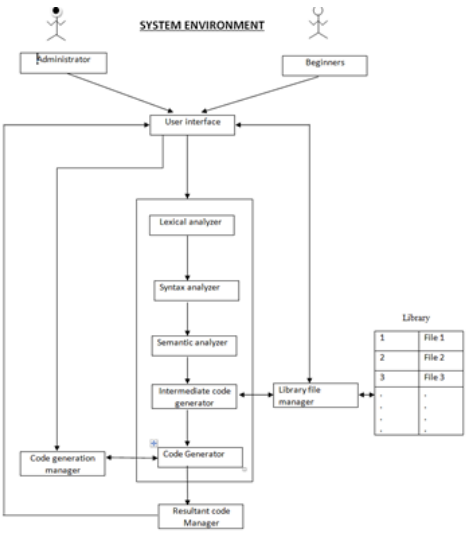


Fig:-1 System environment of the proposed software tool

The proposed software tool consists of several modules which are used to process the input

4.1. INPUT DESIGN

This is a process of converting user inputs into computer based formats. The data is fed into system using simple interactive forms. The forms have been supplied with messages so that user can enter data without facing any difficulty. The data is validated wherever it requires in the project. This ensures that only the correct data have been incorporated into the system. It also includes determining the recording media methods of input, speed of capture and entry into the system. The input design or user interface design is very important for any application. The interface design defines how the software communicates with in itself, to system that interpreted with it and with humans who use.

The main objectives that guide input design are as follows:

User friendly code editor- Providing line numbers and shaded graphical rows to easily identify each line of code.

Arise that lead to processing delays- Input is designed so that it does not lead to bottlenecks and thus avoid processing delays.

Dynamic check for errors in data-errors in data can lead to delays. Input design should be such that the data being entered should be free from error to the maximum possible limit.

Avoided extra steps-more the number of steps more is the chance of an error. Thus the number of steps is kept to a minimum possible.

Kept the process simple-the process should be kept as simple as possible to avoid errors.

4.2. OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In the output design, it is determined how the information is to be displayed for immediate need and also the hard copy output. The output design should be understandable to the user and it must offer great convenience. The output of the proposed software tool is designed as opening the text file containing the translated code.

The main objectives that guide the output design are as follows:

a. User can copy the code and run it in any of the IDE available.
b. Since the output is written in a standard text file, the user can directly call the file in the host program.
c. User can add some more code to the existing output and edit it easily.

4.3. DATA STRUCTURES USED

4.3.1. DATA BANK, TOKEN AND TOKEN ID

Here a table with all the tokens and there ID codes used in lexical analysis are tabulated. These tokens and there IDs are stored using Hash Table while implementing.

4.3.2. LIBRARY FILE, FOR A PARTICULAR PROGRAMMING LANGUAGE:

This is the data in the library file stored in the Library in the software tool. The file consists of all the keywords and header files in the language.
For example:-

Library File:For 'C':-

It includes the data in the library file stored in the Library of the software tool for all the keywords and header files in the language 'C'.

6. EXPERIMENT ANALYSIS

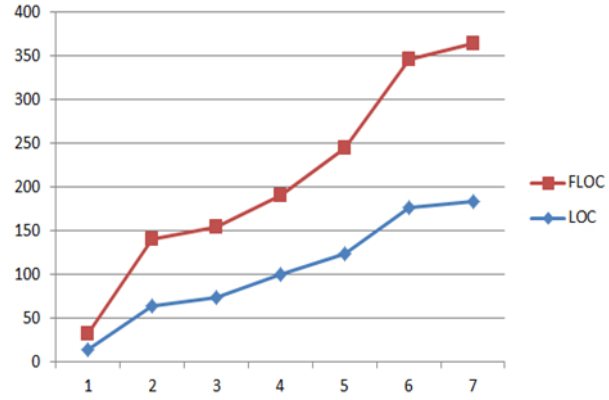


Figure 1. Comparison of Final Lines of Code(FLOC)and Lines of code(LOC)

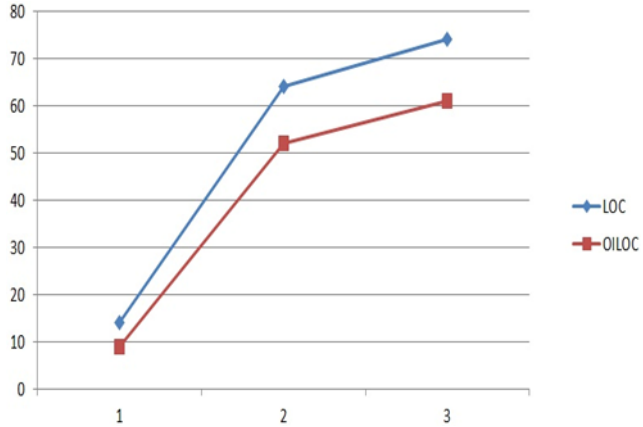


Figure 2. Comparison of Lines of Code(LOC)and Optimized Intermediate Lines of code(LOC)

***Analysis*:** The graphs are plotted with the Lines of Code (LOC) against the number of experiments. From the plots, it is clear that the initial LOC of the pseudocode given by the user is reduced proportionally in the optimized intermediate generated codes (OILOC).Then the final

LOC of the generated code is comparatively larger in proportion of the LOC of the pseudocode (see fig. 1).This measures indicates the efficiency of the tool in the generation of the code of the specified programming language. This measures depends on the efficiency and compatibility of the new developed tool.

7. CONCLUSIONS

This paper is focused on providing a user friendly environment for the beginners in programming. They can easily build a code in specified language from a pseudocode without considering the factor of knowledge about the syntax of the particular language. A beginner level programmer familiar with writing pseudocode can implement his logic in any particular language, simply by using this tool. The main advantage of this tool is that, user can build program code in any language from a single pseudocode. For a beginner in programming, it is difficult to learn the syntax of a particular language at the very first time. The user can implement his logic in a pseudocode and the pseudocode required for this software tool requires simple syntax. A formulated pseudocode is simple to be generated by a beginner. Then this pseudocode is simply submitted to the text area in our tool. Then specify the language of the output required. Then after processing he will get the resultant programming code in a file, which is much simpler with user friendly interface. Then the resultant code can be executed in its programming platform. The library files in the software tool can be manipulated to add more syntaxes into the database. Future versions can be built with giving support to more languages. We can develop this software tool to a universal programming tool, which can be used to build programming code in any of the programming language, from simple, single pseudocode generated by the user. It reduces the user's overhead to be known about the syntax of various languages.

REFERENCES

[1] G Alfred V.Aho,Monica S.Lam,Ravi Sethi,Jeffrey D.Ullman,Compilers Principles,Techniques and Tools, Second edition 2007

[2] Allen Holub, "Compiler Design in C", Prentice Hall of India, 1993.

[3] Kenneth C Louden, "Compiler Construction Principles and Practice",Cenage Learning Indian Edition..

[4] V Raghavan, "Priniples of Compiler Design",Tata McGraw Hill,India, 2010

[5] Arthur B. Pyster, "Compiler design and construction: tools and techniques with C and Pascal", 2nd Edition, Van Nostrand Reinhold Co. New York, NY, USA.

[6] D M Dhamdhare, System programming and operating system, Tata McGraw Hill & Company

[7] Tremblay and Sorenson, The Theory and Practice of Compiler Writing - Tata McGraw Hill & Company.

[8] Steven S. Muchnick, "Advanced Compiler Design & plementation", Morgan Kaufmann Pulishers, 2000.

[9] Dhamdhere, "System Programming & Operating Systems", 2nd edition, Tata McGraw Hill, India.

[10] John Hopcroft, Rajeev Motwani & Jeffry Ullman: Introduction to Automata Theory Languages & Computation , Pearson Edn.

[11] Raymond Greenlaw,H. James Hoover, Fundamentals of Theory of Computation,Elsevier,Gurgaon,Haryana,2009

[12] John C Martin, Introducing to languages and The Theory of Computation, 3rd Edition, Tata McGraw Hill,New Delhi,2010

[13] Kamala Krithivasan, Rama R, Introduction to Formal Languages,Automata Theory and Computation, Pearson Education Asia,2009.

[14] Rajesh K. Shukla, Theory of Computation, Cengage Learning, New Delhi,2009.

[15] K V N Sunitha, N Kalyani: Formal Languages and Automata Theory, Tata McGraw Hill,New Delhi,2010.

[16] S. P. Eugene Xavier, Theory of Automata Formal Language &Computation,New Age International, New Delhi ,2004.

[17] K.L.P. Mishra, N. Chandrashekharan , Theory of Computer Science , Prentice Hall of India.

[18] Michael Sipser, Introduction to the Theory of Computation, Cengage Learning,New Delhi,2007.
[19] Harry R Lewis, Christos H Papadimitriou, Elements of the theory of computation, Pearson Education Asia.
[20] Bernard M Moret: The Theory of Computation, Pearson Education.
[21] Rajendra Kumar,Theory of Automata Language & Computation,Tata McGraw Hill,New Delhi,2010.
[22] Wayne Goddard, Introducing Theory of Computation, Jones & Bartlett India,New Delhi2010.

AUTHORS

Amal M R is currently pursuing M.Tech in Computer Science and Engineering Mar Athanasius College of Engineering, Kothamangalam. He completed his B.Tech from Lourdes Matha College of Science and Technology Thiruvananthapuram. His areas of research are Compiler and Cloud Computing.

Jamsheedh C V is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering, Kothamangalam. He completed his B.Tech from Govt. Engineering College Idukki. His areas of research are Networking and Cloud Computing

Linda Sara Mathew received her B.Tech degree in Computer Science and Engineering from Mar Athanasius College of Engineering ,Kothamangaam,Kerala in 2002 and ME degree in Computer Science And Engineering Coimbatore in 2011. She is currently, working as Assistant Professor, with Department of Computer Science and Engineering in Mar Athanasius College of Engineering, Kothamangalam and has a teaching experience of 8 years. Her area of interests include digital signal processing, Image Processing and Soft Computing.

INTENTIONAL BLANK

A COMPARATIVE STUDY ON IMAGE COMPRESSION USING HALFTONING BASED BLOCK TRUNCATION CODING FOR COLOR IMAGE

Meharban M.S[1]and Priya S[2]

[1]M.Tech Student, Dept. of Computer Science, Model Engineering College
[2]Associate Professor, Dept. of Computer Science, Model Engineering College

ABSTRACT

In this paper scrutinizes image compression using Halftoning Based Block Truncation Coding for color image. Many algorithms were selected likely the original Block Truncation coding, Ordered Dither Block Truncation Coding, Error Diffusion Block Truncation Coding , and Dot Diffused Block Truncation Coding. These above techniques are divided image into non overlapping blocks. BTC acts as the basic compression technique but it exhibits two disadvantages such as the false contour and blocking effect. Hence halftoning based block truncation coding (HBTC) is used to overcome the two issues. Objective measures are used to evaluate the image degree of excellence such as Peak Signal to Noise Ratio, Mean Square Error, Structural Similarity Index and Compression Ratio. At the end, conclusions have shown that the Dot Diffused Block Truncation Coding algorithm outperforms the Block Truncation Coding as well as Error Diffusion Block Truncation Coding.

KEYWORDS

Halftoning, Image Compression, Block Truncation Coding (BTC), Error Diffusion, Dot Diffusion

1. INTRODUCTION

In recent years, the development of multimedia product is rapidly growing which contributes to insufficient bandwidth of network and storage. Thus the theory of image compression gains more importance for reducing the storage space and transmission bandwidth needed. Digital halftoning is a technique for converting continuous-tone images into two-tone image[8]. The results can resemble the original images when viewed from a distance by involving the low-pass nature of the Human Visual System (HVS). Today, digital halftoning plays a key role in almost every discipline that involves printing and displaying. All newspapers, magazines, and books are printed with digital halftoning [15].For color image separate halftone is generated for cyan, magenta, yellow and black .Each halftone screen is rotated to form a pat-tern called rosette. Effective digital halftoning can substantially improve the quality of rendered images at minimal cost [8].The major issues in choosing a halftoning technique are image quality and amount of computation. Some of the major Halftoning method that has been developed so far includes the ordered dithering, Error diffusion and Dot diffusion.
Block Truncation Coding (BTC) is a lossy image compression technique which uses moment

DOI: 10.5121/ijci.2016.5210

preserving quantization method for compressing digital gray scale images as well as color image. BTC has been used for many years for compressing digital monochrome images. BTC has been used for many years for compressing digital monochrome images. It is a simple and lossy image compression technique.

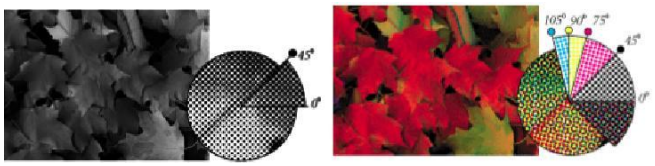


Fig .1 Enlarged details of halftone dot pattern for

(a) Grayscale image (b) Color image

The BTC method preserves the block mean and standard deviation [1].The simplest way to extend BTC to color image is to apply BTC to each color plane independently[10][2].The disadvantage of this method is that three bit plane are needed hence the compression ratios achievable are low.

2. BLOCK TRUNCATION CODING

Block Truncation Coding (BTC) is a lossy image compression technique which uses moment preserving quantization method for compressing digital gray scale images as well as color image [2]. In block truncation coding (BTC), the original image is divided into fixed-size non overlapping blocks of size M N [1] .The block size chosen is usually small to avoid the edge blurring and blocking effect. Each block is independently coded using a two level (1-bit) quantizer. The two values preserve the first and the second moment characteristic of the original block. BTC does not provide a higher gain than any of the modern image compressing algorithms like JPEG or JPEG-2000, but it is much lesser complex [5] .While BTC based image compression method provide low computational complexity ,the method also has the issue of degradation of the image quality when compared to other compression technique. However BTC based image compression also suffer from two major issues namely blocking effect and false contours [6].

3. HALFTONING METHODS

Three common methods for generating digital halftoning images are

1. **Dithering:**Common technique used for generating digital halftoning images is dithering. Dithering creates an output image with the same number of dots as the number of pixels in the source image. Dithering can be thought of as thresholding the source image with a dither matrix. The matrix is laid repeatedly over the source image. Wherever the pixel value of the image is greater than the value in the matrix, a dot on the output image is filled.

2. **Error diffusion:**Error diffusion is another technique used for generating digital half toned images. It is often called spatial dithering. Error diffusion sequentially traverses each pixel of the source image. Each pixel is compared to a threshold. If the pixel value is

higher than the threshold, a 255 is outputted; otherwise, a 0 is outputted. The error, the difference between the input pixelvalue and the output value is dispersed to nearby neighbors.

3. **Dot diffusion:**Dot diffusion method for halftoning, is an attractive method which attempts to retain the good features of error diffusion while offering substantial parallelism. The dot diffusion method for halftoning has only one design parameter called the class matrix.

4. HALFTONING BASED BLOCK TRUNCATION CODING

Halftoning based block truncation coding is an extended compression technique derived from BTC scheme in which the BTC bitmap image is replaced with halftone image. The main difference between BTC and HBTC is on the image block quantizer determination .In contrast to the BTC scheme which tries to maintain its mean value and standard deviation in image block. The HBTCquantizer is simply obtained from the minimum and maximum value found in an image block. Error diffusion based BTC offers an improved image quality. Ordered dithering based BTC is used when the main requirement is performance efficiency .Dot diffusion based method provide a balance of the above two requirements, better visual quality as well as a better performance efficiency [5][9][10].The three methods described above all provide a proper solution to the problems in a traditional BTC, such us blocking effect and false contours.

4.1 ORDERED DITHER BLOCK TRUNCATION CODING

The dithering-based BTC, namely Ordered Dither Block Truncation Coding (ODBTC) is an example of HBTC. In which bit pattern configuration of the bitmap is generated from the dithering approach (void-and-cluster Halftoning) .In encoding stage, the ODBTC scheme utilizes the dither array Look-Up-Table (LUT) to speed up the processing speed. The dither array in ODBTC method substitutes the fixed average value as the threshold value for the generation of bitmap image [6]. The extreme values in ODBTC are simply obtained from the minimum and maximum value found in the image blocks. ODBTC offers high efficiency and low computational complexity. The quantization error cannot be compensated with the ordered dithering halftoning and thus the ODBTC yields lower image quality compared to that of the EDBTC.

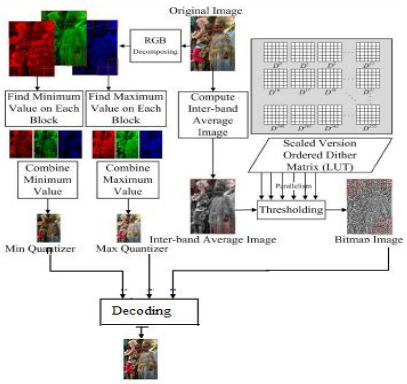


Fig 2. Block or ODBTC

4.2 ERROR DIFFUSION BLOCK TRUNCATION CODING

Error diffusion is a method that provides better visual quality of image .Error diffusion based BTC does this while also compressing the image .In this method, the inherent dithering property of error diffusion is to deal with the problem of false contour. Similar to the BTC scheme, EDBTC looks for a new representation (two quantizer and bitmap image) for reducing the storage requirement. The EDBTC bitmap image is constructed by considering the quantized error which diffuses to the nearby pixels to compensate the overall brightness .EDBTC employs the error kernel to generate the representative bitmap image. Fig 4 ,5 shows the error diffusion kernels for Floyd-Steinberg, Stucki, Jarvis, and Stevenson .Different error kernel yield different halftoning pattern. Error diffusion strategy effectively removes the annoying blocking effect and false contour, while maintaining the low computational complexity. The prolonged processing time is still an issue [5][10].

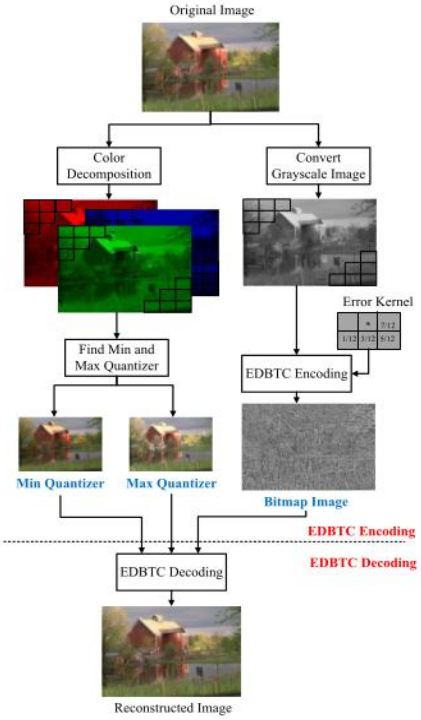


Fig. 4 Error Kernels (a) Floyd-Steinberg (b) Stuck

		*	7/48	5/48
3/48	5/48	7/48	5/48	3/48
1/48	3/48	5/48	3/48	1/48

		*	5/32	3/32
2/32	4/32	5/32	4/32	2/32
0/32	2/32	3/32	2/32	0/32

	*	7/16
3/16	5/16	1/16

		*	8/42	4/42
2/42	4/42	8/42	4/42	2/42
1/42	2/42	4/42	2/42	1/42

Fig. 5 Error Kernels (c) Jarvis (d) Sierr

Fig. 5.Image quality comparison on different EDBTC kernel
(a) Original (b) Floyd (c) Stucki

```
for each y from top to bottom
   for each x from left to right
      oldpixel  := pixel[x][y]
      newpixel  := find_closest_palette_color(oldpixel)
      pixel[x][y]  := newpixel
      quant_error  := oldpixel - newpixel
      pixel[x+1][y  ] := pixel[x+1][y  ] + quant_error * 7/16
      pixel[x-1][y+1] := pixel[x-1][y+1] + quant_error * 3/16
      pixel[x  ][y+1] := pixel[x  ][y+1] + quant_error * 5/16
      pixel[x+1][y+1] := pixel[x+1][y+1] + quant_error * 1/16
```

Fig. 6 Algorithm for error diffusion BTC

4.3.DOT DIFFUSED BLOCK TRUNCATION CODING

Dot diffusion based BTC provides better visual quality and performance efficiency compared to Error diffusion based BTC [4].Here the natural parallelism of dot diffusion is utilized to obtain better processing efficiency .Better image quality is achieved by co-optimizing the diffused matrix and class matrix of the dot diffusion. It also provides better image quality compared to Ordered Dithering based BTC as well. The DDBTC effectively compresses an image by decomposing an image into two quantizers and a bitmap image. The DDBTC diffuses the quantization error of the current processed pixel into its neighboring pixels using the diffused matrix and class matrix concurrently to generate the bitmap image .High dynamic range can easily destroy the blocking effect and false contour.

Class Matrix of size 4*4

.2716	1	.2716
1	X	1
.2716	1	.2716

8	11	6	1
14	2	4	12
5	9	7	3
13	5	0	10

5. EXPERIMENTAL RESULT

5.1 OBJECTIVE QUALITY MEASURES COMPARISON

Objective measures are used to evaluate the image degree of excellence such as Peak Signal to Noise Ratio, Mean Square Error, Compression Ratio and Structural Similarity Index (SSIM).

The PSNR is the ratio between a signal's maximum power and the power of the signal's noise. By using the PSNR values the quality of reconstructed images is best described. Mathematically, PSNR is defined as:

$$PSNR=10*\log 10\left[\frac{255^2}{\frac{1}{M*N}\sum\sum(f(m,n)-g(m,n)^2}\right] \quad (1)$$

Mean square error is a very useful measures as it gives an average value of the energy lost in the lossy compression of the original image .A very small MSE means the image is very closer to the original, f(m,n) is the original image ,g(m,n) is the reconstructed image MSE is defined as.

$$MSE=\frac{1}{M*N}\sum\sum(f(m,n)-g(m,n)^2) \quad (2)$$

The compression ratio is used to measure the ability of data compression by comparing the size of the image being compressed to the size of the original image.

Rate is another metric used to evaluate the performance of the compression algorithm. Rate gives the number of bits per pixel used to encode an image rather than abstract percentage.

Another category of image quality measures is based on the assumption that the human visual system is highly adapted to extract structural information from the viewing field [10]. The error sensitivity approach estimates perceived errors to quantify image degradations, while this approach considers image degradations as perceived structural information variation. The structural similarity index (SSIM) can be calculated as a function of three components: luminance, contrast and structure.

$$SSIM(x; y) = [l(x; y)]^{\alpha}[c(x; y)]^{\beta}[s(x; y)]^{\gamma}$$

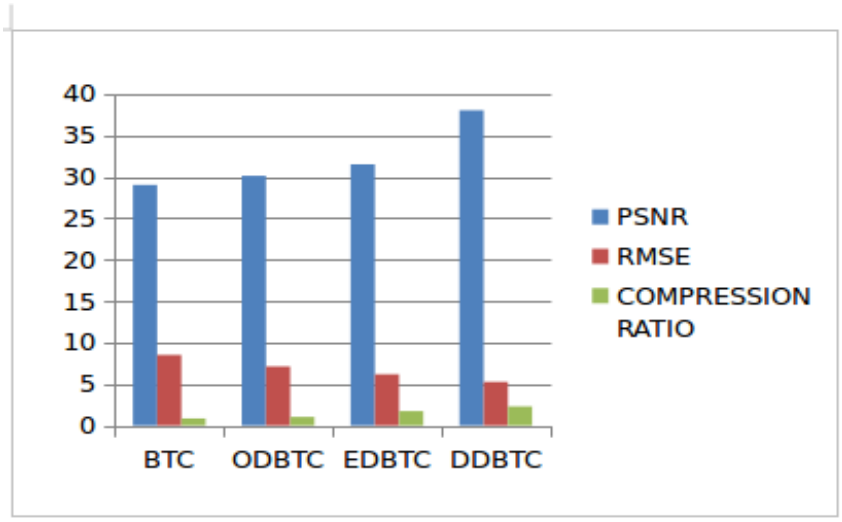


Fig .7 Objective image quality comparisons

Table I. Objective Image Quality Comparison

Type of BTC	RMSE	PSNR	Compression Ratio	SSIM
BTC	7.52	29.03	0.8607	0.8734
ODBTC	7.13	30.13	1.03611	0.93245
EDBTC	6.17	31.52	1.73611	0.94076
DDBTC	5.27	38.03	2.3000	0.95518

The table illustrates the PSNR, MSE, SSIM and Compression Ratio values obtained for different halftoning based BTC. From the table as well as through visual inspection of the result images, it can be seen that Dot Diffused BTC provide better visual quality compared to other halftoning based BTC.

Fig.8.Image quality comparison on different HBTC (a)original,(b)BTC,(c)ODBTC,(d)EDBTC,(e)DDBTC

6. CONCLUSION

In this literature survey, image compression using Halftoning based Block Truncation coding for color image has been scrutinized, which can provide an excellent image quality and artifact free result such us inherent blocking effect and false contour artifact of the traditional BTC simultaneously. Four algorithms were selected specifically, the original block truncation coding (BTC), Ordered dither Block truncation coding (ODBTC), Error diffused block truncation coding (EDBTC) and Dot diffused Block truncation coding (DDBTC). Objectives measures are used to evaluate the image degree of excellence such as PSNR, MSE, SSIM and Compression Ratio, In this survey find out that halftoning based BTC not only applied for gray scale image it can be extended for color image .For future study, other color spaces can be explored for the image compression such as YCbCr, Lab color space, HSI etc.

REFERENCES

[1] E.J Delp and O.R .Mitchell. Image coding using block truncation coding. IEEE Transactions on Image Processing, 27(1):1335–1342, Sept 1979.
[2] G. Qiu. Color Image Indexing Using BTC IEEE Transactions on Image Processing, 12(1) , Jan. 2003.
[3] J.M Guo High efficiency ordered dither block truncation with dither array LUT and its scalable coding application IEEE Transactions on Image Processing, 20(1):97–110, Jan 2010.
[4] J.M Guo and Y.F.Liu. Improved Block Truncation Coding using Optimized Dot Diffusion. IEEE Transactions on Image Processing, 2(1):1269–1275, Jan 2014.

[5] J.M Guo and Y.F.Liu. Improved Block Truncation Coding using modified error diffusion. IEEE Transactions on Image Processing, 44(7):462–464, Mar 2008.
[6] J.M Guo and Y.F.Liu. Improved Block Truncation Coding Based on the Void-and-Cluter Dithering Approach. IEEE Transactions on Image Processing, 18(1):211–213, Jan 2009.
[7] David Saloman. Data Compression the Complete Reference. Fourth Edition.
[8]R Ulichney. Digital Halftoning. Cambridge,USA:: MIT Press,1987
[9] M.Kamel,C.T.Sun and G.Lian. Image compression by variable block truncation coding with optimal threshold. IEEE Transactions on Signal Processing, 39(1):208–212, Jan 1991.
[10] Jing-Ming Guo and Yun-Fu Liu. High Capacity Data Hiding for Error-Diffused Block Truncation Coding.. IEEE Transactions on Signal Processing, 21(12):4808–4818, December 2012.
[11] S.Vimala, P.Uma, B. Abidha. Improved Adaptive Block Truncation Coding for Image Compression International Journal of Computer Applications, 19(7):975–988, April 2011 .
[12] M. D. Lema and O. R. Mitchell. High Absolute moments block truncation coding and its application to color images, IEEE Trans. Commun, 4(32):1148–1157, Oct. 1984
[13] V. R. Udpikar and J. P. Raina. High BTC image coding using vector quantization, IEEE Trans. Commun, 4(35):352–358, Sept. 1987.
[14] H. R. Kang Digital Color Halftoning New York: 1999.
[15] S.Vimala, P.Uma, B. Abidha. Improved Adaptive Block Truncation Coding for Image Compression International Journal of Computer Applications, 19(7):975–988, April 2011
[16] V. R. Udpikar and J. P. Raina. High BTC image coding using vector quantization, IEEE Trans. Commun, 4(35):352–358, Sept. 1987.
[17] P. Franti, and T. Kaukoranta, Binary vector quantizer design using soft centroids, Signal Proc Image Comm , vol. 14, no. 9, pp. 677-681, 1999.
[18] M.Kamel,C.T.Sun and G.Lian. Image compression by variable block truncation coding with optimal threshold. IEEE Transactions on Signal Processing, 39(1):208–212, Jan 1991.
[19] Jing-Ming Guo and Yun-Fu Liu. High Capacity Data Hiding for Error-Diffused Block Truncation Coding. IEEE Transactions on Signal Processing, 21(12):4808–4818, December 2012.
[20] S.Vimala, P.Uma, B. Abidha. Improved Adaptive Block Truncation Coding for Image Compression. International Journal of Computer Applications, 19(7):975–988, April 2011
[21] C. S. Huang and Y. Lin. Hybrid block truncation coding, IEEE Trans. Signal Process, 4(12):328–330, Dec 1997
[22] M. D. Lema and O. R. Mitchell. High Absolute moment blocks truncation coding and its application to color images, IEEE Trans. Commun, 4(32):1148–1157, Oct. 1984.
[23] H. R. Kang. Digital Color Halftoning, New York: 1999.
[24] Smen Forchhammer and Kim S. Jensen. Data Compression of Scanned Halftone Images, IEEE Trans. Commun, 4(42):213–238, Mar. 1997.
[25] Arup Kumar Pal. An efficient codebook initialization approach for LBG algorithm, International Journal of Computer Science, Engineering and Applications (IJCSEA), Vol.1, No.4, August 2011.

AUTHORS

Meharban M S was born in Perumbavoor in 1993. She received BE in Information technology from Cochin University Of science and technology in 2014. She is currently an M.Tech Candidate in computer science from model engineering college, Thrikkakara in 2016. She is working as part time faculty at IGNOU.

Dr. Priya S Obtained her BTech in Computer Science & Engineering from Kerala University, MTech in Computer Science & Engineering from Pondicherry University, Pondicherry India and PhD in Information & Communication Engineering from Anna University, Chennai, India. She is currently working as Associate Professor in the Department of Computer Science & Engineering in Govt Model Engineering College, Thrikkakara, Ernakulam, and Kerala, India. Her experience as a faculty is more than 18 years as of now. She is a life member of IST.

A SECURE SCHEMA FOR RECOMMENDATION SYSTEMS

Asny P.A[1] and Susanna M. Santhosh [2]

[1] Student, Department of Computer Science and Engineering, MBITS Nellimattom and
[2]Assistant Professor, Department of Computer Science and Engineering,MBITS Nellimattom

ABSTRACT

Recommender systems have become an important tool for personalization of online services. Generating recommendations in online services depends on privacy-sensitive data collected from the users. Traditional data protection mechanisms focus on access control and secure transmission, which provide security only against malicious third parties, but not the service provider. This creates a serious privacy risk for the users. This paper aims to protect the private data against the service provider while preserving the functionality of the system. This paper provides a general framework that, with the help of a preprocessing phase that is independent of the inputs of the users, allows an arbitrary number of users to securely outsource a computation to two non-colluding external servers. This paper use these techniques to implement a secure recommender system based on collaborative filtering that becomes more secure, and significantly more efficient than previously known implementations of such systems.

KEYWORDS

Secure multi-party computation, privacy, recommender systems,secret sharing.

1.INTRODUCTION

Recommendation systems are an important part of the information and e-commerce ecosystems. They represents a powerful method for enabling users to filter through large information and product spaces. Recommendation systems consist of a processor together with multitude of users, where the processor provides recommendations to requesting users, which are deduced from personal ratings that were initially submitted by all the users. Recommender systems based on collaborative filtering method that collect and process personal user data constitute an essential part of the service. On one hand, people benefit from online services.On the other hand, direct access to private data by the service provider has potential privacy risks for the users since the data can be processed for other purposes, transferred to third parties without user knowledge, or even stolen. Recent studies show that the privacy considerations in online services seems to be one of the most important factors that threaten the healthy growth of the e-business. In a non-cryptographic setup of such a system, the processor is both able to learn all the data submitted by the users and spoof arbitrary, incorrect recommendations. Therefore, it is important to protect the privacy of the users of online services for the benefit of both individuals and business concern.

DOI: 10.5121/ijci.2016.5211

In this work, replaces the recommendation processor by a general two-server processor that satisfies following conditions,

1) The privacy of the ratings and recommendations of the users is maintained.

2) A server that is under adversarial control is unable to disrupt the recommendation process.

In this model the computation is ongoing and outsourced to two external servers that do not collude. This approach allows for the involvement of many users that need only be online for very short time periods in order to provide input data to, or request output data from, the servers. One of the two servers could be the service provider(SP) that wishes to recommend particular services to users, and the other server could be a governmental organisation guarding the privacy protection of users. The role of the second server could also be commercially exploited by a privacy service provider(PSP), supporting service providers in protecting the privacy of the customers.

The major goals of securing the recommendation system is that,

1) Do not want the servers to learn the personal data of users.

2) The correctness of the user outputs is better preserved, because outputs cannot be corrupted by one server on his own.

3) A malicious server might introduce a couple of new dummy users. These dummy users might help him deduce more personal data than is available through the protocol outputs.

In this work used the SPDZ framework,which enables secure multi-party computations[16] in the malicious model, extended it to the client-server model, and worked out a secure recommendation system within this setting. Not only did this lead to a recommendation system that is secure in the malicious model, but also the online phase became very efficient.To extend SPDZ to the client-server model, developed secure protocols that enable users (clients) to upload their datas to the servers, and afterwards obtain the computed outputs from the servers. This required a subprotocol for generating duplicate sharings in the system. To securely compute a recommendation within SPDZ, had to develop secure comparison protocol and secure integer division protocol.

2.RELATED WORKS

Most of the related works on privacy preserving recommendation is secure in the semi-honest model, so parties are assumed to follow rules of the protocol. As mentioned by Lagendijk et al.[1], “Against malicious adversaries achieving security is a hard problem that has not yet been studied widely in the context of privacy-protected signal processing.”

Nikolaenko et al.[5] securely computed collaborative filtering by means of matrix factorization. They used both homomorphic encryption and garbled circuits in a semi-honest security model. In another paper [6], these authors use similar techniques to securely implement the Ridge regression, a different approach of collaborative filtering.

Erkin et al.[2] securely computed recommendations based on collaborative filtering method. They used homomorphic encryption within semi-honest security model just like Bunn[3] and Ostrovsky[4]. Goethals et al stated that although such techniques can be made secure in malicious model, will make them unsuitable for real life applications because of the increased computational and communication costs.

Some schemes with controllable and revocable anonymity provide linkability by adding a tag to a signature. Using the tag associated with a signature, one can check the linkability on the signatures easily and explicitly. For example, a linkable democratic GS scheme is a variant of a democratic GS scheme to support the tag-based link ability. A message-linkable GS scheme was suggested to resist Sybil attacks in the VANET.

In the last several years, a couple of computation protocols have been developed, which are both practical and secure in the malicious model. The idea is to use public-key techniques in a data-independent pre-processing phase, such that cheap information-theoretic primitives can be exploited in the online phase, which makes the online phase efficient. In 2011, Bendlin et al.[7] presented such a framework with a somewhat homomorphic encryption scheme for implementing the pre-processing phase. This has been improved lately by Damgård et al.[8], which has become known as SPDZ (pronounced "Speedz"). Last year, Damgård et al.[9] showed how to further reduce the precomputation effort.

3.PROBLEM DEFINITION

Recommender systems have become an important tool for the personalization of online services. Generating recommendations in online services depends on privacy-sensitive data collected from the user. Traditional data protection mechanisms focus only on access control and secure transmission, which provide security against malicious third parties, but not the service provider. This creates a serious privacy risk for the users.Most of the system is only secure in semi-honest model. So aim to protect the private data against the service provider while preserving the functionality of the systems.Propose amodified version of the standard model for the secure multi-party computation, which is a cryptologic paradigm in which the players jointly perform a single secure computation and then abort.By introducing general two-server processor in such a way that, as long as one of the two servers is not controlled by an adversary and behave correctly.

4.SECURE RECOMMENDATION SYSTEM

4.1.User-Based Collaborative Filtering

Collaborativefiltering(CF) is a popular recommendation algorithm that based its predictions and recommendations on the ratings or behavior of other users in the system. The fundamental assumption behind this method is that other users opinions can be selected and aggregated in such a way that to provide a reasonable prediction of the active user's preference.To generate recommendations for a particular user in a group of users and items, uses a system based on collaborative filtering, which has the following three steps.

1)Similarities are computed between that particular user and all others.

2) The most similar users are selected by comparing their similarity values with a threshold.

3) The recommendations on all of the items are generated as the average rating of the most similar users.

In collaborative filtering there is one processor R, with N users, and M different predefined items. A small subset of sizeS ($1 \leq S < M$) of these items is assumed to have been rated by each user,reflecting his personal taste. The remaining M – S items have only been rated by a small subset of users that haveexperienced by particular item before. A user that is lookingfor new, unrated items, can ask the processor to produceestimated ratings for theM−S items. The number Nof users can be large, M is in the order ofhundreds, and S usually is a few tens [12].

During initialisation, each user n uploads to processor a list of at most M ratings of items, where each rating V(n,m) is represented by a value within a pre-specified interval. Users can update their rating at any time during the lifetime of the system. A user can, at any time after the initialization, request a recommendation from processor. When the processor receives such a request from a user, computes a recommendation for this user as follows. First, it uses the initial S ratings in each list to determine which other users are considered to be similar to the requesting user, have similarly rated the first S items. The remaining M−S entries in the lists are then used to compute and return a recommendation for the requesting user, consisting of M −S ratings average over all similar users.

To get an idea of required computation we describe the required computational steps. Let Um be the set of users that have rated item m, $S < m \leq M$.

1) Each user uploads his ratings to the processor[1], we assume the first S ratings have been normalized and scaled beforehand. A rating is normalized by dividing it by the length of the vector (V(n,1), . . , V(n,S)), yielding a real number s between 0 and 1. Next, this real number is scaled and rounded to a positive integer consisting of a few bits. The remaining ratings should only be scaled and rounded to an integer with same maximal number of bits.
2) When user A asks for a recommendation, processor computes M − S estimated ratings for A. The similarities $SimA,n = \sum^{s} m = 1 V(n,m) \cdot V(A,m)$ are computedfor each user n.
3) Each similarity value is compared with a public threshold $t \in N+$, and outcome is presented by the bit $\delta n = (t < SimA,n)$.
4) The recommendation for user A consists of M – S estimated ratings, the estimated rating for item m, $S < m \leq M$, simply being an average of the ratingsof the similar users: $Recm = (\sum_{n} \in Um\ \delta n \cdot V(n,m)) \div (\sum_{n} \in Um\ \delta n)$, where ÷ denotes integer division.
5) The processor send back the recommendation RecS+1 . . . RecM to user A.

4.2.SECURE MODEL

4.2.1.SecretSharing

An external dealer distributes shares of a secret value x to the two servers as follows:

1) The dealer selects a value r uniformly at random.
2) The dealer sends the value r to server 1(SP) and the value $x - r$ to server 2(PSP).

The values x1 = r and x2 = x −r are considered to be the share of SP and the share of PSP, respectively. It should be clear from the description above that the shares x1 and x2 are both individually statistically independent of the secret x, while they together allow to determine the value of x, by adding these shares together.

In addition to the distribution of the shares, the dealer distributes authentication tags on the shares with respect to the authentication code C , defined as C(x, (α,β)) = α ・x + β. Here the value (α, β) is called the authentication key and the value α ・x + β the authenticationtag for the share x.

For every share x1 for SP, the dealer generates a random authentication key (α2, β2), computes the corresponding authentication tag m1 = α2 ・x1 + β2 and sends the key (α2, β2) to PSP, and the share x1 and tag m1 to server SP.

4.2.2.Operations for the Computation Phase

In this describes the operations that are needed for the recommender system. Suppose that servers 1 and 2 hold fixed partial authentication keys $\alpha 1, \alpha 2$.

1) Linear Operations:

Let [x] and [y] be secret sharings of arbitrary values x, y and let c be a public constants. Show how to non-interactively compute secret sharings for [$x + y$], [cx] and [$x + c$] respectively. An authenticated secret sharing of the sum $z = x + y$ is computed by locally adding the shares, keys, and tags of x and y. From a sharing [x], an authenticated secret sharing of cx is computed by local multiplication of the shares, keys, and tags of x. To add a public constant to a secret sharing [x], one party adds the constant to its share, and the other party adjusts its authentication key.

2) Multiplication:

Let [x] and [y] denotes the secret sharings of arbitrary values x, y, and [a], [b], [c]be a given multiplication triplet such that c = ab. The following sequence of local linear operations and interactions is used to compute a sharing [z] where z = xy, making use of the precomputed multiplication triplets:

- The servers locally compute the secret sharing [v] = [x −a] from [x] and [a], and open it towards each other.

- The servers locally compute the secret sharing [w] = [y −b] from [y] and [b], and open it towards each other.

- The servers locally computes the secret sharing [z] = [xy] = w[a] + v[b] + [c] + vw.

4.2.3.Secure Architecture for Recommendation Systems

This secure framework relies on techniques from the cryptologic area[15] of secure multi-party computation[14]. This model have the following structures. First,inputphasethat use in the secure computation enables the parties to encrypt their respective inputs.Next a computation phase that takes place during which an encrypted output of the function f is computed from the encrypted inputs. Last, an output phasetakes place where the output is decrypted, and then sent to the appropriate parties. Consider secure multi-party computation in the preprocessingmodel,where at some point in time prior to the selection of the inputs, a preprocessing phase takes place that establishes the distribution of an arbitrary amount of correlated data between the parties involved in the computation phase.This data is completely independent of the input data of the parties in the system. The goal of the preprocessing is to remove as much of the complexity and interaction from the actual computation as much as possible, which as a result makes this computation extremely efficient.

Every computation s corresponds with a function *f*, which represented via an arithmetic circuit consisting of basic operations like addition and multiplication. For considering the recommender application, it suffices to consider these basic operations together with the more complex operations of comparison and integer division, which are composed of basic operations.

The outsourcing aspect of the secure computation, the input phase is non-standard in the sense that the inputs are not provided by the two servers. The input phase results in encryptions of the inputs that are suitable for the computation phase of the two-party computation with preprocessing.The figure 1.shows the architecture of a secure recommendation system.
Although the users providing the inputs could in principle take care of the share distribution, these users cannot be trusted to provide the authentication keys and tags, they might be under control of one of the servers. Here introduces an additional structure to the authentication keys in order to enable a secure two-party computation approach.

Once the encryption of the inputs has been established, the computation recursively handles the operations in the circuit while maintaining the encryption structure as constant.Every operation in the circuit is initiated with two encrypted inputs, and produces an encrypted output without leaking any information on the encrypted values provided. So, at the end of the circuit, an encryption of the final output becomes a vailable.

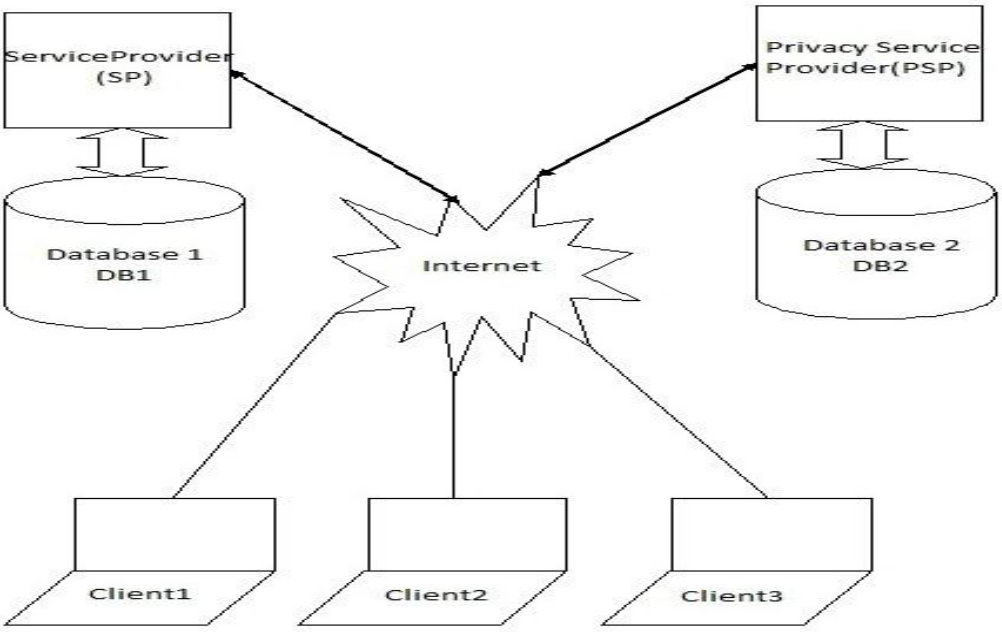


Figure 1. Architecture for A Secure Recommendation systems

The output phase is also non-standard, as the output needs to be revealed to an external parties. Here the idea is that all data related to the encryption of the output is sent back to the relevant users in the system, so that this user can verify the correctness of the shares using the authentication keys and tags provided, and then decrypt the output using the shares.

1)The Input Phase

This framework allows multiple clients to upload any values to the processor,focus on this recommender application[13]. Initially, each user n ($1 \leq n \leq N$) will have to upload his ratings $V(n,m)$ ($1 \leq m \leq M$) once, before the recommendations can be requested. A user could easily act as a dealer by splitting his rating into two shares, and sending each servers(service provider and privacy service provider) a share accompanied by proper authentication tags.

2)The Output Phase

Although this framework allows the clients to download any value from the processor, focus on our recommender application. After a recommendation has been computed, the outputs Rec$m$ $\in (S < m \leq M)$ have to be sent to the requesting users in the system.

5.CONCLUSIONS

In this work provide a highly efficient, privacy-preserving general framework for securing a recommendation system. This framework is then applied to the problem of secure recommendation and, given a sufficient amount of precomputed data, leads to extremely efficient implementations.

While this work focuses on the application of secure recommendation systems, the underlying framework is sufficiently generic for use in other, similar applications, and also easily extends to model variations involving more than two servers.

REFERENCES

[1] Thijs Veugen, Robbert de Haan, Ronald Cramer, and Frank Muller," A Framework for Secure Computations With Two Non-Colluding Servers and Multiple Clients,Applied to Recommendations", Ieee Transactions On Information Forensics And Security, Vol. 10, No. 3, March 2015.

[2] R. Lagendijk, Z. Erkin, and M. Barni, "Encrypted signal processing for privacy protection: Conveying the utility of homomorphic encryption and multiparty computation," *IEEE Signal Process. Mag.*, vol. 30, no. 1,pp. 82–105, Jan. 2013.

[3] Z. Erkin, T. Veugen, T. Toft, and R. L. Lagendijk, "Generating private recommendations efficiently using homomorphic encryption and data packing," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 1053–1066, Jun. 2012.

[4] P. Bunn and R. Ostrovsky, "Secure two-party k-means clustering," in *Proc. 14th ACM Conf. Comput. Commun. Secur.*, 2007, pp. 486–497.

[5] B. Goethals, S. Laur, H. Lipmaa, and T. Mielikäinen, "On private scalar product computation for privacy-preserving data mining," in *Proc. 7th Int. Conf. Inf. Secur. Cryptol.*, 2004, pp. 104–120.

[6] V. Nikolaenko, S. Ioannidis, U. Weinsberg, M. Joye, N. Taft, and D. Boneh, "Privacy-preserving matrix factorization," in *Proc. ACMSIGSAC Conf. Comput. Commun. Secur.*, 2013, pp. 801–812.

[7] V. Nikolaenk, U.Weinsberg, S. Ionnidis, M. Joye, D. Boneh, and N. Taft, "Privacy-preserving ridge regression on hundreds of millions of records," in *Proc. IEEE Symp. Secur. Privacy*, May 2013, pp. 334–348.

[8] R. Bendlin, I. Damgård, C. Orlandi, and S. Zakarias, "Semihomomorphic encryption and multiparty computation," in *Advances in Cryptology* (Lecture Notes in Computer Science), vol. 6632. Berlin,Germany: Springer-Verlag, 2011, pp. 169–188.

[9] I. Damgård, V. Pastro, N. Smart, and S. Zacharias, "Multiparty computation from somewhat homomorphism encryption," in *Advances inCryptology* (Lecture Notes in Computer Science), vol. 7417. Berlin, Germany: Springer-Verlag, 2012, pp. 643–662.

[10] I. Damgård, M. Keller, E. Larraia, V. Pastro, P. Scholl, and N. P. Smart, "Practical covertly secure MPC for dishonest majority—Or: Breaking the SPDZ limits," in *Computer Security* (Lecture Notes in Computer Science), vol. 8134. Berlin, Germany: Springer-Verlag, 2013.

[11] M. Atallah, M. Bykova, J. Li, K. Frikken, and M. Topkara, "Private collaborative forecasting and benchmarking," in *Proc. ACM WorkshopPrivacy Electron. Soc. (WPES)*, 2004, pp. 103–114.

[12] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, *Recommender Systems Handbook*. New York, NY, USA: Springer-Verlag, 2011.

[13] T. Veugen, "Encrypted integer division and secure comparison," *Int J. Appl. Cryptograph.*, vol. 3, no. 2, pp. 166–180, 2014.

[14] BI. Damgård, M. Fitzi, E. Kiltz, J. B. Nielsen, and T. Toft, "Unconditionally secure constant-rounds multi-party computation for equality, comparison, bits and exponentiation," in *Theory of Cryptography*. Berlin, Germany: Springer-Verlag, 2006, pp. 285–304.

[15] R. Canetti, "Universally composable security: A new paradigm for cryptographic protocols," in *Proc. 42nd IEEE Symp. Found. Comput.Sci.*, Oct. 2001, pp. 136–145.

[16] T. Nishide and K. Ohta, "Multiparty computation for interval, equality, and comparison without bit-decomposition protocol," in *PublicKey Cryptography* (Lecture Notes in Computer Science), vol. 4450, T. Okamoto and X. Wang, Eds. Berlin, Germany: Springer-Verlag, 2007, pp. 343–360.

AUTHORS

AsnyP.A. is currently pursuing M.Tech in Computer Science and Engineering in Mar Baselios Institute of Technology and Science,Nellimattom.She completed her B.Tech from Ilahia College of En gineering and Technology,Muvattupuzha. Her specialization in Cyber Security.

Susanna M. Santhosh is currently assistant professor of the Department of Computer Science and EngineeringMar Baselios Institute of Technology and Science,Nellimattom., Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 2010 from College of Engineering, Chengannur and M-Tech from Federal Institute of Science and Technology,Angamaly in 2012. She has around 3 years of teaching experience. Her specializayion in InformationS

INTENTIONAL BLANK

CASSANDRA A DISTRIBUTED NOSQL DATABASE FOR HOTEL MANAGEMENT SYSTEM

Varalakshmi P.[1], Hima S.[2] and Surekha Mariam Varghese[3]

Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

Apache Cassandra is a distributed storage system for managing very large amounts of structured data. Cassandra provides highly available service with no single point of failure. Cassandra aims to run on top of an infrastructure of hundreds of nodes possibly spread across different data centers with small and large components fail continuously. Cassandra manages the persistent state in the face of the failures which drives the reliability and scalability of the software systems. Cassandra does not support a full relational data model because it resembles a database and shares many design and implementation strategies. In this paper, discuss an implementation of Cassandra as Hotel Management System application. Cassandra system was designed to run on cheap commodity hardware. Cassandra provides high write throughput and read efficiency.

KEYWORDS

Cassandra, Data model.

1.INTRODUCTION

Apache Cassandra is an open source, distributed, highly available, decentralized, elastically scalable, fault-tolerant, consistent, column-oriented database. Cassandra's distribution design is based on Amazon's Dynamo and its data model on Google's Bigtable. Cassandra was introduced at Facebook; it is now used at some of the most popular sites on the Web [1].

Apache Cassandra is a type of NoSQL database designed to handle large amounts of data across many servers. This database provides high availability and no single point of failure.

Some of the important points of Apache Cassandra: (1) It is scalable, consistent and fault-tolerant, (2) It is key-value as well as column-oriented database,(3) Its data model is based on Google's Bigtable and distribution design is based on Amazon's Dynamo, (4) Introduced at Facebook, it differs sharply from relational database management systems,(5) Cassandra implements a Dynamo-style replication model, also adds a more powerful "column family" data model, and (6) Cassandra is being used by some of the biggest companies such as Facebook, Twitter, Cisco, Rackspace, ebay, Twitter, Netflix, and more.

Cassandra has become so popular because of its outstanding technical features. Given below are some of the features of Cassandra:

- Elastic scalability: Cassandra allows adding more hardware to accommodate more customers and more data as per requirement.
- Always on architecture: Cassandra is continuously available for critical business applications that cannot afford single point of failure.

DOI: 10.5121/ijci.2016.5212

- Fast linear-scale performance: Cassandra increases throughput as the number of nodes in the cluster is increased. Therefore it provides a quick response time.
- Flexible data storage: Cassandra handles all possible data formats including: structured, semi-structured, and unstructured. It can dynamically provide changes to data structures according to user need.
- Easy data distribution: Cassandra provides the flexibility to distribute data where user need by replicating data across multiple data centers.
- Transaction support: Cassandra supports properties like Atomicity, Consistency, Isolation, and Durability (ACID).
- Fast writes: Cassandra was designed to run on cheap commodity hardware. It performs fast writes and can store hundreds of terabytes of data, without sacrificing the read efficiency.

The rest of this paper is organized as follows. Section 2 discusses NoSQL database. Section 3 presents the Cassandra Architecture. Section 4 describes the data model of Cassandra. Section 5 describes the implementation details of Hotel Management System. The conclusion is given in Section 6.

2.NOSQL DATABASE

A NoSQL database (also called as Not Only SQL) is a database that provides a mechanism to store and retrieve data other than the tabular relations used in relational databases. These databases are schema-free, support easy replication, have simple API, eventually consistent, and can handle huge amounts of data.

The primary objective of a NoSQL database is to have

- simplicity of design,
- horizontal scaling, and
- finer control over availability.

NoSql databases use different data structures compared to relational databases. It makes some operations faster in NoSQL. The suitability of a given NoSQL database depends on the problem it must solve.

3.CASSANDRA ARCHITECTURE

The design goal of Cassandra is to handle big data workloads across multiple nodes without any single point of failure. Cassandra has peer-to-peer distributed system, and data is distributed among all the nodes in a cluster [2].

- All the nodes in a cluster play the same role. Each node is independent and at the same time interconnected to other nodes.
- Each node in a cluster can accept read and write requests, regardless of where the data is actually located in the cluster.
- When a node goes down, read/write requests can be served from other nodes in the network.

3.1.Data Replication In Cassandra

In Cassandra, one or more of the nodes in a cluster act as replicas for a given piece of data. If it is detected that some of the nodes responded with an out-of-date value, Cassandra will return the most recent value to the client. After returning the most recent value, Cassandra performs a read repair in the background to update the stale values.

The figure 1 shows a schematic view of how Cassandra uses data replication among the nodes in a cluster to ensure no single point of failure. Cassandra uses the Gossip Protocol to allow the nodes to communicate with each other and detect any faulty nodes in the cluster.

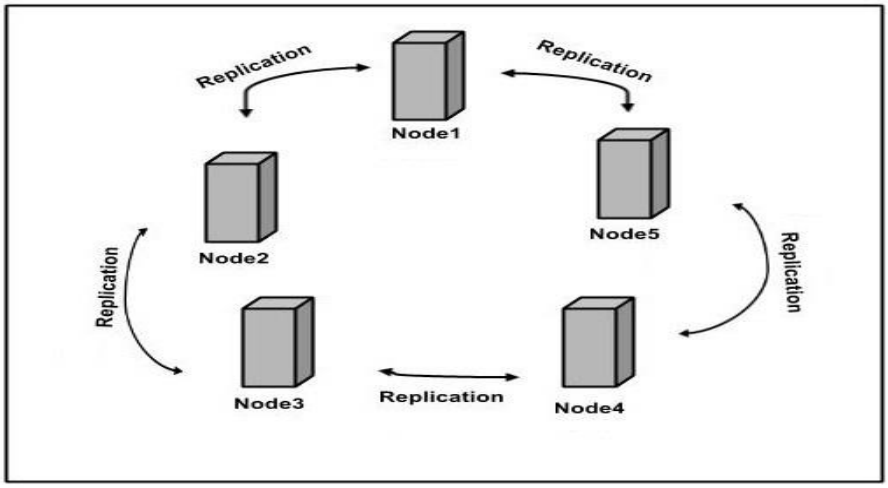


Figure. 1 Schematic view of Cassandra

3.2.Components of Cassandra

The key components of Cassandra are as follows:

- Node: It is the place where data is stored.
- Data center: It is a collection of related nodes.
- Cluster: A cluster is a component that contains one or more data centers.
- Commit log: The commit log is a crash-recovery mechanism in Cassandra. Every write operation is written to the commit log.
- Mem-table: A mem-table is a memory-resident data structure. After commit log, the data will be written to the mem-table. Sometimes, for a single-column family, there will be multiple mem-tables.
- SSTable: It is a disk file to which the data is flushed from the mem-table when its contents reach a threshold value.
- Bloom filter: These are quick, nondeterministic, algorithms for testing whether an element is a member of a set. It is a special kind of cache. Bloom filters are accessed after every query.

3.3.Cassandra Query Language

Users can access Cassandra through its nodes using Cassandra Query Language (CQL). CQL treats the database (Keyspace) as a container of tables. Programmers use cqlsh: a prompt to work with CQL or separate application language drivers.

3.4.Write Operations

Every write activity of nodes is captured by the commit logs written in the nodes. Then the data will be captured and stored in the mem-table. Whenever the mem-table is full, data will be written into the SStable data file. All writes are automatically partitioned and replicated throughout the cluster. Cassandra periodically consolidates the SSTables, deleting unnecessary data.

3.5.Read Operations

During read operations, Cassandra gets values from the mem-table. It checks the bloom filter to find the appropriate SSTable that holds the required data.

4.DATA MODEL

The data model of Cassandra is significantly different from the normal RDBMS [2].

4.1.Cluster

Cassandra database is distributed over several machines that operate together [3]. The outermost container is known as the Cluster. For failure handling, every node contains a replica. In case of a failure, the replica takes charge. Cassandra arranges the nodes in a cluster, in a ring manner, and assigns data to them.

4.2.Keyspace

Keyspace is the outermost container for data in Cassandra. The basic attributes of a Keyspace in Cassandra are:

- Replication factor: It is the number of machines in the cluster that will receive copies of the same data.
- Replica placement strategy: It is the strategy to place replicas in the ring. The different strategies such as simple strategy (rack-aware strategy), old network topology strategy (rack-aware strategy), and network topology strategy (data center-shared strategy) are available.
- Column families: Keyspace is a container for a list of one or more column families. A column family is a container of a collection of rows. Each row contains ordered columns. Column families represent the structure of data. Each keyspace has at least one and often many column families.

5.IMPLEMENTATION DETAILS

The implementation of Apache Cassandra includes installing and configuring Cassandra. Initially download Cassandra from cassandra.apache.org. Copy the folder named cassandra. Move to bin folder. Open the Cassandra.yaml file which is available in the bin folder of the Cassandra folder. Verify that the following configurations.

- data_file_directories**"/var/lib/cassandra/data"**
- commitlog_directory**"/var/lib/cassandra/commitlog"**
- saved_caches_directory**"/var/lib/cassandra/saved_caches"**

Setting the path
Set the path as Cassandra_Home=C:\apache-cassandra-1.2.19
Starting Cassandra
$ cd $CASSANDRA_HOME
$./bin/cassandra –f
Starting cqlsh
Start cqlsh using the command **cqlsh**as shown below. It gives the Cassandra cqlsh prompt as output.
$ cqlsh Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 2.1.2 | CQL spec 3.2.0 | Native protocol v3]
cqlsh>

An application of Cassandra implementation is Hotel Management System (HMS) [5].Cassandra database is chosen for this application because of its increasing throughput as the number of nodes increases, continuous availability for critical business applications and elastic scalability. Moreover Cassandra handles all possible data formats and distribution of data by replicating data across multiple data centres. Cassandra supports ACID properties and it works on cheap commodity hardware.

In the keyspace of Hotel Management System Figure 2 we have the followingcolumn families: Hotel, HotelByCity, Guest, Reservation, PointOfInterest, Room, Room Availability.

In this design, transferred some of the tables, such as Hotel and Guest, to column families. Other tables, such as PointOfInterest, have been denormalized into a super column family. We have created an index in the form of the HotelByCity column family.

We have combined room and amenities into a single column family, Room. The columns such as type and rate will have corresponding values; other columns, such as hot tub, will just use the presence of the column name itself as the value, and be otherwise empty.

Hotel Management System includes details about different hotels, guests who stay in the hotels, availability of rooms for each hotel, and a record of the reservation, which is a certain guest in a certain room for a certain period of time (called the "stay"). Hotels typically also maintain a collection of "points of interest," which are shopping galleries, monuments, museums, parks, or other places near the hotel that guests might like to visit during their stay.

Our application Hotel Management System designed with Cassandra includes the following characteristics:

- Find hotels in a given area.
- Find information about a specific hotel, such as its name, location, room availability etc.
- Find interesting locations near to a given hotel.
- Find availability of rooms in a given date range.
- Find the amenities and rate for a room.
- Possible to book the selected rooms by entering guest information.

The database in Cassandra is created using keyspace. A keyspace in Cassandra is a namespace which defines data replication on nodes. A cluster contains one keyspace per node.

The application we're building will do the following things:

1. Create the database structure.
2. Prepopulate the database with hotel and point of interest data. The hotels are stored in standard column families, and the points of interest are in super column families.
3. Search for a list of hotels in a given city. This uses a secondary index.
4. Select one of the hotels returned in the search, and then search for a list of points of interest near the chosen hotel.
5. Booking the hotel by doing an insert into the Reservation column family should be straightforward at this point, and is left to the reader.

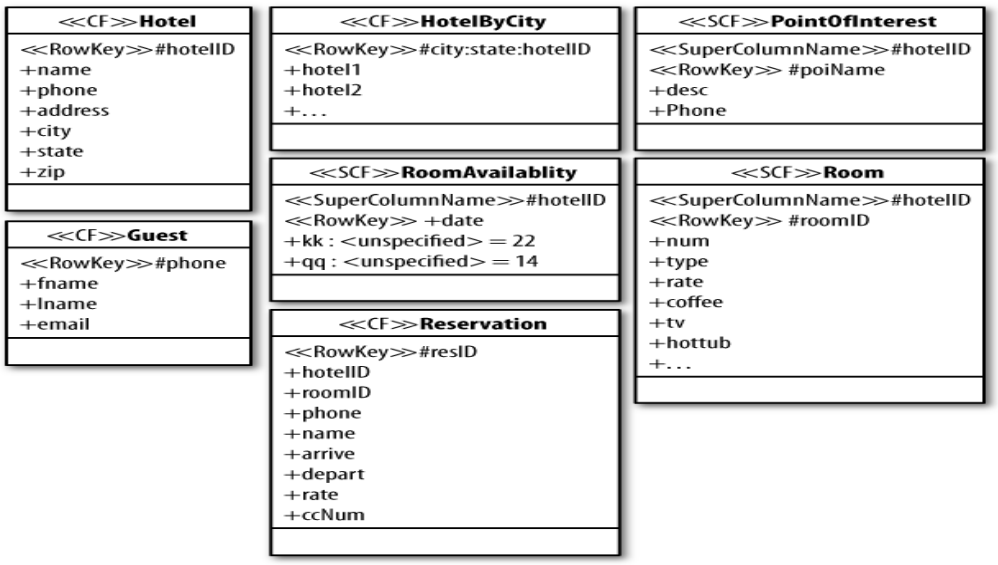


Figure. 2 Hotel Management System

5.1.Table Operations

To create a table use the command CREATE TABLE. The tables required for the Hotel Management System application can be created using this command. The syntax is
CREATE (TABLE | COLUMNFAMILY) <tablename> ('<column-definition>', '<column-definition>')

The primary key is represented by a column that is used to uniquely identify a row. Therefore, defining a primary key is mandatory while creating a table. A primary key is also made of one or more columns of a table [4].

5.2.CURD Operations

To create data in a table use the command INSERT. The syntax for creating data in a table is
INSERT INTO <tablename> (<column1 name>, <column2 name>....) VALUES (<value1>, <value2>....)
UPDATE is the command used to update data in a table. The syntax of update is
UPDATE <tablename> SET <column name> = <new value>
<column name> = <value>.... WHERE <condition>
Reading Data using SELECT Clause from a table in Cassandra. Using this clause we can read a whole table, a single column, or a particular cell. The syntax of SELECT is
SELECT FROM <table name> WHERE <condition>
Delete data from a table using the command DELETE. Its syntax is
DELETE FROM <identifier> WHERE <condition>

5.3.Performance Evaluation

One of the hallmarks of Cassandra is its high performance, for both reads and writes operations. When new nodes are added to a cluster, Cassandra scales it linearly. The performance of Hotel Management System application is evaluated with various hardware requirements such as Intel core CPU @ 1.80 GHz, 64-bit operating system, x64 based processor, 4.00GB RAM. The software specifications include Apache Cassandra version 1.2.19. Figure 3 gives performance of Cassandra operations.

In the graph of performance evaluation of Cassandra database X axis represents the throughput in ops/sec and Y axis represents average latency in ms. Here three operations such as update, insert and read are evaluated for performance. In the graph it is clear that update operation has very high throughput while it is in low latency. Similarly insert operation has high throughput [6] while it is in low latency which is greater than latency of update operation. In the case of read operation which has low throughput while it is in high latency.

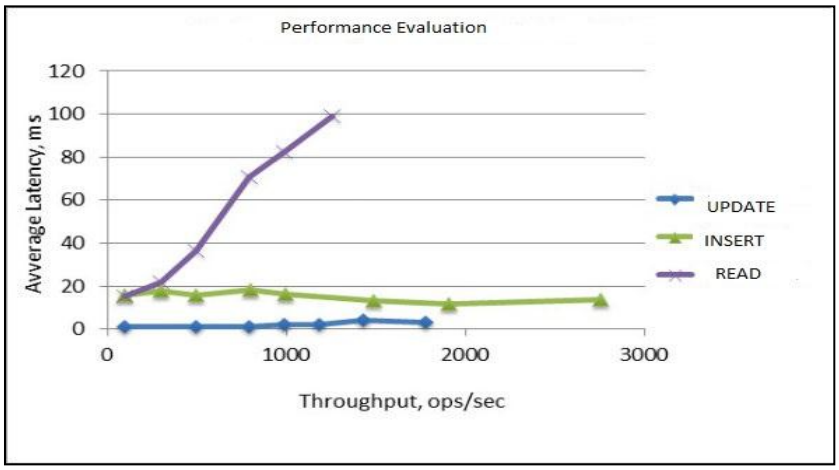


Figure.3.Performance Evaluation

6. CONCLUSION

NoSQL database: Cassandra is built, implemented, and operated a scalable storage system providing high performance, and wide applicability. Demonstrated that Cassandra can support a very high update throughput while delivering low latency. It is very efficient as compared with other databases.

REFERENCES

[1] http://cassandra.apache.org
[2] http://www.tutorialspoint.com/cassandra/
[3] Dietrich Featherston, Cassandra: Principles and Application, 2010
[4] A. Lakshman, P. Malik, Cassandra - A Decentralized Structured Storage System, Cornell, 2009.
[5] https://www.safaribooksonline.com/library/view/cassandra-thedefinitive/9781449399764/ch04.html
[6] Matthias Nicola and Matthias Jarke. Performance modeling of distributed and replicated databases.IEEE Trans. on Knowl.and Data Eng.,12(4):645–672, July 2000.

Authors

Varalakshmi P. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from P.R.S. College of Engineering and Technology, Thiruvananthapuram. Her areas of research are Data Mining, Databases and Image Processing.

Hima S. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from Mohandas College of Engineering and Technology, Thiruvananthapuram. Her areas of research are Image Processing, Database and Data Mining.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems and Distributed Computing, Machine learning. She has published 17 papers in international journals and international conference proceedings. She has been in the chair for many international conferences and journals.

DETECTING PACKET DROPPING ATTACK IN WIRELESS AD HOC NETWORK

Sneha C.S[1] and Bonia Jose[2]

[1]Student, Department of Computer Science and Engineering, MBITS Nellimattom and
[2]Assistant Professor, Department of Computer Science and Engineering, MBITS Nellimattom

ABSTRACT

In wireless ad hoc network, packet loss is a serious issue. Either it is caused by link errors or by malicious packet dropping. The malicious nodes in a route can intentionally drop the packets during the transmission from source to destination. It is difficult to distinct the packet loss due to link errors and malicious dropping. Here is a mechanism which will detect the malicious packet dropping by using the correlation between packets. An auditing architecture based on homomorphic linear authenticator can be used to ensure the proof of reception of packets at each node. Also to ensure the forwarding of packets at each node, a reputation mechanism based on indirect reciprocity can be used.

KEYWORDS

Packet dropping, Homomorphic linear authenticator, Auditor, Indirect Reciprocity

1.INTRODUCTION

In a wireless ad hoc network, nodes communicate with each other via wireless links either directly or relying on other nodes as routers. The nodes in the network not only act as hosts but also as routers that route data to/from other nodes in network. An adversary may misbehave by agreeing to forward packets and then failing to do so. Once being included in a route, the adversary starts dropping packets. That means it stop forwarding the packet to the next node. The malicious node can exploit its knowledge about the protocol to perform an insider attack. It can analyze the importance of the transmitting packet and can selectively drop those packets. Thus it can completely control the performance of the network.

If the attacker continuously dropping packets, it can be detect and mitigate easily. Because even if the malicious node is unknown, one can use the randomized multi-path routing algorithms to circumvent the black holes generated by the attack. If the malicious nodes get identified, the node can be deleted from the routing table of network. The detection of selective packet dropping is highly difficult. Sometimes the dropping of packets may not be intentional. It can be occurred as a result of channel errors. So the detection mechanism should be capable of differentiating the malicious packet dropping and the dropping due to link errors.

The algorithm introduced here provides an efficient mechanism to detect the selective packet dropping. It improves the detection accuracy by calculating the correlation between lost packets with the help of Auto Correlation Function of the bitmaps at each node in the route. Bitmap describes the lost/received status of each packet in the transmission. The basic idea is that even

DOI: 10.5121/ijci.2016.5213

though malicious dropping may result in a packet loss rate that is comparable to normal channel losses, the correlation pattern is different.

To get the correct correlation, the truthfulness of the packet loss bitmaps is essential. In order to ensure the correctness the system uses a public auditing mechanism. The auditor uses a variation of the cryptographic primitive called homomorphic linear authenticator (HLA) [2].It is a signature scheme widely used in cloud computing and storage server systems, which allows client that has storeddata at an untrusted server to verify that the server possesses the original data without retrievingit [3]. Indirect reciprocity is a powerful mechanism for the evolution of cooperation between nodes. The essential concept of indirect reciprocity is "I help you not because you have helped me but because you have helped others" [12].

The remainder of this paper is organized as follows. In Section 2 we review the related work. The system models and problem statement are described in Section3. We present the proposed mechanism in Section 4 and we conclude the paper in Section 5.

2.RELATED WORKS

Based on how much weight a detection algorithm gives to link errors relative to malicious packet drops, the works had been done to detect the malicious packet dropping can be broadly classified into two.

First category focuses on the detection with high malicious dropping rates, where the link errors are ignored. Based on the nature of the detection algorithm, this can be further classified into four. The first sub-category is based on credit systems [9].In this node gets incentive for its cooperation in transmission. When the node correctly transmits the packets to the next hop, it gets credit. Based on the credit value, the node gets priority during the transmission of its own packets. Thus, when the attacker continuously drops packets, its credit decreases and automatically gets expelled from the network. But when the attacker performs a selective dropping, it gets enough credits and can continue as a part of the network. The second sub category is based on reputation systems [4], [5], [6], [7]. In this mechanism the neighbour nodes monitor the activity of all nodes. For a node that drops packets maliciously gets a bad reputation. The reputation is the determining factor while selecting a route for transmission. Thus malicious nodes get excluded from a route. In this mechanism also, if the attacker selectively drop packets and forward some packets, then it can have a better reputation. The third sub category of works focus on the hop to hop acknowledgement, by which it can directly find out the misbehaving node. The fourth sub category uses cryptographic methods for the detection purpose. For example, the work in [8] utilizes Bloom filters to construct proofs for the forwarding of packets at each node. By examining the relayed packets at successive hops along a route, one can identify suspicious hops that exhibit high packet loss rates. But the incorrect proofs will reduce the detection accuracy of this mechanism.

The second category of works focus on the scenario where the number of maliciously dropped packets is significantly higher than that caused by link errors, but the impact of link errors is non-negligible. This type of mechanisms requires the knowledge of the wireless channel. The works in [9] and [10] proposed to detect malicious packet dropping by counting the number of lost packets. If the number of lost packets is significantly larger than the expected packet loss rate made by link errors, then with high probability a malicious node is contributing to packet losses. But counting the number of lost packets is not sufficient to detect the attacker. That is, if the

attacker selectively drop packet then the count of lost packet due to malicious node and the link may get equal.

All methods mentioned above do not perform well when malicious packet dropping is highly selective. But the detection of packet dropping using the correlation between lost packets gives better solution for selective packet dropping.

The methods in [14] delay a jammer from recognizing the significance of a packet after the packet has been successfully transmitted, so that there is no time for the jammer to conduct jamming based on the content/importance of the packet. Instead of trying to detect any malicious behavior, the approach in [14] is proactive, and hence incurs overheads regardless of the presence or absence of attackers.

3.SYSTEM MODEL AND PROBLEM STATEMENT

3.1.System Model

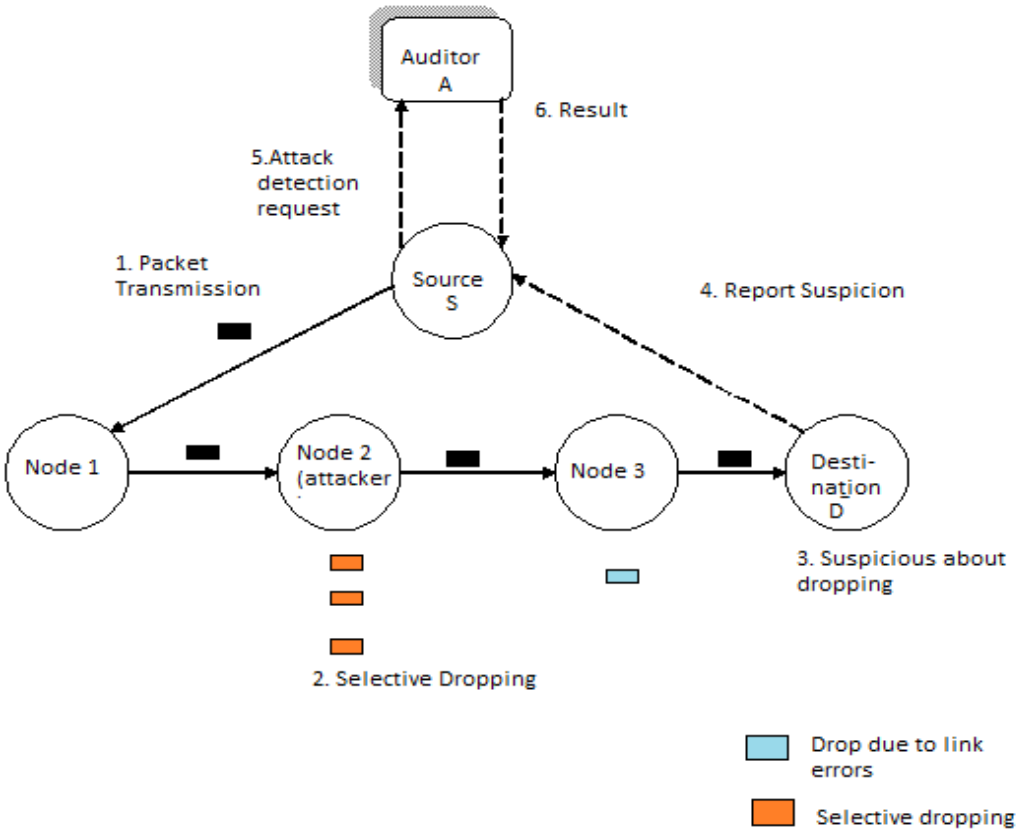


Figure1. System Model

Let P_{SD} be an arbitrary route in a wireless ad hoc network. The source S is aware of the path and it sends packets continuously to the destination D through P_{SD}. Consider that the network is quasi-static type. That means the network topology and link characteristics are constant for a relatively long period of time. Each hop that constitutes the path alternates between good and bad states. Packets transmitted during the good state are successful, and packets transmitted during the bad state are lost. By observing whether the transmissions are successful or not, the receiver obtains a realization of the channel state, which is a combination of zeros and ones. In that "1" denotes the packet was successfully received, and "0" denotes the packet was dropped.

When the receiver notifies some suspicious packet loss, it reports a feedback to the sender. The detection of malicious dropping is performed by an independent auditor module. After receiving the feedback from the receiver, sender requests the auditor to perform detection. The auditor module identifies the malicious dropping by checking the correlation between lost packets at each node. The correlation between lost packet in selective dropping condition and link error condition is different [1]. For this, the information collected by the auditor will be accurate. In order to ensure that the packet received by a node, the mechanism proposed here uses a homomorphic linear authenticator. Also, to ensure the packet forwarding, it uses a reputation based mechanism which uses an indirect reciprocity framework based on evolutionary game theory, described in [11].

3.2.Problem Statement

The adversary, which is a node in the path, may try to degrade the performance of the system by dropping the packets send by the source. The node can perform the dropping selectively or randomly. The detection should be done by an independent auditor module. While performing detection it should verify the correctness of collected information. Also, should produce a publically verifiable proof of the misbehaviour of the node.

Besides this there is a chance for collusion between two nodes. A covert communication channel may exist between any two malicious nodes, in addition to the path connecting them on PSD. As a result, malicious nodes can exchange any information without being detected by Ad or any other nodes in PSD. Malicious nodes can take advantage of this covert channel to hide their misbehavior and reduce the chance of being detected.

4.DETECTION OF PACKET DROPPING

4.1.Overview

The detection mechanism focuses on the correlation between the lost packets at every node in the transmission route. While the sender S transmitting the packets consecutively, each hop in the path will keep a transmission bitmap for every packets. The bitmap is a pattern of 0 and 1, where 1 represents the successfully transmitted packet and 0 represents the unsuccessfully transmitted packets. By using an Auto Correlation Function (ACF), the correlation between these bitmaps can be calculated. Under different packet dropping conditions the correlation function will generate different values. Thus by observing the correlations between lost packets, one can decide whether the packet loss is purely due to regular link errors, or is a combined effect of link error and malicious drop.

But the main challenge is that the packet-loss bitmaps reported by individual nodes along the route may not be correct. For the correct calculation of the correlation between lost packets the truthfulness of bitmap is necessary. This can be achieved by auditing functionality. Auditing can be done by using a cryptographic primitive called homomorphic linear authenticator (HLA), which is a signature scheme to provide a proof of storage from the server to entrusting clients in cloud computing and storage server systems. Besides this to ensure the forwarding, a reputation based mechanism can be used. When a node relays packet successfully, it gets a good reputation from the receiving node. That means, in a path from sender to receiver, the node with minimum reputation dropped more packets.

4.2.System Architecture

In a wireless ad hoc network, the source S is supposed to send the packets to the destination D continuously, through the wireless channel P_{SD}. Here we are considering the quasi-static networks. So the path P_{SD} remains unchanged for a long time. While receiving a sequence of packet, the receiver gets a realization of channel state simply by observing whether the transmissions are successful or not. Successfully received packets are denoted by 1s and others are denoted by 0s. Each node in P_{SD} will also provide a reputation to the relying node when it gets a packet.

There is an auditor A_d in the network. It is not associated with any node and kept as independent. It is totally unaware of the secrets shared between nodes in the path P_{SD}. The detection of malicious packet dropping is performed by this auditor. When the receiver finds out some abnormality in the reception of packets, it will report the suspicion to the source. Once being notified the source send submits an attack-detection request (ADR) to the Auditor.

For the detection of attack, the auditor will collect the information about transmission from each node on the path P_{SD}. The auditor needs to verify authenticity of the collected information. Once the truthful information is collected from every node in the route, the auditor calculates the correlation between them. From this information, it can detect the attack.

4.3.Scheme Details

The system consists of four Phases:

i. Setup Phase

ii. Packet Transmission Phase

iii. Audit Phase

iv. Detection Phase

4.3.1.Setup Phase

Immediately after establishing the route, the setup phase gets started. The source decides on symmetric key crypto system for encryption the packet during the transmission phase. Source securely distributes a decryption key and a symmetric key to each node on the path. Key distribution may be based on the public-key crypto-system. The source also announces two hash functions to every node in the route. Besides this, source also needs to set up its HLA keys.

4.3.2.Packet Transmission Phase

After the successful completion of Setup phase, source enters into the transmission phase. In this phase, before the transmission of packets source computes the hash value of each packet and generates HLA signatures of the hash value for each node. These signatures are then sent together with the packets to the route by using a one-way chained encryption. This prevents the deciphering of the signatures for downstream nodes by the upstream node. When a node in the route receives the packet from source it extracts packets and signature. Then it verifies the integrity of received packet. A database is maintained at every node on P_{SD}. It can be considered as a FIFO queue which records the reception status for the packets sent by source. Every node stores the received hash value and signature in the database as a proof of reception.

To ensure the relying at each node an indirect reciprocity framework based on evolutionary game theory can be used. In this method each node is considered as a player. Generally, helping someone establishes a good reputation, and will be rewarded by others. In this paper, we adopt the reputation updating rule of indirect reciprocity in [12], i.e., the reputation of relay is updated according to the following rule:

	G	B
F	G	G
D	B	G

where a relay who takes the choice X($X \in \{F,D\}$) towards a provider with reputation R ($R \in \{G,B\}$) will be assigned a new reputation R(R;X) ($R \in \{G,B\}$). Here, we adopt the reputation updating such that cooperation leads to a good reputation, whereas defection leads to a bad reputation unless the opponent is a bad player. The total value of reputation can be calculated by subtracting bad reputation from good.Nodes will also keep another database to keep the reputation value.

4.3.3.Audit Phase

When the source issues an attack detection request (ADR), the audit phase gets started. The ADR message includes the id of the nodes on the route, source' s HLA public key information, the sequence numbers of the packets sent by source, and the sequence numbers packets that were received by destination . The auditor requests the packet bitmap information from each node in the route by issuing a challenge. From the information stored on the database, every node generates this bitmap. Auditor checks the validity of bitmaps and accepts if it is valid. Otherwise it rejects the bitmap and considers the node as a malicious one.

This mechanism only guarantees that anode cannot understate its packet loss, i.e., it cannot claim the reception of a packet that it actually did not receive. This mechanism cannot prevent a node from overly stating its packet loss by claiming that it did not receive a packet that it actually received. This latter case is prevented by the mechanism based on reputation which is discussed in the detection phase

4.3.4.Detection Phase

After auditing the reply to the challenge issued by the auditor, it enters into the detection phase. Auditor constructs per hop bitmaps and by using an auto correlation function (ACF) it will find out the correlation between the lost packets. Then it finds out the difference between the calculated value and correlation value of wireless channel. Based on the relative difference, it decides whether the packet loss is due to the malicious node or link error. When it finds out malicious drop, it can consider both ends of the hop as suspicious. That means either the transmitter did not send the packet or receiver did not receive.

After identifying these two suspicious nodes, the detector needs to find out the actual culprit. For this, it can check the reputation value. Now the Auditor module will collect the reputation value for the two suspicious nodes. When a node fails to forward the packet it, it will get minimum reputation. By checking this, the detector can easily distinguish the attacker.

5.CONCLUSIONS

In order to detect the malicious node that drops the packets intentionally, the technique described here utilizes the correlation between the lost packets at each node in the route from source to destination. . For this, uses a public auditing architecture. This mechanism will give a satisfactory improvement in the detection accuracy of selective packet dropping. To correctly calculate the correlation between lost packets, it requires truthful packet loss information from every node in the route. Auditor ensures the integrity of packet loss information of each individual node by using Homomorphic Linear Authenticator (HLA). HLA-based public auditing architecture ensures truthful packet-loss reporting by individual nodes. This architecture is collusion proof, requires relatively high computational capacity at the source node, but incurs low communication and storage overheads over the route.

Based on the indirect reciprocity mechanism, we have theoretically analyzed the evolutionary dynamics of cooperative strategies. The reputation mechanism will ensure the correct forwarding process. Due to the evolutionarily stable strategies based on indirect reciprocity is effective and robust against packet loss and imperfect estimation of reputation.

REFERENCES

[1] Tao Shu, Marwan Krunz, "Privacy – Preserving and Truthfull Detection of Packet Dropping Attacks in Wireless Ad Hoc Networks",April 2015.

[2] C. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores," in Proc. ACM Conf. Comput. andCommun. Secur., Oct. 2007,pp. 598–610.

[3] C. Wang, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for data storage security in cloud computing," in Proc. IEEE INFOCOM Conf., Mar. 2010, pp. 1–9

[4] B. Awerbuch, R. Curtmola, D. Holmer, C. Nita-Rotaru, and H.Rubens, "ODSBR: An on-demand secure byzantine resilient routing protocol for wireless ad hoc networks," ACM Trans. Inform. Syst. Security, vol. 10, no. 4, pp. 1–35, 2008.

[5] S. Buchegger and J. Y. L. Boudec, "Performance analysis of the confidant protocol (cooperation of nodes: Fairness in dynamic adhoc networks)," in Proc. 3rd ACM Int. Symp. Mobile Ad Hoc Netw.Comput. Conf., 2002, pp. 226–236.

[6] W. Galuba, P. Papadimitratos, M. Poturalski, K. Aberer, Z. Despotovic, and W. Kellerer, "Castor: Scalable secure routing for ad hoc networks," in Proc. IEEE INFOCOM, Mar. 2010, pp. 1 –9.

[7] Q. He, D. Wu, and P. Khosla, "Sori: A secure and objective reputation-based incentive scheme for ad hoc networks," in Proc. IEEE Wireless Commun. Netw. Conf., 2004, pp. 825–830.

[8] W. Kozma Jr., and L. Lazos, "REAct: Resource-efficient accountability for node misbehavior in ad hoc networks based on random audits," in Proc. ACM Conf. Wireless Netw. Secur., 2009, pp. 103–110.
[9] T. Hayajneh, P. Krishnamurthy, D. Tipper, and T. Kim, "Detecting malicious packet dropping in the presence of collisions and channel errors in wireless ad hoc networks," in Proc. IEEE Int. Conf. Commun., 2009, pp. 1062–1067.
[10] S. Zhong, J. Chen, and Y. R. Yang, "Sprite: A simple cheat-proof,credit-based system for mobile ad-hoc networks," in Proc. IEEE INFOCOM Conf., 2003, pp. 1987–1997.
[11] Changbing Tang, Ang Li, and Xiang Li, "When reputation enforces evolutionary cooperation in unreliable MANETs", Nov. 2014
[12] M. A. Nowak, and K. Sigmund, "Evolution of indirect reciprocity," *Nature*, vol. 437, pp. 1291 1298, Oct. 2005.
[13] H. Ohtsuki, Y. Iwasa, and M. A. Nowak, "Indirect reciprocity provides only a narrow margin of efficiency for costly punishment," *Nature*, vol.457, pp. 79-82, Jan. 2009.
[14] A. Proano and L. Lazos, "Selective jamming attacks in wireless networks," in Proc. IEEE ICC Conf., 2010, pp. 1–6.

AUTHORS

Sneha C.S currently pursuing M.Tech in Computer Science and Engineering in MBITS Nellimattom. She received B.Tech in Computer Science and Engineering from MBITS, Nellimat tom, M.G University, Kottayam,India in 2013. Her area of interest includes cyber security.

Bonia Jose received **B.Tech**Degree in Computer Science and Engineering from Viswajyothy College of Engineering and Technology, Vazhakkulam and **M.Tech** fromKarunya University, C oimbatore. She is currently working as Assistant professor at MBITS Nellimattom. She is specialized in Networking and Internet Engineering.

DIGITAL INVESTIGATION USING HASH-BASED CARVING

Isabel Maria Sebastian, Noushida A, SafaSaifudeen and Surekha Mariam Varghese

Department of Computer Science and Engineering,
Mar Athanasius College of Engineering, Kothamangalam,Kochi, India

ABSTRACT

File carving is a popular method used for digital investigations for detecting the presence of specific target files on digital media. Hash based sector hashing helps to identify the presence of a target file. The hashes of physical sectors of the media is compared to the database of hashes created by hashing every block of the target files. To enable this, instead of evaluating the hashes of entire files, the hashes of individual data blocks is used for evaluation. Hash-based carving helps to identify fragmented files, files that are incomplete or that have been partially modified. To address the problem of High false identification rate and non-probative blocks, a HASH-SETS algorithm that can help in identification of files and the HASH-RUNS algorithm that helps in reassembling the files is used. This technique is demonstrated using the forensic tool: bulk_extractor along with a hash database: the has hdb and an algorithm implementation written in Python.

KEYWORDS

Forensics investigation, hash-based carving, HASH-SETS, HASH-RUNS algorithms, Sector hashing

1.INTRODUCTION

Digital forensics is a branch of forensic science in corporating the investigation and recovery of evidences found in digital devices, often in relation to computer crime[1].Crime scene investigations involve a scientific examination of the evidences. Digital evidences are mainly used in cases where direct evidences such as eye witnesses are not available or not feasible. File creation and modification dates, internet history cache, emails, chats,windows registery,recyclebin, shadow copies, log files etc. are some of the artifacts that can be extracted and used for forensic investigations[2]. A common method in digital forensics is to search for known files in an accused media. Brute force approach is to compare the known file contents with sectors of disk. Forensic practitioners' often use hash database of file segments to locate the content of sectors on disk with the known content.

Locating files containing child-pornography is a use-case for digital forensics using file carving. This paper presents hash based file carving method for detecting fragments of movies or images in the media.

DOI: 10.5121/ijci.2016.5214

2.RELATED WORK

Hash based file carving was introduced by Garfinkel as part of the solution to the DFRWS 2006 Carving Challenge. Using "the MD5 trick." the extracted text from the carving challenge was used to identify the original target documents on the Internet. He used a manual method to identify the location of the target files i.e. block-hashed the target files and sector-hashed the Carving Challenge and manually matched the two sets[1]'

The term "hash-based carving" was introduced by Collangeet al.in 2009in a pair of publications [3, 4] that explored the use of GPUs to speed the hashing load. But the question of what kind of database would be required to look up hashes produced at such a rate, and how to match files given the fact that the same block hashes appear in many different files was not discussed[1].
Later in 2013,Keydeveloped the File Block Hash Map Analysis (FBHMA) EnScript. This was a tool that createsa hash-map of file blocks from a master file list and selected areas of a target drive was searched for the blocks. But only a few files can be searched by FBHMA at a time.
To serve the purpose of enabling research on file-based digital forensics,Garfinkel et al. (2009) created the GOVDOCS corpus. Also the M57-Patents scenario, a collection of disk images, memory dumps, and network packet captures from a fictional company called M57 was created by Garfinkel et al. (2009).

Shortly, distinct blocks were used to detect the presence of known files on target media. This technique, though discussed by Garfinkel et al. (2010)presented no algorithms for doing so. So in 2012, Foster attempted to determine the reason for the repetition of 50 randomly chosen blocks that were shared between different files in the GOVDOCS corpus.

This idea was expanded by Young et al. (2012) by using a database of distinct block hashes of target files to detect the presence of those files on media being searched. The initial implementation of the hashdb hash database was described in this article [1].

The coverage/time trade-offs for different sample sizes were examined by Taguchi while using random sampling and sector hashing for drive triage. He concluded thatin a wide variety of circumstances, 64 KiB is an optimal read size.

Different file carving taxonomy was proposed by SimsonGarfinkel and Joachim Metz. These methods differed in the way they analyzed their inputs. In block based carving method, the input is analyzed block-by-block to determine if a block is part of a possible output file, on the assumption that each block can only be part of a single file (or embedded file). In statistical carving, the input is analyzed on basis of characteristics and statistics. Other relevant carving methods are Fragment Recovery Carving, File structure based Carving, Header/Footer Carving etc.

3.BACKGROUND

"Hash-based carving" describes the process of recognizing a target file on a piece of searched media by hashing same-sized blocks of data from both the file and the media and looking for hash matches [1]. To search for known contents, the MD5 hash algorithm is used because of its speed.

3.1.Tools Used

Hashdb is the tool that is used for finding previously identified blocks of data in a media such as disk images [9]. The following services are provided by hashdb:

- By importing block hashes and providing lookup services, the hashdb manages block hash databases.
- Hash databases can be created or scanned by other programs using the hashdb library.
- libhashdb of the hashid scanner which is a bulk_extractor plugin,can be used to search for previously identified blocks of data .

Cryptographic hashes (along with their source information) that have been calculated from hash blocks are stored by hashdb. Instead of full file hashing, it relies on block hashing.In block hashing, artifacts are identified at the block scale (usually 4096 bytes). Many of the capabilities of hashdb are best utilized in connection with the bulk_extractor program.

Bulk Extractor (bulk_extractor) is a feature extraction tool written in C++ for extracting features from media images [8].The input which can be a disk image, files or a directory of files, is split into pages and processed by one or more scanners. Bulk Extractor parallel-processes 16MiByte pages of media on multiple cores. 4KiB sectors are recommended for hashing if they are directly supported by the drive. Otherwise combine 8 adjacent 512B sectors into a single 4096-byte super-sector for hashing and for feature extraction.The feature files stores the extracted features so that they can be easily inspected, parsed, or processed with automated tools. These feature files are processed using Python programs.

Python which is a general-purpose, high-level programming language is used for the implementation of "hash-sets" and "hash-runs" algorithm. It is an open-source software.Python is designed to be highly readable. Major advantage of this scripting language is that it has fewer syntactical constructions and best suits a beginner.

Another important tool is md5deep. It is a set of programs used to compute MD5 message digests on an arbitrary number of files. The MD5 hash algorithm (developed by Ronald Rivest) which is an updated version of MD4, is used to generate block hashes because of its speed.It helps to compare and store these small hashes more easily. Alsoin cryptography, verification is done using one-way hashes.

MD5 is just one of many hashing algorithms. There's also SHA-1 (Secure Hash Algorithm), developed by NIST in conjunction with the NSA and produces 160 bit digest. It's more collision resistant than MD5 and is the most commonly used hash for cryptographic purposes these days. Though other hashing algorithms are available, MD5 and SHA-1 are commonly used in cryptography.

3.2.A hash-based carving process

Hash-based carving is a four-step process:

- DATABASE BUILDING: A database of file block hashes is created from the target files.

- SCANNING THE MEDIA: A set of hash values that matched the target files is produced byscanning the searched media. Here we hash 4KiB sectors and searchfor those hashes in the database.
- PROBATIVE BLOCK SELECTION: Most probative (likely) target files on the searched media is determined.
- TARGET ASSEMBLY: Runs of the matching candidate blocks on the searched media are identified and mapped to the corresponding target files.

4.MATCHING SCENARIOS

The basic idea of hash-based carving is to compare the hashes of 4KiB sector clusters from a storage device ("search hashes") with the hashes of every 4KiB block from a file block hash database and identify files based on hashes that the two sets have in common.

- Hash matches can be observed in a variety of scenarios, including: The presence of a copy of an target file, in a complete form on the searched media.
- The presence of a copy of the target file on the searched media at some time in the past which has been deleted at a later point, which may or may not be partially overwritten.
- The presence of a file on the searched media that may have many sectors in common with the required target file.
- The presence of a target file which is embedded in a larger carrying file, detected only if that the file is embedded on an even sector boundary.

5.IMPLEMENTATION

First step in the implementation of hash based carving process is the construction of a database. Here the overlapping 4KiB sectors are hashed using the MD5 tool. Then we scan the disk image of the media using bulk extractor. The output of this stage produces different .txt files. From the obtained output, identify the candidate blocks using probative block selection tests like the ramp test, the white-space test, the 4-Byte histogram test etc. On these selected blocks, apply the HASH-SETS and HASH-RUNS algorithm to determine whether the media contains the required target files. Detailed description of each stage is given below.

5.1.Database Building: creating the target hashdb

A hashdb database that contains the 4KiB block's hash values is created. The output database is renamed with .hdb extension. To build the database, run hashdb from the command line:

- hashdb create sample.hdb

Here sample.hdb is an empty database. ADFXMLfilecontainingsector hash values is required to import data into the database. To populate the hash database with the hashes from the DFXML file called sample.xml:

- hashdb import sample.xml sample.hdb

This command, if executed successfully, will print the number of hash values inserted. For example:

hashdb changes (insert): hashes inserted: 2595

5.2.Media Scanning: Finding Instances Of Known Content On The Searched Media

Here the database is used to search the disk image using the bulk extractor has hdb scanner. Bulk_extractor breaks the disk image into 16MiB "pages" and only processes pages that are not blank. Each page is broken into overlapping 4KiB blocks, each block is hashed with the MD5 algorithm and the resulting hash is used to query the block hash database. Matching sector hashes are reported in bulk_extractor's identified_blocks.txt file. Count value indicates whether each hash matches one or more target files. Matched hashes are used in both the "candidate selection" and "target assembly" phases.

To run bulk_extractor from the command line, type the following command:

- bulk_extractor -o output mydisk.raw

In the above command, output is the directory that will be created to store bulk_extractor results. It cannot already exist. The input mydisk.raw is the disk image to be processed. After the run the resulting output files will be contained in the specified output directory. Open that directory and verify files have been created. There should be 15-25 files. Some will be empty and others will be populated with data. These two steps are combinable using a batch file that incorporates the command line commands for running the hashdb and the bulk extractor.

5.3.Candidate Selection: Identifying Probative Blocks

The hashdb'sexplain_identified_blocks command is used to determine the files to which these block hashes correspond. This is implemented using a data reduction algorithm. For each block inidentified_blocks.txt file, if the block maps to fewer than N files (the default is 20), those files are added to the set of candidate files. To run it from the command line, type the following instructions

- hashdbexplain_identified_blockssample.hdb out/identified_blocks.txt> out/explain_identified_blocks.txt.

The program writes out a new file called explain_identified_blocks.txt with a list of the deduplicated sector hashes, the number of times that the hash appeared in the original file, and all of the source files in which the hash appears.

Next, run the program report_identified_runs.py program which writes theidentified_blocks_explained.txt file.This is implemented using the explain_identified_blocks.txt output file: if the hash is not flagged by any of the ad hoc tests like the ramp test, the white-space test, the 4-Byte histogram test (they identify the non-probative blocks), the hash's sources are added to the set of candidate files.

5.4 .Target Matching:

After selecting the candidate files, the block hashes corresponding to candidates are grouped into source files. They are eventually tallied or reassembled into runs of matching blocks with the HASH-SETS and HASH-RUNS algorithms.

5.4.1.Hash-Sets: Reporting The Fraction Of Target Files

HASH-SETS is a simple, memory efficient algorithm that employs block hashes to generate a report of the fraction of blocks associated with each target file that is found on the searched media [1].

It is implemented as follows:

1. A list of candidate targets are determined by employing the candidate selection algorithm.
2. Repeat for each block hash H in the file identified_blocks_explained.txt:
    (a) Repeat for each target T matching against H:
        i). If T is a Candidate target, add 1 to the score of that particular target.
3. Repeat for each target T
    (a) To compute the fraction of the file present, divide the score by the count of number of blocks in the targetfile.
4. Sort the targets in inverse order of the fraction of the file present so that a list with highest fraction first is obtained.
5. If the fraction recovered exceeds the threshold, report the target file name, number of recovered blocks, and the fraction of the file recovered. The number of blocks recovered is counted by the algorithm and so the score is solely a function of the target file and the searched drive.

Here the number of blocks is counted by the algorithm so that the score would be solely a function of the target file and the searched drive.

5.4.2.Hash-runs: locating target files

The presence of target files is detected by the HASH-SETS algorithm and the HASH-RUNS algorithm is used to report the location.

- Handling of the case when the target file is on the searched media in multiple locations.
- Employs the placement of adjacent logical sectors in a file in adjacent physical sectors of the searched media to its advantage.
- Accounts for different blocks in a file having the same (mod 8) value.
- Detects runs of recognized blocks being separated by null blocks and combines them for efficiency.

In this algorithm, first the data structures are created by the HASH-SETS implementation. It then identifies all of the block runs on the physical disk that correspond to logical runs of the target files [1]. Using the logical block number in the target file, these blocks are sorted by and reported to the analyst.

1. First the identified_blocks _explained.txt file is read by the algorithm (previously produced by the hashdb program). Then, an in-memory database is built that maintains the set of sector hashes associated with each target file.
2. A second database is built by the algorithm for each (target file, sector hash) pair to record the set of block numbers in the target file where the block appears. So if target file A has 6 blocks, and both blocks 1 and 4 match sector hash H1, then the element (A,H1) of the database contains the set {1,4}.
3. Next, the algorithm reads the identified_blocks.txt file. For each sector hash that was found in a target file, it records the disk block at which the hash was found.
4. Finally, there comes the target matching step. For each (target file, mod8) combination:

a) The algorithm builds an array consisting of elements in the form:

[diskblock,{fileblocks},count] Where disk block is the physical disk block, { file blocks} is a set of blocks within the target file where the hash was found, and the number of times in the sector hash database that the sector hash was found.

b) The array is sorted by diskblock.
c) The algorithm runs a sliding window over the rows of the array. It helps to identify

rows that represent sequential disk blocks and file blocks.. A new array of block runs is created, where each element in the array has the values:

- Identified File (from the hash database)
- Score (The number of identified blocks)
- Physical sector start
- Logical block start
- Logical block end

d) If the number of bytes in the sector gap between the two runs matches the number of bytes in the logical blocks, and if all of the sectors on the drive corresponding to the gap contain only NULLs, then the two block run elements are combined.
e) Drop the block runs that are smaller than a predetermined threshold.
f) Finally, for every reported run, use the SleuthKit to determine the allocated or deleted file that corresponds to the first block in the run.

The program's output is obtained in the form of a CSV file that can be readily imported into Microsoft Excel.

6.CONCLUSION

A hash-based carving system implementation is presented here. The bulk_extractor program is used by the carving system to create a block hash database of target files. Also, a sector hash record of searched media, supported by the hashdb hash database is created. Individually recognized block hashes are assembled into carved runs which is performed by a post-processing Python script.

7.IMPROVEMENTS

During the database construction using hashdb, the non-probative blocks can be flagged. These flag status can be stored in a database to help in the probative-block selection process. This improves the overall efficiency. Also, files composed entirely of non-probative blocks can be avoided in the earlier stages itself. File allocation status can be considered along with the (mod 8) value in the target matching. Though this is an expensive filtering step, it contribute a little to the accuracy rate. For an encrypted file system, the media is first mounted on an appropriate decrypting driver and later the unencrypted file is accessed. Also, other hashing techniques like SHA-1 can be used for block hashing.

REFERENCES

[1] SimsonL.Garfinkel& Michael McCarrin (2015), "Hash-based carving: Searching media for complete files and file fragments with sector hashing and hashdb", Digital Investigation 14(2015) 895e8105
[2] https://forensicswiki.org
[3] Collange S, Dandass YS, Daumas M, Defour D. Using graphics processors for parallelizing hashbaseddata carving. CoRR abs/0901.1307. 2009.,http://arxiv.org/abs/0901.1307.
[4] Collange S, Daumas M, Dandass YS, Defour D. Using graphics processors for parallelizing hash-based data carving. In: Proceedings of the 42nd Hawaii International Conference on System Sciences; 2009. Last accessed 03.12.11, http://hal.archives-ouvertes.fr/docs/00/35/09/62/ PDF/ColDanDauDef09.pdf.
[5] Allen B. hashdb. 2014. https://github.com/simsong/hashdb.git.
[6] http://booksite.elsevier.com/samplechapters/9780123742681/Chapter_6.pdf
[7] http://simson.net/ clips/academic/2012.IEEE.SectorHashing.pdf.
[8] https://forensicswiki.org/bulk_extractor
[9] https://forensicswiki.org/hashdb
[10] https://en.wikipedia.org/wiki/Digital_forensics
[11] http://www.tutorialspoint.com/python/
[12] https://github.com/NPS-DEEP/hashdb/wiki
[13] http://sourceforge.net/p/guymager/wiki
[14] http://forensicswiki.org/wiki/Famous_Cases_Involving_Digital_Forensics#

Authors

Isabel Maria Sebastian is currently pursuing B.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering.

Noushida A is currently pursuing B.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering.

SafaSaifudeen is currently pursuing B.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems, Machine Learning and Distributed Computing. She has published 17 papers in international journals and international conference proceedings. She has been in the chair for many international conferences and journals.

INTENTIONAL BLANK

DOUBLE PRECISION FLOATING POINT CORE IN VERILOG

Aparna CV[1] and Mary Joseph[2]

[1]Department of electronics and communication engineering, MACE, Kothamangalam, A P J Abdul Kalam Technological University, Kerala, India

[2]Associate Professor, Department of Electronics and Communication, M. A College of Engineering, Kothamangalam

ABSTRACT

A floating-point unit (FPU) is a math coprocessor, a part of a computer system specially designed to carry out operations on floating point numbers. The term floating point refers to the fact that the radix point can "float"; that is, it can placed anywhere with respect to the significant digits of the number. Double precision floating point, also known as double, is a commonly used format on PCs due to its wider range over single precision in spite of its performance and bandwidth cost. This paper aims at developing the verilog version of the double precision floating point core designed to meet the IEEE 754 standard .This standard defines a double as sign bit, exponent and mantissa. The aim is to build an efficient FPU that performs basic functions with reduced complexity of the logic used and also reduces the memory requirement as far as possible.

KEYWORDS

IEEE 754, ModelSim, Double precision floating point format, Verilog

1.INTRODUCTION

In computing, double precision is a computer number format that occupies two adjacent storage locations in memory of computer. A double precision number, sometimes simply called the double may be defined as integer, fixed point or floating point. 32 bit modern computers use two memory locations to store 64 bit double precision number. Double precision floating point is an IEEE 754 standard used to encode binary or decimal floating point numbers in 64 bits (8 bytes).[1].

In computing floating point is a method of used to represent real numbers in a way that can support a wide range of values. Numbers in general represented as a fixed number of significant digits and scaled using an exponent. The base for this scaling is normally 2, 10 or 16.The typical number that can be represented exactly is of the form

$$\text{Significant digit} * base^{exponent}.$$

The term floating point refers to the fact that the radix point can float i.e., it can be placed anywhere to the significant digits of the number. This position is indicated separately in the

internal representation and the floating point representation can thus be thought of as a computer realization of scientific notations. Since the 1990s, the most commonly encountered floating point representation is that defined by IEEE 754.

Verilog is standardized as IEEE 1364.it is a hardware description language used to model electronic systems. It is most commonly used for digital circuit design and verification at the register transfer level of abstraction. It is used in the verification of analog circuits and mixed signal circuits. [2] A subset of statement in verilog language is synthesizable. Verilog modules that obey a synthesizable coding style known as RTL (Register Transfer Level) .A synthesis software can be used to physically realize RTL. The Synthesis software algorithmically transforms the verilog source into a netlist, a logically equivalent description including only of elementary logic primitives [AND, OR etc] that are available in a specific FPGA or VLSI technology. Further manipulations to the netlist ultimately lead to a circuit fabrication blue print.

The Double precision floating point core in verilog was designed with three objectives in mind. First, develop efficient algorithms for Floating Point operations like addition, subtraction, division, multiplication, rounding and exception handling. Second, implement the proposed algorithm using Verilog. Third synthesize the above proposed algorithm.

2.PROPOSED SYSTEM

In the early days of digital computers, it was quite common that machines from different vendors have different word lengths and unique floating-point formats. This caused many problems, especially in the porting of programs between different machines (designs). The IEEE-754 floating point standard, formally named NSI/IEEE Std 754-1985, introduced in 1985 tried to solve these problems.

This paper implements a floating-point system confirming to this standard can be realized in software, entirely in hardware, and combination of hardware and software. The standard specifies two formats for floating-point numbers, basic (single precision) and extended (double precision), it also specifies the basic operations for both formats which are addition and subtraction of operations. Then different conversions are needed, as integer to floating-point, basic to extend and vice versa. Finally, it describes the different floating-point exceptions and their handling, including non-numbers (NaNs).

2.1.IEEE 754 Standard

The IEEE 754 standard is a technical standard established by IEEE and the most widely standard for floating point computation. It was created in the early 1980s after the introduction of word sizes of 32 bits (or 16 or 64). it based on a proposal from Intel who were designing the 8087 numerical coprocessor. Prof. W. Kahan was the first architect behind this proposal, for which he was awarded the 1989 Turing award. Almost all the modern machines follows IEEE 754 standard. Notable exceptions include IBM, main frames and Cray vector machines. Where the T90 series add an IEEE version but the sv1 still uses Cray floating point format.

2.2.Verilog

Verilog, standardized as IEEE 1364, is a hardware description language used for modeling electronic systems. It is most commonly used for digital circuits design and verification at the register transfer level of abstraction. It also used in the verification of analog circuits and mixed signal circuits. A subset of statement in verilog language is synthesizable. Verilog modules that conform to a synthesizable coding style known as RTL. RTL can physically realized by synthesis software. Synthesis software algorithmically transforms the verilog source into netlist; it is a logically equivalent description consisting of elementary logic primitives [AND, OR etc] which are available in FPGA and VLSI technology. Further manipulations to the netlist ultimately lead to a circuit fabrication blue print (such as photo, mask set for an ASIC or a bit stream fill for an FPGA).Verilog is very much compatible with C language.Its control flow keywords (if/else, for, while, case, etc.) are equivalent, and its operator precedence is compatible. So it's easy to work with verilog by knowing C language. And the main reason to choose verilog as HDL language is to learn new language.

2.3.ModelSim

ModelSim is a multi-language HDL simulation environment by Mentor Graphics, for simulation of hardware description languages such as VHDL, Verilog etc.ModelSim can be used independently, or in conjunction with Altera Quartus or Xilinx ISE. Simulation is performed using the graphical user interface (GUI), or automatically using scripts.ModelSim is offered in multiple editions, such as ModelSimPE, ModelSimSE and ModelSimPE.ModelSimcan also be used with MATLAB/Simulink, using Link for ModelSim.

2.4.Double precision floating-point format

Double precision is a computer numbering format in which it uses two adjacent storage locations in computer memory. The IEEE 754 standard definition of a double is:

--Sign bit: 1 bit

--Exponent width: 11 bits

--Significant precision: 53 bits (52 explicitly stored)

The real values are assumed by a given 64-bit double-precision data with a given biased exponent e and a 52-bit fraction is:

$$-1^{sign} * \ 2^{\exp - 1023} * 1.mantissa$$

Using floating-point variables and mathematical functions (sin (), cos (), log (),

Exp (), sqrt () are the most popular ones) of double precision as opposed to single precision comes at the execution cost: the operations with the double precision are usually slower.

Typically, a floating-point operation takes two inputs with p bits of precision and returns a p-bit result. The ideal algorithm would compute this by first performing the operation exactly, and then rounding the result to p bit (using the current rounding mode).

The algorithms for various operations are given below. According to which the program coded in Verilog.

2.4.Implementation

A floating-point operation uses two inputs with p bits of precision. The ideal algorithm compute the result by first performing the operation, and after that rounding the result to p bit (using the rounding mode).As an example, the operation of the system is explained in next section with addition and subtraction operation.

2.4.1.Block diagram

The double precision floating point adder / subtractor performs addition, subtraction operations. The block diagram of the current system (floating point unit double) is shown in figure (Fig.1)

For addition, the input operands are separated into their mantissa and exponent components, and the larger operand stored into mantissa large and exponent large, the other operand, i.e. the smaller operand populating mantissa small and exponent small. The comparison of the operands to find out which is larger operand only compares the exponents of the two operands, if the exponents are equal, the smaller operand should populate the mantissa large and exponent large registers. It is not an issue because the operands are compared to find the operand with the larger exponent, so that the operand with smaller exponent's mantissa can be right shifted before performing the addition. If the exponents are equal, the mantissas can be added without shifting
The block diagram can be explained more elaborately with two operands naming A and B. Such explanation is provided with the Algorithm part in forthcoming section.

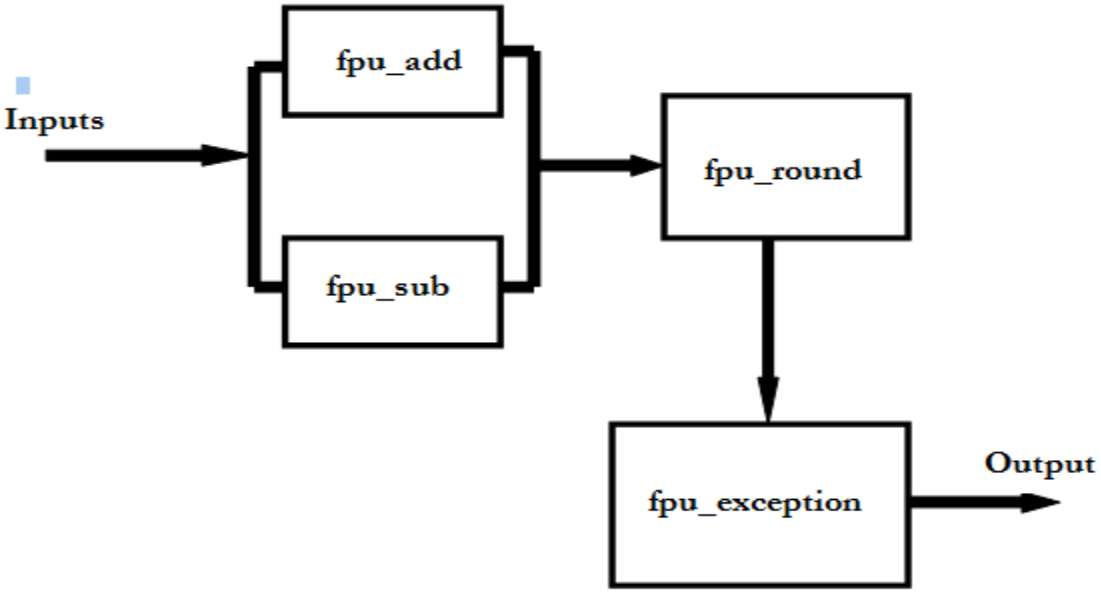


Fig.1Block Diagram of Floating Point Adder/Subtractor

2.4.2.Exceptions

Exception is an event that occurs when an operation on some particular operands have no outcome suitable for the reasonable application. That operation should signal one or more exceptions by invoking the default or, if explicitly requested, a language-defined alternate handling.

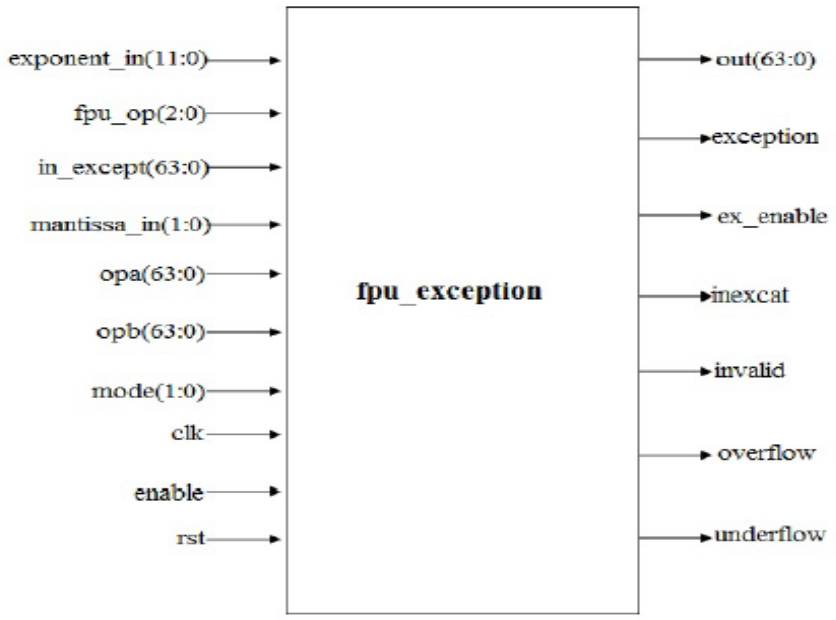


Fig.2 Exception Module

The exception signal will be asserted for some special cases, which are included with the exception algorithm. If the output is positive infinity, and the rounding mode is round to zero or round to negative infinity, then the output will be rounded down to the largest positive number. Likewise, if the output is negative infinity, and the rounding mode is round to zero or round to positive infinity, then the output will be rounded down to the largest negative number. The rounding of infinity occurs in the exceptions module, not in the rounding module. QNaN is defined as Quiet Not a Number. SNaN is defined as Signalling Not a Number. If either input is a SNaN, then the operation is invalid. The output in that case will be a QNaN. For all other invalid operations, the output will be a SNaN. If either input is a QNaN, the operation will not be performed, and the output will be a QNaN. The output in that case will be the same QNaN as the input QNaN. If both inputs are QNaNs, the output will be the QNaN in operand A.

2.5.Design

The algorithms to design the system are included for each operation.

2.5.1.Addition Algorithm

1. Feed two operands (say A and B), each 64 bits long, as the inputs.
2. The sign bit of operand A is the sign of the result.
3. Store the mantissa (bits 51-00) and exponent (62-52) of each operand.
4. Compare the 2 exponents and choose the largest exponent.
5. Find the difference of the 2 exponents (say S).
6. Write the mantissa of the operands as 1.mantissa.
7. Shift the mantissa of the smaller number to the right as many times as equal to the difference calculated in the step 4.
8. Add the 2 mantissas to obtain the final mantissa.

9. The result is obtained by combining sign, exponent and mantissa.

2.5.2.Subtraction Algorithm

1. Feed two operands (say A and B),each 64 bits long, as the inputs.
2. The sign bit of the result is 1 if the difference A-B is positive, else 0.
3. Store the mantissa (bits 51-00) and exponent (62-52) of each operand.
4. Compare the 2 exponents and choose the largest exponent.
5. Find the difference of the 2 exponents (say S).
6. Write the mantissa of the operands as 1.mantissa.
7. Shift the mantissa of the smaller number to the right as many times as equal to the difference calculated in the step 4.
8. Subtract the 2 mantissas to obtain the final mantissa.
9. The result is obtained by combining sign, exponent and mantissa.

2.5.3.Multiplication Algorithm

1. Calculate the sign = XOR of the sign bits of the operands.
2. The exponent is evaluated as: e1 + e21022
3. To evaluate the mantissa:

Consider the example given below: Say, M1= 101100101 M2= 100101100 Multiplication can be carried out in the following manner:

(a) Split the operands to smaller operands each 3 bits long.
For M1: Let M3=101, M4=100 and M5=101.
For M2: Let M6 = 100, M7 =101 and M8= 100.
(b) To evaluate the value of M1 * M8 (say the result is S1):

a. Find M3 * M8 (say A1), M4 * M8(say A2) and M5 * M8(say A3).
b. Add A3, A2 shifted by 3 bits and A1 shifted by 6 bits to get S1.
4. Find S2 (M1 * M7) and 3(M1* M6) following the procedure used to find S1.
5. Add S1, S2 shifted by 3 bits and S3 shifted by 6 bits.

2.5.4.Division Algorithm

1. Sign of the result = XOR of the sign bits of the operands.
2. The exponent is obtained using: e1 e2 + 1023
3. Division operation on the mantissas can be performed as:

Consider the operation:
Mantissa 1/ mantissa 2.
Initially, remainder = mantissa 1.

4. Perform remainder = remainder mantissa2.
5. If remainder > mantissa2, set the resultant bit (say mantissa 3) as 1, otherwise 0.
6. Perform the steps 4 and 5 until the process is completed.

2.5.5.Rounding Algorithm

1. Round to nearest: If first extra remainder bit is a 1, and the LSB of the mantissa is a 1 rounding is done. For rounding, rounding amount is added to the resultant mantissa.
2. Round to zero: No rounding is performed, unless the output is positive or negative infinity. The final output will be the largest positive or negative number in case of rounding in positive or negative infinity.
3. Round to positive infinity: The two extra remainder bits are checked, and if there is a 1 in either bit, and the sign bit is 0, then the rounding amount will be added to the resultant mantissa.
4. Round to negative infinity: Check the two extra remainder bits, and if there is a 1 in either bit, and the sign bit is a 1, then the rounding amount will be added to the resultant mantissa.

2.5.6.Exception Algorithm

The exception signal will be asserted for following special cases.

The exception signal will be asserted for given special cases.

1. Divide by 0: The result is infinity, positive or negative, depends upon the sign of operand A.
2. Divide 0 by 0: Result is SNaN, and the signal asserted for invalid.
3. Divide infinity by infinity: Result will be SNaN, and the invalid signal will be asserted.
4. Divide by infinity: Result will be 0, positive or negative; depending on the sign of operand A the underflow signal will be asserted.
5. Multiply 0 by infinity: Result is SNaN, and the signal will asserted as invalid.
6. Add, subtract, multiply, or divide overflow: Result will be infinity, and the overflow signal will be asserted.
7. Add, subtract, multiply, or divide underflow: Result is 0, and the underflow signal will assert.
8. Add positive infinity with the negative infinity: Result is SNaN, and the invalid signal asserted.
9. Subtract positive infinity from the positive infinity: Result is SNaN, and the invalid signal will be asserted.
10. Subtract negative infinity from negative infinity: Result is SNaN, and the invalid signal asserted.
11. Any one or both inputs are QNaN: Output is QNaN.
12. One or both inputs are SNaN: Output will be QNaN, and the invalid signal will be asserted.
13. If either of the two remainder bits is 1: Inexact signal is asserted.

2.5.7.Mode Algorithm

For some cases, the modes have to be varied. These have been include as a new module called mode. The conditions under this module include:

1. If subtraction of 2 numbers is involved where the second number is negative, the mode of operation has to be switched to subtraction.
2. If subtraction is the operation and the first operand is negative, then the mode is made addition.

3. If second operand is negative and the operation assigned is addition, then the mode has to be changed to subtraction.
4. If the first operand is negative when an operation of addition is called, change the mode to subtraction. In all other cases, the mode is applied as such.

The various modules like addition, subtraction, multiplication, division, mode, multiplexer, rounding and exception were separately done and compiled. They are finally called into the main program and compiled.

2.6.Results

The results were simulated using ModelSim. The obtained wave forms are given below.

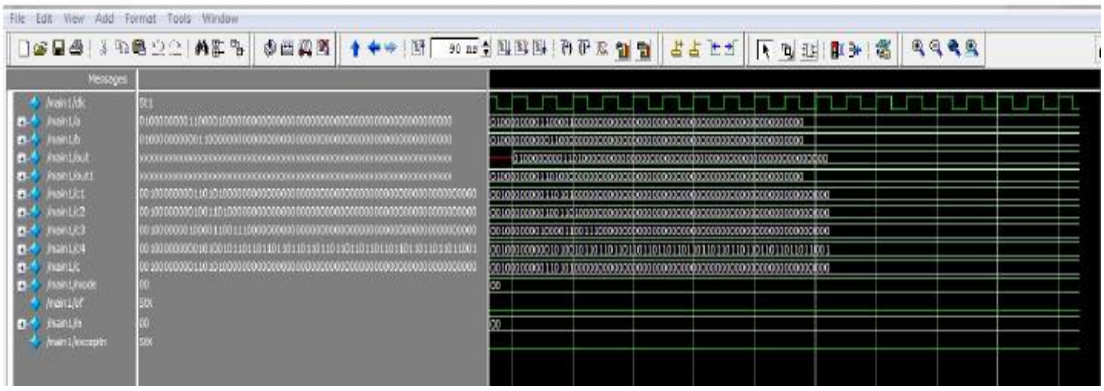


Fig.3.Simulated result of addition

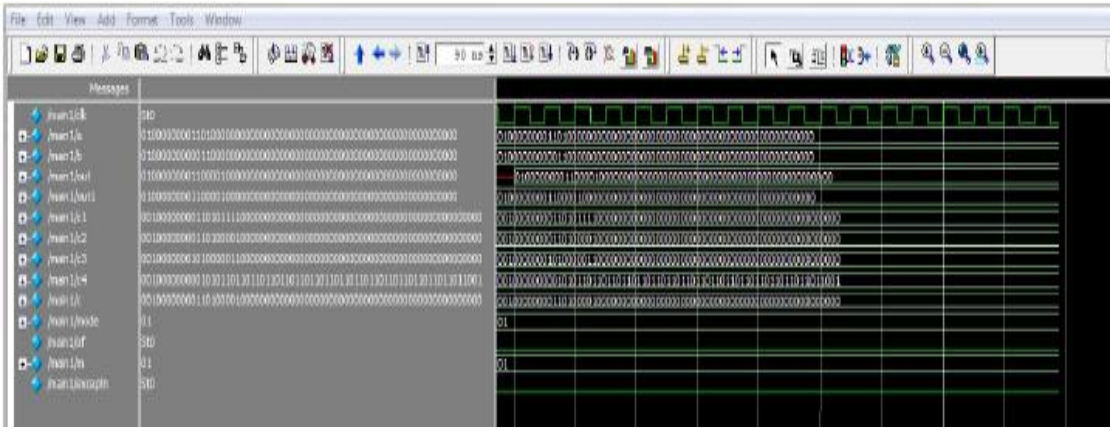


Fig.4.Simulated result of subtraction

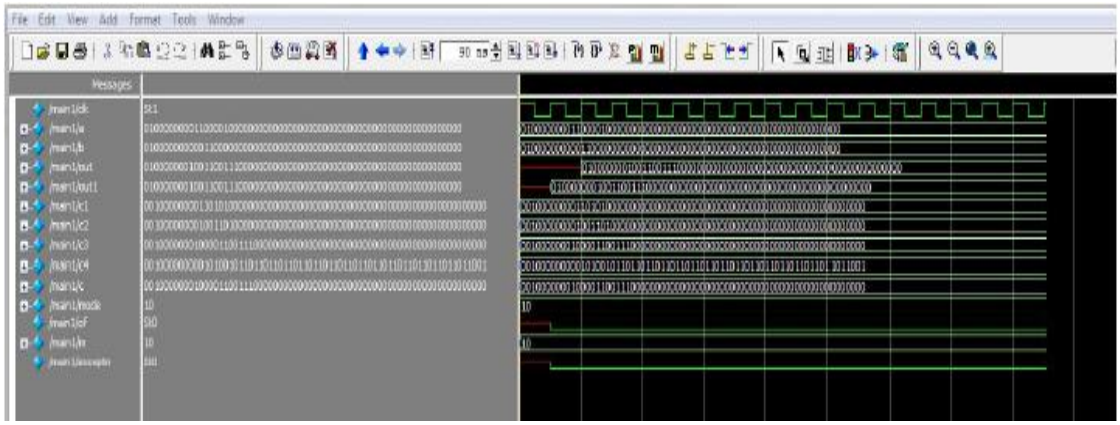


Fig.5.Simulated result of multiplication

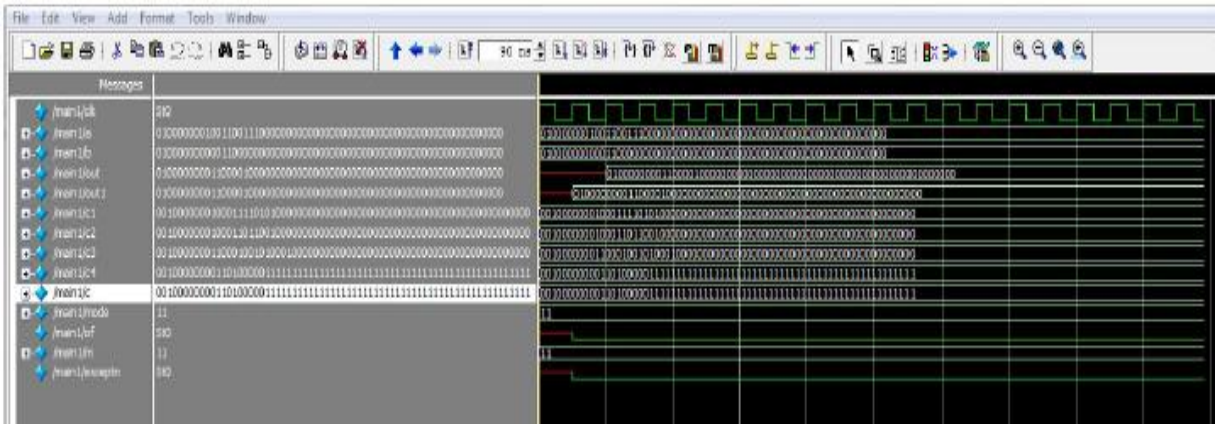


Fig.6.Simulated result of division

3.CONCLUSIONS

Double precision floating point core has been developed using verilog and simulated using ModelSim. The resulted simulations show an efficient double precision floating point arithmetic operation.

The project can be extended by including various other algorithms, say for calculating sine, cosine, tan functions and so on. The project has been implemented for double precision format which may be further extended using suitable algorithms, if available, for extended double and other formats. Similarly, faster algorithms may be utilized to provide a more efficient floating point core.

ACKNOWLEDGEMENTS

This project was done as a curriculum work. I express my profound sense of gratitude and sincere thanks to my project guide Ass. Prof. Vijitha .S for her advice, encouragement, guidance and supervision for making this work a reality. And also would like to thank the Project co-

ordinator Mr.Vinod. G, Associate Professor, Department of Electronics and Communication Engineering, for the worthy advices towards improving the work. Here I conveying my heartfelt thanks to all who have been directly or indirectly involved in making this project. Above all I thank God Almighty for His bountiful blessings for the successful completion of this work.

REFERENCES

[1] Charles Severance, 20th February 1998, "*An Interview with the Old Man of Floating-Point*". http://www.eecs.berkeley.edu/wkahan/ieee754status/754story.ht

[2] International Journal of Emerging Technology and Advanced Engineering, Volume 2, Issue 7, July 2012, FPGA based implementation of double precision floating point adder/ subtractor using verilog.

[3] IEEE XPLORE digital library.

[4] Implementation of IEEE-754 Addition and Subtraction for Floating Point Arithmetic Logic Unit, Volume 3, No. 1, 2010, pp. 131-140, International Transactions in Mathematical Sciences and Computer.

AUTHOR

Aparna C V Graduated (B.Tech) in Electronics and Communication Engineering from Calicut University and currently pursuing M.Tech in VLSI and Embedded Systems from APJ Abdul Kalam Technological University, Kerala, India.

Mary Joseph received M.Tech Degree in Microwave and Radar from Cochin University of Science and Technology (CUSAT), Kochi, India, in 1997. Currently she is working as Associate Professor in M. A. College of Engineering, Kothamangalam. She has joined in M. A. College of Engineering in 1991 as Assistant Professor. In between she worked at Birla Institute of Science & Technology-Pilani's (BITS-PILANI) Dubai Campus for 9 years as Assistant Professor during 2000-2008. Her Research interests include Microstrip Antennas and Uniplanar Antennas.

INTENTIONAL BLANK

International Journal on Cybernetics & Informatics (IJCI) Vol. 5, No. 2, April 2016

DYNAMIC PRIVACY PROTECTING SHORT GROUP SIGNATURE SCHEME

Ashy Eldhose[1] and Thushara Sukumar[2]

[1] Student, Department of Computer Science and Engineering, MBITS Nellimattom
[2]Assistant Professor, Department of Computer Science and Engineering, MBITS Nellimattom

ABSTRACT

Group Signature, extension of digital signature, allows members of a group to sign messages on behalf of the group, such that the resulting signature does not reveal the identity of the signer. The controllable linkability of group signatures enables an entity who has a linking key to find whether or not two group signatures were generated by the same signer, while preserving the anonymity. This functionality is very useful in many applications that require the linkability but still need the anonymity, such as sybil attack detection in a vehicular ad hoc network and privacy preserving data mining. This paper presents a new signature scheme supporting controllable linkability.The major advantage of this scheme is that the signature length is very short, even shorter than this in the best-known group signature scheme without supporting the linkability. A valid signer is able to create signatures that hide his or her identity as normal group signatures but can be anonymously linked regardless of changes to the membership status of the signer and without exposure of the history of the joining and revocation. From signatures, only linkage information can be disclosed, with a special linking key. Using this controllable linkability and the controllable anonymity of a group signature, anonymity may be flexibly or elaborately controlled according to a desired level.

KEYWORDS

Anonymity, Privacy, Group Signature, Opening, Linkability

1. INTRODUCTION

Personal information is more and more publicly accessible due to modern technologies and accordingly privacy is increasingly becoming an important security property. Privacy is characterized by two fundamental notions, anonymity and unlink ability [1]. Anonymity means that a user's identity or identifiable information is concealed in authentication messages. Unlink ability means that given two authentication messages, an unauthorized entity cannot tell whether they are generated by the same user or not..Generally speaking, for accessing a service, users prefer to preserve their privacy, but the service provider may want to relax their privacy to gain sufficient user information.

Extending the idea of digital signature schemes into groups, a new signature scheme i.e. group signature scheme, provides authority to any group member to sign messages anonymously on behalf of the group. A client can verify the authenticity of the signature by using only the group's public key parameters. It must be computationally hard to identity of the group member so that he cannot be linked from a signed message or his signature. However, in the case of a legal dispute, the identity of a signer or member can be revealed by a designated entity i.e. the group manager. The major feature of group signature is the security of the information or the data that makes it more important as well as attractive for many real time applications, such as e-commerce, e-

DOI: 10.5121/ijci.2016.5216

auction and e-voting, where the priority is privacy and anonymity of signer which is very much high and important for any organization.

For an application environment, privacy needs to be adjusted according to the desired policy or reasonable expectation of profit. If the requirements of privacy for both the users and service providers are properly balanced, privacy will be attractive for both of them. Linkability is the key feature required in data mining. However, anonymity is necessary for privacy. It is possible to hide an identity or identifiable information from transactions while revealing still linkable information. For example recommendation system such as the one at Amazon.com[3]. Customers might be happy to participate in the system only if their anonymity is kept and the linkability is given only to their service provider. Customers will feel assured if their buying pattern is revealed only to the service provider and their identities have not been revealed to anyone.

A privacy-protecting signature scheme was recently introduced to provide elaborate privacy controls. Conceptually, it resides between pseudonym systems and normal GS schemes[11]. Neither information identifying a signer nor information linking signatures is revealed explicitly from signatures. However, the anonymity and unlinkability can be controlled by keys. That is, the corresponding signer identity and linkage information can be revealed by an opening key and a linking key, respectively[2]. Using a trapdoor-based approach on these two privacy notions, one can establish a two-level access hierarchy on signer privacy. To be more descriptive, this **P**rivacy-protecting **S**ignature scheme with both **O**pening and **L**inking capabilities in a controllable manner is referred to as a PS-OL scheme for short. A PS-OL scheme supports two seemingly-incompatible properties, that is, privacy and data mining versatility by selectively providing linkability and anonymity.

2. RELATED WORKS

As we know of digital signature and facilities it has provided regarding information security, so extending the idea of digital signature to group where we can parallely authorize multiple information or documents and save time. Group Signatures have vital role in day to day corporate organizations' ecommerce applications. Extending the idea of digital signature schemes into groups, a new signature scheme i.e. group signature scheme, first introduced by Chaum and Heyst in 1991, provides authority to any group member to sign messages anonymously on behalf of the group [19]. A client can verify the authenticity of the signature by using only the group's public key parameters. It must be computationally hard to identity of the group member so that he cannot be linked from a signed message or his signature. However, in the case of a legal dispute, the identity of a signer or member can be revealed by a designated entity i.e. the group manager. GS schemes provide controllable anonymity such that a signer can be identified from a signature by a trusted group manager. It provides unlinkability on signatures against all users except the group manager. A number of GS schemes have been presented to address various features [7].

Direct Anonymous Attestation (DAA) has been proposed for the remote anonymous authentication of a trusted platform module. While DAA guarantees complete anonymity, i.e., no party can reveal a signer's identity from a signature, it provides signer-controllable linkability, i.e., a signer can generate an anonymous signature with a tag, which is linkable to another signature from the same signer. There are variants of a GS scheme to alleviate the centralized group manager's rights that can reveal a signer's identity[4].

In the Democratic GS scheme, the group membership is controlled jointly and equally by all group members. A signer of a signature can be identified by a member, in other words, the

signer's anonymity can be provided only against non members. In the tracing-by-linking GS scheme, no signer can be identified by any authority if he or she signs only once per event.

Some schemes with controllable or revocable anonymity provide linkability by adding a tag to a signature [9]. Using the tag associated with a signature, one can check the linkability on signatures easily and explicitly. For example, a linkable democratic GS scheme is a variant of a democratic GS scheme to support the tag-based linkability. A message-linkable GS scheme was suggested to resist Sybil attacks in VANET[10].

The message-linkable property means that given two anonymous signatures on the same message, one can easily decide whether they are generated by the same signer or not. To provide this property, the scheme uses a static tag generated with a message and a secret key[11].

3. PROBLEM DEFINITION

The secret signing key of a group member includes a key-pair for a standard digital signature scheme that is certified by the group manager. The group member's signature is an encryption, under a public encryption key held by the group manager, of a standard signature of the message together with certificate and identity information, and accompanied by a non-interactive zero-knowledge proof that encryption contains what it should. Previous works did not try to achieve security notions as strong as this paper target, nor to pin down what properties of the building blocks suffice to actually prove security. It is well-known in the literature that two cryptographic solutions have been widely used to preserve privacy, a pseudonym system and group signatures (GS) [12]. The pseudonym system supports anonymity, but a signer cannot avoid being linked by anyone who obtains their signatures. A group signature (GS) scheme is considered as one of the most versatile primitives for anonymity. However, following the concept of a traditional GS (or referred to as a normal GS), the linkability is given only to an opener, who is not usually a service provider but a special group manager. The definitions and results of previous paper are for the setting in which the group is static, meaning the number and identities of members is decided at the time the group is set up and new members cannot be added later[5]. An also proper definition for security has not been provided even for the basic static-group case.

The objective of this paper is to implement a group signature scheme based on following assumptions:

- Group signature scheme based upon hard computational assumptions, such as, elliptic curve cryptography (ECC) and a honest verifier ZKPK Protocol.
- Group signature scheme should be unaffected by joining or leaving of any member.
- Group signature scheme must satisfy all basic security requirements like anonymity, traceability, and unlinkability.

EC cryptography schemes are public-key mechanisms which are able to give the same facilities as the schemes of RSA or Elgamal. But the security of ECC is based on a hardness of another problem, known as the elliptic curve discrete logarithm problem (ECDLP). The best algorithms to solve ECDLP have full exponential-time (unlike RSA's algorithms which have the sub exponential-time)[17]. Thus, required security level can by achieved with significantly smaller keys in elliptic curve system than in its rival- RSA system. Zero-knowledge is defined by means of a distinguisher Dwhich essentially tries to distinguish between proofs produced by a prover (with respect to a real common random string), or a simulator (with respect to a simulated common random string)[8].

4. PRIVACY PROTECTING SIGNATURE SCHEME WITH BOTH OPENING AND LINKING CAPABILITY

Privacy-protecting **S**ignature scheme with both **O**pening and **L**inking capabilities in a controllable manner is referred to as a PS-OL scheme for short. A PS-OL scheme supports two seemingly-incompatible properties, that is, privacy and data mining versatility by selectively providing linkability and anonymity. A PS-OL scheme has benefits in flexibly organizing participants over a normal GS, considering that a linker can be built up separately from an opener. This separation enables a bottleneck (strong trusted relationship and on-line processing) to be removed in an anonymous system. A PS-OL scheme is constructed from a linear combination encryption with many parameters. Since the underlying structure is quite complex, the system requires heavy operations, and its signature length is also relatively long. In this paper, we construct a PS-OL scheme for a dynamic membership, where group signatures can be anonymously linked, but the corresponding linkage information can only be revealed with a linking key. The linking key is secretly managed by a privileged party called a linker who is delegated the link capability by the opener. Note that the capability of linking signatures is placed below the capability of opening the signer identity of the signatures. We can achieve a stepwise access control on anonymity by adding this Controllable Linkability (CL) to the controllable anonymity that can identify a signer from signatures using an opening key. The linking capability of this dynamic group signature differs from the tracing capability of a traceable signature scheme. The traceable signature scheme enables a tracer to trace only a specific user's signatures, not other users. In contrast, a linker of our scheme can deal with every user's linkage information with a key. Though a traceable signature scheme can be used for our linkability, it involves complex computation. For example, for n signatures and m tracing keys, $n \times m$ computation is required for a traceable signature scheme while n computation is required for this scheme.

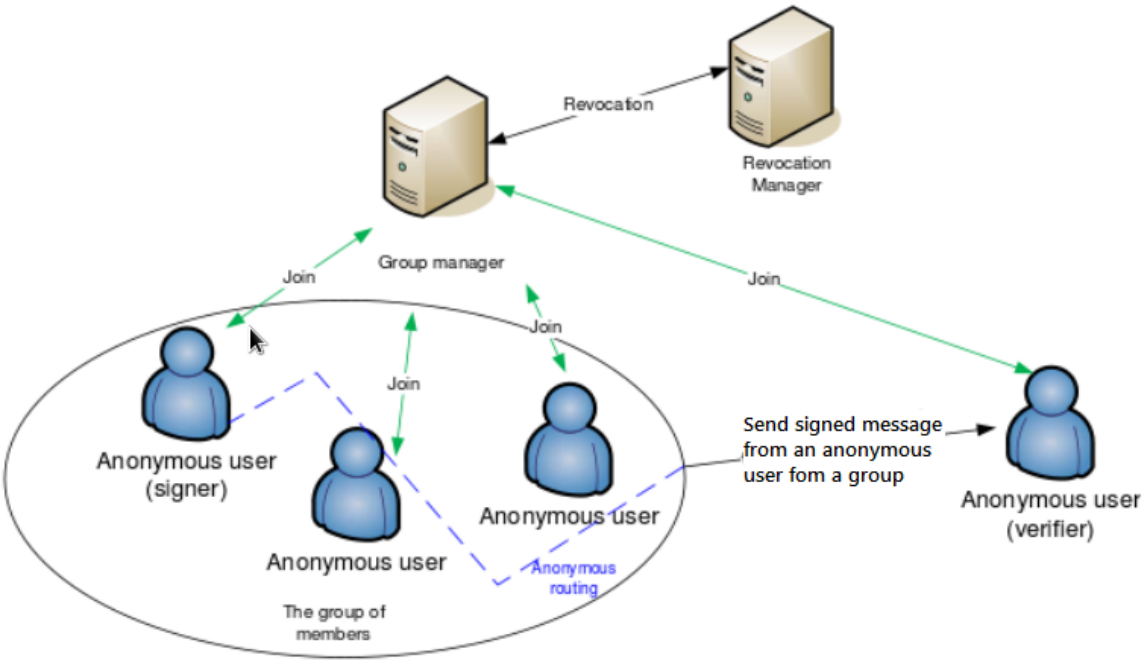


Figure 1 Principles of Group Signature Scheme

The proposed PS-OL scheme supports a dynamic group membership where a user can join or leave a group. Leaving a group is also referred as to be revoked. However, the linking capability can be consistently preserved regardless of changes to the membership status of the signer. In addition, the CL property does not expose the history of the joining and revocation. Despite the

additional functionality of CL, our scheme has a compact structure to yield a very short signature that is one group element shorter than the best-known GS.Early works of GSSs considered only a static setting [6], where the group is fixed at the time of the setup, whereas more recent constructions consider dynamic groups [7], i.e., new members may be added and possibly deleted to and from the group over time. Moreover, in some cases it is also desirable to have distributed authorities, i.e., one party only receives the opening key and a distinct party receives the issuing key required to add new members or to revoke existing members.

4.1 Model

This section presents a security model for a PS-OL scheme. This model assumes three authorities

- Issuer
- Opener and
- Linker

who have their independent privileges and a certain level of trust. A linker is assumed to behave honestly but curiously, and so it can try to find passively user's identity only with signatures collected.

This model explicitly considers a revocation algorithm that performs the update of keys. For the revocation, it makes use of a revocation list, denoted by **RL**. An entry of **RL** consists of an index and private information for a user who has been revoked. It is managed by the Issuer and initially set to be empty [14]. The list is used to update a user signature key and a group public key. In this model, whenever the information of keys to be revoked are given according to a pre-defined policy, **RL** is immediately updated to include them and entries are arranged to the latest revocation index. One can publicly access the list. A user signature key includes a non-negative integer λ, called a revocation index, to indicate that the key has been updated up to the λ^{th} entry of **RL**. Let $\lambda\hat{} > \lambda$ be the most up-to-date number of revoked keys in **RL**. For generation of a signature, the user signature key is updated up to the $\lambda\hat{}^{th}$ entry of **RL**. The generated signature includes $\lambda\hat{}$ to indicate that the signature was generated by the key which has been updated up to the $\lambda\hat{}^{th}$ entry of **RL**. It can be verified with the group public key that has been updated up to the $\lambda\hat{}^{th}$ entry of **RL**.

The PS-OL scheme uses a registration list **REG** = *(*REG[1], . . . ,REG[*n*]*)*. REG[*i*] contains private informationforthe *i*-th registered user. The registered usersareall different. **REG** is managed by Issuer and can be accessedby Opener to identify a signer.
A PS-OL scheme consists of the following algorithms.

- Setup phase: group manager computes the public key and the secret key in this phase by implementing the algorithm for group key generation. He inputs a security parameter to the algorithm and it returns the group public key and also the secret key of group manager. The secret key is kept with him and the group public key is circulated among the members.

- Issue phase: an interactive protocol is established in this phase between the group manager and the to-be-member after which the user becomes a valid group member. A secret key is chosen by the Group member using which another parameter is generated by the member. This generated parameter is sent to the group manager. Then using his own secret key the group manager generates the group member's signing key and returns it to newly joined group member.

- Sign phase: This is the signing phase in which an protocol is established between the group member and the verifier where he has to verify a group signature whether it is generated by a valid group member or not. Group member uses the signing key pairs to sign the message. The generated group member signature of knowledge is sent by the member to the verifier for verification.

- Verify phase: This phase implements a deterministic algorithm using given group public key and the signed message to verify the validity of the group signature. Signer sends his signature to the verifier, i.e. the signature generated by the signature of knowledge. The message is accepted if true value is returned by the verification phase else the message is rejected if false value is returned by the verification phase.

- Open phase: This phase implements a deterministic algorithm to reveal the identity of the signer, by taking input a signed message and the secret key of group manager. The signature is taken as input by the group manager and using the private parameters outputs the identity of the signer as return value. This open algorithm is implemented when a incident of a legal dispute arises.

- Judge phase: This phase implements a judge algorithm to check the user produced the signature on the message using the secret key.

- Link phase:This phase deals with the linkage information of every user with a key. The linkage information can only be revealed with a linking key [17].

4.2 Security Notions

- Anonymity: Given a sign which is valid must be difficult for anyone to discover the identity of the signer computationally. As the constant differs every time, the same member generates different signature for every new message to be signed. The group manager only can determine the identity of the signing member using his secret key. For a nonmember it is almost not possible to discover the secret parameters of the signing group member as the knowledge of the secret key of the group manager is required and so without the secret key of the group manager it is almost impossible to determine the secret parameters of the signer and hence an outsider cannot cannot determined the identity of the signer. In this property we conclude that if neither group manager's secret key nor group member's secret key is exposed then it is infeasible to reveal the signer of a authorized valid signature.

- Unforgeability: Only a valid authorized member belonging to the group can produce a valid signature i.e. a valid member only can produce a signature on behalf of his group.

- Unlinkability: This property states that deciding if two valid signatures were generated by the same group member is difficult. According to this property one cannot conclude that both signatures are from the same member or not if he's provided with two signatures.

- Traceability: Using only open algorithm and the group manager's secret key, the group manager can track the identity of the signing member if given any valid signature. Like in case of any legal dispute or emergencies, any signer's identity can be traced by the group manager only. It is not possible for an outsider to track the signer because open algorithm, which used to trace a signing group member, requires the knowledge of the secret key of the group manager.

- Exculpability: The group members even along with the group manager are not able to sign a document on behalf of any other group member. The knowledge of the secret parameters of the group member is required to generate a valid signature. And every member has his own unique secret key that are used to generate the signature. Even a group manager cannot sign on behalf of any group member because the group manager does not have the members' secret keys[18].

5. CONCLUSION

A dynamic PS-OL scheme is constructed which yields a short signature. The constructed scheme achieves anonymity, traceability, non-frameability, and also three security requirements for (controllable) linkability. Also this scheme outperforms the best-known anonymous signature schemes. This scheme will be very versatile and useful in many privacy-enhancing applications with limited resources.

REFERENCES

[1] G. Ateniese, J. Camenisch, M. Joye, and G. Tsudik, "A practical and provably secure coalition-resistant group signature scheme," in Advancesin Cryptology (Lecture Notes in Computer Science), vol. 1880. Berlin, Germany: Springer-Verlag, 2000, pp. 255–270.

[2] D. Boneh and X. Boyen, "Short signatures without random oracles and the SDH assumption in bilinear groups," J. Cryptol., vol. 21, no. 2, pp. 149–177, 2008.

[3] D. Boneh, X. Boyen, and H. Shacham, "Short group signatures," in Advances in Cryptology (Lecture Notes in Computer Science), vol. 3152. Berlin, Germany: Springer Verlag, 2004, pp. 41–55.

[4] E.Brickell, J. Camenisch, and L. Chen, "Direct anonymous attestation," in Proc. ACM CCS, 2004,pp.132–145.

[5] P. Bichsel, J. Camenisch, T. Grosß, and V. Shoup, "Anonymous credentials on a standard java card," in Proc. ACM CCS, 2009, pp. 600–610.

[6] E. Brickell, L. Chen, and J. Li, "Simplified security notions of direct anonymous attestation and a concrete scheme from pairings," Int. J. Inf.Security, vol. 8, no. 5, pp. 315–330, 2009.

[7] P. Bichsel, J. Camenisch, G. Neven, B. Warinschi, and N. P. Smart, "Get short via group signatures without encryption," in Security andCryptography for Networks. Berlin, Germany: Springer-Verlag, 2010, pp. 381–398.

[8] X. Boyen and C. Delerablee, "Expressive subgroup signatures," in Security and Cryptography for Networks (Lecture Notes in Computer Science), vol. 5229. Berlin, Germany: Springer-Verlag, 2008, pp. 185–200.

[9] E. Brickell and J. Li, "A pairing-based DAA scheme further reducing TPM resources," in Proc. 3rd TRUST, 2010, pp. 181–195.

[10] M. Bellare, D. Micciancio, and B. Warinschi, "Foundations of group signatures: Formal definitions, simplified requirements, and a construction based on general assumptions," in Advances in Cryptology (Lecture Notes in Computer Science), vol. 2656. Berlin, Germany: Springer-Verlag, 2003, pp. 614–629.

[11] J.-M. Bohli and A. Pashalidis, "Relations among privacy notions," ACM Trans. Inf. Syst.Security, vol. 14, no. 1, 2011, Art. ID 4.

[12] M. Bellare, H. Shi, and C. Zang, "Foundations of group signatures: The case of dynamic groups," in Topics in Cryptology (Lecture Notes in Computer Science), vol. 3376. Berlin, Germany: Springer-Verlag, 2004, pp. 136–153.

[13] D. Boneh and H. Shacham, "Group signatures with verifier-local revocation," in Proc. ACM CCS, 2004, pp. 168–177.

[14] X. Boyen and B. Waters, "Compact group signatures without random oracles," in Advances in Cryptology, vol. 4004. Berlin, Germany: Springer-Verlag, 2006, pp. 427–444.

[15] X. Boyen and B. Waters, "Full-domain subgroup hiding and constant size group signatures," in Public Key Cryptography, vol. 4450. Berlin, Germany: Springer-Verlag, 2007, pp. 1–15.

[16] J. Camenisch and T. Groß, "Efficient attributes for anonymous credentials," in Proc. ACM CCS, 2004, pp. 345–356.
[17] J. Camenisch, S. Hohenberger, and A. Lysyanskaya, "Compact E-cash," in Advances in Cryptology (Lecture Notes in Computer Science), vol. 3494. Berlin, Germany: Springer-Verlag, 2005, pp. 302–321.
[18] J. Camenisch and A. Lysyanskaya, "Dynamic accumulators and application to efficient revocation of anonymous credentials," in Advancesin Cryptology (Lecture Notes in Computer Science), vol. 2442. Berlin, Germany: Springer-Verlag, 2002, pp. 61–76.

Authors

Ashy Eldhose is currently pursuing M.Tech in Cyber Security in MBITS, Nellimattom. She completed her B.Tech from MBITS, Nellimattom, Kerala, India. Her areas of research are Network Security and Information Forensics.

ThusharaSukumar is currently working as the Assistant Professor in Department of Computer Science and Engineering at MBITS, Nellimattom, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering from College of Engineering, Kidangoor and ME in Computer Science and Engineering from PSNACET, Dindigul. Her research interest include Image Mining.

EFFICIENT FEATURE SUBSET SELECTION MODEL FOR HIGH DIMENSIONAL DATA

Chinnu C Georgel[1] and Abdul Ali[2]

[1]Department of Computer Science, Ilahia College of Engineering and Technology
chinnulaby@gmail.com
[2] Department of Computer Science, Ilahia College of Engineering and Technology
abdulali@icet.ac.in

ABSTRACT

This paper proposes a new method that intends on reducing the size of high dimensional dataset by identifying and removing irrelevant and redundant features. Dataset reduction is important in the case of machine learning and data mining. The measure of dependence is used to evaluate the relationship between feature and target concept and or between features for irrelevant and redundant feature removal. The proposed work initially removes all the irrelevant features and then a minimum spanning tree of relevant features is constructed using Prim's algorithm. Splitting the minimum spanning tree based on the dependency between features leads to the generation of forests. A representative feature from each of the forests is taken to form the final feature subset.

KEYWORDS

Feature subset selection, filter technique, feature clustering, feature reduction

1. INTRODUCTION

Data mining is the process of automatically discovering useful information from large data repositories [2]. The rapid advances in technologies led to accumulation of vast amount data. As the number of organizations grows day by day the breeding of new technologies and new styles of data also increases. It is difficult to handle them and store them in an efficient manner and of course retrieval of data is extremely challenging [3]. Nowadays we cannot use traditional methods to explore the data analysis because of the size of dataset. So it is very important to make study on data analysis. Most of the technologies are blended with data mining or we can say that data mining is vital and indispensible concept for every technology. Traditional machine learning algorithms like decision trees or artificial neural networks are examples of embedded approaches [9][10].Data mining tasks are mainly divided into predictive and descriptive. Predictive refers to predict the particular attribute based on other attributes. Descriptive task is to derive patterns like correlations, trends, clusters, trajectories and anomalies. Association analysis is used to discover patterns from correlated data and the output of analysis is represented using implication rules. Cluster analysis is the method of partitioning the datasets into different clusters and each cluster data is strongly correlated with intra-manner and inter clusters shows strong repulsion. There are several clustering methods but it is difficult to give a crisp categorization. Anomaly detection is the opposite of cluster analysis. Anomaly detection is the process of identifying significantly different data from rest of data. Example: Fraud detection.

DOI: 10.5121/ijci.2016.5217

Choosing a subset of good features with respect to the target concepts, feature subset selection is an effective way for reducing dimensionality, removing irrelevant data, increasing learning accuracy, and improving result comprehensibility [9]. Many feature subset selection methods have been proposed and studied for machine learning applications. They can be divided into four broad categories: the Embedded, Wrapper, Filter, and Hybrid approaches.The embedded methods incorporate feature selection as a part of the training process and therefore may be more efficient than the other three categories [23] decision trees or artificial neural networks are examples. The wrapper methods use the predictive accuracy of a predetermined learning algorithm to determine the goodness of the selected subsets, the accuracy of the learning algorithms is usually high. However, the generality of the selected features is limited and the computational complexity is large. The filter methods are independent of learning algorithms, with good generality. Their computational complexity is low. The hybrid methods are a combination of filter and wrapper methods [25] by using a filter method to reduce search space that will be considered by the subsequent wrapper.

2. RELATED WORKS

Dimensionality can be defined as the number of attributes or features that an object posses. Data objects compose of several features. Most of the features give unique characteristics to the object. As the number of features increases, processing will be difficult [11][3]. But lesser number of features will also, only give a vague idea about the object. So it is important to reduce the number of features into a number which represents the objects effectively [13][18].

Feature selection is an important concept of reducing dimensions [19][15]. Each feature in a dataset represents a particular characteristic of object. But sometimes some of them are irrelevant and or correlated. Finding such features and removing it is known as dimensionality reduction [20][19].Feature subset selection involves identifying and removing as much as irrelevant and redundant features as possible. This is because 1) irrelevant features do not contribute to the predictive accuracy [14], and 2) redundant features do not redound to getting a better predictor for that they provide mostly information which is already present in other feature(s). Most of the feature subset algorithms removes irrelevant features successfully but not able to handle redundant features [5],[18]. Traditionally, feature subset selection research has focused on searching for relevant features. A well-known example is Relief [11], which weighs each feature according to its ability to discriminate instances under different targets based on distance-based criteria function. This will not eliminate redundant features. But with irrelevant features redundant features will affect the speed and accuracy of machine learning algorithms. So the redundant features also have to be eliminated.

CFS [18], FCBF [20], and CMIM [15] are examples that take into consideration the redundant features. In CFS correlation based heuristic evaluation function is used. Feature subsets contain features that are highly correlated with the class and uncorrelated with each other. Irrelevant features are ignored and redundant features are screened out CFS uses symmetrical uncertainty to measure correlations. In FCBF predominant correlation is used. Fast filter method can identify relevant features as well as redundancy among relevant features without pairwise correlation analysis. CMIM iteratively picks features which maximize their mutual information with the class to predict conditionally to the answer of any feature already picked. The proposed algorithm eliminates both irrelevant and redundant features efficiently using clustering based feature selection.

Cluster analysis is a method of classifying the given data. There are several cluster methods exists. It is difficult to say that one cluster method is efficient than other. Each of the cluster method overrides other based on at least one condition. Most of the cluster methods help to find irrelevant and correlated features. This implies that clustered data is efficient for removing relevant or irrelevant data. Hierarchical clustering has been adopted in word selection in the context of text classification (e.g., [22] and [1]). Distributional clustering has been used to cluster words into groups based either on their participation in particular grammatical relations with other words by Pereira et al. [22] or on the distribution of class labels associated with each word.

3. PROPOSED METHOD

Irrelevant features, together with redundant features, will affect the correctness of the learning machines [6]. Thus, feature subset selection should be able to identify and remove as much of the irrelevant and redundant information as possible. Moreover, "good feature subsets hold features highly correlated with (predictive of) the class, yet uncorrelated with (not predictive of) each other." [7].

Based on these aspects a novel algorithm which can efficiently and effectively deal with both irrelevant and redundant features, and obtain a good feature subset. The algorithm works in four steps. First step is to calculate the symmetric uncertainty of features between target classes and other features. In this paper the algorithm calculates symmetric uncertainty not only considering the co-occurrence of feature but also the change of feature when other changes. Then the features whose symmetric uncertainty values less than a threshold is considered irrelevant and are removed. For the remaining features minimum spanning tree is constructed using prim's algorithm. The third step splits the minimum spanning tree into several forests. Each forest is considered as a single cluster. As we know that the elements within the cluster is strongly correlated. So we can select the representative features from it. Fourth step takes representative features from every cluster.

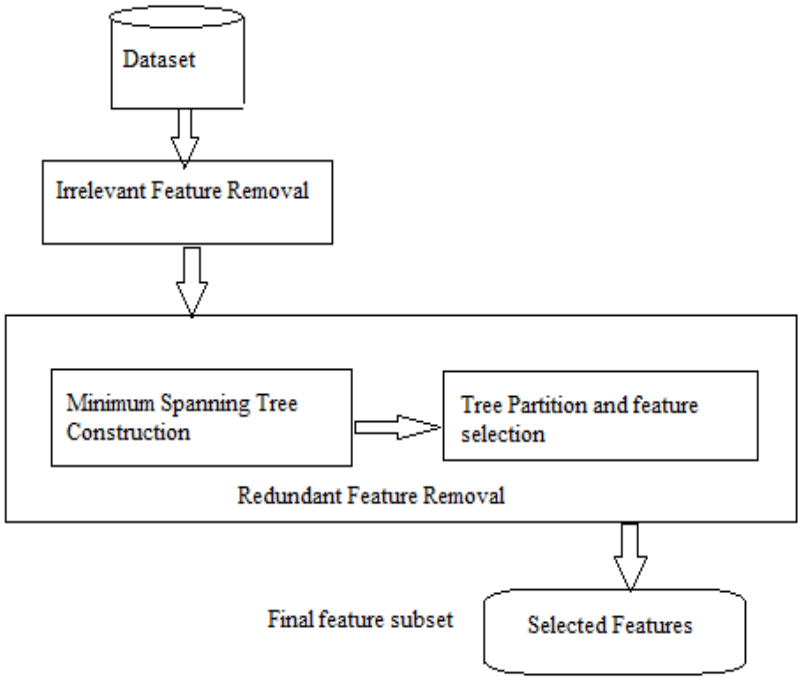


Figure 1.Framework of the proposed feature subset selection algorithm.

Relevant features have strong correlation with target concept so are always necessary for a best subset, while redundant features are not because their values are completely correlated with each other. Most of the information contained in redundant features is already present in other features. As a result, unneeded features will affect the accuracy. Symmetric uncertainty is a nonlinear estimation of correlation between feature values. The symmetric uncertainty is derived from mutual information by normalizing it to the entropies of feature values. The symmetric uncertainty can be defined as the measure of correlation between either two features or a feature and target concept.

$$SU(X,Y) = \frac{(2 \text{x} \text{Information Gain}(X,Y))}{H(X) + H(Y)}$$

Where, H(X) is the entropy of a discrete random variable X. Suppose p(x) is the prior probabilities for all values of X, H(X) is defined by

$$H(X) = -\sum p(x) \log 2\ p(x)$$

Gain(X/Y) is the amount by which the entropy of Y decreases. It reflects the additional information about Y provided by X and is called the information gain [23] which is given by

$$\text{Gain}(X/Y) = H(X) - H(X/Y)$$
$$= H(Y) - H(X/Y)$$

Where H(X) is the conditional entropy which quantifies the remaining entropy (i.e., uncertainty) of a random variable X given that the value of another random variable Y is known.

Information gain is a symmetrical measure. That is the amount of information gained about X after observing Y is equal to the amount of information gained about Y after observing X. This ensures that the order of two variables will not affect the value of the measure. Symmetric uncertainty treats a pair of variables symmetrically, it compensates for information gain's bias toward variables with more values and normalizes its value to the range [0,1].

Given SU(X,Y) the symmetric uncertainty of variables X and Y , the relevance T-Relevance between a feature and the target concept C, the correlation F-Correlation between a pair of features, the feature redundancy F-Redundancy and the representative feature R-Feature of a feature cluster can be defined as follows.

Definition 1 (T-Relevance): The relevance between the feature Fi and the target concept C is referred to as the T-Relevance of Fi and C, and denoted by SU(Fi,C). If SU(Fi,C) is greater than a predetermined threshold value we say that Fi is a strong T-Relevance feature.

Definition 2(F-Correlation): The correlation between any pair of features Fi and Fj is called the FCorrelation of Fi and Fj, and denoted by SU(Fi,Fj).

Definition 3 (F-Redundancy): Let S {F1, F2, . . . , Fi, . . . Fk } be a cluster of features. if SU(Fj,C) ≥ SU(Fi,C) ^ SU(Fi,Fj)> SU(Fi,C)is always corrected for each Fi, then Fi are redundant features with respect to the given Fj (i.e., each Fi is a F-Redundancy).

Definition 4 (R-Feature).A feature is a representative feature of the cluster S (i.e., Fi is R-Feature) if and only if, Fi = argmax SU(Fj, C). This means the feature, which has the strongest TRelevance, can act as a R-Feature for all the features in the cluster.

According to the above definitions, feature subset selection can be the process that identifies and retains the strong T-Relevance features and selects R-Features from feature clusters. The behind heuristics are that

1. irrelevant features have no/weak correlation with target concept;
2. redundant features are assembled in a cluster and a representative feature can be taken out of the cluster.

Proposed Algorithm

Inputs: D (F1, F2, …,Fm, C) – the given dataset
T-Relevance threshold.
Output: S - Final feature subset
1. for i=1 to m do
2. T-Relevance = SU(Fi,C)
3. if T-Relevance > T-Relevance threshold then
4. S = S ∪ {Fi};
5. G = NULL; //G is a complete graph
6. for each pair of features {Fi, Fj } ⊂ S do

7. F-Correlation = SU (Fi', Fj')

8. While calculating F-Correlation for each feature check change of a feature with respect to another

8. Add Features Fi and / or Fj to G with F-Correlation as the

weight of the corresponding edge

9. minSpanTree = prim(G); // Generate minimum spanning

tree using Prim's algorithm

10. Forest = minSpanTree

11. For each edge Eij∈ Forest do

12. if SU(Fi, Fj) <SU(Fi, C) ^ SU(Fi, Fj) < SU(Fj, C) then

13. Forest = Forest - Eij

14. Select representative features from each tree (one with maximum symmetric uncertainty value)

17. S = S ∪ {Representative features};

18. return S

Consider a data set D consists of m features f= {f1, f2…fm} and class C, the T-Relevance SU (fi, C) value for each feature is calculated in the first step. The features whose T-Relevance value greater than a predefined threshold value forms the relevant feature subset F= {F1, F2 …Fk}.In the second step F-correlation values for each pair of features is calculated considering the whether a feature changes when other feature changes along with checking co-occurrence of the features. A graph G=(V,E) is constructed with each target-relevant feature as node and the F-Correlation value as the weight of the edge. Graph G reflects the correlations among all the target-relevant features .Then a minimum spanning tree is constructed which connects all vertices such that the sum of the weights of the edges is the minimum, using the well-known Prim algorithm [21].After constructing minimum spanning tree edged with weights smaller than both of the T-Relevance SU(Fi, C) and SU(Fj, C), are deleted from the MST. This deletion will result in many disconnected trees T1,T2 etc. . Each tree T represents a cluster and features in a cluster are strongly correlated so are redundant. Finally a representative feature is selected from each disconnected tree to form the final feature subset.

4. EXPERIMENTAL RESULTS

We present the experimental results in terms of proportion of selected features.

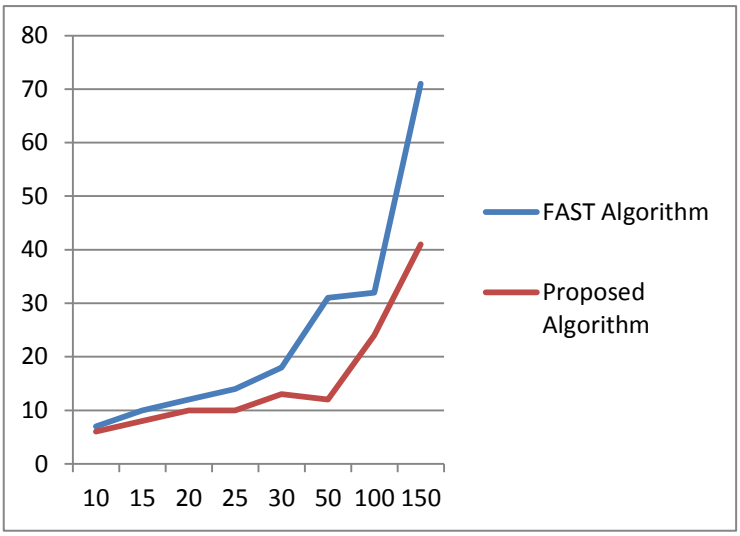


Figure 2. Comparison of New algorithm and FAST algorithm

Both algorithms FAST and Proposed algorithm achieve significant reduction of dimensionality by selecting only a small portion of the original features. But the experimental results shows that the proposed algorithm achieves more reduction in dimensionality. We get only the most important features which is capable to classify it accurately.

5. CONCLUSION

In this paper we have presented a feature selection algorithm for high dimensional data in which a clustering-based approach is used. A dependency measure is used to find out the correlation between features .The model involves the following steps removing irrelevant features, constructing minimum spanning tree, partitioning the minimum spanning tree and selecting representative features. In the algorithm partitioning of minimum spanning tree results in clusters and each cluster is treated as a single feature. Thus the dimensionality is drastically reduced.

We have compared the performance of the proposed algorithm with FAST algorithm and found that the new algorithm achieve significant reduction of dimensionality by selecting only a small portion of the original features.

For the future work, we plan to explore different types of correlation measures.

ACKNOWLEDGEMENTS

I wish to thank the Management, the Principal and Head of the Department (CSE) of ICET for the support and help in completing the work.

REFERENCES

[1] Qinbao Song, Jingjie Ni, and Guangtao Wang, “A Fast Clustering-Based Feature Subset Selection Algorithm for High-Dimensional Data” IEEE Transactions on knowledge and data engineering, Vol. 25, No. 1, January 2013

[2] Data Mining: Concepts and Techniques 2nd ed. Jiawei Han and Micheline Kamber

[3] H. Almuallim and T.G. Dietterich, “Learning Boolean Concepts in the Presence of Many Irrelevant Features,” Artificial Intelligence, vol. 69, nos. 1/2, pp. 279-305, 1994.

[4] Top 10 algorithms in data mining XindongWu • Vipin Kumar • J. Ross Quinlan • Joydeep Ghosh • Qiang Yang Hiroshi Motoda • Geoffrey J. McLachlan • Angus Ng • Bing Liu • Philip S. Yu •Zhi-Hua Zhou • Michael Steinbach • David J. Hand • Dan Steinberg

[5] R. agarwal and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," Proc, 20th Int'l Conf. Very large Data Bases, pp.487-499,1994.

[6] R. Battiti, "Using Mutual Information for Selecting Features in Supervised Neural Net Learning," IEEE Trans. Neural Networks, vol. 5,no. 4, pp. 537-550, July 1994.

[7] R. Kohavi and G.H. John, "Wrappers for Feature Subset Selection," Artificial Intelligence, vol. 97,nos.1/2, pp. 273-324, 1997.

[8] M.A. Hall and L.A. Smith, "Feature Selection for Machine Learning:Comparing a Correlation-Based Filter Approach to the Wrapper," Proc.12th Int'l Florida Artificial Intelligence Research Soc. Conf., pp. 235-239, 1999.

[9] T.M. Mitchell, "Generalization as Search," Artificial Intelligence, vol. 18, no. 2, pp. 203-226, 1982.I.Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," J. Machine Learning Research, vol 3, pp. 1157-1182, 2003.

[10] K. Kira and L.A. Rendell, "The Feature Selection Problem: Traditional Methods and a New Algorithm," Proc. 10th Nat'l Conf.Artificial Intelligence, pp. 129-134, 1992.

[11] H. Liu, H. Motoda, and L. Yu, "Selective Sampling Approach to Active Feature Selection," Artificial Intelligence, vol. 159, nos. 1/2, pp.49-74,2004.

[12] L.D. Baker and A.K. McCallum, "Distributional Clustering of Words for Text Classification," Proc. 21st Ann. Int'l ACM SIGIR Conf. Research and Development in information Retrieval, pp. 96-103, 1998.

[13] G.H. John, R. Kohavi, and K. Pfleger, "Irrelevant Features and theSubset Selection Problem," Proc. 11th Int'l Conf. Machine Learning, pp.121-129, 1994

[14] F. Fleuret, "Fast Binary Feature Selection with Conditional Mutual Information," J. Machine LearningResearch, vol. 5, pp. 1531-1555, 2004.

[15] A.Y. Ng, "On Feature Selection: Learning with Exponentially Many Irrelevant Features as Training Examples," Proc. 15th Int'l Conf. Machine Learning, pp. 404-412, 1998.

[16] L. Yu and H. Liu, "Feature Selection for High-Dimensional Data: A FastCorrelation-Based Filter Solution," Proc. 20th Int'l Conf. Machine Leaning, vol. 20, no. 2, pp. 856-863, 2003.

[17] M.A. Hall, "Correlation-Based Feature Subset Selection for Machine Learning," PhD dissertation, Univ. of Waikato, 1999.

[18] J.R. Quinlan, C4.5: Programs for Machine Learning. Morgan Kaufman, 1993.

[19] L. Yu and H. Liu, "Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution," Proc. 20th Int'l Conf. Machine Leaning, vol. 20, no. 2, pp. 856-863, 2003.

[20] R.C. Prim, "Shortest Connection Networks and Some Generalizations," Bell System Technical J., vol. 36, pp. 1389-1401, 1957

[21] F. Pereira, N. Tishby, and L. Lee, "Distributional Clustering of English Words," Proc. 31st Ann. Meeting on Assoc. for Computational Linguistics, pp. 183-190, 1993. I.Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection," J. Machine Learning Research, vol 3, pp. 1157-1182, 2003

[22] .U. Fayyad and K. Irani, "Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning," Proc. 13th Int'l Joint Conf. Artificial Intelligence, pp. 1022-1027, 1993.

[23] J. Yu, S.S.R. Abidi, and P.H. Artes, "A Hybrid Feature Selection Strategy for Image Defining Features: Towards Interpretation of Optic Nerve Images," Proc. Int'l Conf. Machine Learning and Cybernetics, vol. 8, pp. 5127-5132, 2005.

AUTHORS

Chinnu C. George, is a PG Scholar in Ilahia college of Engineering and Technology, Muvattupuzha, Kerala. She received B-Tech Degree in Computer Science from Viswajyothi College of Enginnering and Technology (M G University) Vazhakulam, Kerala in 2007. Her research interest includedatamining and Natural Language Processing.

Abdul Ali, Assistant Professor, is an Assistant Profersor of Computer Science Department, in Ilahia college of Engineering, Muvattupuzha, Kerala. He received M-Tech Degree in Computer and Information Technology from Center for information technology and engineering University Campus,M S University, Tirunelveli, Tamilnadu in 2010. He received B-Tech Degree in Computer Science from M G University College Thodupuzha, Kerala in 2007 .His research interest include image processing and networking.

INTENTIONAL BLANK

International Journal on Cybernetics & Informatics (IJCI) Vol. 5, No. 2, April 2016

ENHANCING THE PERFORMANCE OF E-COMMERCE SOLUTIONS BY FRIENDS RECOMMENDATION SYSTEM AND NEO4J DATABASE

Shahina C P, Bindu P S and Surekha Mariam Varghese

[1]Department of Computer Science and Engineering, M. A. College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

In the past, selling needs brick and mortar store and it is limited to a local customer base. But today, ecommerce websites make it easy for the customers around the world to shop by virtual store. Performance of ecommerce websites depend not only the quality and price of the product but also the customer centric suggestions about the product and services and retrieval speed of websites. There is more possibility to buy a product brand suggested by friends or relatives than by viewing commends from unknown people. We can implement such a customer centric approach using Neo4j database, which provides easy and fast retrieval of information based on multiple customer attributes and relations than relational databases. Experiment shows the enhancement of performance in terms of time and complexity.

KEYWORDS

RDBMS, NOSQL, GRAPH DATABASE, NEO4J, E-COMMERCE, SOCIAL NETWORK, CYPHER QUERY LANGUAGE, MYSQL.

1. INTRODUCTION

Humans are social creature, so social recommendation within an ecommerce platform is more effective mainly because of the following reasons. Friends of friends recommendation helps customers to connect and build network faster and more organically. Product recommendations [5] help business to maximize their revenue. But most of the reviews are not reliable, the friends or related people's recommendations ensures reliability about a product.

This paper presents an ecommerce social network that contains a social community [2] of customers having relationships with other customers and thus with the products. Nowadays we have various social communities. A person may have different social community [2] membership like community of friends during school education, college education, professional community, other friend's community etc. E-commerce websites can make use of these social communities for implementing friends' recommendation system [3]. A customer who is aware of a particular product makes a review of that product telling that it is good or not. Based on this comment, a friend or a person related to him or a friend of a friend can take a decision to buy a product or not. Neo4j graph [3] database can be used to implement this system, since we can assure reliability, response time, efficiency and accuracy for large datasets and relations compared to relational database.

2. PROBLEM DEFINITION

Ecommerce is an application used for online shopping. It provides more efficient and cheaper distribution of products. It is very helpful for customers because of convenience, selection of product from more varieties than local shops. But sometimes, while accessing ecommerce platforms, we face issues such as

poor customer service after purchasing a product, poor product quality, delay in delivery of the product etc. So while purchasing a product via ecommerce website, we prefer to get feedback of other customers who had already bought that product earlier or who had already accessed that specific ecommerce platform. Some ecommerce websites provide recommendations, but there is no guarantee that the suggestions posted by ecommerce platform itself are reliable. So to provide reliable feedback / suggestions, we propose an ecommerce solution with an integrated friend's recommendation system. In this application a customer will get a review of a product which is suggested by a friend or a friend of friend via various social community [2] networks.

The first choice for building the ecommerce solution integrated with social networks is by using RDBMS. It is commonly used solutions to store the data in almost all applications. Sql is a language used for querying and maintaining the relational database. MySQL database is an open source relational database management system which stores data in the forms of tables. In our application different tables are required for storing personal attributes, products, relation etc. Each table has specified schema and relationship between different tables are represented using foreign keys. Many constraints are also required for this integrated ecommerce - social network solution. Considering all these points, we came to the conclusion that MySQL is not suitable to store large amount of semi structured data with dynamically growing behaviour.

The integrated solution is very powerful, but at the same time very challenging because the ecommerce websites and friend connections / relations are highly dynamic in nature. This solution will also have to handle the high density connections and may need complex queries to handle features such as customer centric personalized shopping. Traditional databases are very good in complex query processing but it cannot handle big data with dynamic schema changes [1]. So this application has to handle huge amount of data and relations with different varieties / categories without compromising the retrieval speed / response time because of the semantic checks [1].

Neo4j database are nosql, open source graph database which also supports ACID transaction [3]. The developers describe Neo4j as "embedded, disk-based, fully transactional that stores data in graph rather than in tables".Neo4j is highly advisable for managing complex relationship between heterogeneous data such as many views of customer problem is being addressed for various clients. In Neo4j, everything is stored in form of an edge, a node or an attribute. Data is represented as nodes; relationships between data can be represented as edges; both nodes and edges can have properties. Neo4j can scale up to billions of nodes and relationships.Neo4j graph databases supports complex querying because it stores data as relation which helps faster retrieval of information using graph traversals. Instead of using relational tables, neo4j uses graph nodes to store data. Similar to attributes in RDBMS, Neo4j uses properties for each node. Each node can have different number and types of properties. For example a node representing a person can have properties like name, DOB, designation etc. and node representing product can have properties like product code, product name etc. Similarly properties may vary person to person and product to product. In RDBMS relationship between different tables are represented as foreign key but here relationships between different nodes can be represented as edges. Similar to nodes, each edge can have properties. For example relationship between person to person may have friend, known since 10 years etc. and relationship between person and product may be buy, suggestions etc.

Another commonly used graph database is flock DB. Flock DB is simpler than neo4j but not capable of doing graph traversal in depth. Flock DB is intended to handle twitters single depth followers[9]. Ecommerce social network requires a complex graph consisting of many customers and item. This, in turn, requires in depth traversals while querying. Since flock DB is capable of only single depth traversal, neo4j is best suited for this application.

Cypher is a declarative graph query language for the Neo4j which allows expressive and efficient querying and updating the graph data. Cypher is a relatively simple but still very powerful language. Very complicated database queries can easily be represented using Cypher. Figure 1 shows ecommerce

application with friends' recommendation system including nodes with different person and products with different relationships between them. The relationships between different nodes are represented by edges.

Name	Neo4j	Mysql
Description	Open source RDBMS	Widely used Open source RDBMS
Database model	Graph DBMS	Relational DBMS
Implementation language	Java	C and C++
License	Open source	Open source
Developer	Neo Technology	Oracle
Transaction concepts	ACID	ACID
Server Operating System	Linux OS X Windows	FreeBSD Linux OS X Solaris Windows
Data scheme	schema-free	Yes
Typing	Yes	Yes
Secondary indexes	Yes	Yes
SQL	No	Yes

3. METHODOLOGY

By using cypher query our application can be implemented in following steps-

1. Creation of graph of social community networks of customers with relationships including friends, family colleagues, and other acquaintances along with their properties.
2. Creation of product nodes with their properties.

3. Associate customer nodes with product nodes with relationship like buy, suggest etc. Figure 1 show the graph generated.

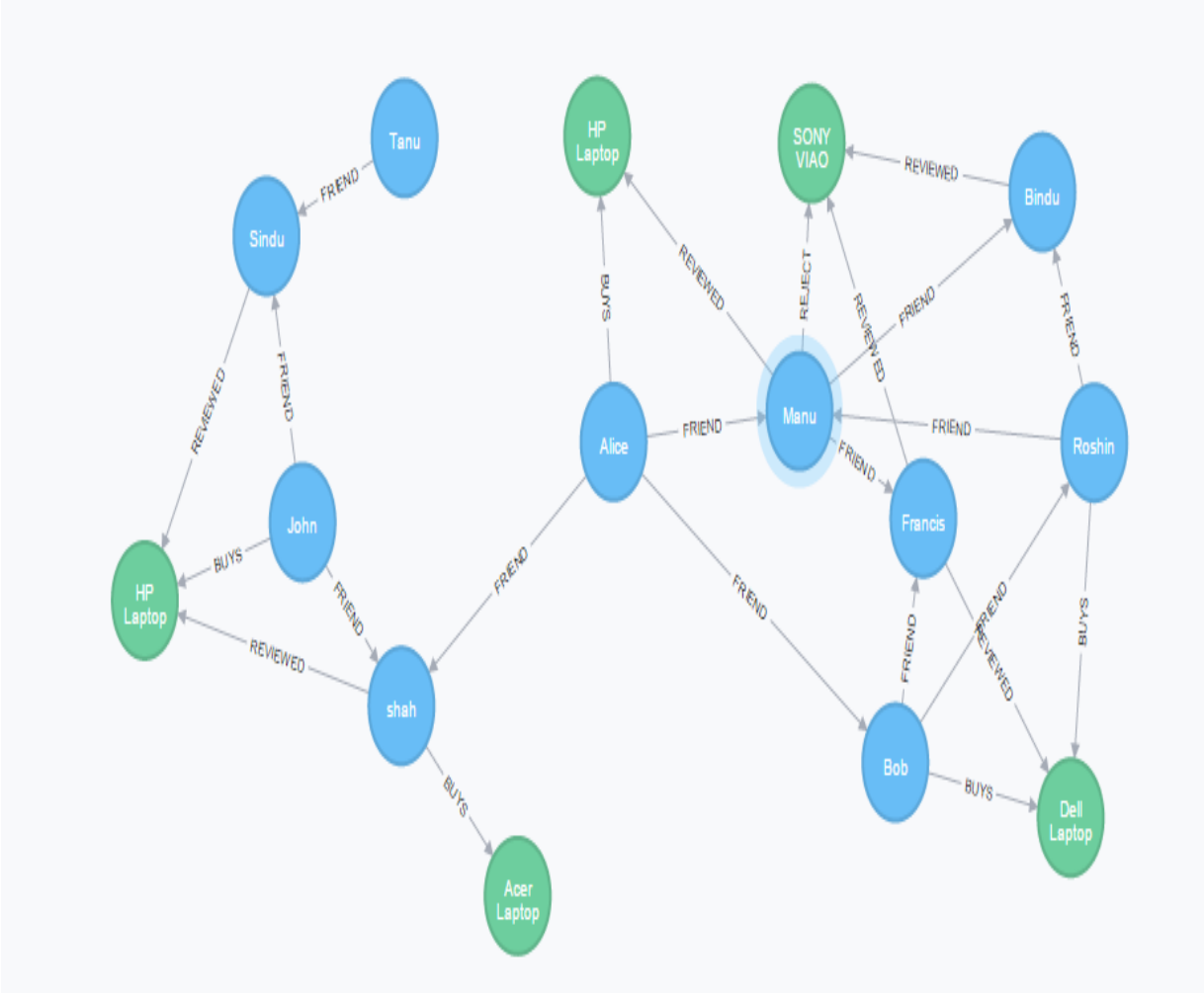


Figure 1. Example of ecommerce social network

Step 1 – Creation of Social Community

A set of customer nodes were created with properties like name, designation, age group, place etc. The edges representing relationships as friend were created. The properties for relationships were expressed as period of friendship and type of friendship such as close, very close, family friend.

Step 2 – Creation of Product Nodes

A set of product nodes with properties like name, price, model were created. The relationships between products and customers were represented with properties like Reviewed, Buys, and Reject. The properties

of Buys can be date of purchase, feedback, purpose etc. The property of Reviewed can be a comment such as Excellent, Bad service, Not good etc. The property of Reject can be Decision such as buy later.

For recommendation of a product, there are 3 factors [6].
1. Model data and data relationship to know how recommendation can be made
2. Make recommendation in real time by querying the relationship
3. Make it richer by adding data and relationship dynamically

4. PERFORMANCE ANALYSIS

4.1. WRITING QUERIES IN NEO4J

Graph data queries are straightforward to write and easy to understand. Graph databases have their own syntax for such queries. Cypher queries are much simpler than SQL queries; a long SQL statement can be divided to many queries using Cypher with much less number of lines.

4.1.1 How does Neo4j traverse a graph during a Query Execution?

Neo4j takes the cypher query as a metadata description of what we want and depending on the statistics, available indices it uses a combination of operations to perform the query. The main operations are look ups nodes, expand, expand into (between two nodes), hash-join, apply and semi apply [7].

4.2. QUERY PERFORMANCE

Query performance in a relational database is impacted by data growth and the number of JOINs [1]. As tables get bigger, so do indexes, which means that joining the same number of entities requires more and more time to execute. As questions / queries get more complex, the challenge increases as there is a huge number of entities to join. Even if the data volume stays constant, computational complexity explodes, which impacts query performance [1].

Neo4j is scalable and in this solution, it shows very small increase in the time to execute the query as data volume grows. This is because it doesn't compute relationships at query time but stores them at insertion time. In addition, graph queries look at the neighbourhood of starting points, so regardless of the total amount of data in the database, the amount of data that is examined remains roughly the same.

For example this project uses Neo4j to make real-time product recommendations by using information about what customer is looking for. Customers can view their friends recommendations and suggestions so can decide whether to purchase that product. In RDBMS, different tables are needed to store customer, product, order, friends etc. and different foreign keys are needed to connect these tables to provide the product recommendations.

Despite their name, relational databases do not store relationships between data elements in the same table i.e. relationship between two customers within a customer table is difficult in RDBMS. More over they are not well suited for ecommerce and social network systems which is having highly connected data with no specific schema. So as the volume, velocity and variety of data increases, the relationships grows even faster. So relational databases do not adapt well to changes.

Unlike a relational database, a graph database is structured entirely around data relationships. Graph databases treat relationships not as schema structure but as data. In Neo4j, the data is structured around data relationships so real-time query performance can be achieved no matter how large or connected the dataset.

For example we have three queries. First we evaluated these queries by 500 objects. The queries and results are shown in Figure 2. The three queries defined were:

q0: Find all friends of Bob.
q1: Find the product bought by Bob's friends.
q2: Find the suggestions of products that Bob's friend bought.

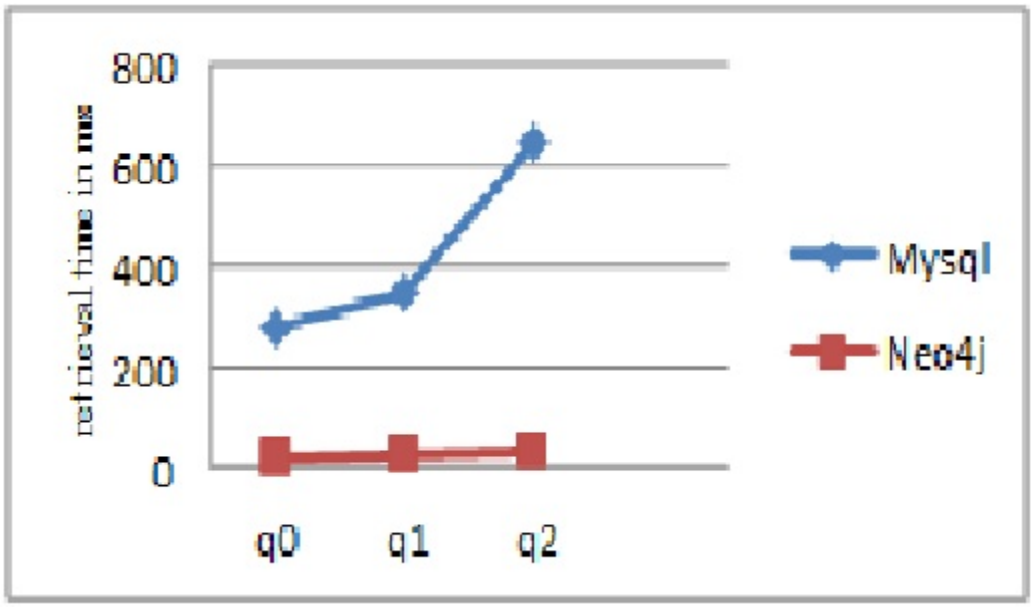


Figure 2. Retrieval time of queries having 500 objects by Neo4j and MySQL

5. CONCLUSIONS

In this paper, we implement an integrated solution to connect friends' recommendation system with ecommerce platforms using Neo4j graph database. We came to a conclusion that when compared with relational database, performance of Neo4j is better in terms of time and complexity. On comparison overall performance of graph databases exceeds the relational database

REFERENCES

[1] "Overcoming SQL Strain and SQL Pain" http://neo4j.com/resources/wp-overcoming-sql-strain

[2] "Social network-Ne4j graph database" http://neo4j.com/use-cases/social-network/

[3] "Nosql Databases", http://nosql-database.org/

[4] "Neo4j graph database" http://neo4j.com/developer/graph-database/

[5] "Real time recommendation with Neo4j" http://neo4j.com/use-cases/real-time-recommendation-engine/

[6] "Building recommendation engine with cypher in two minutes"- http://neo4j.com/resources/wp-recommendations-bu

[7] " Ne4j graph database" http://neo4j.com/blog/introducing-new-cypher-query-optimizer/

[8] "Neo4j" http://neo4j.com

[9] "Flockdb"https://en.wikipedia.org/wiki/FlockDB

Author

Shahina C P. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from MES College of Engineering, Kuttippuram. Her areas of research are Modern Database System and Machine Learning .

Bindu.P.S. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from SRM College of Engineering, Chennai. Her areas of research are Modern Database System and Data Mining .She is a principal of Govt.Polytechnic College, presently doing M Tech , on deputation. She has an experience of 21 years in various capacities as Lecturer, Head of Department and Principal in the department of Technical Education, Kerala.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology.Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Machine learning, Network Security, Database Management, Data Structures and Algorithms, Operating Systems and Distributed Computing. She has published 17 papers in international journals and international conference proceedings. She has been in the chair for many international conferences and journals.

INTENTIONAL BLANK

HEXAGONAL CIRCULARLY POLARIZED PATCH ANTENNA FOR RFID APPLICATIONS

Prakash K.C, Vinesh P.V., Jayakrishnan M.P., Dinesh R., Mohammad Ameen and Vasudevan K.

Center for Research in Electromagnetics and Antennas (CREMA), Dept. of Electronics, Cochin University of Science and Technology, Kochi-22, Kerala, India.
keyceepee@gmail.com

ABSTRACT

A compact design of a hexagonal single feed circularly polarized microstrip antenna for RFID applications is proposed. This structure fabricated on FR4 substrate offers compactness, good axial ratio bandwidth with a broadside radiation characteristic in the entire band, better gain, good impedance bandwidth at a resonant frequency of 2.45 GHz. The structure is suitable for RFID reader antenna applications.

KEYWORDS

RFID, Hexagonal Patch, CP, Elliptical Slot & RHCP.

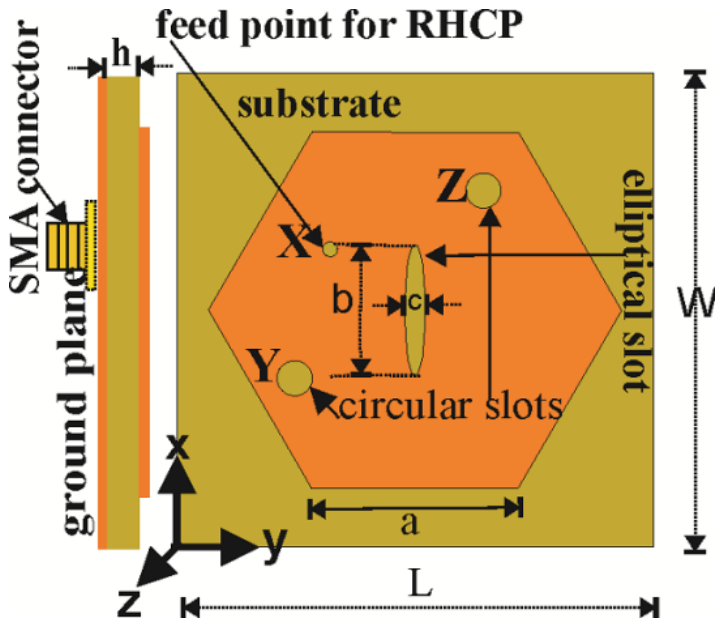
1. INTRODUCTION

There are a lot of advantages for circularly polarized (CP) antenna which make it attractive for many wireless systems. As strict orientation between receiving and transmitting antenna is not required, this avoids false identification of targeted objects by RFID readers and hence no polarization mismatch losses. The CP antenna is very effective in combating multi path effects. A single co-axial fed circularly polarized antenna is proposed in this work. It is simple in its structure than dual-fed ones and the generated mode is usually excited in an electrically thin cavity region of the microstrip antenna. The circularly polarized radiation is generated by the effect of asymmetries or perturbation segments in the patch. Various perturbation methods have been adopted in literature for designing a single feed circularly polarized microstrip antenna. Truncating the corners of the hexagonal patch and inserting slits[1], One 'V' shaped slot embedded into a cross shaped rectangular patch antenna[2], asymmetric-circular shaped slotted microstrip patch antennas with slits[3], two crossed PIFA[4], a cross-strip embedded along the X-shaped slot for a proximity-fed technique [5], antenna using a set of slits and slots[6], a cross slot embedded on the radiating patch[7], truncating corners of a square patch and inserting slits of different lengths at the edges of a square patch[8, 9], two slits inserted at the annular-ring patch[10], embedding a circular central elliptical slot on a circular patch [11], truncating the corners of a polygonal patch are some of the techniques employed to radiate circularly polarized waves. In this structure, a hexagonal microstrip patch antenna with the perturbation technique of embedding a central elliptical slot and two circular slots is proposed. The dimensions of the patch and slots and the feed point location are extremely important and optimized for good performance. It exhibits good circular polarization radiation characteristics.

2.ANTENNA GEOMETRY

The proposed antenna geometry consists of regular hexagonal patch with a central elliptical slot and two circular slots; one on the right side and the other on the left side of the elliptical slot. It is fabricated on an FR4 substrate with a dielectric constant 4.4, thickness 1.6 mm and a tan δ = 0.02. As depicted in Figure

1, the elliptical slot is situated on the center of the patch oriented along the X axis. The dimensions of the elliptical slot are designated as ***b*** the major radius (base radius), the secondary radius ***c*** and the aspect ratio ***k*** the ratio of ***c*** to ***b***. The radius of the circular slot r_a is chosen to be 3mm.



(a) (b)

Figure 1. Geometry of the proposed antenna. 1(a) side view and 1(b) Top view, a = 17.52mm, b = 11.2mm, c = 0.84mm, k = 0.15 , L = W = 40.4mm, guide wavelength λg = 0.0606 m, ε_{reff} = 4.087,X(25.4mm,13mm),ra = 3mm, Y(14.4mm,10 mm,1.6mm) Z(30.4mm,26mm,1.6mm)

In this structure, elliptical and the circular slots are chosen as the detuning elements to split the fundamental TM11 mode to two orthogonal modes TM01 and TM10 [11]. The dimensions of the elliptical and circular slots are so crucial that it affects the radiation of circularly polarized waves. The antenna is excited through co-axial probes at the feed location X to give circularly polarized electromagnetic radiation in the clockwise direction, i.e., Right Handed Circular Polarization (RHCP). The feed point is chosen along the locus of 50Ω characteristic impedance and is adjusted for good matching. The photograph of the proposed antenna is depicted in Figure 2.

Figure 2. Photograph of the antenna prototype

3. DESIGN

As the circular patch and hexagonal patch are closely related to each other, the design of a circular patch is considered for the design of a hexagonal patch antenna [12].

Then by equating the areas of a circle and a regular hexagon, the edge or side length of the regular hexagon a may be found [12] out as,

$$a = r_e \sqrt{\frac{2\pi}{3\sqrt{3}}} \quad (1)$$

By the proper selection of the feed point, the dominant mode is detuned into two equal amplitude orthogonal degenerate modes, orthogonal even mode TM_e and orthogonal odd mode TM_o, without considering the higher order modes. By introducing the asymmetry, the resonant frequencies of two degenerate modes are made nearly equal.

The normalized impedance ratio of the detuned modes is a complex number, which has a magnitude part and phase part. The phase part corresponds to the phase difference of the CP field components, while the magnitude is proportional to the axial ratio. The angle is designed to be 900 and the magnitude to be less than 3 dB. Thus the criterion for circular polarization is satisfied. By selecting appropriate dimensions for the slot, the two degenerate mode frequencies are made nearly equal.

By introducing the asymmetry due to the central elliptical slot and the two circular slots, the dominant mode of the patch is detuned into two orthogonal degenerate modes, the odd mode and the even mode. The odd mode is oriented about the X axis and the even mode is oriented about the Y axis. The odd mode suffers minimum perturbation due to the elliptical slot. As the fundamental mode exhibits a zero in this area, the fields will not be affected significantly. If the slot had been a rectangular geometry with no change in orientation, the odd mode surface current would not have been affected by this perturbation and the resonant angular frequency would have been approximately same as that of the fundamental resonant frequency. In the even mode (oriented about the Y axis), the surface current is forced to traverse more length around the elliptical and circular slots and hence the resonant frequency ω_e is reduced. If a second orthogonal cut is made at the center [13], it will enhance the electrical path of the odd mode and the corresponding resonant frequency will be reduced. Here in this proposed structure, the effect of more electrical length is achieved by the two circular slots and the appropriate dimension of secondary radius for the elliptical slot. Thus the odd mode resonant frequency ω_o is reduced. Thus surface currents in both odd mode and even mode are reduced. The dimensions of the slots have a major role in determining the resonant frequency for both the modes. The overall dimension of the antenna decreases, which results in the same central frequency for the two modes. It was demonstrated in [14] that with annular shapes it is possible to enhance the bandwidth because less amount of stored energy is stored beneath the patch metallization and quality factor is thereby reduced. Both the effects of splitting of TM11 mode into orthogonal modes TM01 and TM10 and enhancement of bandwidth due to annular shape can be achieved by introducing an elliptical slot on the center of the patch and two circular slots on both sides. A mandatory condition for achieving circular polarization is equal amplitude and 900 out of phase for the degenerate modes. By selecting proper slot dimensions and feed location, the two modes are made equal in amplitude and a phase difference of 900 between them. This results in circular polarized radiation. Thus the perturbation produced by the elliptical slot and circular slots gives rise to mode degeneration. The overall resonant frequency is reduced. The edge of the hexagonal patch is designed for a resonant frequency of 2.63 GHz. Due to the elliptical and circular slots the resonant frequency is reduced to 2.45 GHz and thus the overall dimension is reduced by a factor of 18.10%.

To design the hexagonal patch, design equations are derived based on the geometry and the frequency of operation after computing the guide wavelength λg, that is the wavelength in the dielectric given by,

$$\lambda_g = \frac{\lambda_0}{\sqrt{\varepsilon_{reff}}} \quad (2)$$

where λ_0 is the free space wavelength at 2.45 GHz and the effective permittivity of the substrate ε_{reff} is computed as 4.106, using the formula in [15] as,

$$\varepsilon_{reff} = \varepsilon_r - \frac{C_r \varepsilon_r}{2}\left(\frac{2h}{x} + \frac{h^2}{x^2}\right) \quad (3)$$

where $C_r = 0.7$, being a thin substrate, x is the radius of the circumscribed circle - the circle in which the hexagon is inscribed- and h is the height of the substrate.

4. RESULTS AND DISCUSSION

4.1. Reflection characteristics

Using the computed parameters, the antenna was simulated using Ansys HFSS 13.0 and tested using ZVB20 vector network analyzer. Optimum feed point was chosen for the best impedance matching of the antenna. The measured and simulated reflection characteristics of the proposed antenna at the co- axial feed point X are plotted in Fig.3. The fundamental resonant frequency of the antenna without slots is 2.63 GHz, whereas with slot the simulated resonant frequency is 2.45 GHz.

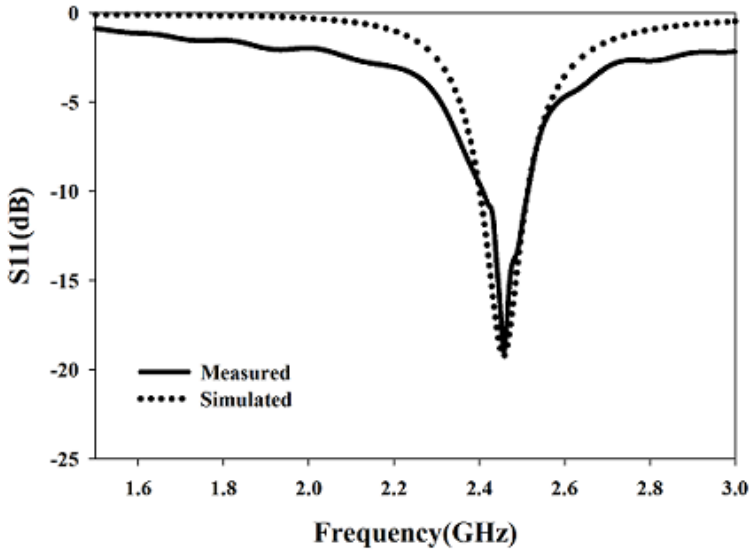


Figure 3. S11 plot of the proposed structure.

4.2. Axial ratio

The axial ratio is the magnitude of the normalized impedance ratio (the ratio of two detuned modes). Simulation results show that for various values of aspect ratio **k** (the ratio of secondary radius to base radius), the value of minimum axial ratio varies. It is observed that when **k** exceeds 0.08, axial ratio becomes less than 3 dB and hence circular polarization radiation is exhibited. The CP radiation at **k**= 0.15 is confirmed from the simulated axial ratio graphs shown in Figure 4, where the necessary criterion for CP, the axial ratio <3 dB is satisfied.

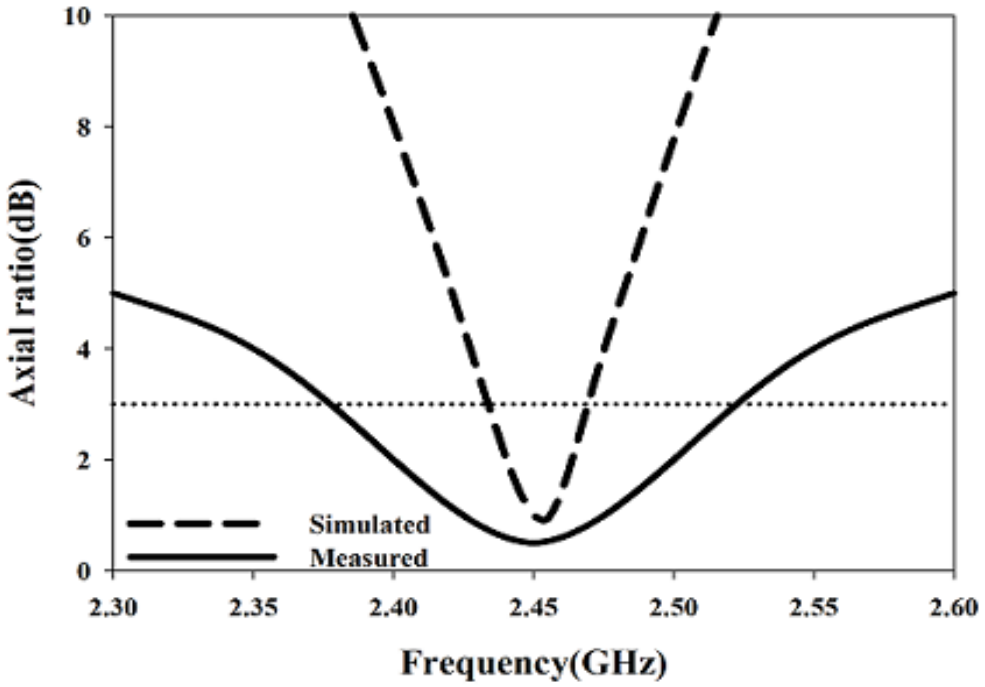


Figure 4. Axial ratio plot

4.3. Surface current distribution

The surface current distribution simulated at the center frequency of 2.45 GHz is plotted in Figure 5. Surface current distribution at $\varphi = 0^0$ is equal in magnitude and opposite in direction to that at $\varphi = 180^0$. Same is the case of surface current distribution at $\varphi = 90^0$ and $\varphi = 270^0$ and hence the criterion for CP is satisfied. The direction of rotation of current is clockwise in the +Z axis and the sense of polarization is confirmed as right handed circular polarization.

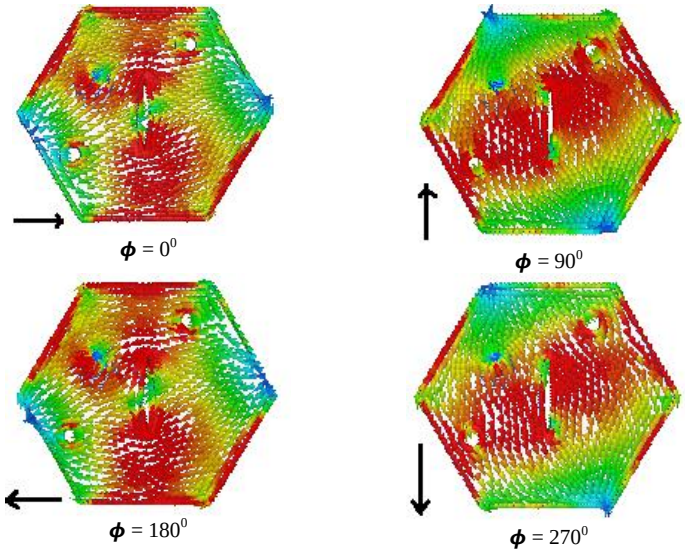
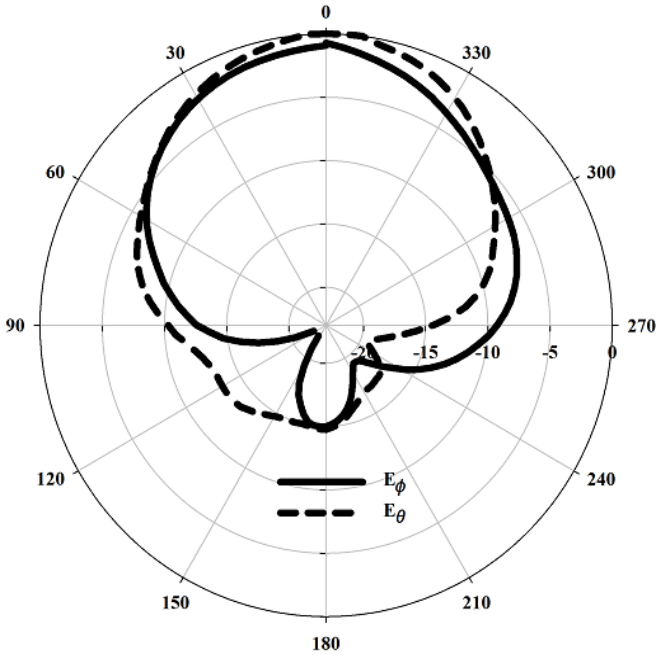


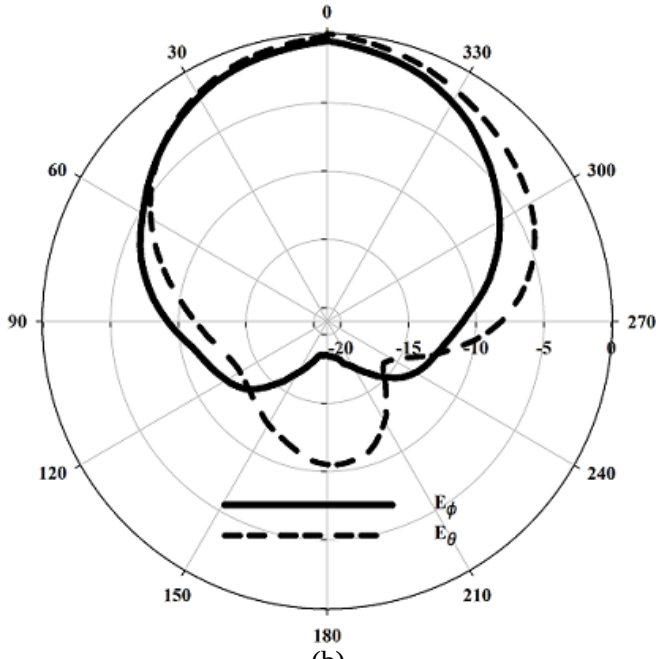
Figure 5. Surface current distribution of the proposed antenna

4.4. Radiation pattern

The measured radiation patterns of the proposed antenna in XZ and YZ planes for $\varphi = 0^0$ and 90^0 are plotted in Figure 6. The pattern gives a 3 dB beam width of 80^0.



(a)



(b)

Figure 6. Radiation pattern of the antenna in (a) XZ plane (b) YZ plane

4.5. Gain

The gain of the antenna is measured using gain comparison method and the measured value is 4.5 dBi at 2.45 GHz.

4.6. Smith chart

The simulated Smith chart is depicted in Figure7.The dip in it corresponds to the center frequency 2.45 GHz indicates the excitation of two orthogonal degenerate modes, justifying the circularly polarized radiation.

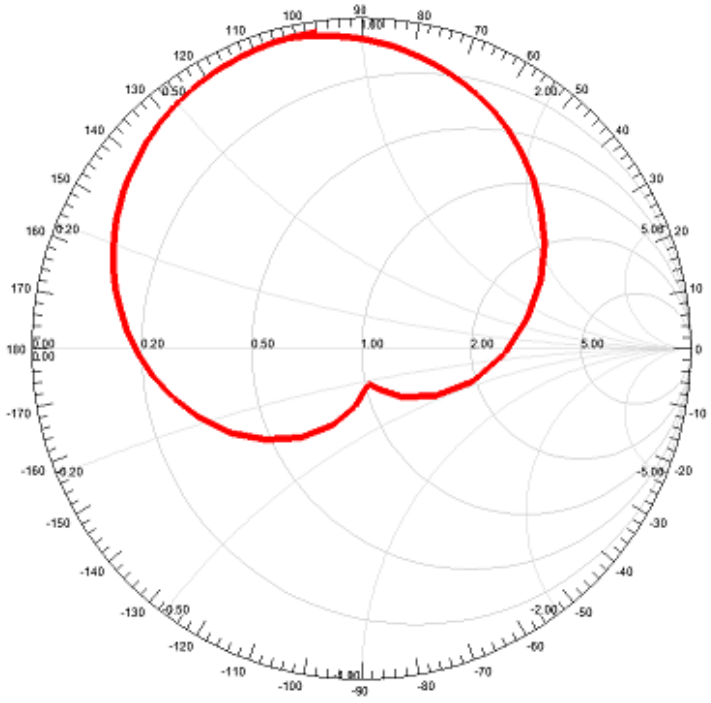


Figure 7. Smith chart

The measured values of the antenna are compared with the circular patch antenna with elliptical slot [11] and tabulated in Table 1.

Table 1. Comparison of measured values

Parameter	**Circular patch antenna with elliptical slot [11]**	**Proposed antenna**
Center frequency	2.45 GHz	2.45GHz
Return loss	19 dB	19dB
Gain	3.85dBi	4.5 dBi
Minimum axial ratio	0 dB	0.5dB
Axial ratio bandwidth	30 MHz (1.22%)	143 MHz(5.8%)
10dBimpedance bandwidth	130 MHz (5.3%)	100MHz(4.08 %)

3 dB beamwidth	-	80^0
Area reduction	9.5%	18.06 %

5. CONCLUSIONS

A novel design of a single feed circularly polarized hexagonal patch antenna has been designed, simulated and experimentally investigated. The elliptical and circular slots perturbation method yields better circular polarization characteristics. The experimental results confirm that it is suitable for RFID applications in the 2.45 GHz band.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the University Grants Commission (UGC), Sree Ayyappa College, Travancore Devaswom Board Management, University of Kerala and Cochin University of Science and Technology for facilitating an opportunity for research under the Faculty Development Programme of UGC.

REFERENCES

[1] K. Qian & X. Tang (2011), " Compact LTCC dual – band circularly polarized perturbed hexagonal microstrip antenna," *IEEE Antennas and Wireless Propag. Lett.,* vol. 10, ,pp. 1212 –1215.

[2] M. S. Nishamol, V. P. Sarin, D. Tony, C. K. Aanandan, P. Mohanan, & K. Vasudevan, (2011) "Design Of A Circularly Polarized Rectangular Microstrip Antenna for GPS Applications," *Microwave and Optical Technology Letters*, vol. 53, no. 2, pp. 468–470.

[3] Nasimuddin, Z. N. Chen & X. Qing, (2010), "Asymmetric-circular shaped slotted microstrip antennas for circular polarization and RFID applications," IEEE Trans. Antennas Propag., vol. 58, no. 12, 2010, pp. 3821–3828.

[4] S.Pflaum, F. Canneva, P. Le Thuc, G. Kossiavas & R. Staraj, (2013) " PIFA antenna with tilted circular polarization angle for RFID readers," *IEEE Antennas Propag. Soc. AP-S Int. Symp*.,,pp. 1734–1735.

[5] Horng-Dean Chen, Shang-Huang Kuo, Chow- Yen- Desmond Sim & Ching-Han Tsai (2012) "Coupling-Feed Circularly Polarized RFID Tag Antenna Mountable on Metallic Surface," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 5.

[6] D. L. Nguyen, K. S. Paulson & N. G. Riley, (2012) "Reduced-size circularly polarised square microstrip antenna for 2.45 GHz RFID applications," *IET Microwaves, Antennas Propag.*, vol. 6, no. 1, p. 94.

[7] J.S. Row & C.Y. Ai,(2004) "Compact Design Of Single Feed Circularly Polarized Microstrip Antenna",*Electronic Lett* 401093– 94.

[8] W.S. Chen, C.K. Wu & K.L. Wong, (2001) " Novel Compact Circularly Polarized Square Microstrip Antenna", IEEE Trans on Antennas and Propagation, 49, 340–342.

[9] J.S. Roy & M. Thomas, (2008), "Design of a circularly polarized microstrip antenna for WLAN", *Prog Electroma*gnet*ics*, Res 3, pp: 79–84.

[10] S. S. Gao, Q. Luo, & F. Zhu, (2014) *Circularly Polarized Antennas* John Wiley & Sons Ltd, United Kingdom..

[11] S.Maddio, A. Cidronali & G. Manes, (2011) "A New Design Method For Single- Feed Circular Polarization Microstrip Antenna With An Arbitrary Impedance Matching Condition", *IEEE Transactions on Antennas and Propagation.*, vol. 59, no. 2, pp. 379 –389.

[12] Nagendra Kushwaha & Raj Kumar,(2013) " Design Of Slotted Ground Hexagonal Microstrip Patch Antenna And Gain Improvement With FSS Screen" *Progress In Electromagnetics Research B,* vol. 51, pp: 177–199.

[13] Horng- Dean Chen, Shang-Huang Kuo, Chow- Yen- Desmond Sim & Ching-Han Tsai, (2012) "Coupling-Feed Circularly Polarized RFID Tag antenna Mountable on Metallic Surface," *IEEE Transactions on Antennas and Propagation,* vol. 60, no. 5,Pages: 2166 - 2174,

[14] Bhattacharyya, A. & L. Shafai, "A Wider Band Microstrip Antenna For Circular Polarization," (1988) , *IEEE Transactions on Antennas and Propagation,* vol. 36, no. 2, pp 157–163.

[15] Ramesh Garg, Prakash Bhartia, Inder Bahl & Apisak Ittipiboon, (2001) *Microstrip Antenna Design Handbook ,* Artech House, London.

AUTHORS

Prakash K.C was born in the year1970. He passed his M.Sc Degree in Electronics Science from Cochin University of Science and Technology, Kerala, India, in 1993, with third rank.with third rank. He has got 22 years of teaching experience in Electronics and holds the designation of Associate Professor in Sree Ayyappa College, Eramallikkara; a Kerala Govt. Aided college affiliated to the University of Kerala. Currently he is engaged in research under the faculty development programme of UGC, in the area of microwave antennas for RFID applicatios. He has got 1 international journal and four international conference publications to his credit. He is a Fellow of IETE. His areas of interest include RFID CP antennas, metamaterial, energy harvesting etc.

Vinesh P V received the B.Sc. degree in electronics from the University of Kannur, India,and the M.Sc. degree in Electronics fromthe MES College Erumely, Kottayam, India, in 2004 and 2006,respectively. He is currently working towards the Ph.D.degree at Cochin University of Science andTechnology (CUSAT), Kochi, India. His research interests include designing of multiband antennas, planar inverted F antennas, ZOR antenna etc

Jayakrishnan M.P reeceived the B.Sc. degree in Electronics from the Mahatma Gandhi University, Kottayam, India and M.Sc. degree in Electronics Science from the Cochin University of Science and Technology, Kochi, India, in 2012 and 2014,respectively. He is currently working towards Ph.D. degree in Microwave Electronics at the Cochin University of Science and Technology (CUSAT), Kochi, India. His research interests include Microwave based Bio-Sensors, Implantable Antennas, Frequency Selective surfaces, Metamaterials etc.

Dinesh .R was born in the year1985. He passed his M.Sc Degree in Electronic Sciencefrom Cochin University of Science and Technology in 2005. He has got 6 years of research experience in Microwave field. Currently he is an Assistant Professor at NSS college, Rajakumari, Idukki; a Kerala Govt. Aided college affiliated to the MG University , Kerala. He has got 13 international journal and 15 International conference publications to his credit.

Mohammad Ameen was born in the year 1990. He received the B.Tech degree and M.Tech degree in Electronics and Communication Engineering from Mahatma Gandhi University, kerala, in 2012 and 2014, respectively. He has 1 year of research experience in Microwave field. Currently he is a Senior Research Fellow in CSIR (New Delhi) funded project at Department of Electronics, Cochin University of Science and Technology, Cochin, Kerala. He has got one international journal and four international conference publications to his credit. His research interests include high gain antennas and MIMO antennas.

K.Vasudevan took PhD in 1982 in Microwave antennas and joined as lecturer in Cochin University of science and technology in 1985. He is a professor in this University since 1995. He was head of department of Electronics from 2004 to 2010 and Dean, faculty of technology during 2010 to 2013. Currently, he is a CSIR Emeritus scientist in Cochin University. He has more than 220 publications in International journals and conference proceedings and has more than 1200 citations and an h- index of 19. He has been an investigator for about 20 major projects from various agencies having a total outlay of 30 crores of rupees. He has chaired and presented papers in several IEEE international conferencesabroad.. He is a senior member of IEEE and Fellow of IETE.

IMPLEMENTATION OF LINEAR DETECTION TECHNIQUES TO OVERCOME CHANNEL EFFECTS IN MIMO

Gopika k [1] and M Mathurakani[2]

[1]Student in M.Tech, ECE Department, Toc H Institute of Science and Technology, Ernakulam, India
[2]Professor, ECE Department, Toc H Institute of Science and Technology
Formerly Scientist, DRDO, NPOL, Kochi,Ernakulam, India

ABSTRACT

Spatial diversity technique enables improvement in quality and reliability of wireless link. Antenna diversity along with understanding effects of channel on transmitted signal and methods to overcome the channel impairment plays an important role in wireless communication where sharing of channel occurs between users. In this paper single input single output system (SISO) is compared with multiple input multiple output system (MIMO) in terms of bit error rate performance. Bit error rate performance is also evaluated for MIMO with least squares (LS) and Minimum mean square error (MMSE) linear detection. Further analysis and simulation is done to understand the effect of channel imperfections on BER.

KEYWORDS

Spatial diversity; SISO; MIMO; MMSE; LS

1. INTRODUCTION

Today's and future wireless communication devices are expected to support various multimedia services and demands a huge bandwidth, but it is a scarce resource and hence needed to be managed and utilized carefully. To effectively utilize this limited spectrum different users have to share the channel. The communication takes place in same space and uses same spectrum, interference will occur as a result and is a major limitation in wireless communication. So interference needs to be managed effectively, otherwise it will affect the system performance and limits the capacity that the system can achieve.

Evolution of antenna terminology started from transmitter and receiver equipped with single antennas for transmitting and receiving [1]. Achievable capacity was limited and the system known as SISO worked best for only line of sight distances. As a communication scenario is concerned not only line of sight but non line of sight communication also takes place. To achieve better capacity, spatial diversity techniques were implemented thus evolved multiple antenna terminals at transmitter called as multiple input single output (MISO) system enabling transmit diversity and multiple antenna terminals at receiver called single input multiple output (SIMO) system enabling receive diversity. These systems even though worked better when comparing

DOI: 10.5121/ijci.2016.5220

with SISO, capacity achieved was poor. It necessitates the importance of achieving diversity at both transmitter and receiver leading to the MIMO systems, where in this system the capacity achieved depends on the minimum number of transmit or receive antennas [2]. Significant advantages of MIMO systems are increase in both system capacity and spectral efficiency. The capacity of a wireless link increases linearly with the minimum number of transmitter or receiver antennas. The data rate can be increased by spatial multiplexing without consuming more frequency resources and without increasing total transmitter power. Reduction of effects of fading due to increased diversity is particularly beneficial when different channels fade independently. MIMO system enabled joint processing or combining of signals and system integrity. A single MIMO system is called single user MIMO, because the single user terminal with which both transmit and receive diversity is enabled. In communication space not a single user but large numbers of users are present leading to the evolution of multi user MIMO. Therefore future works in this project will be focusing towards multiuser MIMO and interference alignment techniques.

The quality of a wireless link can be described by three parameters, namely the transmission rate, transmission range and transmission reliability. Transmission rate can be increased by reducing transmission range and reliability. By reducing transmission rate and reliability transmission range can be increased, while transmission reliability can be increased by reducing transmission rate and range. By combining two important technologies MIMO technology and orthogonal frequency division multiplexing (OFDM) above parameters can be simultaneously improved.

2. ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING

In single carrier modulation the transmitted pulse width has to be reduced as it is occupying more bandwidth. If width of transmitted pulse is reduced channel which was narrow band channel initially starts behaving like a wideband channel resulting in severe inter symbol interference, so single carrier modulation is not effective. When a signal propagates through a mobile radio channel the transmitted signal will undergo a variation in its characteristics like amplitude, phase etc. referred to as fading of a signal. Multipath phenomenon will generate two effects mainly, frequency selective fading and intersymbol interference. Frequency selective fading occurs when the signal is transmitted through a constant gain channel with linear phase response over a bandwidth that is smaller than the bandwidth of transmitted signal. When a signal undergoes frequency selective fading at the receiver the received signal will get distorted and multiple versions of transmitted signal will be received which is attenuated and delayed in time i.e. for some frequencies in the bandwidth the channel does not allow any information to go through and thus deep fades occurs to particular frequencies. It does not occur uniformly across the band but occurs at selected frequencies. OFDM is implemented to overcome these impairments that the signal is suffering from when transmitted through a shared channel [3].

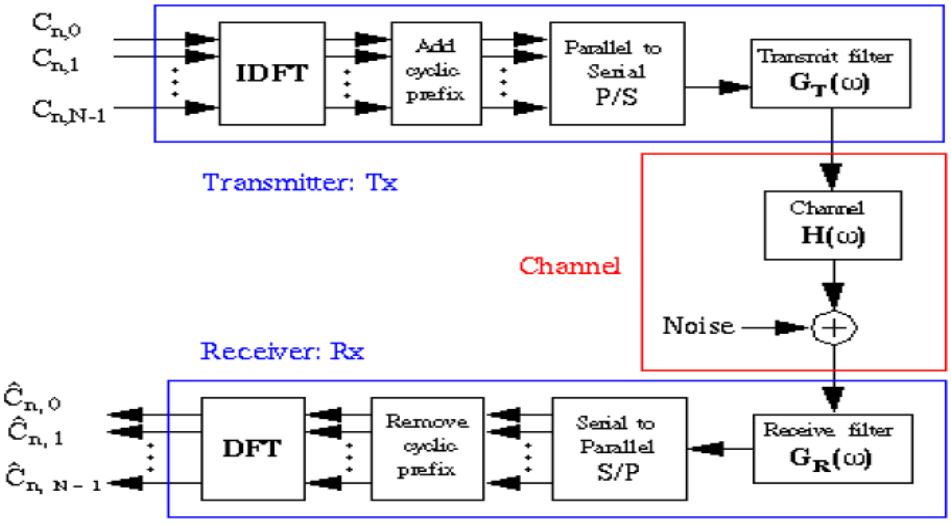


Figure 1. OFDM system

OFDM is a discrete implementation of multicarrier modulation; it divides the entire bit stream to be transmitted in to sub streams and sends them over different subchannels. The subchannels are orthogonal and the number of subchannels is designed such that each subchannel has a bandwidth less than the coherence bandwidth of channel.

Figure 1.shows the block diagram of OFDM system. The incoming bits are mapped to symbols according to some modulation scheme like quadrature phase shift keying (QPSK). The serial data is converted in to parallel blocks, then each block of symbols is forwarded to inverse fast fourier transform (IFFT) block and OFDM modulated. Then the OFDM signal will be appended with a cyclic prefix by copying last portion of OFDM signal. The cyclic prefix length is chosen such that its length should be larger than the maximum path delay of the channel, to eliminate intersymbol interference (ISI). Then the serially converted OFDM signal is transmitted. At the channel this signal will undergo some transformation like convolution with the channel; the linear convolution at the channel will be converted in to circular convolution due to the presence of cyclic prefix. At the receiver a reverse process that took place in transmitter occurs, the signal is converted in to parallel signals and cyclic prefix is removed, fast fourier transform (FFT) of signal is taken and channel effect can be removed by simply dividing this signal FFT with channel FFT, an advantage of cyclic prefix and hence circular convolution is that circular convolution in time domain will be converted in to multiplication in frequency domain, by doing this frequency selective fading channel is converted in to flat fading channel in subchannel perspective. OFDM is easy to implement in digital domain, its bandwidth efficient, robust to fading and is flexible in resource allocation.

3. MIMO SYSTEM MODEL

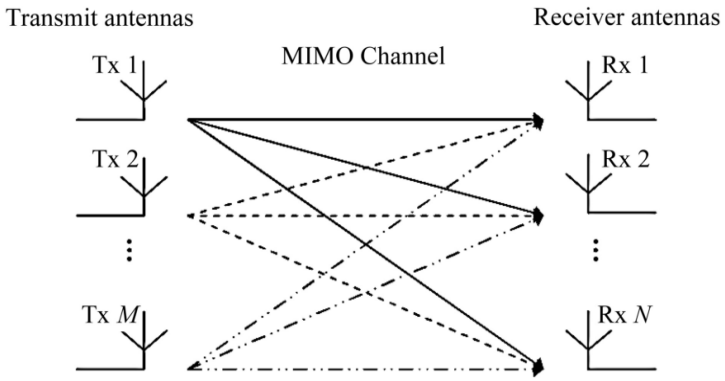


Figure 2. MIMO system

3.1 MIMO

Multiple input multiple output is a method implemented to improve the system capacity and capacity of a radio link using multiple antenna terminals at both transmitter and receiver side by utilizing multipath propagation. Through the multiple transmitting antennas either multiple copies of transmitted data streams to enhance the diversity gain or multiple data streams to improve the spatial multiplexing gain can be transmitted. When the multiple copies of data streams are transmitted from different terminals each stream will choose different paths to its receivers, during its propagation some of the data streams undergoes deep fading but other streams through different path survives and can be received at the receiver antenna with diversity enabled and datas can be recovered. Figure 2.illustrates a MIMO system with M transmitting and N receiving antennas and a MIMO channel in between them, the size of channel matrix is NxM

3.2 Channel State Information

As the data streams S passes through the channel H, signal undergo some variation and noise N will get added in to it. So the received stream Y will be,

$$Y=HxS+N. \quad (1)$$

Receiver can suppress the effects of noise by increasing signal to noise ratio. But to deal with H receiver needs to have the knowledge about channel. Receiver has to be simple in terms of cost and size, but if the channel is predicted at the receiver and thus detecting data will increase the complexity of receiver. So the transmitter will do the hard work of predicting the channel and sending it along with the data streams through precoding. If the channel assumed is imperfect, error rate will increase and received signal will be distorted. Channel state information can be acquired through assuming reciprocity of alignment, it's the signalling dimensions along which a receiving node sees the least interference from other users are the signalling dimensions along

which this node will cause least interference to other nodes in reciprocal network where all transmitters and receivers switch roles.

4. SPACE TIME CODING AND SPACE FREQUENCY CODING

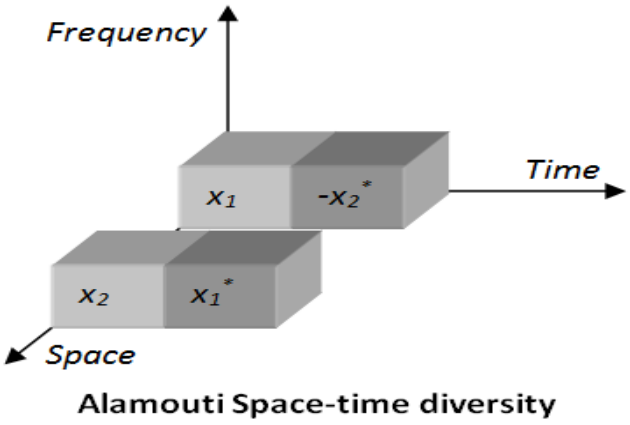


Figure 3. STBC illustration

Space time block codes (STBC) are spatial temporal codes that use diversity enabled in transmitter and receiver, multiple copies of data are transmitted in the hope that at least some of them may survive the physical transmission path. Transmit diversity proposed by alamouti was the first space time block codes [4]. Figure 3 shows the space time coding implementation using two transmit and two receive antennas. STBC in particular, the data stream to be transmitted is encoded in blocks, which are distributed among spaced antennas and across time. While it is necessary to have multiple transmit antennas, it is not necessary to have multiple receive antennas, although to do so improves performance. This process of receiving diverse copies of the data is known as diversity reception. An STBC is usually represented by a matrix. Each row represents a time slot and each column represents one antenna's transmissions over time. Two symbols X1 and X2 are transmitted from transmitter 1 and 2 at time slot 2n then at timeslot 2n+1,-X2*,X1* are transmitted from two transmitters to enable diversity. Simple linear operations are performed at the receiver such that X1 should be received at receiver1 and X2 at receiver2.

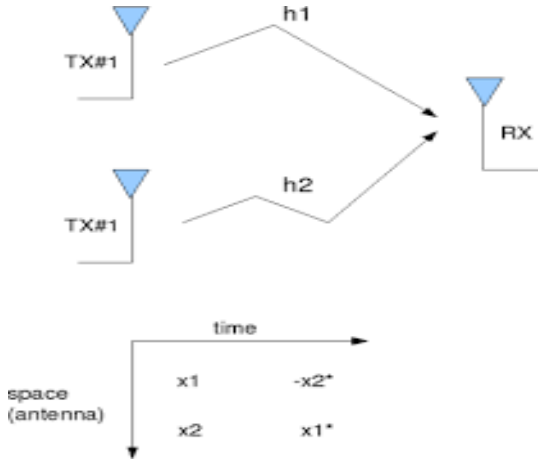


Figure 4. Alamouti encoding

Figure 4.shows the signals received at two receivers at two different slots 2n and 2n+1 can be represented by set of equations

Y1(2n)=H11xX1+H12xX2. (2)
Y2(2n)=H21xX1+H22xX2. (3)
Y1(2n+1)=-H11xconj(X2)+H12xconj (X1). (4)
Y2(2n+1)=-H21xX2+H22xconj (X1). (5)

Hij is the channel response between j^{th} transmitter and i^{th} receiver. Y1 and Y2 are the symbols received at receiver 1 and 2 respectively. Here two transmit and receive antennas are used for simple implementation; higher order diversity can be enabled by increasing the number of antennas. The symbols can be retrieved at receiver

X1=conj(H11)xY1(2n)+H12xconj(Y1(2n+1)) +conj(H21)xY2(2n)+H22xconj(2n+1). (6)
X2=conj(H12)xY1(2n)+H11xconj(Y1(2n+1)) +conj(H22)xY2(2n)-H22xconj(2n+1). (7)

Space frequency coded alamouti transmission scheme is over different frequency rather than over different time slots as in space time coding, where a symbol goes through four different paths at two different frequencies thus achieving frequency and space diversity [5].

5. LEAST SQUARES AND MMSE LINEAR DETECTION

The transmitted signal after passing through the channel will undergo Rayleigh fading and additive white Gaussian noise will get added to the signal. So to remove the channel effect LS or MMSE based linear detection can be used.

In LS the linear detection is done as
$S=(HxH^{H})^{-1}xH^{H}xY.$ (8)

In MMSE the linear detection is done as

$$S=(HxH^{H}+\sigma^{2})^{-1}xH^{H}xY. \quad (9)$$

H is the channel matrix; Y is the received symbol vector; σ^2 is the noise variance. MMSE considers effect of noise in linear detection procedure and is assumed to perform better than LS based linear detection.

6. SIMULATION RESULTS

To understand the effect of OFDM on transmitted symbol, coding is done for SISO-OFDM and plotted the BER vs. SNR graph. For the simulation 51200 bits are transmitted during each iteration, size of FFT taken is 512, cyclic prefix length is 10 and channel length is 6. The simulation result is shown in figure 4.

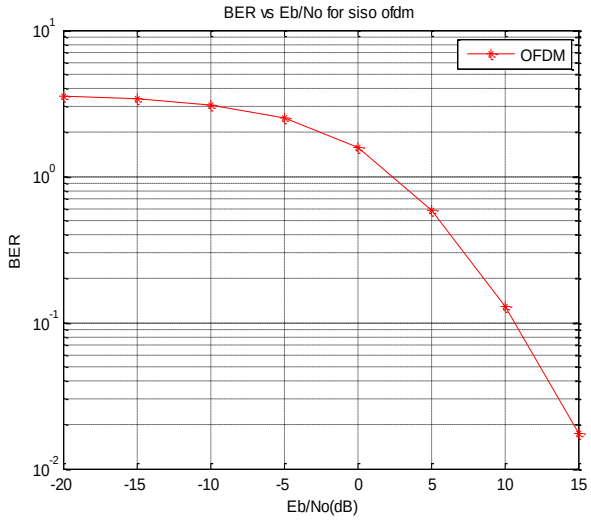


Figure 5. BER performance for SISO OFDM

SFBC for two transmit and receive antennas using LS and MMSE linear detection is simulated, considering 51200 bits transmitted for each SNR values and taking channel effect and AWGN, the result obtained is shown in figure 5.

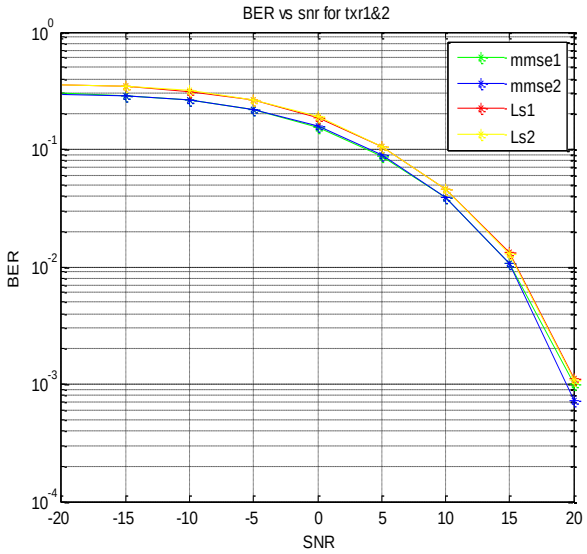


Figure 6. BER performance for 2x2 MIMO using SFBC MMSE and LS comparison

Comparing SNR value and corresponding BER value of figure 4 and figure 5 conclude that MIMO shows better performance. In figure 5, BER vs. SNR for LS and MMSE is compared; better performance is shown by MMSE. Further the effect of channel imperfections on BER is evaluated and result is that as the channel imperfection increases BER also increases. Channel is assumed to be constant during entire transmission but noise variance is kept increasing hence the effect of channel imperfection achieved and simulated the result, shown in figure 6.

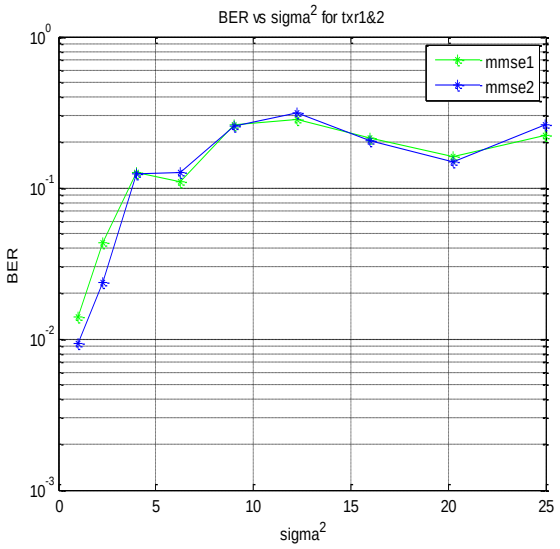


Figure 7. Channel imperfections in MIMO SFBC MMSE linear detection

7. FUTURE WORKS

For effective utilization of bandwidth, several users have to share same bandwidth. This leads to the evolution of multi user MIMO (MU-MIMO). When different users share same spectrum co-channel interference occurs and the effects needs to be mitigated. Several approaches like interference avoidance and treating interference as noise have several limitations, so to overcome that and to manage interference, interference alignment techniques will be introduced, adaptive beamformer design for interference alignment and cancellation its performance estimation and comparison of various beamformers will be evaluated under perfect and imperfect channel state information (CSI) in MATLAB tool [6].

8. CONCLUSION

As wireless communication is concerned, effects of channel and noise on transmitted signal are to be considered during the detection procedure, knowledge of channel variations or methods to overcome its effects is essential for a designer to predict and achieve higher efficiency. In this paper channel effects are considered and MMSE, LS linear detection methods are adopted to recover original data. Simulation results are plotted and results obtained, proves MMSE works best.

REFERENCES

[1] Balanis P, Ioannides P Introduction to smart antennas [MC 2007]
[2] Proakis,J., Digital Communications, McGraw-Hill
[3] M.Jiang and L. Hanzo,"Multiuser MIMO-OFDM for next-generation wireless systems," Proc. IEEE, vol.95, no. 7,pp.1430-1469,Jul. 2007
[4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," IEEE J. Select. Areas Commun., vol. 16,pp.1451-1458, Oct.1998.
[5] C.Ciochina, D.Castelain, D. Mottier and H. Sari;"Single Carrier Space-Frequency Block Coding Performance Evaluation",Vehicular Technology Conference,2007,VTC-2007 Fall,2007 IEEE 66th,pp.715-719,September 30,2007-October 3,2007
[6] S. Morteza razavi and Tharmalingam Ratnarajah, "Adaptive LS and MMSE based Beamformer Design for Multiuser MIMO Interference Channels," IEEE Trans. Wireless Commun.

Authors

Gopika k has graduated from Ilahia College of Engineering and Technology of Mahatma Gandhi University in Electronics & Communication Engineering in 2014. She is currently pursuing her M.Tech Degree in Wireless Technology from Toc H Institute of Science & Technology, Arakunnam. Her research interest includes Signal Processing and Wireless communication.

M. Mathurakani has graduated from AlagappaChettiar College of Engineering and Technology of Madurai University and completed his masters from PSG college of Technology of Madras University. He has worked as a Scientist in Defence Research and development organization (DRDO) in the area of signal processing and embedded system design and implementation. He was honoured with the DRDO Scientist of the year award in 2003.Currently he is a professor in Toc H Institute of Science and Technology, Arakunnam. His area of research interest includes signal processing algorithms, embedded system modeling and synthesis, reusable software architectures and MIMO and OFDM based communication systems.

INFORMATION SATURATION IN MULTISPECTRAL PIXEL LEVEL IMAGE FUSION

Preema Mole [1] and M Mathurakani[2]

[1]Student in M.Tech, ECE Department, Toc H Institute of Science and Technology, Ernakulam, India
[2]Professor, ECE Department, Toc H Institute of Science and Technology
Formerly Scientist, DRDO, NPOL, Kochi, Ernakulam, India

ABSTRACT

The availability of imaging sensors operating in multiple spectral bands has led to the requirement of image fusion algorithms that would combine the image from these sensors in an efficient way to give an image that is more informative as well as perceptible to human eye. Multispectral image fusion is the process of combining images from different spectral bands that are optically acquired. In this paper, we used a pixel-level image fusion based on principal component analysis that combines satellite images of the same scene from seven different spectral bands. The purpose of using principal component analysis technique is that it is best method for Grayscale image fusion and gives better results. The main aim of PCA technique is to reduce a large set of variables into a small set which still contains most of the information that was present in the large set. The paper compares different parameters namely, entropy, standard deviation, correlation coefficient etc. for different number of images fused from two to seven. Finally, the paper shows that the information content in an image gets saturated after fusing four images.

KEYWORDS

Multispectral image fusion, pixel-level image fusion, principal component analysis, Grayscale image.

1. INTRODUCTION

The deployment of multiple number of sensors operating in different spectral bands has led to the availability of multiple data. To educe information from these data there is a need to combine all data from different sensors. This can be accomplished by image fusion algorithms. Image fusion is generally defined as the process of combining images of a scene from different spectral bands into a single composite image which is more informative. This fused image will be suitable for human vision and computer processing. Image fusion aims at improving the geometric precision, spatial resolution, classification accuracy and enhance capability of spatial display. The main objective of image fusion is to Extract all the useful information from different input source images,Reduce redundancy and improve quality,Eliminate artifacts or any inconsistencies that distract human observers.

The fusion of multispectral images used in remote sensing and other applications yields better recognition results since the narrowband images highlights salient features which is sometimes neglected in captured images. Based on the level of processing where fusion is performed, the image fusion techniques are classified into three main levels: pixel level, feature level, and

DOI: 10.5121/ijci.2016.5221

decision level. Fusion at pixel level means processing at lowest level based on information extracted from set of pixels present in the input source images. These information are the originally measured quantities that are directly involved in the fusion process. The advantage of using pixel level image fusion is that it is easy to implement, simple and efficient with respect to time. In feature level, the main idea is to extract the feature sets from each input source images, and then perform an appropriate fusion rule to generate the fused image. The feature includes pixel intensities, edges, textures etc. Finally, in decision level fusion involves processing at highest level. Each classifier applies a threshold on the match score of input source images and transmits the ensuing decision. Both feature and decision have got a disadvantage that they generates inaccurate and incomplete transfer of the fused information. But, on the other hand pixel level improves the content of the final fused image. In recent years, many pixel level image fusion methods have been proposed. Some of the well known image fusion methods includes high pass filtering technique, Laplacian pyramid method, weighted average method , IHS transform-based image fusion, Discrete wavelet transform(DWT), Gradient pyramid, PCA based fusion, Stationary wavelet transform(SWT), Dual tree complex wavelet transform (DTCWT), etc. The basic strategy is to perform certain multiscale decomposition on each input source images and then to combine all this decompositions to achieve one merged representation based on the fusion rule. Then to this combined representation, the inverse transformation is applied to construct the final fused image.

1.1. ENTROPY

The entropy feature yields a measure of randomness in the intensity values of an image. This measure of randomness can be used to characterize the texture of input image. Also entropy defines the measure of an image's smoothness in terms of Gray level values. The higher the value of entropy, higher will be the number of Gray levels and lower the energy.

$$\text{Entropy is given as: } -\Sigma(p*\log p) \qquad (1)$$

Where p is the probability associated with the Gray level. For an image, this p is obtained by dividing number of pixels with Gray level by total number of pixels in an image.

1.2. STANDARD DEVIATION

It is a measure that provides an understanding of the spread intensities across the image. Standard deviation also indicates the contrast in an image. The contrast of each image is defined as an unbiased estimate of the standard deviation

It is given as:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(Xi-\mu)^2}{N-1}} \qquad (2)$$

Where N is the total number of pixels and μ is the mean. In image processing, standard deviation can also be taken as an estimate of underlying brightness probability istribution of an image.

1.2. CORRELATION COEFFICIENT

Correlation is an approach that comes from analyzing the displacement between two consecutive images. To find a characteristic feature from two images, the first image is compared with the second within a certain search range. Within this range the position of optimum similarity between two images is found.

Correlation coefficient is defined as the degree of correlation between two images. The value of this coefficient remains between -1 and +1.

The correlation coefficient between two random variables X and Y with expected values μ_x and μ_y and standard deviation SD, is defined as

$$r = \frac{Cov(X, Y)}{SD(X).SD(Y)} \qquad (3)$$

2. SURVEY OF RELATED RESEARCH

We examine some of the salient features of related research that has been reported. These works in image fusion can be traced back to mid eighties. Burt[1] was one of the first to report the use of Laplacian pyramid techniques. In this method, several copies of images was constructed at increasing scale, then each copy was convolved with original image. The advantages of this method was in terms of both computational cost and complexity. In 1985, P.J.Burt et.al and E.H.Adelson in[2] analysed that the essential problem in image merging is pattern conservation that must be preserved in composite image. In this paper, authors proposed an approach called Merging images through pattern decomposition. At about 1988, Alexander Toet et. Al proposed composite visible/thermal-infrared imaging apparatus[3]. R.D. Lillquist et. al in [4] presented a , Composite visible/thermal-infrared imaging apparatus Alexander Toet et al(1989),introduced ROLP(Ratio Of Low Pass) pyramid method that fits models of the human visual system. In this approach, judgments on the relative importance of pattern segments were based in their local luminance contrast values[5]. Alexander Toet et. al in [6], introduced a new approach to image fusion based on hierarchical image decomposition. This approach produced images that appeared to be more crispy than the images produced by other linear fusion scheme. H. Li et.al in [7] presented an image fusion scheme based on wavelet transforms. In this paper, the fusion took place in different resolution levels and more dominant features at each scale were preserved in the new multiresolution representation. I. Koren et.al in 1995,proposed a method of image fusion using steerable dyadic wavelet transform[8], which executed low level fusion on registered images by the use of steerable dyadic wavelet transform. Shutao Li et.al in [9], proposed pixel level image fusion algorithm for merging Landsat thematic mapper (TM) images and SPOT panchromatic images. V.S. Petrovoic et.al in [10] introduced a novel approach to multi resolution signal level image fusion for accurately transferring visual information from any number of input image signals, into a single fused image without the loss of information or the introduction of distortion. This new Gradient fusion reduced the amount of distortion, artifacts and the loss of contrast information. V.Tsagaris et.al in 2005 came up with the method based on partitioning the hyperspectral data into subgroups of bands[11]. V. Tsagaris and V. Anastassopoulos proposed

Multispectral image fusion for improved RGB representation based on perceptual attributes in [12]. Q. Du, N. Raksuntorn, S. Cai, and R. J. Moorhead, et al in 2008, investigated RGB colour composition schemes for hyperspectral imagery. In their paper, they proposed to display the useful information as distinctively as possible for high-class seperability. The work also demonstrated that the use of data processing step can significantly improve the quality of colour display, whereas data classification generally outperforms data transformation, although the implementation is more complicated [13].

3. IMAGE FUSION ALGORITHM

There are various methods that have been developed to perform image fusion. The method focussed in this paper is principal component analysis. In this section, the necessary background for vector representation of multidimensional remotely sensed data is provided. Also, the introduction to Principal Component Analysis (PCA) and the basic principles of pixel level fusion method are also provided.

3.1. PRINCIPAL COMPONENT ANALYSIS

Principal Component Analysis has been called as one of the most valuable results from applied Linear algebra. It is a mathematical procedure that transforms a number of correlated variables into a number of uncorrelated variables called principal components. These components captures as much of variance in data as possible. PCA provides an efficient way to reduce the dimensionality (for e.g. 10 dimensional data to 2 dimensional data). The principal component is taken to be along the direction with the maximum variance. The second principal component will be orthogonal to the first. The third principal component is taken in the maximum variance direction in the subspace perpendicular to the first two and so on. The PCA is also known as Karhunen-Loeve transform or the Hotelling transform.

3.2. MULTIDIMENSIONAL SPACE VECTOR REPRESENTATION AND DIMENSIONALITY REDUCTION

The properties of multispectral data set with K different channels and MxN number of pixels per channel can be examined if each pixel is described by a vector whose components are individual spectral responses to each multispectral channel.

$$X = [X_1, X_2, X_3, \ldots\ldots\ldots X_k]^T \quad (4)$$

The mean for this vector is given by,

$$\overline{X} = E[X] = 1/M.N \sum_{i=1}^{M.N} X_i \quad (5)$$

This mean vector defines the average position of the pixel in vector space, while the covariance matrix describes their scatter.

$$C_x = 1/ M.N \sum_{i=1}^{M.N} X_i X_i^T - \overline{X}\overline{X}^T \quad (6)$$

The covariance matrix is used to measure the correlation between multispectral band images. In case if there is high degree f correlation between multispectral band images, the corresponding off-diagonal elements in the covariance matrix will be large. The correlation between different multispectral images can also be described by means of correlation coefficient. The correlation coefficient r is related to covariance matrix as covariance matrix divided by the standard deviation of corresponding multispectral components. ie $r=C_{ij}/ SD(i).SD(j)$. According to the property of covariance matrix, C_x will be symmetric and all diagonal elements will be 1.

Among several linear transformations, Karhunen-Loeve also known as Principal Component Analysis is an important one. For this transform the covariance matrix is real and symmetric thereby making it possible to find a set of orthonormal eigen values and corresponding eigen vectors. Let e_i and λ_i for i =1, 2,……K (where K is the number of multispectral band images) be the eigen vectors and corresponding eigen values of C_x arranged in descending order. Next an another matrix A is formed in such a way that the rows of A are formed by the eigen vectors of C_x corresponding to largest eigen value and last row of A with the eigen vectors corresponding to the smallest eigen value. Thus this matrix is known as the transformation matrix that maps vector X into Y

$$Y= A^T(X - \bar{X}) \qquad (7)$$

From this transformation, the mean of Y will result in zero and covariance matrix for Y is given as,

$$C_y= AC_xA^T \qquad (8)$$

The resulting covariance matrix C_y will be the diagonal matrix and the elements along the main diagonal will be the eigen values of C_x . The value zero in the off-diagonal elements of the matrix denotes that the vector population Y are uncorrelated. This transformation will establish a new coordinate system whose origin is at the centroid of the population and the axes are in the direction of eigen vectors of C_x . This is explained in the figure 1, where we can see that there is an obvious correlation between X_1 and X_2 .

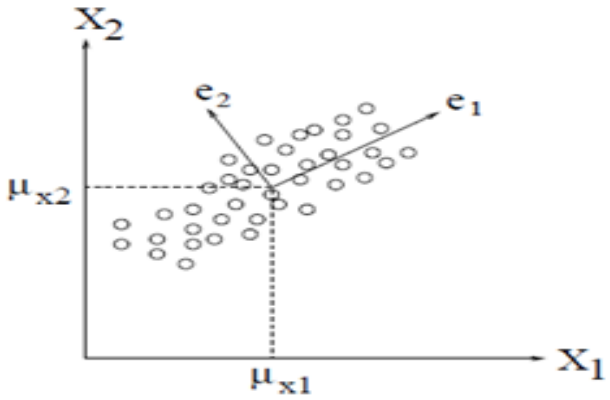


Figure 1. Data distribution before transformation

Now to decorrelate them, X_1 and X_2 are transformed to new variables Y_1 and Y_2 .

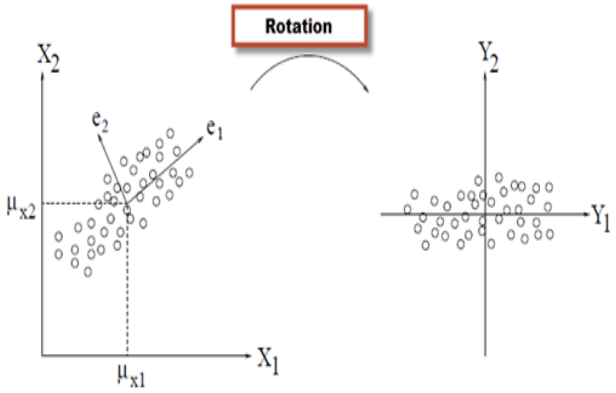


Figure 2. Data distribution after transformation

Here in figure 2, it can be seen that after transformation most of the variance in the data is along the variable Y_1 and hence the variables are decorrelated. Same mechanism is used for decorrelating the data in PCA technique also. The PCA transformation is one of the best method for Grayscale only.

3.3. PCA ALGORITHM

Let the images to be fused be arranged in two-column vectors. The steps followed to project this data on to 2-D subspaces are:

1. Organise the input images into column vectors. i.e. X
2. Compute the empirical mean vector M_e along each column.
3. Subtract the empirical mean vector M_e from each column of the data matrix.
4. Find the covariance matrix C_x as C_x = (1/no. of pixels)XX^T
5. Compute the eigenvectors V and eigenvalue D of C_x and sort them in decreasing order of eigenvalue.
6.Consider the first column of V which corresponds to larger eigenvalue to compute the components P_1 and P_2 as:

$$P_1 = \frac{V(1)}{\sum V} \quad \text{and} \quad P_2 = \frac{V(2)}{\sum V}$$

3.4. IMAGE FUSION USING PCA

The fusion of source images using PCA is shown in Fig.2. Any number of images can be fused using PCA. Here, the images to be fused are $I_1(x,y)$and $I_2(x,y)$. The PCA is applied to these images resulting in the components P_1 and P_2. Then final fused image is given as:

$$I_f(x,y) = P_1I_1(x,y) + P_2I_2(x,y) \qquad (9)$$

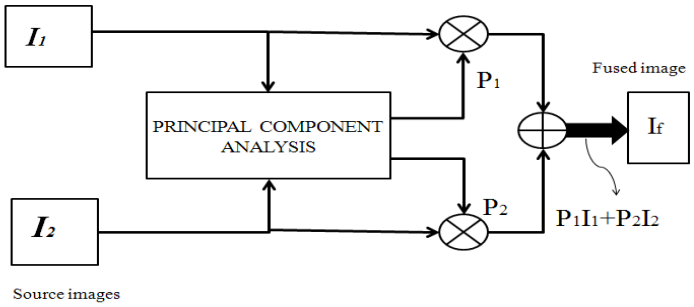


Figure 3. Image fusion using PCA

4. APPLICATIONS

Remote sensing technique have proven to be powerful tool for monitoring the earth's surface and atmosphere. The application of image fusion can be divided into Military and Non Military applications. Military applications include Detection, location tracking, identification of military entries, ocean surveillance, etc. Image fusion has also been extensively used in Non military applications that include interpretation and classification of aerial and satellite images.

5. EXPERIMENTAL PROCEDURES AND RESULTS

The multispectral data set used in this work consists of 7multispectral bands images acquired from Landsat Thematic Mapper (TM)sensor. The size of each image is 850 x 1100 pixels. The average orbital height for these images is 700km and spatial resolution is 30meters except the band 6 which is 90meters. The spectral range of sensor is depicted in table 1.

Table 1. Spectral range of data

Band number	Spectral range(μm)
1 Blue	0.45-0.52
2 Green	0.51-0.60
3 Red	0.63-0.70
4 Near infrared	0.76-0.85
5 Mid infrared	1.55-1.75
6 Thermal infrared	10.4-12.5
7 Mid infrared 2	2.08-2.35

The experiments conducted in this work aim to demonstrate the consistency of information and other parameters with increasing number of images. Fusion was performed for different number of images. Parameters namely, entropy, standard deviation, energy and correlation coefficient for different number of images was examined.

The fusion results are demonstrated in figure 4. The image in figure 4(d) is derived from the PCA analysis.

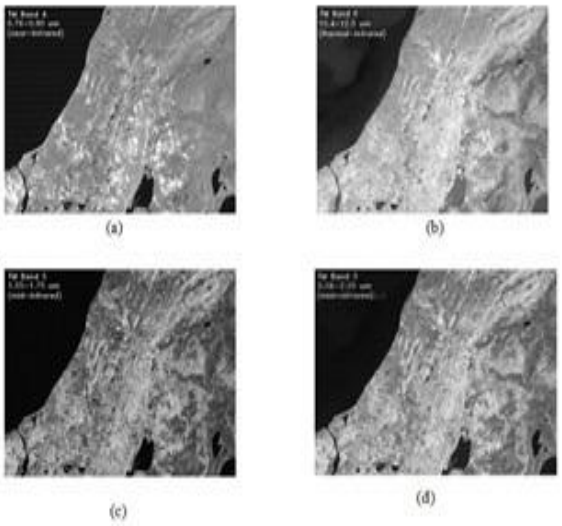


Figure 4.Detailed image (a)band 4 Near infrared (b)band 5 Mid infrared(c)band 6 Thermal infrared (d)fused image using PCA

Fusion result shown in figure 5 shows the entropy of the fused image with different number of images fused. It can be seen that fusing more images increases the entropy but after fusing four images the entropy remains constant. This shows that maximum entropy in fused image is attained when 4 images are fused.

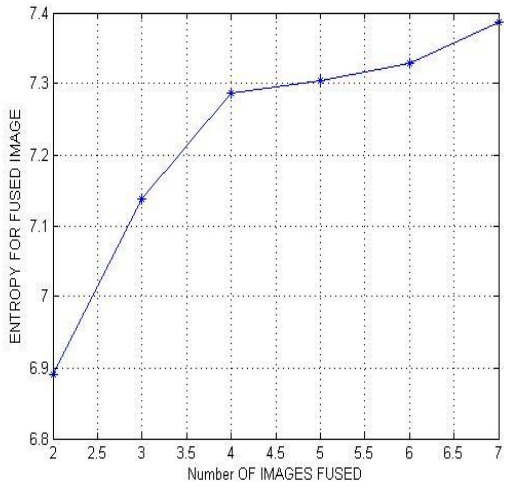


Figure 5.Entropy for different number of images fused

Figure 6 shows the result of standard deviation on fusing different number of images. In this figure also it is clear that the maximum standard deviation is obtained when four images are

fused. Thereafter, standard deviation which represents the contrast of an image decreases and remains constant.

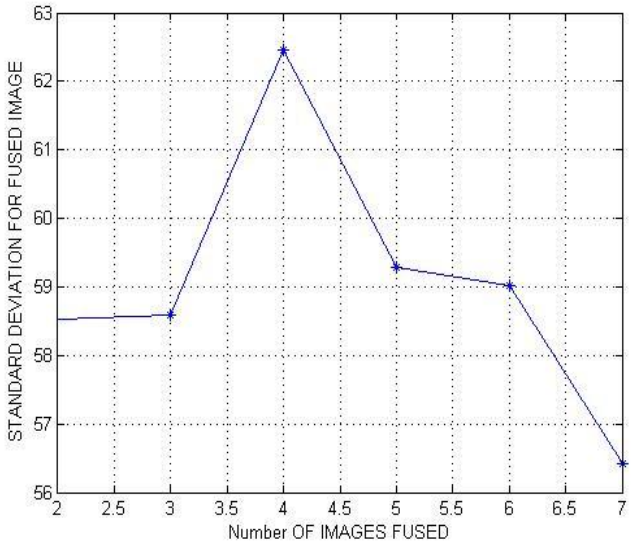


Figure 6.Standard deviation for different number of images fused

The energy of the fused image for different number of images fused is shown in figure 7

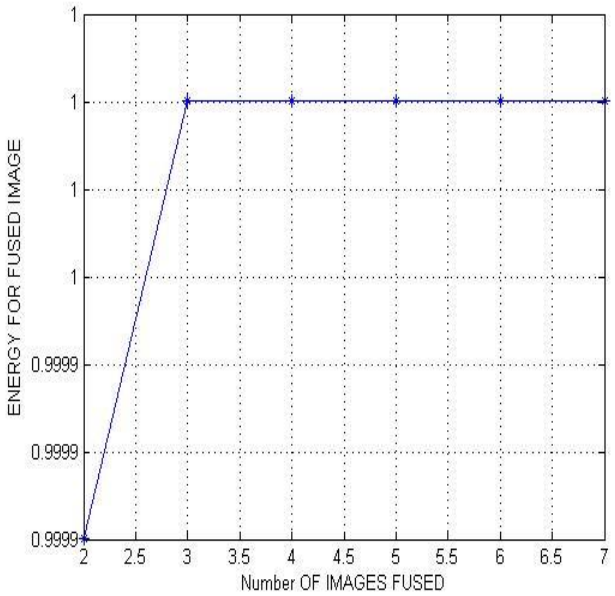


Figure 7.Energy for different number of images fused

The values given in table 2 represent the correlation coefficient between the source images. These values when compared to the correlation coefficients between the final fused image and input

source images for different number of images fused given in table 3 proves that the fused images have got more information and correlation with the input images.

Table 2. Correlation coefficient between source images

Source images	Correlation coefficient
Image 1 and 2	0.8501
Image 2 and 3	0.8633
Image 3 and 4	0.8609
Image 4 and 5	0.8649
Image 5 and 6	0.9355
Image 6 and 7	0.8315

Thus, comparison shows that the fused image contains all the information present in the individual input source images.

Table 3. Correlation Coefficients between Fused and source images for different number of images fused

IMAGES	2 IMAGE FUSION	3 IMAGE FUSION	4 IMAGE FUSION	5 IMAGE FUSION	6 IMAGE FUSION	7 IMAGE FUSION
IMAGE 1	0.9633	0.9473	0.9164	0.9129	0.9005	0.8989
IMAGE 2	0.9602	0.9493	0.9627	0.9637	0.9631	0.9600
IMAGE 3		0.9549	0.9409	0.9339	0.9244	0.9174
IMAGE 4			0.9624	0.9646	0.9666	0.9634
IMAGE 5				0.8919	09096	0.9259
IMAGE 6					0.9415	09440
IMAGE 7						0.8647

6. CONCLUSIONS

This paper analyses maximum number of images from different spectral bands that can be fused to get more information using Principal Component Analysis method. In this paper it is concluded that as the number of images to be fused increases correspondingly information, standard deviation and energy also increases. But after fusing four images, these parameters remains constant since maximum value has been attained on fusing just four images. The values of correlation coefficient for the fused and the source images shows that the fused image have more interdependence to the source images than the interdependence between the source images itself. This proves that the fused image contains all the information contained in the input spectral source images.

REFERENCES

[1] P.J. Burt, The pyramid as a structure for efficient computation, in: A.Rosenfeld (Ed.), Multiresolution Image Processing and Analysis,Springer-Verlag, Berlin, 1984, pp. 635.

[2] P.J. Burt, E.H. Adelson, Merging images through pattern decomposition,Proc. SPIE 575 (1985) 173181.

[3] E.H. Adelson, Depth-of-Focus Imaging Process Method, United States Patent 4,661,986 (1987).

[4] R.D. Lillquist, Composite visible/thermal-infrared imaging apparatus,United States Patent 4,751,571 (1988).

[5] A.Toet, Image fusion by a ratio of low-pass pyramid, Pattern Recognition Letters 9 (4) (1989) 245253.

[6] A.Toet, Hierarchical image fusion, Machine Vision and Applications 3 (1990) 111.

[7] H. Li, S. Manjunath, S. Mitra, Multisensor image fusion using the wavelet transform, Graphical Models and Image Processing 57 (3) (1995) 235245.

[8] I. Koren, A. Laine, F. Taylor, Image fusion using steerable dyadic wavelet transform, in: Proceedings of the International Conference on Image Processing, Washington, DC, USA, 1995, pp. 232235.

[9] S. Li, J.T. Kwok, Y. Wang, Using the discrete wavelet frame transform to merge Landsat TM and SPOT panchromatic images, Information Fusion 3 (2002) 1723.

[10] V.S. Petrovic, C.S. Xydeas, Gradient-based multiresolution image fusion, IEEE Transactions on Image Processing 13 (2) (2004) 228237

[11] V. Tsagaris, V. Anastassopoulos, and G. Lampropoulos, Fusion of hyperspectral data using segmented PCT for enhanced color representation, IEEE Trans. Geosci. Remote Sens., vol. 43, no. 10, pp. 23652375,Oct.2005.

[12] V. Tsagaris and V. Anastassopoulos, Multispectral image fusion for improved RGB representation based on perceptual attributes, Int. J.Remote Sens., vol. 26, no. 15, pp. 32413254, Aug. 2005.

[13] Q. Du, N. Raksuntorn, S. Cai, and R. J. Moorhead, Color display for hyperspectral imagery, IEEE Trans. Geosci. Remote Sens., vol. 46, no. 6,pp. 18581866, Jun. 2008.

AUTHORS

Preema Mole has graduated from Sree Narayana Gurukulam College of Engineering of Mahatma Gandhi University in Electronics & Communication Engineering in 2013. She is currently pursuing her M.Tech Degree in Wireless Technology from Toc H Institute of Science & Technology, Arakunnam. Her research interest includes Signal Processing and Image Processing.

M. Mathurakani has graduated from Alagappa Chettiar College of Engineering and Technology of Madurai University and completed his masters from PSG college of Technology of Madras University. He has worked as a Scientist in Defence Research and development organization (DRDO) in the area of signal processing and embedded system design and implementation. He was honoured with the DRDO Scientist of the year award in 2003.Currently he is a professor in Toc H Institute of Science and Technology, Arakunnam. His area of research interest includes signal processing algorithms, embedded system modeling and synthesis, reusable software architectures and MIMO and OFDM based communication systems.

INTENTIONAL BLANK

LARGE UNIVERSE CP-ABE WITH WHITEBOX TRACEABILITY

Anusha Sivanandhan[1] and Angel M Eldhose[2]

[1] Student,Department of Computer Science and Engineering, MBITS Nellimattom
[2]Assistant Professor, Department of Computer Science and Engineering, MBITS Nellimattom

ABSTRACT

In a ciphertext-policy attribute-based encryption (CP-ABE) system, decryption keys are defined over attributes shared by multiple users. Traceability is the ability of ABE to trace the malicious users or traitors who intentionally leak the partial or modified decryption keys for profits. Nevertheless, due to the nature of CP-ABE, it is difficult to identify the original key owner from an exposed key since the decryption privilege is shared by multiple users who have the same attributes. In this paper, we propose a new CP-ABE system that supportstraceability of malicious users who leaked their decryption privileges.Thistraceable CP-ABE does not weaken the expressivenessor efficiency when compared with the most efficient conventional(non-traceable) CP-ABE systems. In our newsystems attributes need not be fixed at system setup,the attributes' size is not polynomially bounded and thepublic parameters' size does not grow linearly with thenumber of attributes.

KEYWORDS

Attribute-based encryption,ciphertext-policy,large universe ,white-box traceability,malicious user.

1. INTRODUCTION

The notion of Attribute-Based Encryption (ABE) was introduced as a generalization of fuzzy Identity-Based Encryption (IBE)[1],[2]. In a CP-ABE system, each user is issued a decryption key by an authority according to the attributes he possesses, and the encryptor decides what attributes the eligible receivers should have by encrypting the messages with an access policy defined over some attributes. If and only if a user's attributes satisfy the access policy of a ciphertext, he can decrypt the ciphertext. Not only does ABE (especially CP-ABE) provide a new promising tool for implementing fine-grained access control over encrypted data, but also has it attracted much attention in the research community.

In general, an ABE system can be classified to "small universe" and "large universe" constructions. In the "small universe" construction, the attributes are fixed at system setup and the size of the attributes is polynomially bounded, and furthermore the size of public parameters grows linearly with the number of attributes. While in the "large universe" construction, the attributes need not be specified at system setup and the size of the attribute universe is unbounded. The "large universe" construction for ABE system brings an obvious advantage that the designer of the ABE system need not bother to choose a particular bound of the attributes at system setup.

In CP-ABE, each user possesses a set of attributes and can decrypt the ciphertext if his/her attributes satisfy the ciphertext's access policy. This results in an obvious consequence that the

DOI: 10.5121/ijci.2016.5222

encryptor or system does not know who leaks the decryption key to others intentionally. Due to the fact that the attributes are shared by multiple users and different users may have the same subset of attributes, the encryptor or system has no feasible method to trace the suspicious receiver if the decryption key is leaked. We take Alice (with attributes {Alice, Assistant Professor, Computer Science}) and Bob (with attributes {Bob, Assistant Professor, Computer Science}) as an example. They both have the same decryption keys corresponding to attributes {Assistant Professor, Computer Science} and can decrypt such a ciphertext encrypted by the attributes {Assistant Professor, Computer Science}. Suppose no other receiver in the system has both attributes ({Assistant Professor} and {Computer Science}) at the same time. If it happens to exist a user who can decrypt the ciphertext except Alice and Bob, it is significant to find out who leaks such decryption key to him, Alice or Bob?

This drawback should be fixed in practice in case of leaking decryption key. It is necessary to add the property of traceability to the original ABE scheme, to identify who exactly leaks the decryption key. The above traceability is called white-box traceability, which means that any user who leaks his/her decryption key to the third user or device intentionally or unintentionally will be identified. However, up to now, there exists no practical traceable CP-ABE system supporting the property of large universe as the (non-traceable) CP-ABE system. Large universe CP-ABE system with white-box traceability is not yet achieved in practice: (1) The CP-ABE systems supporting traceability so far proposed do not support the property of large universe, the attributes need to be fixed at system setup and the size of the attributes is polynomially bounded. Besides, public parameters' size grows linearly with the number of attributes. (2) The large universe CP-ABE system proposed is secure in the standard model; however, it does not support the property of traceability.

2. RELATED WORKS

Sahai and Waters introduced the notion of Fuzzy Identity- Based Encryption[36]. Goyal *et al.* later formalized two notions of ABE[14]: CP-ABE (where user keys are labeled with sets of attributes and ciphertexts are associated with policies) and KP-ABE (where ciphertexts are labeled with sets of attributes and private keys are associated with access structures). Subsequently, many constructions of selectively secure KP ABE and CP-ABE systems were proposed[6],[7],[8],[5],[4],[3]. Many advances have been made for ABE as the following directions: new proof techniques to obtain fully secure , decentralizing trust by setting multiple authorities and outsourcing computation.

The first large universe KP-ABE construction was proposed in unbounded hierarchical based encryption[9]n. It was built on composite order groups and proved selectively secure in the standard model. Then the first large universe KP-ABE construction on prime order groups proposed[10] was inspired by the dual pairing vector space framework. Recently, the first large universe CP-ABE construction built on prime order bilinear groups was proposed by Rouselakis and Waters. It was proved selectively secure in the standard model under "*q*-type" assumption. Another branch of ABE research considers the problem of traceability. The notion of accountable CP-ABE was first proposed to prevent illegal key sharing among colluding users. Then a multi-authority ciphertext-policy (AND gates with wildcard) ABE scheme with accountability was proposed in, which allowed tracing the identity of a misbehaving user who leaked the decryption key to others. Liu, Cao and Wong lately proposed a white-box and black-box traceability CP-ABE system which supported policies expressed in any monotone access structures.

Black-Box Traceable ABE Systems. In our construction, we target to make decryption key leakage to be traceable in the white-box model, i.e., the decryption keys leaked/sold will be used by the buyers to perform decryption using the ABE decryption algorithm. In practice, a stronger

traceability notion is called black-box traceability, which is analogous to the notion of black-box traitor tracing in broadcast encryption [11], [12]. In particular, given a decryption equipment (where the embedded decryption key or algorithm could be unknown or hidden), the buyers can use it to retrieve plaintexts from ciphertexts. A black-box traceable ABE should allow an authority to find out the identity of the malicious user (i.e., whose decryption keys are used to create this decryption equipment).

3. PROBLEM DEFINITION

In ABE, the decryption privilege of a decryption key is shared by multiple users who possess the corresponding attributes, so that any malicious owner of a decryption key would have the intention or be very willing to leak partial or even his entire decryption privilege for financial interest or any other incentive, especially when there is no risk of getting caught. We refer to this issue as MaliciousKeyDelegation.Nevertheless, due to the nature of CP-ABE, it isdifficult to identify the original key owner from an exposed keysince the decryption privilege is shared by multiple users whohave the same attributes. In general, an ABE system is small universe, In the "small universe" construction, the attributes are fixed at system setup and the size of the attributes is polynomially bounded, and furthermore the size of public parameters grows linearly with the number of attributes.

Consider a commercial application such as a pay-TV system with huge number of users for example. Each user is labeled with lots of related attributes, which are defined as TV channels that the user have ordered. As a versatile one-to-many encryption mechanism, CP-ABE system is quite suitable in this scenario. The pay-TV system provides several TV channels for users, and those who have paid for the TV channels could satisfy the access policy to decrypt the ciphertext and enjoy the ordered TV channels. CPABE enables fine-grained access control to the encrypted data according to attributes in users' ordered lists. However, there are two problems with this approach. First, if someone (who does not have the privilege to access to those TV chat a lower cost, she/he could also get access to the TV channels. It is necessary to find out who is selling the decryption key. Second, as the TV channels of the pay-TV system expand, an increasing number of new attributes need to be added to the system to describe the new channels. If the number of the attributes exceeds the bound set during the initial deployment of the pay-TV system, then the entire system has to be re-deployed and possibly all its data will have to be re-encrypted, which would be a disaster to the pay-TV in the commercial applications. The problems, as described above, are the main obstacles when CP-ABE is implemented in commercial applications such as pay-TV systems and social networks. Due to the nature of CP-ABE, if a malicious user leaks its decryption key to others for profits (such as selling the decryption key on the Internet), it is difficult to find out the original key owner from an exposed key since the decryption key is shared by multiple users who have the same attributes.

Inorder to solve this issue,large universe enhanced traceable system was introduced. The main features of this system are:

1) *White-box traceability*- Our new systems can trace the malicious users or traitors who may leak the partial or modified decryption keys to others for profits.
2) *Large universe*. In our new systems attributes need not be fixed at system setup, the attributes' size is not polynomially bounded and the public parameters' size does not grow linearly with the number of attributes.
3) *Constant storage overhead*- we adopt the Shamir's *(.t, . n)* threshold scheme in tracing the malicious users or traitors, the storage cost for traceability does not grow linearly with the number of the users, it is constant which only depends on the Threshold *.t* .

4) *Dynamical scalability*- It yields another result that the stored data for traceability need not be updated when new users are added into the system or malicious users are removed from the system, which makes the system more practical for applications.

4. ENHANCED TRACEABLE LARGE UNIVERSE CP-ABE SYSTEM

To realize large universe construction, we adopt the "individual randomness" and "layer" technique from [9] and [13]. We use the "layer" technique to encrypt data securely and to be able to decrypt. We employ two "layers": the "attribute" layer and the "secret sharing" layer, and use a "binder term" to connect these two layers securely. In the "attribute" layer, we utilize *u, h* terms to provide a Boneh-Boyen-style [14] hash function *(u Ah)*. As for the "secret sharing" layer, during **KeyGen**and **Encrypt** phases we use *w* term to hold the secret randomness *r* and the secret randomnesss shares respectively. Finally, we use the *v* term to "bind" this two layers together.

To realize traceability, we use the Boneh-Boyen-style signature [14] in both the T-LU-CPABE system and the eT-LU-CPABE system. Furthermore, we find that the identity table *T* with the tuple identity and its randomness used in [26] and the T-LU-CPABE system grows linearly with the number of the users.2 With the number of the users in a system scaling large, the corresponding identity table *T* for traceability will expand as a result, which leads to heavy burden of the storage space for *T*. Besides, the corresponding cost of searching *K_* in *T* during the **Trace** phase is relatively huge. In our eT-LU-CPABE system, we utilize the Shamir's *(.t, .n)* threshold scheme to optimize the property of traceability. We only need store *.t* − 1 points and the value *f (*0*)* on a polynomial *f (x)* at system setup. Consequentially, our storage for traceability does not grow linearly with the number of the users and is a constant.

4.1. MODEL

An enhanced Traceable Large Universe CP-ABE system (eT-LU-CPABE system) is a CP-ABE system where attributes need not be fixed at system setup and can trace the user by his/her decryption key. Moreover, we enhance the T-LU-CPABE system by eliminating the identity table *T*. In this eT-LU-CPABE system, we utilize the Shamir's*(.t, n)* threshold scheme to optimize the property of traceability. In our eT-LU-CPABE system, we utilize the Shamir's *(.t, n)* threshold scheme to optimize the property of traceability.

In a Traceable CP-ABE system (T-CPABE system) it is not required that the attributes for the encryption process be fixed at the setup phase.A Traceable Large Universe CP-ABE system (T-LU-CPABE system) is a CP-ABE system where attributes need not be fixed at system setup and can trace the user by his/her decryption key. We enhance the original large universe CP-ABE system by adding users' identities and a **Trace** algorithm. The identities of the user are added whenever they register into the system.Compared with the T-LU-CPABE System, the significant and remarkable advantage of our new eT-LU-CPABE system is that the system does not need to maintain the identity table *T* and the storage overhead for traitor tracing is constant.

The main idea of our traceability in the eT-LU-CPABE system is as follows.

Firstly, the **Setup** algorithm initializes an instance of Shamir's *(.t ,. n)* threshold scheme *INS(.t, . n)* and keeps a polynomial *f (x)* and *.t* − 1 points {*(x*1, *y*1*),(x*2, *y*2*), . . . , (x.t*−1, *y.t*−1*)*} on *f (x)* secret. Then we insert *c* into the decryption key *sk*during the **KeyGen**phase where *c* = *Enck*.2 *(x||y), x* = *Enck*.1 *(id), y* = *f (x)*. Note that the tuple *(x, y)* is a point on *f (x)*. During the **Trace** phase, the algorithm extracts *(x*∗ = *x_, y*∗ = *y_)* from *x_||y_* = *Dec .k*2 *(K_)* in the decryption key *sk*,and then it checks whether *sk*is issued by system.If*(x*∗ = *x_ , y*∗ = *y_)* ∈ {*(x*1, *y*1*), (x*2, *y*2*), . . .*

, $(x.t-1, y.t-1)\}$, the algorithm computes *Deck*.1 $(x*)$ to get *id* to identify themalicious user directly. Otherwise, the algorithm computesthe secret of *INS* $(.t, .n)$ by interpolating with $.t - 1$ points$\{(x1, y1), (x2, y2), \ldots, (x.t-1, y.t-1)\}$ and $(x*, y*)$. If therecovered secret is equal to $f(0)$, the algorithm computes *Deck*.1 $(x*)$ to get *id* to identify the malicious user. If the equation fails, *sk*is not issued by the system. In this way, we could trace the owner of the decryption key. Meanwhile, it brings the benefit that the system only stores $f(0)$ and $.t - 1$ points on $f(x)$, and thus the storage for traceability is a constant.This system is more secure when compared with others.

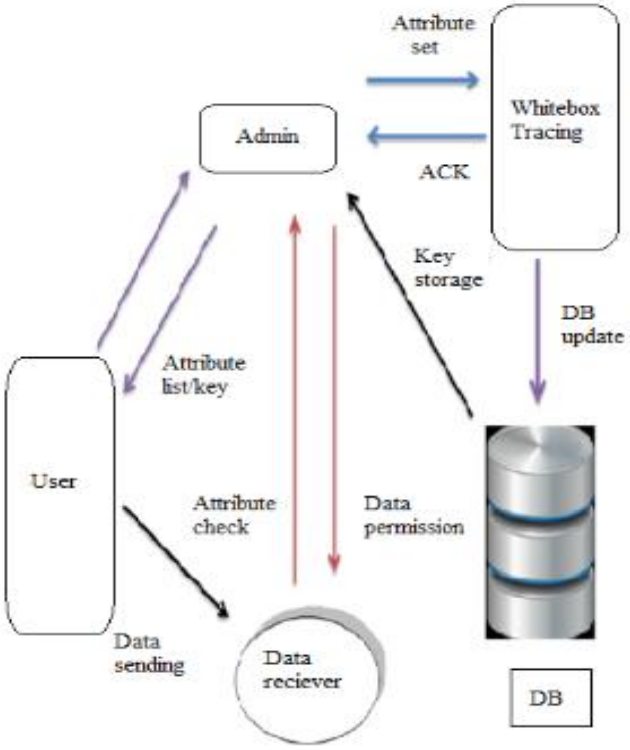


Figure 1. System Architecture

An eT-LU-CP-ABE system consists of six algorithms as follows:

• **Setup**: The algorithm takes as inputs a security parameter $\lambda \in N$ encoded in unary. It outputs the public parameters *pp* and the master secret key *msk*. We assume that the description of the attribute universe *U* is contained in the public parameters. In addition, it initializes an instance of Shamir's $(.t, .n)$ threshold scheme denoted by $INS(.t, . n)$.

• **KeyGen**: The key generation algorithm takes as inputs the public parameters *pp*, the master secret key *msk*and a set of attributes $S \subseteq U$ for a user with identity *id*. The security parameter in the inputs ensures that it is polynomial time in λ. The algorithm outputs a secret key *skid,S*corresponding to *S*.

• **Encrypt**: The encryption algorithm takes as inputs the public parameters *pp*, a plaintext message *m*, and an access structure A over *U*. It outputs the ciphertext*ct*.

• **Decrypt**: The decryption algorithm takes as inputs the public parameters *pp*, a secret key *skid,S*, and a ciphertext*ct*. It outputs the plaintext *m* or $\perp$.

• **KeySanityCheck**: The decryption algorithm takes as inputs the public parameters *pp* and a secret key *sk*. If *sk*passes the key sanity check, it outputs 1. Otherwise, it outputs 0. The key

sanity check is a deterministic algorithm, which is used to guarantee the secret key to be well-formed in the decryption process.

• **Trace**: This algorithm is used to identify the malicious user. From the above algorithm we can conclude whether the key was well-formed or not. The tracingalgorithm takes as inputs the public parameters *pp*,an instance of Shamir's *(.t, . n)* threshold scheme *INS (.t, . n)*, the master secret key *msk*, and a secretkey*sk*. The algorithm first verifies whether *sk*is well formedto determine whether *sk*needs to be traced.If*sk*is well-formed and could recover the secret of *INS (.t, .n)*, the algorithm outputs an identity *id* implyingthat*sk*is linked to *id*. Otherwise, it outputs a specialsymbol _ implying that *sk*does not need to be traced.We define a secret key *sk*is *well-formed* which meansthatKeySanityCheck*(pp, sk)*→ 1.

4.2. SHAMIRS THRESHOLD SCHEME

It is well known for Shamirs (t, n) threshold scheme (or Shamirs secret sharing scheme) in cryptography. The essential idea of that scheme is that t points on a t - 1 degree curve are sufficient to confirm such a curve, that is, t points are enough to determine a t - 1 degree polynomial. For a (t; n) threshold scheme, a secret can be divided into n parts (or even more), which are sent to each participant a unique part. All of them can be used to reconstruct the secret. Suppose that the secret is assumed to be an element in a finite field Fp . Choose t - 1 number of random coefficients a1, a2, at-2 element of Fp and at-1 element of Fp and set the secret in the constant term a0. Every participant is given a point (x, y) on the above curve, that is, the input to the polynomial x and its output y = f(x). Given a subset with any t points, recover the constant term a0 using the Lagrange interpolation. White box traceability is implemented in this paper as a web application.

4.3. PROBABILISTIC EQUATION

Probabilistic encryption is an encryption algorithm with some randomness during the encryption, which leads that encrypting same messages yields different ciphertexts in the various times. The first provably-secure probabilistic public-key encryption scheme was proposed by GoldwasseranMicali, based on the hardness of the quadratic residuosity assumption. Later, some efficient probabilistic encryption schemes appeared including ElGamal, Paillier and various constructions under the random oracle model. In our scheme, we only use the property of probabilism to output a ciphertext that cannot be distinguished from a random number from the view of the adversary. Without loss of generality, we define such a probabilistic encryption *(Enc.k ,Dec.k)* in our scheme where *.k* is the secret key for encryption and decryption. From the point of efficiency, symmetric encryption scheme is quite suitable since encryption and decryption are easy to perform.

The main idea in ABE is that the role of the users is taken by the attributes. Thus, the access structure A will contain the authorized sets of attributes. For CP-ABE, if a user of the system posses an authorized set of attributes then he can decrypt the ciphertext, otherwise, he can't get any information from ciphertext if the set he possed is unauthorized. In our construction, we restrict our attention to monotone access structure.

5. CONCLUSIONS

In this work, we constructed An Enhanced Traceable CP-ABE system which achieved the efficiency and security level as one of the best existing (non-traceable) CP-ABE systems.CPABE systems which include white box traceability of the authorized malicious users have been developed. We can trace the malicious users leaking the partial or modified decryption keys to others for profit. The attribute size is unbounded and the public parameters size does not grow linearly with the number of attributes. The cost of achieving traceability in our system is also very low. In addition, we optimize the system in tracing the malicious users to cut down the storage cost for traceability and to make the system efficient in the revocation of the users. Based on the above advantages, our new systems could be applied to many scenarios such as pay-TV systems and social networks. This system is selectively secure in the standard model, when compared with others.

REFERENCES

[1] A. Shamir, "Identity-based cryptosystems and signature schemes," in Proc. CRYPTO, 1984, pp. 47–53

[2] V. Goyal, "Reducing trust in the PKG in identity based cryptosystems," in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 2007, pp. 430–447.

[3] T. Okamoto and K. Takashima, "Fully secure functional encryption with general relation from,,. the decisional linear assumption," in Advances inCryptology. Berlin, Germany: Springer-Verlag, 2010, pp. 191–208

[4] A. Lewko, T. Okamoto, A. Sahai, K. Takashima, and B. Waters, "Fully secure functional encryption: Attribute-based encryption and (hierarchical) inner product encryption".

[5] V. Goyal, A. Jain, O. Pandey, and A. Sahai, "Bounded ciphertext policy attribute based ,.encryption," in Automata, Languages and Programming .Berlin, Germany: Springer-Verlag, 2008, pp. 579–591.

[6] N. Attrapadung, B. Libert, and E. de Panafieu,"Expressive key-policy attribute-based encryption with constant-size ciphertexts," in Public KeyCryptography. Berlin, Germany: Springer-Verlag, 2011, pp. 90–108.

[7] J. Bethencourt, A. Sahai, and B. Waters, "Ciphertext-policy attribute based ., encryption," in Proc.and EEE Symp. Secure. Privacy (SP), May 2007, pp. 321–334.

[8] L. Cheung and C. Newport, "Provably secure cipher text policy ABE,"in,Proc. 14th ACM Conf. Comput. Commun.Secur., 2007, pp. 456–465.

[9] A. Lewko and B. Waters, "Unbounded HIBE and attribute-based encryption ,,(HABE)"in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 2011, pp. 547–567.

[10] T. ElGamal, "A public key crypto system and a signature scheme based on discrete logarithms," and in Advances in Cryptology. Berlin, Germany: Springer-Verlag, 1985, pp. 10–18..

[11] B. Chor, A. Fiat, and M. Naor, "Tracing traitors," in Proc. CRYPTO, Y. Desmedt, Ed.,1994, vol.839 257–270, ser. Lecture Notes in Computer Science, Springer.

[12] D. Boneh, A. Sahai,and B. Waters,"Fully collusion resistant traitor tracing with short Ciphertexts and Private keys," in Proc. EUROCRYPT,,. S. Vaudenay, Ed., 2006, vol. 4004, pp. 573– 592, ser. Lecture Notes in Computer Science, Springer.

[13] Y.Rouselakis and B Waters "Practical constructions and new proof methods for large universe attribute based encryption," in Proc. ACMSIGSAC Conf. Comput. Commun.Secur., 2013, pp 463–474.

[14] D. Boneh and X. Boyen, "Short signatures without random oracles, in Advances in Cryptology (Lecture Notes in Computer Science),, vol.,, 3027, C. Cach in and J. L.Camenisch, Eds and also Berlin ,Germany: Springer-Verlag, 2004, pp. 56–73.

[15] Z. Liu, Z.. Cao, and D. S. Wong, "White-box traceable cipher text- policy attribute- based encryption Supporting any monotone access structures," IEEE Trans. Inf. Forensics Security, vol,... 8, no. 1, pp. 76–88, Jan. 2013.

[16] A. Sahai, H. Seyalioglu,,,.and B. Waters, "Dynamic credentials and cipher text delegation for attribute based encryption," in Advances inCryptology. Berlin, Germany: Springer-Verlag, 2012, pp. 199217.

[17] A. Sahai.., and B. Waters, "Fuzzy identity-based encryption," in Advances in Cryptology. Berlin.., and Germany: Springer-Verlag, 2005, pp. 457–473.
[18] T. Oka moto and K. Takashima, "Homomorphic encryption.., and,. signatures from vector., and their decomposition," in Pairing-Based Cryptography. Berlin, Germany: Springer-Verlag, 2008, pp. 57– 74.

AUTHORS

AnushaSivanandhan. is currently pursuing M.Tech in Cyber Security in Mar Baselios Institute of Technology and Science, Nellimattom, Kerala, India. She completed her B.Tech from Mar Baselios Institute of Technology and Science, Nellimattom. Her areas of research are Network Security and Information Forensics.

Angel M Eldhose is currenty working as the Assistant Professor in Department of Computer Science and Engineering at mar baselios institute of technology and science, Nellimattom, Kerala, India. He recieved his B-Tech Degree in Computer Science and Engineering from Illahia College of Engineering and Technology and M-Tech in Computer Science and Engineering from KarunyaUniversity.His research interest include Image Processing and Database Security.

Mobile Tracker

Shirin Salim[1], Dipina Damodaran B[2] and Surekha Marium Vargese [3]

Department of Computer Engineering, M A College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

In today's fast world, mobile has become one of the important commodities of a human being. It has become a necessity rather than a luxurious commodity. The Mobile Tracking helps to track the current location of the mobile. The security of the devices powered with this is most challenging problem existing now. Almost everyone might have experienced misplacing or losing their mobile phones. Hence, it is necessary to create an app in all smart phones to detecting thefted phone. This paper discuss a tracking application including SIM detection, call monitoring, image capture, contact capture, profile switching etc for thefted mobile phone. These features are quite different from the existing tracker applications which would be helpful in tracing the lost mobile without the help of any protecting agency. The application installed will be running in the background and won't be shown in the task manager as well. Once the mobile phone is lost, this application enables the user to track a mobile device and to receive notification via SMS to a predefined number. Some specified formatted messages can be used to control the thefted mobile phone.

KEYWORDS

Mobile Tracking, GSM, GPS.

1. INTRODUCTION

The introduction of smartphones in the mobile market brings a tremendous change. They have many features and applications that is very useful to us. Different type of OS are available in the market, among them Android OS become widespread now a days because of its excellent features.

The Mobile Tracker , An android application to locate and track mobile phones is a unique and efficient application which has a variety of features that enhances the existing mobile tracking system [2] . The app stands different from the existing system as it is not only the GPS value it makes use of but it works on GSM/text messaging services which make it a simple and unique one. The app is able to enable the GPS when a non authorised SIM card is detected in the device by comparing the Integrated Circuit Chip Card Identification (ICC ID). The ICC ID number is unique for each SIM card. The app is filled on with features like changing profile, call monitoring, Sim card detection, and location fetching through GPS [3][4]. All these features works on SMS basis. So, incoming SMS format plays an important role. The android application running in the smartphone monitors all the incoming messages. If the SMS is received in a predefined format it reads and performs the expected task. The application installed will be running in the background and won't be shown in task manager. Many one of us have been experienced the missing of your mobile phones, if it is an android phone this application helps you to track it without the help of cyber cell. Once it is lost, this application enables the user to track a mobile device notification via SMS to a predefined number saved in that application at the time of installation itself. This application uses Android OS which demonstrates a system that uses a regular mobile phone equipped with a GPS receptor and connected to a global system for

mobile (GSM) network that takes advantage of these technologies in behalf of the user safety [5]. The app is a useful mobile application that combines several features which aims at the user's security.

The app is directed to two user profiles, the client and server to be tracked. The server side requires any android based Smartphone starting from version Android 2.2 having the app installed in it with GPRS and GPS enabled. The client side requires any other OS based mobile phones for sending and receiving SMS. If there is any error in sending the message from the operator, there won't be any message sends to the operator by the application, instead no action takes place at the server side. This application uses Android OS which demonstrates a system that uses a regular mobile phone equipped with a GPS receptor and connected to a Global System for Mobile(GSM) network that takes advantage of these technologies in behalf of user safety. The app is a useful mobile application that combines several features which aims at the user's security.

The app is filled on with features like changing profile, call monitoring, SIM card detection, location fetching through GPS and transfer of images to email address. All these features work on the SMS basis. So, incoming SMS format plays an important role. The android application running in the smartphone running monitors all the incoming messages. If the SMS is received in a predefined format it reads the SMS it reads the SMS and performs the expected task [5][6].

The paper is organized in the following sections. Section II describes the application development including its Requirements, features and the technologies employed. Section III describes the application functionality along with experimental evaluation and results. Finally, section IV presents the conclusion and some possible future work.

2. PROBLEM DEFINITION

Almost everyone might have experienced misplacing or losing their mobile phones. Hence, it is necessary to create an app in all smart phones to detecting thefted phone. This paper discuss a tracking application including SIM detection, call monitoring, image capture, contact capture, profile switching etc for tracking mobile phone.

2.1 ARCHITECTURE

The working for the app contains two user profiles, one android phone starting from version 2.2 with GPS enabled where the app is installed and another is any other OS based mobile phone which is used to control the thefted android phone by sending and receiving SMS.

Figure 1 shows the Mobile Tracker requirements [4].

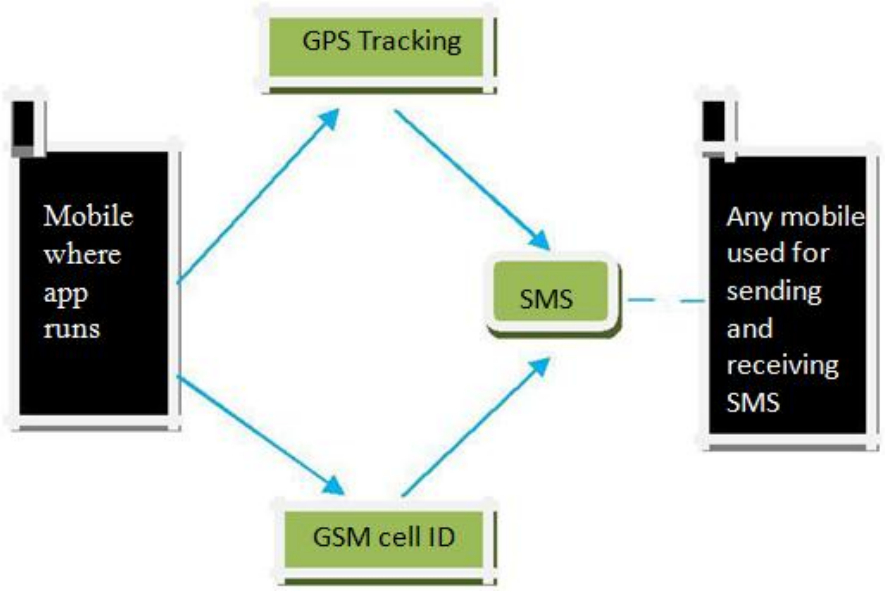


Figure 1: Mobile Tracker Architecture [1]

The SIM card detection feature of the app checks whether the currently used SIM card is authorized or not. If it detects that there is an unauthorized SIM card in the phone suddenly it will send a notification warning message to the predefined emergency number. The warning message contains the current GSM cell ID, phone number of unauthorized SIM card and also network provider.

The camera feature requires the front or back camera and internet connection on the phone.
The application fetches the location through GPS and GPS values along with address is send to the predefined number. To activate this feature the android phone must be GPS enabled. Also the phone must have the internet connection so that the details of the thefted phone can be send through email to a specified email id.

3. METHODOLOGY

SIM cards are identified by its Integrated Circuit Card ID(ICC-ID) and it will be stored in the SIM cards and are also get printed in the body of SIM card during the process personalization. Whenever the application starts, the ICC-ID of the current SIM card is compared with the predefined ICC-ID to detect unauthorized SIM card n the device. After the SIM replacement ,we will get a notificaion about the IMEI number with the details of the inserted new SIM.The call history details like incoming, outgoing calls which are made from the lost antroid mobile are provided by the call monitoring feature. And also it will send all the saved numbers and name, as SMS to the predefined number. Thus we get more contact informations about the cellphone thief .Suppose the GPS data shows that the mobile is very nearer to us but we can't identify it ,so that it is in silent mode or the flicked person make the phone in silent as he is still around.Then we can convert the profile of the phone from silent to general and also from general to silent as per our requirement by just sending an SMS using the predefined format with the help of our profile management feature in the application.

By just sending a predefined SMS ,we can activate the back camera of our mobile phone.whenever the message recieved ,the application activates the back camera and capture image and send it to our email ID auomatically that we had saved in the application earlier. And

we can check our email account ,thus we can extract more information about the thief like who, how, where etc. Using a particular predefined format SMS we can also find out or retrieve the current location of the lost Smart phone. Whenever the SMS get we can retieve the current location of the thefted phone at any time. Also we can retrieve all the inbox details, that is messages that the thief or new owner gets to the predefined number by sending another format of message. Thus we can get more information about his friends and all. If we want to block the thief to making use of our Smart phone that he thefted, we can block him by sending an SMS to the predefined number, so that he cant make call enough. Also we can remove the block whenever we want.

The main advantage of the application is that it automatically deletes the incoming and outgoing SMS from the Smart phone that is thefted as it get installed the app such that the new owner or the thief who using the cell is clueless about it all.

The following is the data flow diagram of the Mobile Tracker application:

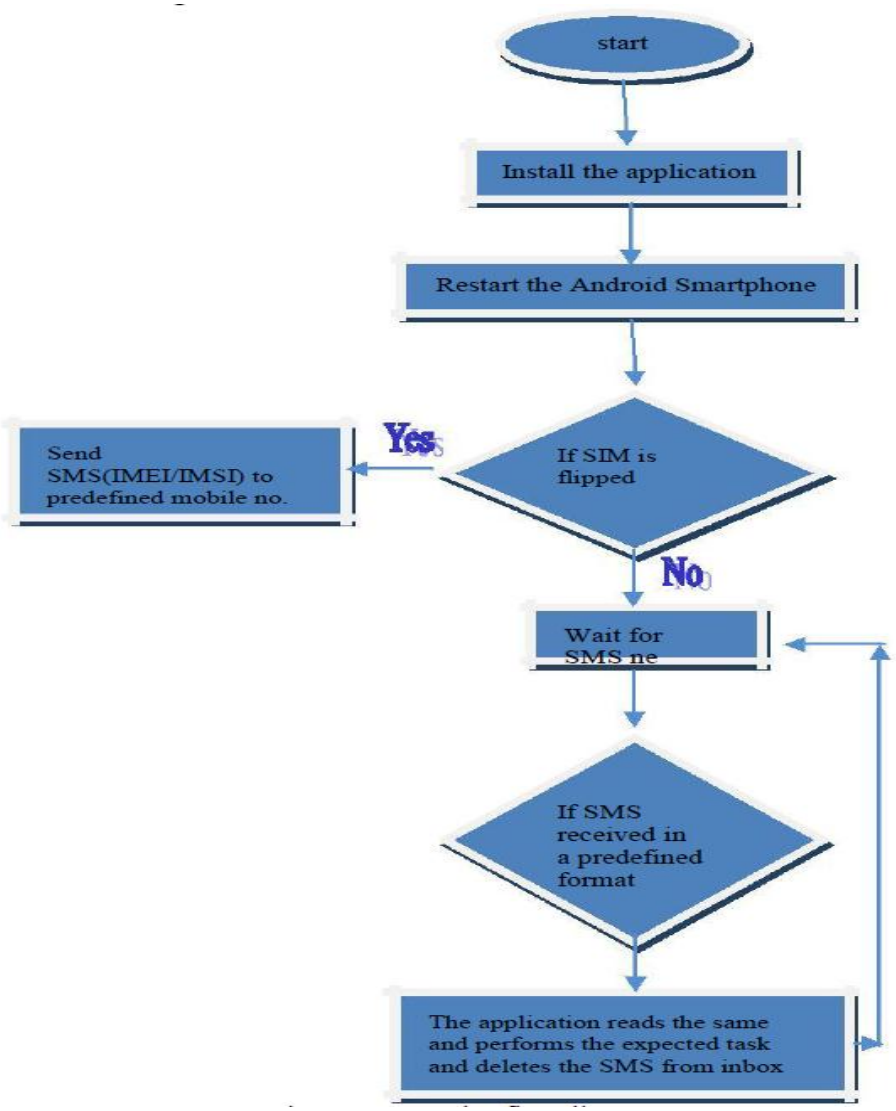


Figure 2. Dataflow Diagram[1]

Algorithm:

Step1:Start the process.
Step2:Install the application. After installing athe application on the Smart phone ,it will be set to start running in background every time the device os restarts.
Step3:Restart the Antroid Smart phone.
Step4:If the SIM is flipped the application sends SMS regarding the details of the new SIM to the predefined mobile no.
Step5:The application auto starts every time the mobile boots up. Then it goes to running mode and will start the main service which continuously listens for the incoming SMS messages.
Step6:Whenever it gets a new SMS ,it checks the content of the message and if the message is in a particular format ,application reads the same ,performs the expected task and replies back to the previously number.

The application installed will be running in the background and won't be shown in the task manager as well. Once the mobile phone is lost ,this application enables the user to track a mobile device and to receive notification via SMS to a predefined number.

4. EXPERIMENTAL RESULTS

User is the person who sends the message to the lost android phone using the application.

1. If SIM is replaced the predefined number will receive the IMEI/IMSI number and details of the newly inserted SIM.
2. If SMS is of the format @tracklocation then can find out the location of the thefted phone.
3. If SMS is of the format @trackcontacts then can findout the contacts of the thefted phone.
4. If SMS is of the format @trackcall log then can find out the call logs of the thefted phone.
5. If SMS is of the format @trackmessages then can find out the messages which are there in the thefted phone
6. If SMS is of the format @trackSR then can convert SILENT mode to the RING mode.
7. If SMS is of the format @trackRS then can convert RING mode to SILENT mode.
8.If SMS is of the format @trackactive then can convert into the blocking mode.
9.If SMS is of the format @tracknonactive then can convert into the nonblocking mode.
10. If SMS is of the format @trackpicture then can take the picture of the person.

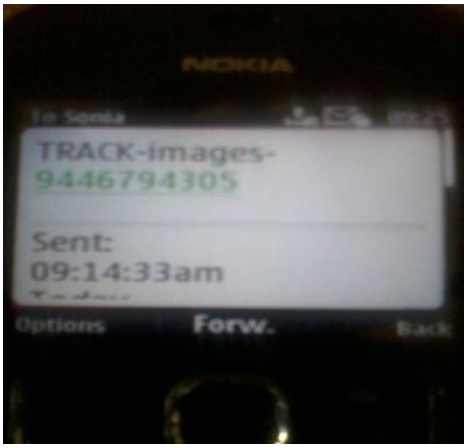


Figure 3. sim change happens

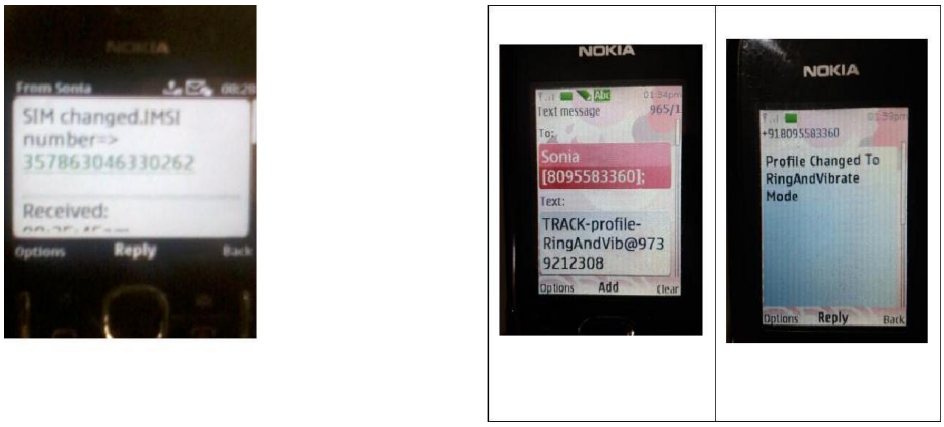


Figure 4.capturing imag

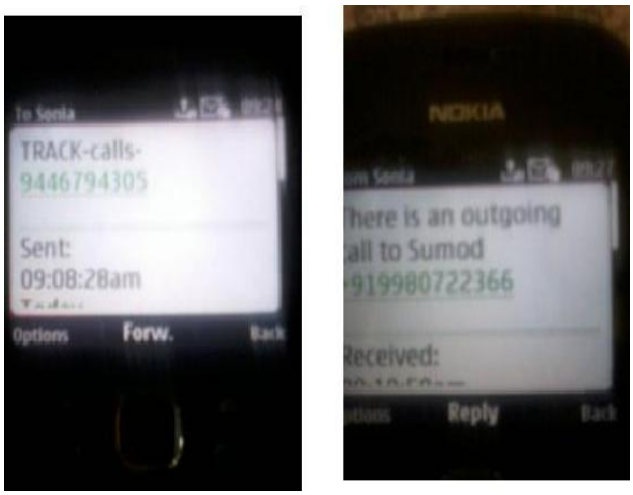


Figure 5: capturing incoming and outgoing calls

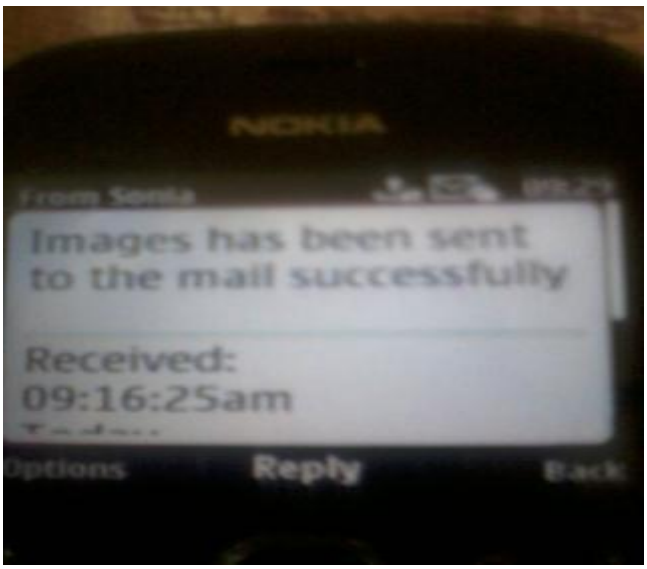


Figure 6. mail image to address

5. CONCLUSION

The mobile tracker is an efficient application for tracking of mobile phone. It uses the GPS and WI-FI for finding the location. When the SIM is replaced it will detect the location, track picture, call log, call details, and profile changing-silent to ring and ring to silent, track picture. This application doesn't work if the phone is switched off. For future work, it is proposed to implement some algorithm where the phone itself identifies that it is being lost. Whenever, the phone is off for more than 48 hours it should make it switch on automatically.

REFERENCES

[1] Alzantot, M. , Youssef, M. "UPTIME: Ubiquitous pedestrian tracking & locating using mobile phones "Wireless Res. Center, Egypt-Japan Univ. of Sc. & Tech. (E-JUST), Alexandria, Egypt Sonia C.V, Dr. A.R.Aswatha M.Tech.(student), Telecommunication Engineering Department DSCE, VTU, Bangalore, India Professor & Head of the Department (Telecommunication Engineering) DSCE, VTU, Bangalore, India.

[2] Luis C.M Varandas;Binod Vaidya;Joel J.P.C Rodrigues; "mTracker: A Mobile Tracking Application for Pervasive Environment" IEEE 24th International Conference on Advanced Information Networking and Applications Workshops,pp.962-967April 2010.

[3] Lin, Ding-Bing B. "Mobile location estimation and tracking for GSM systems" IEEE 15th International Conference on Personal, Indoor and Mobile Radio Communications, vol.4, pp.2835-2839, Sep. 2004.

[4] Bayir, Murat Ali" Track me! a web based location tracking and analysis system for smart phone users" 24th International Symposium on Computer and Information Sciences, pp.117-122,Sep.2009.

[5] Sangwoo Cho; Haekyung Jwa; Joohwan Chun; Jong Heun Lee; Yoon Seok Jung; "Mobile position location with the constrained bootstrap filter in a cellular communicationsystem"Thirty-Fourth Asilomar Conference n the context of near field communication applications", International Conference on Management of Mobile business, 2007, p.5.

[6] Madlmayr, G.; Dillinger, O.; Langer, J.; Schaffer, C.; Kantner, C. "The benefit of using SIM application toolkit i[5]Hellebrandt,Martin ,Mathar,Rudolf "Location tracking of mobiles in cellular radio networks" IEEE Transactions on Vehicular Technology,vol.48,pp.1558-1562,Sep1999.

AUTHORS

Shirin Salim. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from Ilahia college of Engineering and Technology, Muvattupuzha. Her areas of research are Machine Learning and Data Mining.

Dipina Damodaran B is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from Malabar College of Engineering Trissur. Her areas of research are Networking, Data Structures and Data Mining.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Machine Learning Database Management, Data Structures and Algorithm.

INTENTIONAL BLANK

NEO4J AS A SOLUTION TO HOSPITAL LOCALIZATION APPLICATION

Richa Kuriakose[1], AnuSebastian[2] and Surekha Mariam Varghese[3]

Department of Computer Science and Engineering, M.A College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

Graphs are efficient ways to visualize and representreal world data. Solutions to many real time scenarios can be easily provided when there is powerful graph databases like neo4j that can be used to efficiently query the graphs with multiple attributes. For instance, querying a system with medical and hospital data can be used to address the problem of location wise medical decision making. Here in this paper we present a neo4j as a solution to medical query.

KEYWORDS

Graph databases, SQLite, cypher query language, Neo4j, Relational databases.

1. INTRODUCTION

Nowadays people find it difficult to locate hospitals with a required specialization. For example, a patient who needs an appointment at the ophthalmologist, need to first identify those hospitals having eye specialists in his area. The information can be made available through an online application. The choice of database for the application is a determining factor in the speed, efficiency and ease of use of the online application.

The application that address the hospital and medical query can use any type of database. Relational Database Systems implement the relational model to work with the data. Relational databases are commonly used solution for data storage in almost all applications[16],[17]. SQL is the language used for querying and maintaining the relational databases [4],[13],[14]. SQLite is a relational database that is arguably the most widely deployed database engine, as it is used today by several operating system, browser and embedded systems. SQLite has bindings to many programming languages. As a self-contained, file-based database, SQLite offers an amazing set of tools to handle all sorts of data with much less constraint and ease compared to hosted, process based (server) relational databases. These features of SQLite made us to choose SQLite as the representative of relational databases to build Hospital localization application. But the relational databases have restriction in terms of size of the database. The connection between the entities in relational databases is done using the foreign key. But the connections and relations can be easily represented in the graph databases using the nodes and edges.

Graph databases enable us to build sophisticated models that map closely to our problem domain. Graph databases are schema less and efficient storage for semi structured data[12],[15]. They

DOI: 10.5121/ijci.2016.5224

express queries as traversals. They can do fast deep traversals instead of slow SQL joins. They allow ACID transactions with rollback support. The common graph databases includes Neo4j, Flockdb etc. Neo4j is an open-source graph database, implemented in Java. It employs the mathematics of graph and utilizes its huge potential for fast information extraction speeds to store information in the form of nodes and relationships. In this paper, we perform a comparison on both SQLite and Neo4j on the platform of hospital application.

2. PROBLEM DEFINITION

Graph databases are databases that uses graph structures for semantic queries with nodes edges and properties [1], [2], [12], [15]. In a graph database nodes represent entities such as people, business, accounts or any other item. Properties are information that relate to nodes. Edges are lines that connect to nodes or nodes to properties and they represent the relationship between the two. Most of the important information is really stored in the edges. Meaningful patterns emerge when one examines the connections and interconnections of nodes, properties, and edges. Figure 1 shows how data is represented in a graph database.

Neo4j is the most popular graph data base use in today [3]. Neo4j uses declarative query language called cypher query language for querying [5], [6]. Very complicated queries can be expressed through cypher. Cypher is relatively simple but still powerful. Cypher is designed to be a human query language. It makes simple things easy and complex things possible. Its constructs are based on English prose and queries are self-explanatory. Being a declarative language, Cypher focuses on the clarity of expressing what to retrieve from a graph, not on how to retrieve it. Cypher syntax provides a familiar way to match patterns of nodes and relationships in the graph.

Hospital localization is an application designed to find out the primary health centers and the specialized hospitals in an area, having the list of all hospitals, specialization and distance between the hospitals. The application allows fast and efficient best effort search in emergency situations. In situations like accident or other medical emergencies we may have to find the closest primary health center for getting first aid and after that from there it may be necessary to shift to the nearest hospital with required specialization. In such circumstances it is necessary to find the nearest health center and specialized hospital closest to the nearest health center. This type of query requires different levels of traversal, if we represent location wise hospital data and connection between them as a graph. For answering such queries we need a powerful graph database that combines the expressiveness of graph and dependability of database. Neo4j is the best choice. Because Neo4j is a graph database with powerful and simple query language (Cypher Query Language- CQL) capable of expressing complex queries. Relational databases cannot answer such queries that easily. In hospital localization application the distance between the hospitals need to be represented as an attribute to the edge. Neo4j allows to add properties to nodes as well as edges.

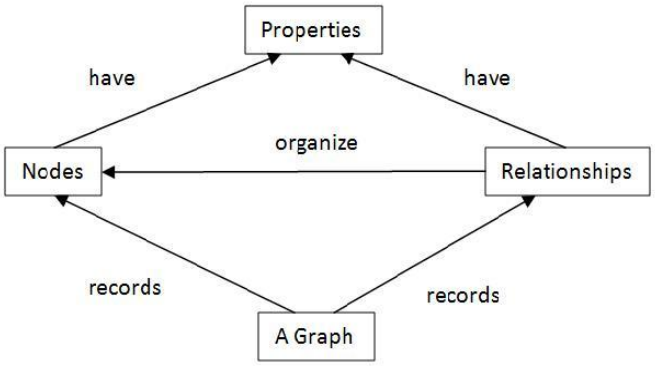


Fig. 1. Data representation in graph database

In relational database like SQLite we need different tables to represent the hospital localization application. If we use Neo4j for the hospital localization application the hospitals can be represented as nodes and the connection between them by road can be represented as the relationship between them. We can have specializations in a hospital as properties for the nodes and the distance between the hospitals as the property for the relationship. Figure.2 shows an instance of the graph for the hospital localization application including the hospitals in the Kothamangalam town within 100 meters from the National Highway 49.

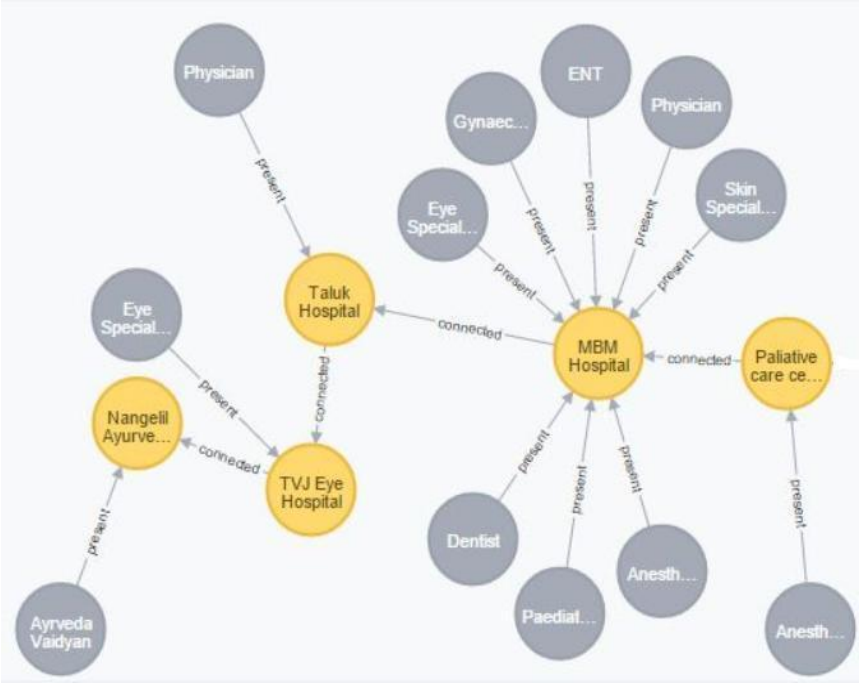


Fig. 2. Instance from Hospital Localization Application

The Figure depicts hospitals with their specialization and distance between them. The application allow different type of queries like, query for required specialization, query for hospitals within a given distance, query for the closest hospital or closest hospital with desired specialization etc.

3. METHODOLOGY

A detailed view of how hospital localization can be implemented with SQLite and Neo4j is discussed in this session. Let's first see the implementation with SQLite. It is necessary to have a table to store hospital information, another table to store distance information. For querying a hospital within a distance from a location it is necessary to perform JOIN operation on the tables. For example, if a patient wants to localize a hospital specialized in ophthalmologist, he could use a simple select query for determining those hospitals from a particular distance from his location Since JOIN operations are used in case of SQLite, it takes more accumulation of time.

In Neo4j we use cypher query language to develop the hospital localization application as a graph database [7]. It follows SQL like syntax, which is very simple and human readable. Let's now see the implementation of Neo4j. The procedure for implementing hospital localization application using Neo4j is given below:

1. Create nodes representing hospitals
2. Associate each hospital with the specializations as properties
3. Associate hospital nodes with relationship and add distance as property to the relationship.

Step1. Create nodes representing hospitals

Nodes for each hospital in an area are created using special queries. Each hospital node have a set of properties or attributes associated. The hospital nodes form the major nodes of a particular locality.

Step2. Associate specializations to hospital nodes

Specializations are added to each hospital nodes using links. Links determine the presence of specialization in a hospital. This association helps in fast checking of specializations in a hospital.

Step3. Associate hospital nodes with relationship

The hospital nodes of a locality are associated using relationship. The edges representing relationships as "connected" are created. This association helps in the faster traversal through hospitals in a locality."Distance" is added as a property for the relationship. These are done using the "CREATE" constructs in CQL and Querying is done mainly using "MATCH", "WHERE", "WITH" and "RETURN".

4. PERFORMANCE EVALUATION

On analyzing the performance of both Neo4j and SQLite on hospital localization application, the performance of Neo4j is more when compared to that of SQLite. An ideal graph database should understand analytic queries that go beyond neighborhood selection. In relational databases like SQLite, the index represents pre-determined knowledge of the structure of the computation without knowing the specific input parameters. A relational query selects a subset of the data and

joins it by known fields [10], [11]. The JOIN operations decelerates the query by pre-computing a portion of the query. In a graph database like Neo4j, the equivalent index is the portion of an analytic or graph algorithm that can be pre-computed or kept updated regardless of the input parameters of the query [8], [9]. Examples include the connected components of the graph, to localize the hospitals. Hence on regard to the efficiency factor, SQLite is not a good choice.

Neo4j has more accuracy than SQLite. Neo4j is much faster than SQLite. It is easy to use Cypher query language when compared to JAVA. Cypher is a declarative graph query language for the graph database, Neo4j that allows for expressive and efficient querying and updating of the graph store. Cypher is a relatively simple but still very powerful language. Very complicated database queries can easily be expressed through Cypher. This allows users to focus on their domain instead of getting lost in database access.

For evaluating the performance of SQLite and Neo4j we took three queries. The results are shown in Fig 3.

Query 1: Find all hospitals in a locality within a distance
Query 2: Find the specializations in a particular hospital
Query 3: Find the ophthalmologists within a restricted area

Table1: Neo4j Vs SQLite

Name	**Neo4j**	**SQLite**
Description	Open source graph Database	Widely used in-process RDMS
Database model	Graph DBMS	Relational DBMS
Implementation Language	Java	C
Server operating Systems	Linux OS X Windows	server-less
Data scheme	Schema-Free	yes
Typing	Yes	Yes
Secondary Indexes	Yes	Yes

SQL	No	yes

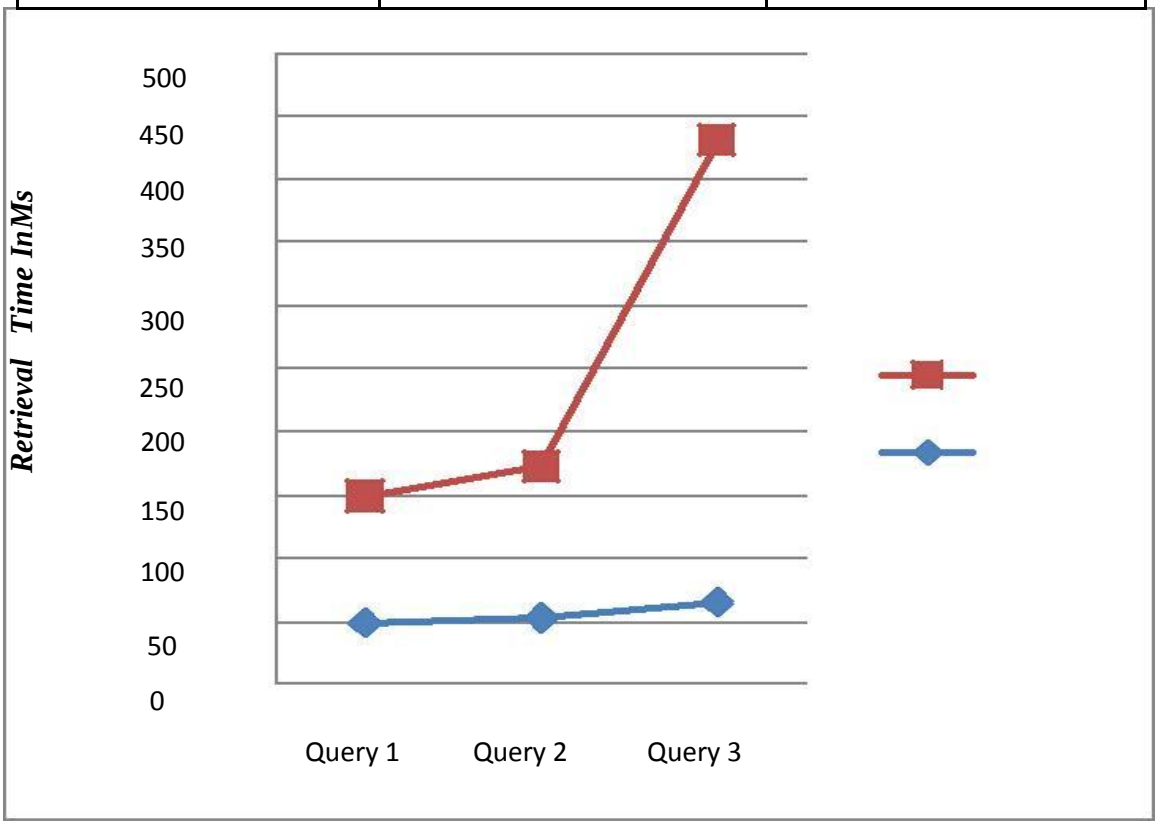


Fig 3. Retrieval time for queries in Neo4j and SQLite

Table 1 shows a comparison of Neo4j and SQLite based on some basic features. Figure.3 shows the retrieval time for the above mentioned queries for Neo4j and SQLite in milliseconds. A comparitive rise can be seen for Query 3 for SQLite when compared to the Query 3 for Neo4j. This is due to the rise in retrieval time to select the ophthalmologists within a resticted area as per Query 3.

5. CONCLUSION

SQLite is a relational database that is arguably the most widely deployed database engine, as it is used today by several operating system, browser and embedded systems. SQLite has bindings to many programming languages. Graph databases enable us to build sophisticated models that map closely to our problem domain. Graph databases are schema less and efficient storage for semi structured data.

In this paper, we undergo performance evaluation between Neo4j and SQLite on hospital localization application. We created hospital localization application using both SQLite and

Neo4j and also plotted the retrieval time for three queries in Neo4j and SQLite. We came to a conclusion that when compared, Neo4j performs better than SQLite. Neo4j has more accuracy than SQLite. Neo4j is much faster than SQLite. It is easy to use Cypher query language when compared to JAVA. Cypher is a declarative graph query language for the graph database, Neo4j that allows for expressive and efficient querying and updating of the graph store. Cypher is a relatively simple but still very powerful language. Very complicated database queries can easily be expressed through Cypher. This allows users to focus on their domain instead of getting lost in database access.

Neo4j is a graph database whereas SQLite is a relational database. On comparison overall performance of graph databases exceeds the relational database. In conclusion, given a traversal of anartificial graph with natural statistics, the graph database Neo4j is more optimal than the relational database SQLite.

REFERENCES

[1] ShaliniBatra, CharuTyagi. Comparative Analysis of Relational And Graph Databases. International Journal of Soft Computing and Engineering (IJSCE), 2(2), May 2012.

[2] Renzo Angles And Claudio Gutierrez." Survey of Graph Database Models"ACM Computing Surveys, Vol. 40, No. 1, Article 1, Publication date: February 2008.

[3] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. The 5th ACM/USENIX Internet Measurement Conference, 2007.

[4] "NOSQL Databases", http://nosql-database .org/

[5] "Neo4j," http://neo4j.org/.

[6] Neo4jmanual, Internet: http://docs.neo4j.org/chunked/stabl e/graphdb-neo4jnodes. html ,2010

[7] J. Paredaens and B. Kuijpers, "Data Models and Query Languages for Spatial Databases," Data & Knowledge Engineering (DKE), vol. 25, no. 1-2, pp. 29–53, 1998

[8] P. Urb ´ on, "Nosql graph database matrix," http://nosql.mypopescu.com/post/619181345 /nosql-graph-databasematrix, May 2010.

[9] "Short overview on the emerging world of graph databases,"http://www.graphdatabase.org/overview.html

[10] Michael Owens, The Definitive Guide to SQLite, USA: Apress, 2006, 341-362.

[11] SQLite homepage [EB/OL] ,http://www.sqlite.org.

[12] Florian Holzschuher, Prof. Dr. René Peinl. Performance of Graph Query Languages: Comparison of Cypher, Gremlin and Native Access in Neo4j.EDBT/ICDT'13March 18 - 22 2013.e

[13] SzabolcsRozsnyai, AleksanderSlominski, YurdaerDoganata.Large-Scale Distributed Storage System for Business Provenance.CloudComputing (CLOUD), IEEE International Conference,2011.

[14] IlyaKatsov. NoSql Data Modeling Techniques. March 2012.

[15] Ian Robinson, Jim Webber, Emil Eifrem. Graph Databases(Early release revision 1).ȮReilly Media, Inc., 02-25 2013.

[16] Apache Software Foundation.Home.http://zookeeper.apache.org, 2013.

[17] YanmeiHuo, Hongyuan Wang, Liang Hu, Hongji Yang. A Cloud Storage Architecture Model for Data-Intentsive Applications. 2011.

AUTHORS

RichaKuriakose is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering, Kothamangalam. She completed her B.Tech from Mar Athanasius College of Engineering, Kothamangalam. Her areas of research are DataMining and Machine Learning.

Anu Sebastian is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering, Kothamangalam. She completed her B.Tech from Pulincunnoo Engineering College, Alappuzha. Her areas of research are DataStructures and Algorithms and Biocomputing.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems, Distributed Computing and Machine Learning. She has published 17 papers in international journals and international conference proceedings. She has served as reviewer, committee member and session chair for many international conferences and journals.

OUTCOME ANALYSIS USING NEO4J GRAPH DATABASE

Mary Femy P.F, Reshma K.R, Surekha Mariam Varghese

Department of Computer Science and Engineering, Mar Athanasius college of engineering, Kothamangalam, Kerala, India

ABSTRACT

Databases are an integral part of a computing system and users heavily rely on the services they provide. When interact with a computing system, we expect that data be stored for future use, that the data is able to be looked up fastly, and we can perform complex queries against the data stored in the database. Many different emerging database types available for use such as relational databases, object databases, key-value databases, graph databases, and RDF databases. Each type of database provides unique qualities that have applications in certain domains. Our work aims to investigate and compare the performance and scalability of relational databases to graph databases in terms of handling multilevel queries such as finding the impact of a particular subject with the working area of pass out students. MySQL was chosen as the relational database, Neo4j as the graph database.

KEYWORDS

Neo4j, NOSQL, Graph database.

1. INTRODUCTION

The relational model has dominated the computer industry since the 1980s mainly for storing and retrieving data. Lately, however, relational database has been losing its importance due to its reliance on a strict schema which makes it difficult to add new relationships between the objects. Another important reason of its failure is that as the available data is growing manifolds, it is becoming complicated to work with relational model as joining a large number of tables is not working efficiently. One of the proposed solutions is to transfer to the Graph databases as they aspire to overcome such type of problems. This paper provides a comparative analysis of a graph database Neo4j with the most widespread relational database MySQL.

Relational databases such as Oracle and MySql excel when it comes to capturing repetitive, tabular data. Despite the word “relational” in their name, relational database are much less effective at storing or expressing relationships between stored data elements.Unlike a relational database, a graph database is structured entirely around data relationships. Graph databases treat relationships not as a schema structure but as data, like other values.

The graph database queries are domain-specific user-friendly and can be considered as "SQL for graphs". The similarity to SQL is intentional and makes transition much easier for developers. When SQL query on the RDBMS is as long as half a novel, the Cypher Query equivalent is much shorter and intuitive. The traverser API in RDBMS is highly resource intensive, since each step to neighboring node has to be depicted with JOIN. In contrast, the graph database hypergraph property allows direct access to neighboring nodes by eliminating the edge attribute.

DOI: 10.5121/ijci.2016.5225

Graph databases support a graph model which allows a direct persistent storing of objects in the database together with the relations between them. A graph database should provide access to query methods that not only deal with the stored objects, but also with graph structure. The best known example of such an operation is traversal, which in its most simple form can be used to obtain the neighbors of an object, that is, the objects that are directly related to.In this paper we investigate the outcome analysis of pass out students in a particular institution using Neo4j and compare it with MySQL.

2. BACKGROUND

The relational database management system was created in the 1970s. Its popularity has skyrocketed, and it has become a primary data storage structure in academic and commercial pursuits. Relational databases range from small, personal databases like Microsoft Access to large database servers like MySQL, Microsoft SQL Server, and Oracle. This paper focuses on MySQL. Graph database researches was popular in the early 1990s, but died out for a series of reasons including the surge of XML research and hypertext. With the rise of the Internet as a tool for the public, data began to increase both in interconnectedness and in volume. The graph model was used to represent huge amounts of data more than it had in the past. Traditional data stores were capable of handling graph data. Yet, they were neither designed to do so nor efficient at it. There was clear desire for data store tailored to the needs of graph data.

In recent years, software developers have been investigating storage alternatives to relational databases. NoSQL is a term for some of those new systems. BigTable, CouchDB, Cassandra, Project Voldemort, and Dynamo are all NoSQL projects, as they are all high-volume data stores that reject the object-relational and relational models.

Atomicity, consistency, isolation, and durability (ACID) are a set of governing principles of the relational model. They guarantee database reliability. NoSQL rejects ACID. The term"NoSQL" as a term for modern web data stores, first began to gain popularity in early 2009. It is a topic that has gained recognition from IT community but has yet to garner large scale academic study. The NoSQL movement has its own discussion groups, conferences and blogs. As the typical database administrator attempts to question whether to move from the relational model to NoSQL model, the NoSQL community presents him or her with potential flags that data might be more suitable for a NoSQL system.

1. Having tables with lots of columns
2. Having attribute tables.
3. Having lots of many-to-many relationships.
4. Having tree-like characteristics.
5. Requiring frequent schema changes.

Data provenance meets several of these criteria, so it would be fitting to investigate NoSQL solutions to the provenance storage problem. An experiment was conducted to test the viability of NoSQL data store, Neo4j, for data provenance needs.

3. MOTIVATION

This paper is a comparison of the relative usefulness of the relational database MySQL and the graph database Neo4j to store graph data. A directed acyclic graph (DAG) is a common data structure to store data provenance information relationships. The goal of this study was to determine whether a traditional relational database system like MySQL, or a graph database, such as Neo4j, would be more effective as the underlying technology for the development of a data provenance system.

Graph database models can be characterized as data structures for the schema. The instances are modeled as graphs or generalizations. Data manipulation is expressed by type constructors and graph oriented operations. One of the motivations towards this paper is to provide a benchmarking mechanism to measure the effectiveness of graph traversal operations. It also motivates us to measure the capabilities of graph databases to perform query like traversal where one searches for topologically related vertices for a given vertex. It also searches the graph analysis/mining operations that require the traversal of the whole graph.

4. APPLICATIONS OF GRAPH DATABASE

Several areas have witnessed the emergence of huge data networks called complex networks. So graph databases are the best database to implement such complex network of relationships having millions of nodes and relationships. The main application areas of graph databases are:

1) Social networks*:* In social networks, nodes are people or groups, while links show relationships or flows among nodes. Some examples are friendships, business relationships, research networks, communication records (mail, telephone calls, email), computer networks, and national security There is growing activity in the area of social network analysis and also in visualization and data processing techniques for these networks.
2) Information networks: Information networks model relations representing information flow, such as citations among academic papers, World Wide Web, peer-to-peer networks, relations among word classes in a thesaurus, and preference networks.

5. ADVANTAGES

The benefits of using a graph data model are given by: the introduction of a level of abstraction which allows a more natural modeling of graph data; query languages and operators for querying directly the graph structure; and ad-hoc structures and algorithms for storing and querying graphs. Graph databases are also somewhat similar to object databases in case where objects and relationships between them are all represented as objects with their own respective sets of attributes. Graph database consists of several advantages:

- It enables very fast queries when the value of the data is the relationships between people/items.
- Use Graph Databases to identify relationships between people/items, even when there are many degrees of separation.
- Where the relationships represent costs, identify the optimal combination of groups of people/items.

6. PROPOSED SYSTEM

Neo4j, like many other graph databases, builds upon the property graph model; labeled nodes (for informational entities) are connected via directed, typed relationships. Both nodes and relationships hold arbitrary properties (key-value pairs). There is no rigid schema, but with node-labels and relationship-types we can have as much meta-information as we like. When importing data into a graph database, the relationships are treated with as much value as the database records themselves. This allows the engine to navigate your connections between nodes in constant time. That compares favourably to the exponential slowdown of many-JOIN SQL-queries in a relational database.

Proposed System consists of research and comparison of two databases such as Neo4j and MySQL databases. A graph database stores data in a graph, the most generic of data structures, capable of elegantly representing any kind of data in a highly accessible way.

MySQL has significant market penetration in the academic and scientific fields. Furthermore MySQL has significant support, both from the manufacturers and from the user community. It is a pure relational database, as opposed to an object-relational database like Oracle and SQL Server.
Neo4j is open source for all noncommercial uses. It has been in production for over five years. It is quickly becoming one of the foremost graph database systems. According to the Neo4j website, Neo4j is "an embedded, disk-based, fully transactional Java persistence engine that stores data structured in graphs rather than in tables". The developers claim it is exceptionally scalable (several billion nodes on a single machine), has an API that is easy to use, and supports efficient traversals. Neo4j is built using Apache's Lucene for indexing and search. Lucene is a text search engine, written in Java, geared toward high performance.

A. Types of Graph Database Models

1) Neo4j Graph Database: As a robust, scalable and high-performance database, Neo4j is suitable for full enterprise deployment or a subset of the full server can be used in lightweight projects.

It features:

- Intuitive using a graph model for data representation
- Reliable
- Durable and fast, using a tradition disk Based native storage engine.
- Extraordinarily scalable, up to several billion nodes/relationships/properties.
- Expressive with a powerful, human readable graph query language
- Fast with a powerful traversal framework for high speed graph queries.
- True ACID transactions
- High availability
- Scales to billions of nodes and relationships
- High speed querying through traversals

Proper ACID behavior is the foundation of data reliability. Neo4j enforces that all operations that modify data occur within a transaction, guaranteeing consistent data. This robustness extends from single instance embedded graphs to multi-server high availability installations. Neo4j is a commercially supported open-source graph database. It was designed and built from the ground-up to be a reliable database, optimized for graph structures instead of tables. Neo4j is based on the data model of a directed multigraph with edge labels and optional node and edge properties. Node and links can be changed but have identity maintained by DBMS. Labels and property keys are strings, property values can be primitive java data types and strings or arrays of both. The fundamental units that form a graph are nodes and relationships. In Neo4j, both nodes and relationships can contain properties. Nodes are often used to represent entities, but depending on the domain relationships may be used for that purpose as well.

Neo4j's query language Cypher aims to be a user-friendly language that is designed to be read and understood easily. It allows you to declare patterns (MATCH) that you want to find in the graph and then apply filters (WHERE), projection (RETURN) and paging (LIMIT,SKIP,ORDER BY) to your result data.

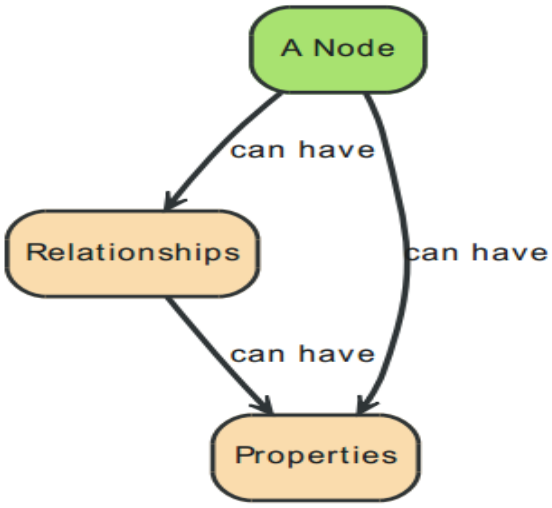


Fig.1. Neo4j Graph Database Nodes and relationships

B. Graph Databases and Their Support for Querying Graphs

1) *Adjacency Queries:* In this type of queries the primary notion is node/edge adjacency. Two nodes are adjacent when there is edge between them.
2) Reachability Queries*:* These queries are characterized by path or traversal problem. The problem causes in reachability test whenever two given nodes are connected to path.
3) Pattern Matching Queries: Pattern matching queries find all sub-graphs of data graph that are isomorphic to pattern graph.
4) Summarization Queries*:* Summarized queries are not related to consult the graph structure. They are based on special functions that allow summarizing or operating on the query results, normally returning a single value.

7. EXPIREMENTAL EVALUATION

A. Setup: Computer Configurations and Datasets

We used core i3 processor, 2 GB of RAM and 10 GB SATA for implementing Neo4j graph database. Here we used institution data for evaluating Neo4j.

B. Global Query

The queries were designed to simulate some of the types of queries used in provenance systems. For example, traversals are necessary to determine data objects (nodes) derived from or affected by some starting object or node. That is, if a data object is determined to be incorrect, that information must be propagated to all “descendants” of the node. Searching for specific values within the payload is another common operation.

The performance of a global constraint based user lookup was constructed to measure the performance of queries typically issued on databases. The intent of the global query was to characterize the performance of queries requiring inspection of all users in the system.

For each of the queries with varying dataset sizes, 100 data points were collected. The resulting data points were then averaged to summarize the data collected for the particular test.

1). MySQL Global Query

The global query that was run against the MySQL database was intended to return all of the nodes in the system that were between a provided age group. The query utilized is as follows (where the two question marks were the lower limit on the age and the upper limit):

SELECT count(*) FROM student_node WHERE student_node.age > ? AND student_node.age < ?;

2). Neo4J Global Query

The global query that was run against the Neo4j database was aimed at attaining the same data as the MySQL global query. That is, users in the database that are within a given age range. The query utilized is as follows (where 'X' stands for the minimum age and 'Y' stands for the maximum age):

START x=node(*) WHERE (x.age? > X and x.age? <Y) return count(*);

C. Query Performance

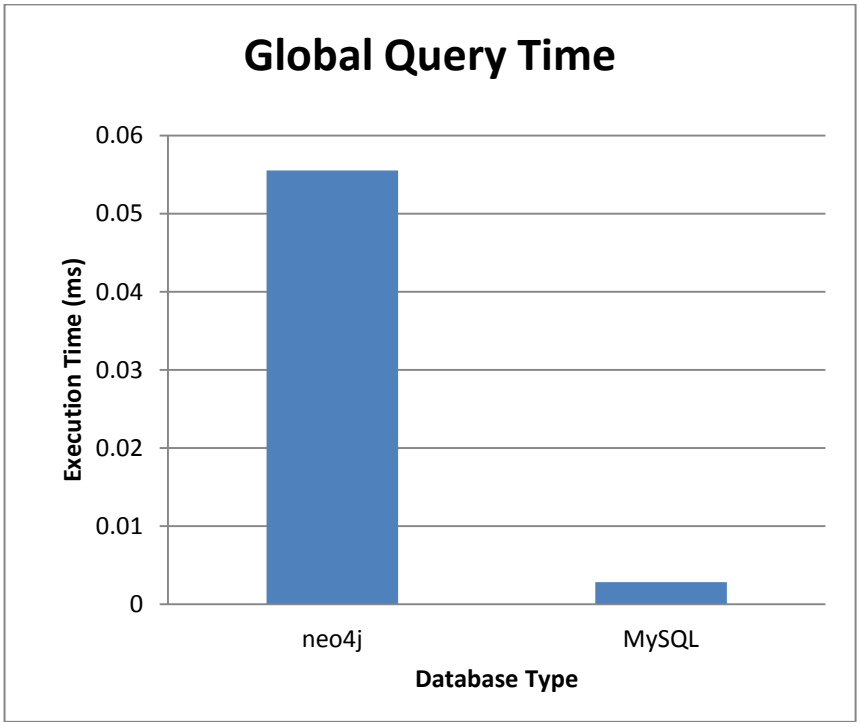


Fig.2. Global Query Performance

In Figure, the global query execution time is given for the neo4j database vs. the MySQL database. For this test, queries were run against the entire underlying database that looked for a range of random ages, which was run a total of 100 times with differing age ranges. The figure illustrates that the average execution time for the neo4j global query and the MySQL global query. The data shows that the neo4j execution time was an order of magnitude larger.

8. CONCLUSIONS

Graph databases are specifically designed to handle graph based data more efficiently than the traditionally adopted relational databases. One such specific type of data is social networking data, or any type of data that is highly dependent on relationships between a collection of nodes. In particular, this study compared the performance of neo4j vs MySQL.

In general, the graph database did better at the structural type queries than the relational database. In full-text character searches, the graph databses performed significantly better than the relational database. The fact that the indexing mechanism used in the graph database was based on strings made the numeric queries less efficient. While a query on non-integer numeric data, such as doubles, was not included in the benchmark tests, the result would have likely been even worse for the graph database.

Graph databases exhibit the ability to scale exceptionally well with large numbers of nodes and/or relationships, whereas MySQL, or presumable any relational database begins to see a performance degradation with input data. The input data utilized in this experiment was comprised of only a few columns that represented friendship relationships between users. However, graph databases are easily transformed to contain many relationships amongst the nodes and also to have many attributes tied to any given node and/or relationship. In order to provide this in a relational database, many more relations (tables) need to be added to the underlying database, which has an impact on the system's performance.

In this paper we investigate and compare the performance and scalability of relational database to graph database by finding the impact of a particular subject with the working area of pass out students. In fact, Neo4j is best suitable to implement complex network of relationships having millions of nodes and relationships.

REFERENCES

1 Renzo Angles And Claudio Gutierrez." Survey of Graph Database Models"ACM Computing Surveys, Vol. 40, No. 1, Article 1, Publication date: February 2008.

2 Renzo Angles. "Comparison of Current Graph Database Models", Department of Computer Science, Talca.

3 "Short overview on the emerging world of graph databases,"http://www.graph-database.org/overview.html.

4 "Neo4j," http://neo4j.org/.

5 P. Urb ´ on, "Nosql graph database matrix," http://nosql.mypopescu.com/post/619181345 /nosql-graph-databasematrix, May 2010.

6 "NOSQL Databases", http://nosql-database .org/

7 D. Dominguez-Sal, P. Urb on-Bayes, A. Gim enez-Va n o, S. G omezVillamor, N. Mart ınez-Baz an, and J. L. Larriba-Pey, "Survey of graph database performance on the hpc scalable graph analysis benchmark," in Proc. of the 2010 international conference on Web-age information management (WAIM). Springer-Verlag, 2010, pp. 37–48.

8 Neo4j B log, Internet: http://blog. neo4j.org/2009/04/current-database-debate-andgraph.html ,201 0.

9 Neo4jmanual, Internet: http://docs. neo4j.org/chunked/stabl e/graphdb- neo4jnodes. html ,2010

10 J. Paredaens and B. Kuijpers, "Data Models and Query Languages for Spatial Databases," Data & Knowledge Engineering (DKE), vol. 25, no. 1-2, pp. 29–53, 1998.

AUTHORS

Mary Femy P F is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from Matha College of Technology, Paravur. Her areas of research are Data Mining, Databases and Image Processing.

Reshma K R is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from MEA Engineering College, Perinthalmanna. Her areas of research are Data Mining, Databases and Image Processing.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems and Distributed Computing, Machine learning. She has published 17 papers in international journals and international conference proceedings. She has been in the chair for many international conferences and journals.

PRIVACY PRESERVING INFORMATION RETRIEVAL OVER UNSYNCHRONIZED DATABASES

Meenu Poulose[1] and Tinku Soman Jacob[2]

[1] Student, Department of Computer Science and Engineering, MBITS, Nellimattom.
[2]Assistant Professor, Department of Computer Science and Engineering, MBITS, Nellimattom

ABSTRACT

A database is a collection of information that is organized so that it can easily be accessed, managed and updated. It may also contain the private and public information about an individual. The user access the information by giving some relevant queries. When a user retrieves any data the server may able to identify which the data is. If the user accesses the x^{th} data then that 'x' must be hidden from the server. Private information retrieval can solve this issue. Here introducing an efficient PIR scheme that uses multiple servers. Each non-colluding server stores the identical copies of the database. When any of the servers contain non-identical copy of database i.e. unsynchronized database the user will get an error. Our proposed system works properly even under such circumstances since it has the mechanism to find the unsynchronized databases. This multi server PIR scheme has the same computational and communication complexity. It also allows multiple PIR queries simultaneously.

KEYWORDS

Private information retrieval, information theoretic privacy, database management, distributed source coding.

1. INTRODUCTION

Most of the web services require the user sign up by giving some personal information like mail id, date of birth etc. But these information are not much safe within those servers. It may be viewed by some other third parties also. For example we have an account in an online shopping site and we used to view a particular product that we want to buy. When we visit any other site some pop-up messages may came. It may be an advertisement from the previous site about our favourite product. Here the identity and the retrieved data are disclosed by the server. It demands the need of private information retrieval.

We have many searching techniques that hide the user identity. But it is not a well protection since the contents we are searching is not hidden. By analyzing this contents can infer a conclusion about the user identity. To address this concern, some private search tools instead mask the contents of web queries. These tools include chaffing and winnowing approaches like TrackMeNot, encrypted database searches, private information retrieval (PIR), oblivious transfer,

DOI: 10.5121/ijci.2016.5226

and private stream searches (PSS). Chaffing and winnowing involves sending bogus queries to a server and extracting only the relevant results. Private information retrieval (PIR) is a way for a client to look up information in an online database without letting the database servers learns the query terms or responses. A simple and inefficient way to do this is for the database server to send a copy of the entire database to the client, and let the client look up the information for herself. This is called trivial download. The goal of PIR is to transmit less data while still protecting the privacy of the query.

PIR protocols can be grouped into two classes corresponding to the security guarantees they provide. One class is computational PIR, in which the database servers can learn the client's query if they can apply sufficient computational power to break a particular cryptographic system. The other class of protocols thoseconsider in this work is information- theoretic PIR , in which no amount of computation will allow the reconstruction of the client's query. In these protocols, the query is protected by splitting it among multiple database servers. As is common in many distributed privacy-enhancing technologies, such as mix networks, Tor , or some forms of electronic voting , we must assume that some fraction of the servers above some threshold are not colluding against the client. It seems unlikely that multiple large-scale servers would maintain identical databases without colluding. Furthermore, there is little incentive for existing service providers like Google or Yahoo to enable private searches.

In this paper we first introduce the basic multi server PIR system with synchronized databases and then explain the proposed system i.e. the PIR over unsynchronized databases. The key idea of our scheme is simple: here first determine which records are unsynchronized, and then construct a PIR query that avoids these problematic records. When the number of unsynchronized database records scales sub linearly in the database size, our scheme has asymptotic communication and online computation costs that are identical to state-of-the-art PIR schemes. In practice, the system incurs slightly higher communication and server- side computation compared to traditional PIR. Our approach also allows multiplequeries to be processed in a single batch of PIR, unlike existing schemes.

2. RELATED WORKS

Since 1995, much work has been done creating protocols for private information retrieval (PIR). Many variants of the basic PIR model have been proposed, including such modifications as computational vs. information-theoretic privacy protection, correctness in the face of servers that fail to respond or that responds incorrectly, and protection of sensitive data against the database servers themselves.We begin with an example of information-theoretic, multi-server PIR proposed by Chor et al [2].

2.1. BASIC PIR SCHEME

Two servers store identical copies of a database of records $f = [f_1 \ldots f_n]^T$, and a client wishes to retrieve the w^{th} record, f_w. In practice, records can be of arbitrary length, but for simplicity, suppose each database element is a single bit, 0 or 1. The user's request can be represented by $e_w \in \{0,1\}^N$, the indicator vector with a 1 at index w and 0's elsewhere. To disguise this query, the user generates a random string $\alpha \in \{0,1\}^N$ with each entry a Bernoulli (1/2) random variable. The queries sent to servers 1 and 2 are $\alpha \oplus e_w$ and α, respectively. Each server computes the inner product of its received query vector with the database x using bitwise addition (XOR) and returns

a single-bit result. The user XOR the results from the servers to get f_w precisely. The scheme is illustrated in Figure 1. In an honest-but-curious adversarial model, this PIR scheme is information theoretically private, since received queries α and $\alpha \oplus e_w$ appear random.

2.2. COLLUSION RESISTANCE

Multi server PIR scheme will fail when the servers were colluded. The client can improve her privacy by querying many randomly-selected servers (e.g. in a P2P network); this reduces the likelihood of sending queries to colluding servers. Algorithmically, there exist PIR schemes that offer the property of κ-collusion-resistance as long as no more than κ of the d servers collude, information-theoretic security is guaranteed.

2.3. DISTRIBUTED SOURCE CODING

Distributed source coding (DSC) is an important problem in information theory and communication. DSC problems regard the compression of multiple correlated information sources that do not communicate with each other. By modelling the correlation between multiple sources at the decoder side together with channel codes DSC is able to shift the computational complexity from encoder side to decoder side. In our problem, the client is the receiver, and the servers are the distributed sources. Assume that the number of unsynchronized database elements s is small, so the servers' contents are highly correlated. The client must learn which database elements are unsynchronized. the difference of the servers' data - to successfully complete PIR.

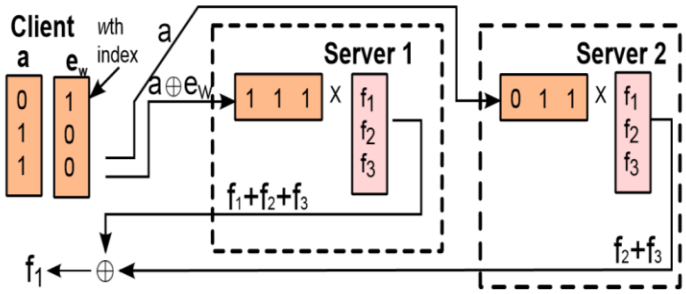


Figure 1: Basic two-server PIR scheme. Each server computes the bitwise sum of a user-specified subset of database records. Because the two user-specified subsets differ only at the w^{th} index, the binary addition of each server's results gives the desired record.

3. PROBLEM DEFINITION

Existing PIR schemes do not have the mechanism to handle the unsynchronized databases. PIR allows a user to retrieve information from a database without revealing the server which data is retrieved. To achieve this privacy mainly two types of PIR schemes were introduced. Among them multi-server PIR is capable to achieve that privacy without more computations. Here the main requirement is that it needs multiple non-colluding servers. The requirements for multi-server PIR are;

1) Multiple servers are available.
2) Each server stores a duplicate copy of the database.
3) The individual servers do not collude.
4) The servers are honest-but-curious.
5) Servers willingly implement PIR algorithms.

Existing schemes only addressed the assumptions 3 and 4. Some PIR schemes are robust to Byzantine servers that return arbitrary , incorrect information [10]. Other schemes allow up to k servers to collude without losing any privacy [8].If any of the servers database is unsynchronized then the existing schemes will either produce an error or output an incorrect result. Figure 2 shows such a scenario. In order to handle such unsynchronized databases here introducing an efficient PIR scheme that can locate the unsynchronized databases and after modification the user can retrieve the desired data.

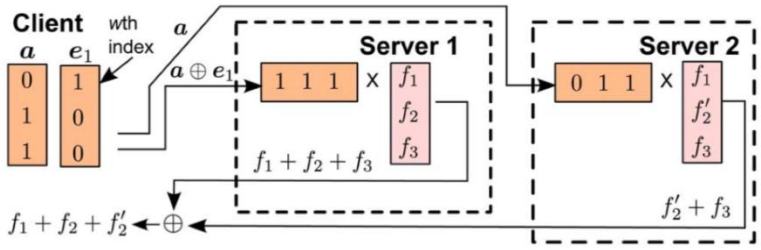


Figure 2: Basic two-server PIR scheme when one of the records contains an outdated data.

4. PRIVACY PRESERVING INFORMATION RETRIEVAL OVER UNSYNCHRONIZED DATABASES

Multi-server PIR schemes require multiple non-colluding servers. If any of the server collude to another then the privacy cannot be guaranteed. If we have strictly non-colluding servers still there exist a problem i.e. the unsynchronized databases. Consider the scenario, that we have three non-colluding servers and each server posses the similar copies of databases. But in the third sever the user failed to update his records. So it has an out-of-date entry. When the user requests for a data he will get the sum of the desired records with an error term. Some alternative methods are there to solve this problem. One is to treat the reply from the third server as an error and query the second or third server [5]. But it will increase the cost of communication since it needs to communicate with a server that is not already connected. The main aim here is to find the perfectly synchronized set of servers. Our proposed system asymptotically has the same computational and communication complexity as state-of-the-art PIR schemes for synchronized databases; this comes at the expense of probabilistic success and two rounds of communication (most existing schemes require only one).

The proposed scheme consists of two phases. Phase 1 find which records are unsynchronized and phase 2 retrieves the desired data to the user. The collusion-resistance scheme is used here for synchronized databases.

4.1. PHASE 1: LOCATE UNSYNCHRONIZED RECORDS

In the multi-server PIR schemes each server will have the same copies of databases. If any of the servers contains a mis-synchronized database we will find the location of that record in this phase. Unlike [4] to find the location this scheme relies on hash functions. Here assume that the hashes of each record are stored within the databases. Suppose one server posses the record f_i and the other non colluding server also has the same record as f_i. Then both the servers will have same hash value H (f_i). If one of these two servers has an out-of-date record f_i then both hash values will be different.

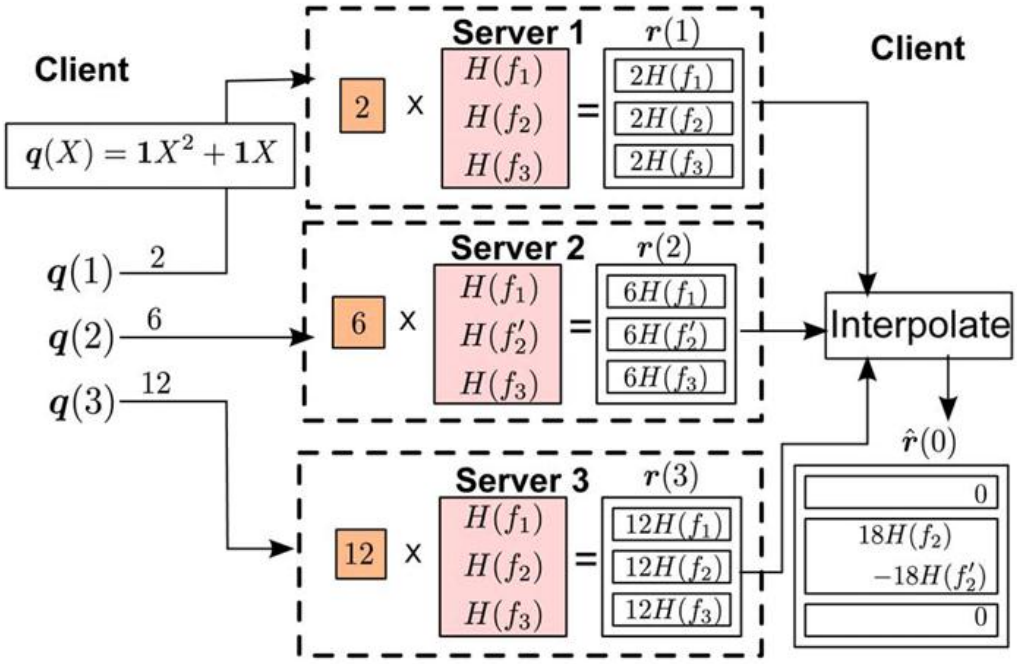


Figure 3: Identifying the location of unsynchronized records. Here f_2 is not synchronized across all servers.

Suppose we have only two servers, and a genie sums the servers' respective views of the database hashes over $GF(2^l)$, giving $H(f^{(1)}) + H(f^{(2)})$. The synchronized (i.e., equal) hashes cancel, giving a sparse vector of length n, with nonzero entries at the unsynchronized records. A parity check matrix could be used to compress this sparse vector for transmission to the client: $\mathbf{A}\cdot(H(f^{(1)}) + H(f^{(2)}))$ By linearity, this is equivalent to the equation $\mathbf{A}\cdot H(f^{(1)}) + \mathbf{A}\cdot H(f^{(2)})$. So to communicate the same information in a distributed fashion, each server can simply compress its own database with a pre-determined $\mathbf{A}$ matrix, and the client can recover the sparse vector from the compressed vectors. This is for a two server architecture and the same idea works for more than two servers i.e. each server S_i individually compresses its view the database by returning $\mathbf{A} \cdot f^{(i)}$ to the client. After this the client can do pair wise reconstruction finding the set of unsynchronized records.

4.2. PHASE 2: RETRIEVE THE DESIRED RECORD(S)

After the first round of communication the client knows the locations of the unsynchronized records. So the servers must ensure that they avoid touching those records. If the record w is

unsynchronized then both server's query vectors should be zero at the w^{th} index i.e. server neither touches the unsynchronized records. The same idea holds for more servers. Here we are not focusing on the data synchronization. Instead of it, this mechanism provides information retrieval even under the circumstance where the data in each server are not synchronized.

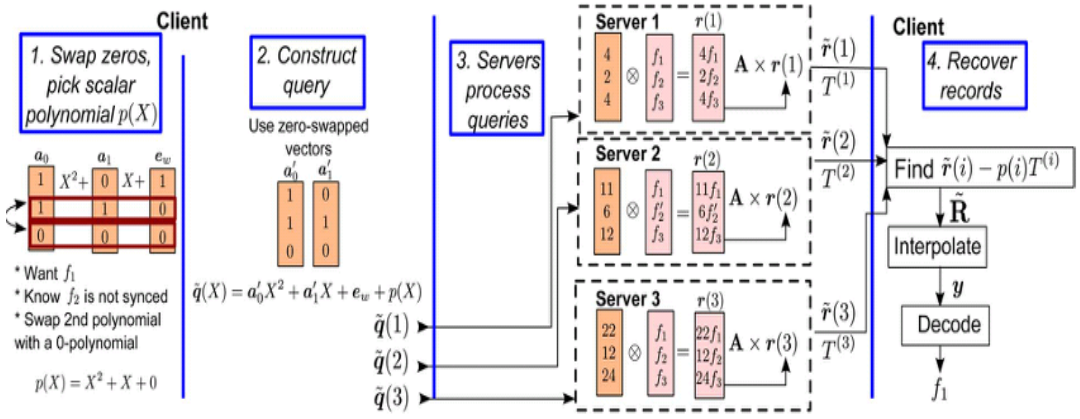


Figure 4: An example for PIR scheme, phase 2.Here the database have 3 records i.e. n=3

4.3. SYNCHRONIZATION

The existing multi-server PIR scheme describes the information retrieval process over unsynchronized database, but it doesn't have a phase to synchronize the outdated database. After retrieving the data there should be some steps to achieve synchronization. For that we can use the information that received after the phase 1. Phase 1 locates where the unsynchronized data records are. Hash functions of each record from each servers can be computed from the phase 1. After comparing the hash values from servers we can identify the unsynchronized server as well as thesynchronized servers. To achieve synchronization just replace the unsynchronized data with the data from synchronized records.

The required amount of communication and computation are similar to the other existing schemes. The main communication overhead in our scheme when compared to an optimal private information retrieval [3] is the downlink communication in phase 1 when locating unsynchronized database records. But the total communication complexity is same since the above overhead is dominated by the uplink in phase 1. So the communication complexity is same between our schema and [3]. Next is computation, our scheme requires the pre-computation of hash values of each record for the synchronization phase. Here the server computation cost is essentially equal to the [3], But client takes some additional computation to find the unsynchronized record's locations and to interpolate the desired records.

To understand the performance of this multi server PIR we can consider two factors i.e. the probability of success and the total query run time. The probability of success relies on the communication. The probability of success can be described as the probability of correctly retrieving the desired record. It increases as a function of communication cost. Some patters of query records and mis-synchronizations may generate decoding errors. That will result a client

being not recovering the requested record. Even when the number of mis-synchronized records is small this PIR scheme returns the desired records with a high probability.

Next is runtime, multi-server PIR requires more time to locate unsynchronized records when compared to other schemes like [3]. Runtime can be expressed as a function of the unsynchronized database records. Our runtime overhead is comparatively small, but it increases with the number of unsynchronized records. The overhead of our scheme does not increase as a function of database size, because the runtime overhead is dominated by locating the synchronization errors. Locating this is a process that depends on the number of unsynchronized records than the database size.

5. CONCLUSIONS

The privacy preserving information retrieval uses the multi server PIR scheme that can work even with the unsynchronized databases. This PIR scheme finds the correct location of the unsynchronized records and then retrieves the desired records. This scheme uses the same computational and communication complexities of the existing multi server PIR. It uses the distributed source coding to achieve the privacy. The proposed method needs the number of unsynchronized records to be small. And it is the first multi server scheme that returns the desired record even when server's databases are not perfectly synchronized.

REFERENCES

[1] Giulia Fanti, Kannan Ramchandran, "Efficient Private Information Retrieval Over Unsynchronized Databases," Ieee Journal Of Selected Topics In Signal Processing, Vol. 9, No. 7, October 2015.

[2] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in Proc. IEEE FOCS, Milwaukee, WI, USA, 1995, pp. 41–50.

[3] C. Devet, I. Goldberg, and N. Heninger, "Optimally robust private information retrieval,"IACRCryptologye PrintArchive, vol.2012,p.83, 2012.

[4] G. Fanti and K. Ramchandran, "Efficient private information retrieval over unsynchronized databases," in Proc. Allerton, 2014.

[5] I. Goldberg, "Improving the robustness of private information retrieval," inProc. IEEESymp.Securityand Privacy, 2007,pp.131–148.

[6] D. Gross, "Yahoo Hacked, 450,000 Passwords Posted Online," CNN Tech., Retrieved from [Online]. Available: http://www.cnn.com/2012/ 07/12/tech/web/yahoo-users-hacked

[7] R.Sion and B. Carbunar,"On the computational practicality of private information retrieval," in Proc. NDSS, 2007, pp. 2006–2016.

[8] A. Beimel, Y. Ishai, and E. Kushilevitz, "General constructions for information-theoretic private information retrieval," J. Comput. Syst. Sci., vol. 71, no. 2, pp. 213–247, 2005.

[9] M. Barbaro and T. Zeller, A Face is Exposed for AOL Searcher no. 4417749 NewYorkTimes, Aug.2006 [Online]. Available:http:// www.nytimes.com/ 2006/08/09/technology/09aol.html

[10] A. Beimel and Y. Stahl, "Robust information-theoretic private information retrieval," in Proc. Security Commun. Netw., 2003, pp. 326–341, Springer.

AUTHORS

Meenu Poulose is currently pursuing M.Tech in Cyber Security in MBITS, Nellimattom. She completed her B.Tech from MG University College of Engineering, Muttom, Kerala, India. Her area of specialization is Cyber Security.

Tinku Soman Jacob is currently working as the Assistant Professor in Department of Computer Science and Engineering at MBITS, Nellimattom, Kerala, India. He received his B.Tech Degree in Computer Science and Engineering from Mar Athanasius College of Engineering Kothamangalam and M.E in Software Engineering from Sree Krishna College of Engineering and Technology, Coimbatore. His area of specialization is Software Engineering.

SHARE MARKET MANAGEMENT SYSTEM BASED KEYWORD QUERY PROCESSING ON XML DATA

Darsana C.S.[1], Roshni P.[2], Chandini K.[3] and Surekha Mariam Varghese[4]

Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala

ABSTRACT

Nowadays, many research has been focused on XML database systems. In share market, the value of a share changes over time. For managing market values of a share, traditional databases are commonly used. Main such example is relational database. But use of such databases are not efficient since it cannot handle more complex applications and do not have the ability to be scalable. In this paper we implement a Share Market Management System using an XML database. The main function of Share Market Management System (SMMS) is to store, retrieve and update the information about share market values. Keyword-based query solution is implemented to obtain the results. Then its performance is evaluated by considering the compilation, parsing and execution metric as total time.

KEYWORDS

XML Data; Native XML Database; XQuery; Keyword query;

1. INTRODUCTION

Stock market is the market in which shares are issued and traded, which is also known as the equity or share market. It is one of the most vital areas of a market economy.The stock market is the important source for companies to raise money. This is where businesses are publicly traded. With the development of online systems, online share marketing has become a trend. These sites provide current share market value details.

XML is a standard for storing and exchanging information over Internet. XML documents contain both the data and the informative relationship structuring of that data in a way that both machines and people can read. All its information can be send from one party to other. So they are self describing. The proposed system use XML database for storing share market information as documents. The information include equity name, open and closing value, rating and offer price. The data can be queried, transformed and exported. XML is a good method to handle sparse data.

In traditional share-market applications, details of equities are usually kept in relational database. Even-though they provides flat storage and retrieval, they possess many problems when dealing with redundant and sparse. The details include equity name, open, high and low values, number of shares traded, total turnover. Sometimes some features may be absent. In a relational database

DOI: 10.5121/ijci.2016.5227

store, values are also stored redundantly. This can cause null values in database and a lot of space wastage. To overcome these problems, we use an XML database system named BaseX. BaseX is a native and light-weight XML database management system which is specialized in storing, visualizing and querying large XML documents.

A brief description of XML database is presented in section 2. Section 3 explains BaseX, a tool to process XML documents. In section 4 the proposed method is explained. Performance analysis of this method is evaluated in Section 5. The paper is concluded in section 6

2. XML DATABASE

An XML database is a data storage software system. It stores data in eXtendedMarkup Language format. These data can then be queried, transformed, exported and returned to a caller. XML databases are document-oriented databases which falls in the category of NoSQL database. A NoSQL database provides a mechanism for storage and retrieval of data in a form other than tabular representation. XML - enabled and Native XML are the main two type of XML database. Attractive feature of XML is it's simplicity. Information coded in XML is easy to read and understand. Computers process it easily. XML is in W3C standard and is extensible, because there is no fixed set of tags. Based on our requirement we can create and use new tags. XML is a self describing language. XML documents can be stored without schemas and they contain meta data. Any XML tag can possess an unlimited number of attributes such as author or version. XML contains machine-readable context information. Tags, attributes and element structure provide context information. It opens up new possibilities for highly efficient search engines, intelligent data mining, agents, etc. XML can embed multiple data types.XML documents can contain any possible data type — from multimedia data (image, sound, video) to active components (Java applets, ActiveX). It is easy to map existing data structures like relational databases or file systems to XML. Null values and multiple values are managed efficiently by XML language. So XML is a good solution for dealing with websites having semi-structure.

2.1. NATIVE XML DATABASE

Native XML database is based on the container format, not in table format. It stores large amount of XML data and document. Native XML database is usually queried by the XPath-expressions. Native XML database has advantage over the XML-enabled database. It is highly capable to store, query and maintain the XML document than XML-enabled database.

A sample for an XML document is shown below. This is only a portion of document in the database we have considered("equity.xml")

```
<?xml version="1.0"?>
<share>
<security id="500510">
        <securityname>LT</securityname>
        <securitygroup>A</securitygroup>
        <open>1,474</open>
        <high>1,498</high>
        <low>1,400</low>
```

```
        <noshares>1,028,765</noshares>
        <turnover>14,840.39</turnover>
        <trades>41,101</trades>
</security>
        <security id="500285">
        <securityname>SPICEJET</securityname>
        <securitygroup>B</securitygroup>
        <open>43.8</open>
        <high>43.1</high>
        <low>46.75</low>
        <noshares>24,055,017</noshares>
        <turnover>10,932.52</turnover>
        <trades>44,857</trades>
        </security>
<security id="503283">
        <securityname>GEOJIT</securityname>
        <securitygroup>B</securitygroup>
        <open>36.40</open>
        <high>54.40</high>
        <low>43.40</low>
        <noshares>44,055,32</noshares>
        <turnover>42,265.74</turnover>
        <trades>44,745</trades>
        </security>
</share>
```

The data considered above is dynamically fetched from the stock exchange site and is transformed as XML Data.

3. BASEX

XML data handling is not an easy task even though it is highly readable. So tools are needed to make the XML content easy to explore, visualize and edit.

BaseX is a native and light-weight XML database management system and XQuery processor. BaseX is used for storing, querying, and visualizing large XML documents and collections. This platform independent DBMS uses standardized query languages such as XPath and XQuery. It also uses a tabular representation of XML tree structures to store XML documents. BaseX offers a Client-Server architecture to handle concurrent read and write operations of multiple users.

3.1. XQUERY

XQuery Processing is the important part of BaseX. This is fast and efficient. XQuery defines a query that transforms collection of unstructured and structured data in XML format. XQuery is an expression oriented programming language and contains a superset of Xpath. Xpath expression addresses specific portions of an XML document. It gives FLWOR expression. That is, FOR,

LET, ORDER BY, WHERE, RETURN. BaseX gives java bindings also.

Let us consider an XQuery for the above XML document

for $v in doc("equity.xml")/share/security
where $v/open>40
return data($v/securityname)

The result for the above XQuery will be,

LT
SPICEJET

The above is an example for a FLWOR expression in XQuery. The names of equities having open values higher than 40 are displayed in the result. This gives user the ability to quickly and easily know about securities or equities. The visualization of XML database is shown in the figure Fig 1

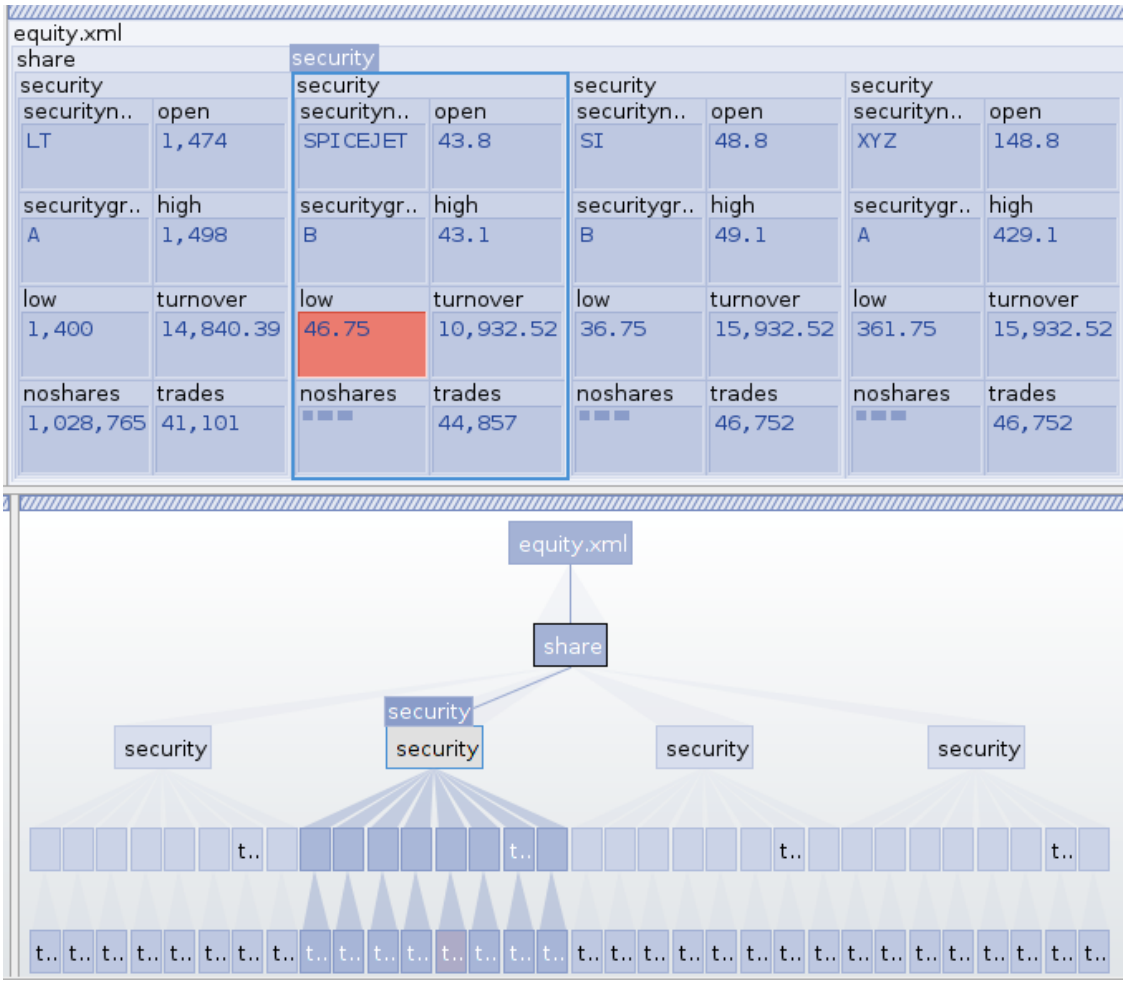


Figure 1. Map and Tree Visualization of XML data

4. THE METHOD

In this paper, we are concentrating on different operations such as storage, retrieval ,updation of shares. In Share Market Management System, the main operation is the retrieval of equity information based on some search queries. For example, if a person wants to know the securities whose open value is above some particular value, he can easily get that by giving the opening value as the search query. The interface accepts a keywords set from user, and then generates a FLWOR expression(Fig 2).[2]. The generated FLWOR expression is then sent to BaseX (Native XML Database). XQuery engine performs query on XML database, then sends user the result (Fig 3).

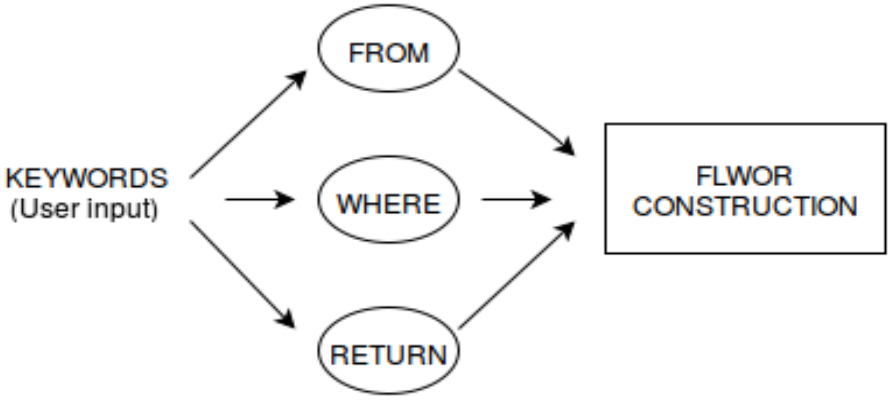


Figure 2. FLWOR Construction for XMLQuery

We have implemented it using the BaseX and Eclipse IDE. The XQuery is compiled and executed within the Eclipse IDE. The client version of the XQuery interpreter is used. The XQuery module is executed and the results are displayed in the window of the Eclipse.

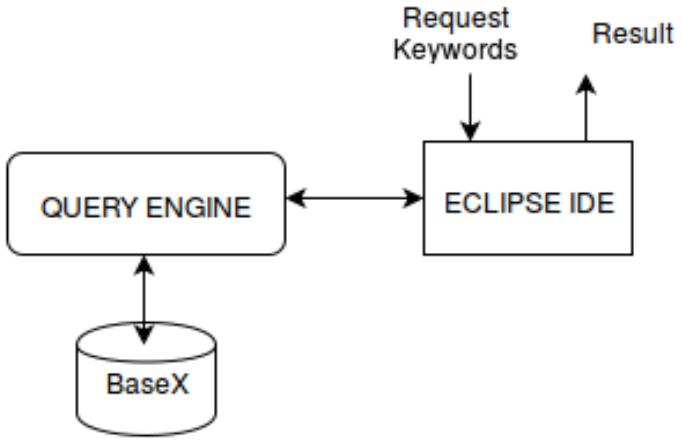


Figure 3. Query Processing

4.1. STEPS

The stock market values are dynamically fetched from the stock market site, transformed and stored in "equity.xml". The details include equity name, open, high and low values, number of shares traded, total turnover. Later, an XML database was created using BaseX tool. After creation, the tree structure view of the XML database was examined.

A client program was written to communicate with the BaseX database. The FLWOR (For, Let, Where, Order by and Return) query is constructed after accepting keywords from the user. The client program in Eclipse IDE (An integrated development environment) was connected to a running BaseX server instance, executing database commands and evaluating XQuery expressions.

5. PERFORMANCE EVALUATION

A series of tests were conducted to evaluate the performance by varying the number of inputs. The general conclusions drawn from these experiments are represented in graph. For simplicity, we have considered small datasets only. It is shown in the Figure.4

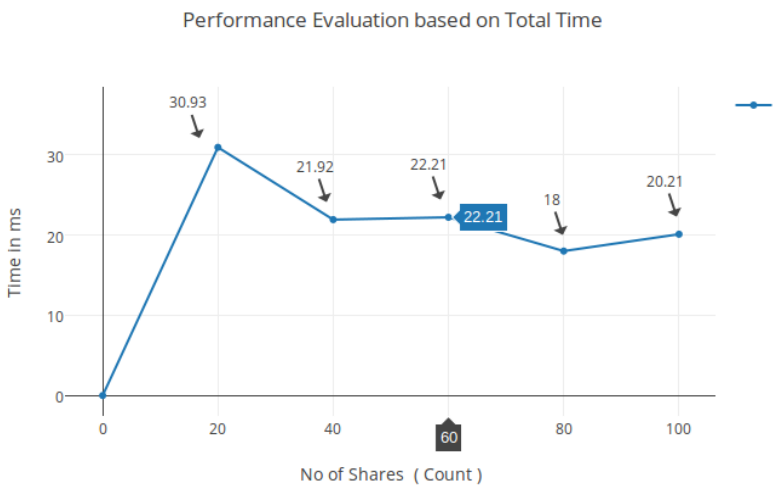


Figure 4. Performance Evaluation

6. CONCLUSION

The main function of Share Market Management System (SMMS) is to store, update and retrieve the information about share market values. Keyword-based query solution is implemented to obtain the results. The FLWOR query is constructed based on keywords from the user. The client program in Eclipse IDE was connected to a running BaseX server instance, executing database commands and evaluating XQuery expressions.

REFERENCES

[1] Zafari, H. ; Dept. of Comput. Eng., Islamic Azad Univ. Malayer Branch, Malayer, Iran ; Hasani, K. ; Shiri, M.E. XLight, An Efficient Relational Schema to Store and Query XML Data, Data Storage and Data Engineering (DSDE), 2010 International Conference.

[2] ZHuaJiang ; Coll. of Comput. Sci. & Technol., Wuhan Univ. of Technol., Wuhan, China ; Qing Yang A Keyword-Based Query Solution for Native XML Database, Internet Technology and Applications (iTAP), 2011 International Conference.

[3] Wahid, N. ; Dept. of Comput. Sci. & Eng., La Trobe Univ.,Melbourne, VIC, Australia ; Pardede, E. Single Transition Constraint for XML Update Validation , Network-Based Information Systems (NBiS), 2012 15th International Conference.

[4] Draxler, S.; Stevens, G.; Boden, A. Keeping the Development Environment Up to Date A Study of the Situated Practices of Appropriating the Eclipse IDE , Software Engineering, IEEE Transactions on, On page(s): 1061 - 1074 Volume: 40, Issue:11, Nov. 1 2014

[5] T. Grust and M. van Keulen.Tree awareness for relational DBMS Kernels: Staircase Join. In H. M. Blanken, T. Grabs,H.-J. Schek, R. Schenkel, and G. Weikum, editors. Intelligent Search on XML Data, volume 2818 of Lecture Notes in Computer Science, , pages 231-245. Springer, 2003

AUTHORS

Darsana CS. is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from Government Engineering College,Waynad , Kerala. Her area of research is Database Transactions.

Roshni P is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from Federal Institute of Science and Tech. Angamaly Kerala. Her areas of research are Database and Big data Analytics.

Chandini K is currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech from LBS Engineering College ,Kasargod Kerala. Her areas of research are Machine Learning and Image Processing.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from College of Engineering, Trivandrum affiliated to Kerala University and M-Tech in Computer and Information Sciences from Cochin University of Science and Technology, Kochi in 1996. She obtained Ph.D in Computer Security from Cochin University of Science and Technology, Kochi in 2009. She has around 25 years of teaching and research experience in various institutions in India. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems, Machine Learning and Distributed Computing. She has published 17 papers in international journals and international conference proceedings. She has been in the chair for many international conferences and journals.

INTENTIONAL BLANK

XOR-BASEDVISUAL CRYPTOGRAPHY

Nidhin Soman[1] and Smruthy Baby[2]

[1] Student,Department of Computer Science And Engineering, MBITS, Nellimattom,
[2]Assistant Professor, Department of Computer Science And Engineering, MBITS, Nellimattom

ABSTRACT

Visual cryptography scheme is a cryptographic technique which allows visual information.It is solve the poor visual quality problem. XOR-based VC. Actually, two XOR-based VC algorithms are proposed, namely XOR-based VC for general access structure (GAS) and adaptive region incrementing XOR-based VC. to be encrypted in such a way that the decryption can be performed by the human visual system,without the help of computers. There are diverse visual cryptography schemes developed based on different factors like pixel expansion, meaningless or meaningful shares, contrast, security, type of secret image and the number of secret images encrypted. This paper discusses most of the visual cryptography schemes and the performance measures used to evaluate them

KEYWORDS

Visual cryptography, random grid, pixel expansion, extended visual cryptography,XOR general access structure, region incrementing, pixel expansion, adaptive security level

1. INTRODUCTION

The world today relies on the internet for information storage, transmission and retrieval. Because of these huge amount of multimedia information is transmitted over the internet. For example, various confidential data such as military maps and commertial identifications are transmitted over the internet. While using secret images security issues should be taken into consideration because hackers may utilize weak link over communication network to steal information that they want. To deal with security problems of secret images, various image secret sharing schemes have been developed.

Visual cryptography is introduced by first in 1994 Noar and Shamir. Visual cryptography is a cryptographic technique which allows visual information to be encrypted in such a way that the decryption can be performed by the human visual system, without the aid of computers. Visual cryptography scheme eliminates complex computation problem in decryption process, and the secret images can be restored by stacking operation. This property makes visual cryptography especially useful for the low computation load requirement.

Visual cryptography is a powerful visual secret sharing scheme in which a secret image is distributed among some participants by dividing the secret image into two or more noise-like shares (or shadow images). When the shares on transparencies are stacked (superimposed) together, the original secret image will be revealed without any mechanical devices like a computer. Decryption can be done using the Human Visual System The process of visual

DOI: 10.5121/ijci.2016.5228

cryptography proposed by Naor and Shamir discusses a technique for encrypting a binary secret image into *n* shares (printed on transparencies), where each pixel is expanded *m* times. Each participant will get a share image but the secret image cannot be revealed with any one share. Any n participants can compute the original secret when any *k* (or more) of them are stacked together. No group of *k-1* (or fewer) participants can compute the original secret.

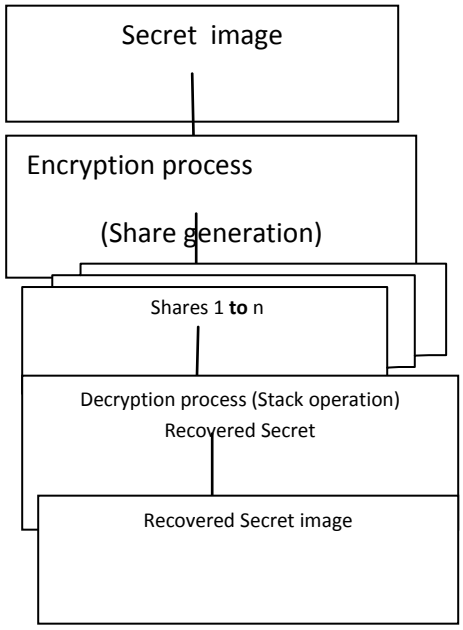


Figure 1: Basic flowchart of Visual Cryptography

The secret image cannot be seen from one transparency, but when *k* or more transparencies are stacked together the image will begin to emerge as the contrast between the black and white pixels becomes sufficient that the human.Initially the secret image is encoded (i.e. shares are generated) and during decoding the k or n shares are stacked together (according to the *(k, n)* or *(n, n)* scheme discussed later) to reveal the secret image. The secret image will get visible to the human visual system. In the *(k, n)* visual cryptography scheme, two collections of *(n* x *m)* Boolean matrices (Basis matrices), C0 and C1 are used. To share a white (black) pixel, the dealer randomly selects one row of the Boolean matrix C0 (C1) and assigns it to the corresponding share image. The gray level and contrast of the m sub-pixels in each of the n share images is defined by the chosen row. The major drawbacks of visual cryptography include pixel Expansion.

2. TRADITIONAL VISUAL CRYPTOGRAPHY

Secret is something which is kept from the knowledge of any but the initiated or privileged. Secret sharing defines a method by which a secret is distributed among a group of Secret pixel Share 1 Share 2 Stacked (OR operation)participants, whereby each participant is allocated a piece of the secret. This piece of the secret is known as a *share*. The secret can only be reconstructed when a sufficient number of shares are combined together. While these shares are separate, no information about the secret can be accessed. That is, the shares are completely useless while they are separated. Within a secret sharing scheme, the secret is divided into a number of shares and

distributed among n persons. When any *k* or more of these persons (where k ≤ n) bring their shares together, the secret can be recovered. However, if *k -1* persons attempt to reconstruct the secret, they will fail.

Due to this threshold scheme, we typically refer to such a secret sharing system as a *(k, n)*-threshold scheme or k-out of- n secret sharing, where n is the number of Total Participant and k is the number of Qualified Participant The basic model for visual sharing of the k out of n secret image. *A (k, n)* VSS scheme is a method by which the shared image (printed text, handwritten notes, pictures, etc.) is visible by k or more participants by stacking their transparencies with the help of an overhead projector. To share a white pixel, the dealer randomly chooses one of the matrices in C0 and to share a black pixel, the dealer randomly chooses one of the matrices in C1. The chosen matrix defines the colour of the *m* sub-pixels in each one of the *n* transparencies. The major drawback is the pixel expansion and low contrast of the reconstructed secret image

3. EXTENDED VISUAL CRYPTOGRAPHY

In visual cryptography, it is also obvious that, while the shares appear to be random (and, in fact, can be shown to contain no informational content that can be used to recover the original secret image on their own), the shares also have no interesting content that could be used to carry other information (such as a biometric image) that might be helpful in a security context. For example, if a share image could be selected to be the finger-print of the share holder, this could be useful in authenticating a user's right to hold that share when the parties meet to combine their share images to reveal the secret. In 1996, Ateniese, Blundo, and Stinson proposed extended visual cryptography (EVC) schemes that can construct meaningful share images. More security is provided for the shares as a cover image is provided for it. For example, if one of the shares of a finger print is covered by another person's finger print then the outsiders may think that the covered share is the original secret image of the finger print

Extended Visual Cryptography (EVC) takes the idea of visual cryptography further by creating shares which are meaningful to anyone who views them. This helps to alleviate suspicion that any encryption has taken place and also presents visually pleasing shares which incorporate all the previously mentioned features of VC. It allows the construction of visual secret sharing schemes within which the shares are meaningful as opposed to having random noise on the shares. After the sets of shares are superimposed, this meaningful information disappears andthe secret is recovered. Meaningful shares avoid attention of hacker considering the security issues over the communication channels. EVCS can also be viewed as a method of steganography. One scenario of the applications of EVCS is to evade the custom inspections, because the shares of EVCS are meaningful images, hence there are fewer chances for the shares to be suspected and detected. In case of EVCS, shares were simply generated by replacing the white and black sub-pixels in a traditional VCS share with transparent pixels and pixels from the cover images, respectively. This scheme provides meaningful share images but endure pixel expansion problem

3.1 RANDOM GRIDS BASED VISUAL CRYPTOGRAPHY

Random grid (RG) is a method to implement visual cryptography (VC) without pixel expansion. RG is defined as a transparency comprising a two-dimensional array of pixels, where each pixel can be fully transparent (white) or totally opaque (black), and the choice between the alternatives is made by a coin-flip procedure. Half of the pixels in a RG are white, and the *remaining pixels* are black. Encoding an image by random grids was introduced initially in 1987 by Kafri and Keren. A binary secret image is encoded into two noise-like transparencies with the same size of

the original secret image, and stacking of the two transparencies reveals the content of the secret. Comparing RGs with basis matrices, one of the major advantages is that the size of generated transparencies is unexpanded. The RG scheme is similar to the probabilistic model of the VC scheme, but the RG scheme is not based on the basis matrices

3.2 COLOUR VISUAL CRYPTOGRAPHY SCHEMES

Up to 1996, visual cryptography schemes were only applied to binary images. Rijmen and Preneel have introduced a visual cryptography scheme for colour images. In their scheme, each pixel of the colour secret image is expanded into a 2×2 block in order to generate two share images. Each 2×2 block on the share image is filled with red, green blue and white respectively, and thus no clue about the secret image can be recognized from any one of these two shares alone. Verheul and Van Tilborg introduced another method for encrypting a coloured image, called *c*-colour (*k*,*n*)-threshold scheme. In this scheme one pixel is expanded into m sub-pixels, and each sub-pixel is partitioned into c colour regions. In each sub-pixel, exactly one colour region will be coloured, and all the remaining colour regions are black. The colour of one pixel is based on the interrelations between colures of the stacked sub-pixels. For this coloured visual cryptography scheme with c colours, the pixel expansion m is c x 3.

Colour Decomposition: In this, every colour on a colour image can be decomposed into three primary colours: C, M,Y (if subtractive model is used) or R, G, B(if additive models used). This method expand every pixel of a colours secret image into a 2 x 2 block in the sharing images and keep two coloured and two transparent pixels in the block

3.3 PROGRESSIVE VISUAL CRYPTOGRAPHY

Progressive Visual Cryptography takes into consideration the premise of perfect secret recovery and high quality secret reconstruction. Many of the schemes do require computational effort in order to perfectly reconstruct the secret. A new sharing concept emerged known as “Progressive Visual Cryptography” which revealed the secret image progressively as more and more number of shares were stacked together.

3.4 REGION INCREMENTING VISUAL CRYPTOGRAPHY

In traditional visual cryptography scheme, one whole image is considered as a single secret and same encoding rule is applied for all pixels of one image. So it reveals either entire image or nothing. It may be the situation that different regions in one image can have different secrecy levels, so we can't apply same encoding rule to all pixels. Ran-Zan Wang developed a scheme Region Incrementing Visual cryptography for sharing visual secrets of multiple secrecy level in a single image . In this scheme, different regions are made of a single image, based on secrecy level and different encoding rules are applied to these regions

3.5 SEGMENT BASED VISUAL CRYPTOGRAPHY SCHEME

Traditional visual cryptography schemes were based on pixels in the input image. The limitation of pixel based visual cryptography scheme is loss in contrast of the reconstructed image, which is directly proportional to pixel expansion. Bernd Borchert proposed a new scheme which is not

pixel-based but segment-based. It is useful to encrypt *messages* consisting of symbols represented by a segment display. For example, the decimal digits 0, 1, 9 can be represented by seven-segment display. The advantage of the segment based encryption is that, it may be easier to adjust the secret images and the symbols are potentially easier to realize for the human eye and it may be easier for a no expert human user of an encryption system to understand the working. The secret, usually in the form of digits is coded into seven segment display before encrypted. Two random share images will be generated during encryption. Decryption process involves the stacking of these two share images

3.6 DYNAMIC VISUAL CRYPTOGRAPHY

The core idea behind dynamic visual cryptography is increasing the overall capacity of a visual cryptography scheme. This means that using a set of two or more shares,we can potentially hide two or more secrets. Multiple secret sharing is very useful when it comes to hiding more than one piece of information within a set of shares.

4. XOR VISUAL CRYPTOGRAPHY

A (k, n) visual cryptographic scheme encrypts a secret image into n share images (printed on transparencies)distributed among n participants. When any k participants stack their shares on an overhead projector (OR operation),the secret image can be visually discovered by a human visual system without the aid of computers (computation).But the monotone property of OR operation reduces the visual quality of reconstructed secret image for OR-based VCS. Generally all the conventional visual cryptography schemes (VCS) uses OR operation for stacking operations and so it is also called OR-based VCS. But it offers a poor visual quality image during decoding (stacking). Major advantage of XOR-based VCS (XVCS), is that since it uses XOR operation for decoding which results into exact recovery of the secret

4.1PROBABILISTIC VISUAL CRYPTOGRAPHY SCHEMES

In this scheme, usually there is no pixel expansion, i.e., m is The reconstruction of the image however is probabilistic, meaning that a secret pixel will be properly reconstructed only with a certain probability. On the other hand, in the deterministic model the reconstruction of an approximation

5. GENERAL ACCESS STRUCTURES (GAS)

In *(k, n)* scheme, using any of the 'k' shares someone can decode the secret image which in turn reduces security . To overcome this issue the basic model is extended to general access Structures by G. Ateniese, C. Blundo, A. De Santis, and D. R. Stinson [, where an access structure is a specification of all qualified and forbidden subsets of 'n' shares. Any subset of 'k' or more qualified shares can decrypt the secret image but no information can be obtained by stacking lesser number of qualified shares or by stacking disqualified share[1].

5.1THE PROPOSED ALGORITHM FOR GAS

Comparing to conventional VC, advanced properties such as good resolution, contrast and colour are provided by XOR-based VC at the expense of utilizing light-weight computational devices. Nowadays, light weight devices such as cell phones and smart devices are popular. XOR-based VC is possible to be widely used in the future. Some state-of-the-art works on XOR-based VC ,, are confined to threshold cases. Designing VC method for GAS becomes more necessary. An access structure, denote as *(TQual,TForb)*, is required in the proposed algorithm. When an access structure is given ,the basis *T*0 can be obtained. We construct the XOR-based VC for GAS based on the basis *T0*. Diagram of the proposed XOR-based VC for GAS is depicted in FIG2:The share generation is a pixel-wise operation, and *n* shared pixels are constructed via the proposed algorithm for every given secret pixel. Simply, the proposed algorithm consists of two components: the generation of *t* pixels and the construction of remaining $n - t$ pixels

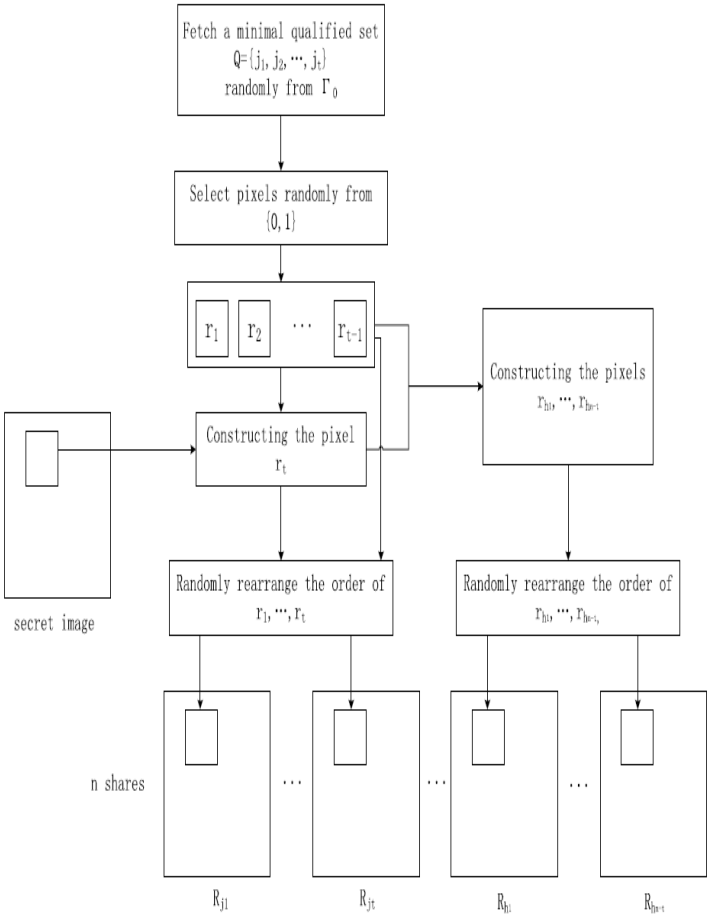


Figure 2.Diagram of the xor based vc Gas

Input: a binary secret image S with $M \times N$ pixels, and an access structure $(\Gamma_{Qual}, \Gamma_{Forb})$.
Output: n shares $R_1, \ldots, R_n$.
Step 1: Obtaining the basis Γ_0 of the access structure $(\Gamma_{Qual}, \Gamma_{Forb})$. Denote L as the number of minimal qualified sets in the basis Γ_0.
Step 2: For each position (i, j) in the secret image, n shared pixels $R_1(i, j), \ldots, R_n(i, j)$ are generated by Steps 3-7.

Step 3: Randomly select a minimal qualified set $Q = \{j_1, \ldots, j_t\}$ in the basis Γ_0.
Step 4: Construct $t - 1$ pixels $r_1, \ldots, r_{t-1}$ by

$$\begin{cases} r_1 = Random(\bullet) \\ \cdots \\ r_{t-1} = Random(\bullet) \end{cases} \tag{3}$$

where procedure *Random* return a value randomly chosen from {0, 1}.
Step 5: Construct the tth pixel by

$$r_t = S(i, j) \oplus r_1 \oplus \cdots \oplus r_{t-1} \tag{4}$$

where symbol $\oplus$ denotes the Boolean XOR operation.
Step 6: Rearrange the order of $r_1, \ldots, r_t$ and assign the values of the rearranged pixels to $R_{j_1}(i, j), \ldots, R_{j_t}(i, j)$.
Step 7: The remaining $n - t$ pixels are constructed by

$$\begin{cases} r_{h_1} = S(i, j) \oplus r_1 \oplus \cdots \oplus r_{t-1} \oplus r_t \\ r_{h_2} = S(i, j) \oplus r_1 \oplus \cdots \oplus r_{t-1} \oplus r_t \oplus r_{h_1} \\ \cdots \\ r_{h_{n-t}} = S(i, j) \oplus r_1 \oplus \cdots \oplus r_{t-1} \oplus r_t \oplus r_{h_1} \oplus \cdots \oplus r_{h_{n-t-1}} \end{cases} \tag{5}$$

where $h_1, \ldots, h_{n-t}$ are the $n - t$ indices in $\{1, \ldots, n\} - \{j_1, \ldots, j_t\}$. Pixels $r_{h_1}, \ldots, r_{h_{n-t}}$ are assigned to the shared pixels $R_{h_1}(i, j), \ldots, R_{h_{n-t}}(i, j)$, respectively.
Step 8: Output the n shares $R_1, \ldots, R_n$.

Figure 3. XOR-based VC for General Access Structure.

5.2 ADAPTIVE REGION INCREMENTING XOR-BASED VC

In the proposed method, the concept of *adaptive security level* is introduced. In the previous methods the security levels in the secret image are revealed in accordance with the quantity of stacked shares. Every share is with the same priority. For the *adaptive security level*, the security levels are reconstructed according to the qualified set in which the members can be specified by the sharing strategy. Different shares are with different priorities. XORbased VC with adaptive security *level* property is named as adaptive region incrementing XOR-based VC.

To construct the adaptive region incrementing XOR-based VC, algorithm proposed in the former section is adopted. For each minimal qualified set Q, Q is assigned an initial security level. Based on the initial security levels of the minimal qualified sets, security levels of the remaining qualified sets are calculated by Security Level Assignment algorithm . When the assignment finishes, shares of the adaptive region incrementing XOR-based VC are constructed. Usually, the XOR-ed result by shares in a qualified set reveals the associated security levels while the XOR-ed result by shares in a forbidden set cannot.

Herein, we describe the generation of shares for the adaptive region incrementing XOR-based VC. Diagram of this method is depicted in Fig. 4. A binary secret image with k security levels $L1$, . . . , Lk are considered as input, as well as minimal qualified sets in t_0 assigned with initial security levels. First of all, the remaining qualified sets, which are not assigned the initial security levels, are automatically given the associated security levels by the assignment algorithm. When the security level assignment completes, the share construction begins. The construction is an approach derived from the XOR-based VC for GAS. Similarly, it comprises two parts: the construction of t pixels and the generation of the remaining $n - t$ pixels.

For each time, a pixel s is constructed based on the security level of given secret pixel. A qualified set, which contains t participants, is randomly chosen from the set of qualified sets.
At the next step, $t - 1$ shared pixels are constructed randomly, and the tth shared pixel is generated based on the $t - 1$ random pixels and pixel s. The remaining $n - t$ shared pixels are iteratively constructed by pixel s and the former shared pixels that have been assigned values. Similarly, the iterative generation of the $n - t$ shared pixels helps enhancing the visual quality. Herein, the security level assignment algorithm is given as follows.

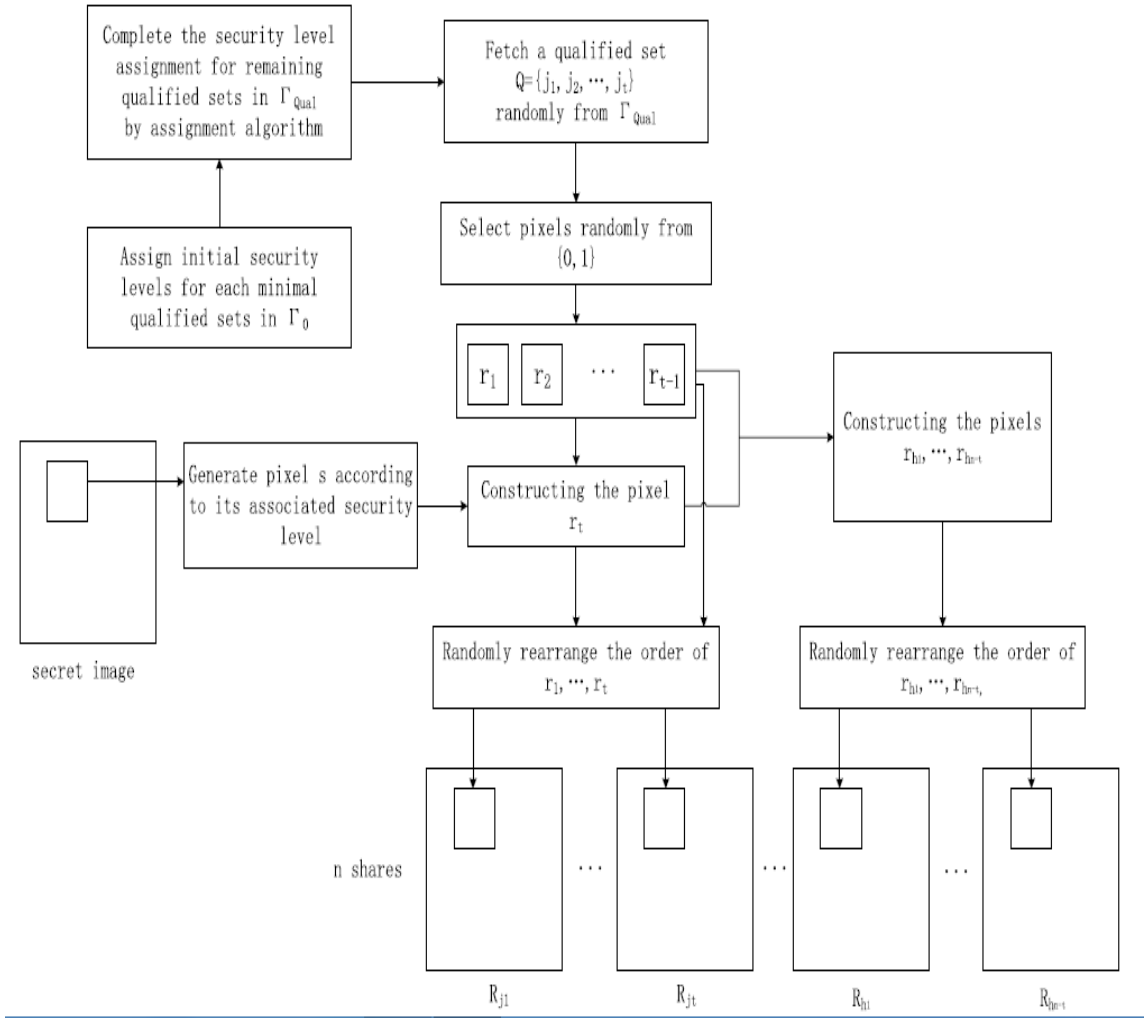


Figure. 4. Diagram of the adaptive region incrementing XOR-based VC.

Input: An access structure $(\Gamma_{Qual}, \Gamma_{Forb})$ whose minimal qualified sets are with initial security levels.
Output: Qualified sets in Γ_{Qual} with assigned security levels.
Step 1: For each qualified set $Q \in \Gamma_{Qual}$ which is not assigned the security level, obtain qualified sets $A_1, \ldots, A_m \subset Q$.
Step 2: If all the qualified set $A_1, \ldots, A_m$ are with assigned security levels, get the highest security level L_t $(1 \leq t \leq k)$ from $A_1, \ldots, A_m$. If not, the assignment for Q is processed later until each qualified set in $A_1, \ldots, A_m$ is assigned a security level.
Step 3: Determine the security level L of Q by

$$L = \begin{cases} L_{t+1}, & \text{if } t < k, \\ L_t, & \text{otherwise.} \end{cases}$$

Step 4: Generate a pixel s according to L_d by

$$s = \begin{cases} 0, & \text{if } S(i, j) \in L_0 \text{ or } y > d, \\ 1, & \text{otherwise.} \end{cases} \quad (8)$$

Step 5: Construct $t - 1$ pixels $r_1, \ldots, r_{t-1}$ by

$$\begin{cases} r_1 = Random(\bullet) \\ \cdots \\ r_{t-1} = Random(\bullet) \end{cases} \quad (9)$$

where procedure *Random* return a value randomly chosen from {0, 1}.
Step 6: Construct the tth pixel by

$$r_t = s \oplus r_1 \oplus \cdots \oplus r_{t-1} \quad (10)$$

where symbol $\oplus$ denotes the Boolean XOR operation.
Step 7: Rearrange the order of $r_1, \ldots, r_t$ and assign the values of the rearranged pixels to $R_{j_1}(i, j), \ldots, R_{j_t}(i, j)$.
Step 8: The remaining $n - t$ pixels are constructed by

$$\begin{cases} r_{h_1} = s \oplus r_1 \oplus \cdots \oplus r_{t-1} \oplus r_t \\ r_{h_2} = s \oplus r_1 \oplus \cdots \oplus r_{t-1} \oplus r_t \oplus r_{h_1} \\ \cdots \\ r_{h_{n-t}} = s \oplus r_1 \oplus \cdots \oplus r_{t-1} \oplus r_t \oplus r_{h_1} \oplus \cdots \oplus r_{h_{n-t-1}} \end{cases} \quad (11)$$

where $h_1, \ldots, h_{n-t}$ are the $n - t$ indices in $\{1, \ldots, n\} - \{j_1, \ldots, j_t\}$. Pixels $r_{h_1}, \ldots, r_{h_{n-t}}$ are assigned to the shared pixels $R_{h_1}(i, j), \ldots, R_{h_{n-t}}(i, j)$, respectively.
Step 8: Output the n shares $R_1, \ldots, R_n$.

Figure 5.Security Level Assignment

6. RELATED WORK

In this paper XOR based cryptography is applied to both text and video.some of the Data hiding algorithms can be applied to secure the image This paper has applications in military, bank etc the construction of the pixel is based on the random value..This is a new technique.For example,

we consider one department head and two department researchers for sharing a secret with two security levels. The department head has the prime priority that he can reveal the whole secret by using his share and one of the two shares held by the researchers. But the two researchers only can recover the first security level in the secret by their two shares. information hiding algorithm can be used for the Security. There could be a chance to get man in middle attack, then shares can be protected using a key. This key can be automatically generated. This key can be protected by another encryption and decryption .In this shares some of the texts are embedded and are hidden and its implementation is very easy.

7. PERFORMANCE ANALYSIS

There are various parameters used to evaluate the performance of visual cryptography scheme.
Pixel expansion- Pixel expansion m refers to the number of sub-pixels in the generated shares that represents a pixel of the original secret image. It represents the loss in resolution from the original secret image to the shared one.
Contrast- Contrast is the relative difference in weight between combined shares that come from a white pixel and a black pixel in the original image. Contrast of the recovered secret image must be adjusted so that it is visible to the human eye.

Security- Security is satisfied when each share individually discloses no information of the original image and the original image cannot be reconstructed with shares fewer than k in(k, n) scheme.

Accuracy- Accuracy is measured to be the quality of the reconstructed secret image and evaluated by peak signal-to-noise ratio (PSNR). Mean Squared Error (MSE) can also be used for accuracy evaluation

8. CONCLUSIONS

Visual cryptography offers perfect security for all the digitally transmitted secret images. This paper discusses various visual cryptography schemes and commonly used performance evaluation parameters. Diverse visual cryptography schemes were developed based on different factors like pixel expansion, meaningless or meaningful shares, contrast, security, type of secret image (either binary or colour) and the number of secret images encrypted. In this paper, we further exploit the extended capabilities for XOR-based VC. In essence, two XOR-based VC algorithms, namely XOR-based VC for GAS and adaptive region incrementing XOR-based VC, are introduced. For the first method, complicated sharing strategy by using GAS can be implemented

REFERENCES

[1] Xiaotian Wu and WeiSun (2015)Extended Capabilities for XOR-Based Visual Cryptography.
[2] G. Ateniese, C. Blundo, A. DeSantis, and D. R. Stinson, “Visualcryptography for general access tructures”, Proc.ICAL96,Springer, Berlin, 1996, pp.416-428.
[3] T. Chen and K. Tsao, “User-friendly random-grid-based visual secret sharing”, IEEE Trans. Circuits Syst. Video Technol., vol.21, no. 11, pp. 1693_1703, Nov. 2011
[4 P. Tuyls, H. Hollmann, J. Lint, and L. Tolhuizen, “XOR-based visual cryptography schemes”, Designs, Codes, Cryptography,vol. 37, no. 1, pp. 169_186, 2005.

[5] C. Blundo, A. De Santis, and M. Naor, "Visual cryptography for grey level images," Inf. Process. Lett., vol. 78, no. 6, pp. 255–259, Nov. 2000
[6] C. Blundo and A. De Santis, "Visual cryptography schemes with perfect reconstruction of black pixels," Comput. Graph., vol. 22, no. 4,pp. 449–455, Aug. 1998.
[7] C.-N. Yang, "New visual secret sharing schemes using probabilistic method," Pattern Recognit. Lett., vol. 25, no. 4, pp. 486–494, Mar. 2004.
[8] X. Wu and W. Sun, "Random grid-based visual secret sharing for general access structures with cheat-preventing ability," J. Syst. Softw., vol. 85,no. 5, pp. 1119–1134, May 2011
[9] S. J. Shyu, "Visual cryptograms of random grids for general access structures," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 3,pp. 414–424, Mar. 2013
[10] C.-N. Yang, H.-W. Shih, C.-C. Wu, and L. Harn, "k out of n region incrementing scheme in visual cryptography," IEEE Trans.Circuits Syst. V
[11] C.-N. Yang and D.-S. Wang, "Property analysis of XOR-based visual cryptography," IEEE Trans. Circuits Syst. Video Technol., vol. 24, no. 2,pp. 189–197, Feb. 2014
[12] M. Naor and A. Shamir, "Visual cryptography," in Advances in Cryptology (Lecture Notes in Computer Science), vol. 950. Berlin, Germany:Springer-Verlag, 1995, pp. 1–12.
[13] F. Liu and C. Wu, "Embedded extended visual cryptography schemes,"IEEE Trans. Inf. Forensics Security, vol. 6, no. 2, pp. 307–322,Jun. 2011
[14] F. Liu, C. Wu, and X. Lin, "Step construction of visual cryptography schemes," IEEE Trans. Inf. Forensics Security, vol. 5, no. 1, pp. 27–38,Mar. 2010
[15] R.-Z. Wang, "Region incrementing visual cryptography," IEEE Signal Process. Lett., vol. 16, no. 8, pp. 659–662, Aug. 2009

AUTHOR

NIdhin Soman. is currently pursuing M.Tech in Computer Science and Engineering in Mar Baselios Institute of Technology And Science Nellimattom. He completed his B.Tech from,. Mar Baselios Institute of Technology And Science, Nellimattom. His areas of research are Network Security and Image Processing.

Smruthy Babyreceived B.TechDegree from Mahatma Gandhi University in 2008 and M.Techin computer science and engineering from Anna.University. Currently working as Assistant professor at Mar Baselios Institute of Technology And Science,.Her research interest include Network Security and Image Processing.

AGE CLASSIFICATION FROM FINGERPRINTS –WAVELET APPROACH

Ajitha T Abraham[1] and Asst. Prof. Yasim Khan M[2]

[1]Department of Electronics and Communication Engineering, College Of Engineering, Poonjar, Kottayam, Kerala

[2]Department of Electronics and Communication Engineering, College Of Engineering, Poonjar, Kottayam, Kerala

ABSTRACT

This research implements a novel and simple method of age classification using fingerprints. Two methods are combined for gender classification. The first method is the Singular Value Decomposition (SVD), employed to extract fingerprint characteristics by doing synthesis and reconstruction. The second method is the analysis for feature extraction by using 2D Bi-orthogonal Wavelet decomposition, up to 4 level decomposition used for the process of gender identification. This method is experimented with the internal database of 250 fingerprints finger prints in which 125 were male fingerprints and 125 were female fingerprints. Tested fingerprint is grouped into any one of the following five groups: 6-7, 8-12, 13-15, 16-19, 20-30, 30-50 and above 50. Overall classification rate of 60% has been achieved. Results of this analysis make this method a prime candidate to utilize in forensic anthropology for age classification in order to minimize the suspects search list by getting a likelihood value for the criminal gender.

KEYWORDS

Fingerprint, SVD, Wavelets,BWT

1. INTRODUCTION

Age of a person can be identified using different biometric traits such as face, iris, retina, speech, gait, hand geometry and fingerprint. Fingerprint is one of the most common traits of human and can be easily obtained. Now a days thumbprints and fingerprint of each finger are taken in order to provide the identity proof to that particular person, e.g. to get a passport or a unique identity card in India, one had to give the impression of his/her thumb and fingerprints. A person's fingerprint is permanent even before they are born.

Around 6-8 weeks after conception the volar pads (ball like structures that make up the contour of the fetal hand) form; by 10-12 weeks after conception the volar pads begin to recede; around the 13th week skin ridges appear and take the shape of the receding volar pad; lastly around the 21st week after conception the fingerprint patterns are complete[1].

A Fingerprint is the representation of the epidermis of a finger; it consists of a pattern of interleaved ridges and valleys. Fingertip ridges evolved over the years to allow humans to grasp and grip objects[1,2]. Like everything in the human body, fingerprint ridges form through a combination of genetic and environmental factors. This is the reason why even the fingerprint of identical twins is different [3].Fingerprint is an impression of friction ridges, from the surface of

DOI: 10.5121/ijci.2016.5229

the finger-tip. Fingerprints have been used for personal identification for many decades; more recently becoming automated due to advancements in the computing capabilities Fingerprints have some important characteristics that make them invaluable evidence in crime scene investigations:

1. A fingerprint is unique to a particular individual, and no two fingerprints possess exactly the same set of characteristics.
2. Fingerprints do not change over the course of person's lifetime (even after superficial injury to the fingers).
3. Fingerprint patterns can be classified, and those classifications then used to narrow the range of suspects .

In this work, age identification is mainly based on image synthesis and analysis. SVD is used for synthesis and BWT is used for analysis. Figure 1 illustrates the BWT and SVD based Age Classification system.

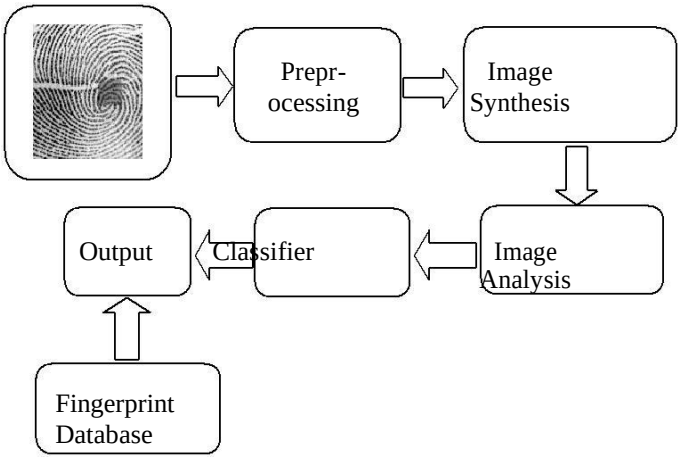


Fig. 1 Block Diagram of Age Classification System.

Features of fingerprints vary with sexes, ethnic groups and age categories. In this case the fingerprint is obtained from the Digital Persona Optical Fingerprint scanner. The paper is aimed in developing an algorithm for classifying the gender through fingerprint.

Wavelet transform is a popular tool in image processing and computer vision because of its complete theoretical framework, the great flexibility for choosing bases and the low computational complexity [10]. As wavelet features has been popularized by the research community for wide range of applications including fingerprint recognition, face recognition and gender identification using face, authors have confirmed the efficiency of the BWT approach [14]for the gender identification using fingerprint.

The SVD approach is selected for the age discrimination because of its good information packing characteristics and potential strengths in demonstrating results. The SVD method is considered as an information oriented technique since it uses principal components analysis procedures (PCA), a form of factor analysis, to concentrate information before examining the primary analytic issues of interest [13]. Threshold gives very strong consistent results. It uses the database which was generated in the learning stage of the proposed system and it classifies genders of the fingerprints.

The remainder of this paper is divided into 4 sections. They are as follows: Section II covers the literature review and comparison on the design research of previous systems. We find a lot of the information from Internet. Section III discusses on the proposed method. Section IV describes the performance evaluation of the developed system. Section V concludes the paper and out lines the contributions of the work. The limitations were highlighted and suggestions are made for further development to improve the system.

2. LITERATURE REVIEW

Human fingerprints have been discovered on a large number of archeological artifacts and histological items. Although these findings provide evidence to show that ancient

people were aware of the individuality of fingerprints, it was not until the late sixteenth century that the modern scientific fingerprint technique was first initiated (Jain, et al, 2003). In 1686, Marcello Malpighi, a professor of anatomy at the University of Bologna noted in his writings the presence of ridges, spirals and loops in fingerprints. Afterwards many studies have been conducted on fingerprints based on its patterns and features.

Fingerprint identification and classification has been extensively researched in times past, however very few researchers have studied the fingerprint gender classification problem. Age classification can be made using the spatial parameters or frequency domain parameters or using the combination of both. Most of the findings are based on the spatial domain analysis and few were based on the frequency domain. Earlier work on age classification based on the spatial domain analysis shows that the ridge thickness is different for each age classification groups. [1].

P.Gnanasivam, et al,[4] in 2012 presented a study.In this paper discrete wavelet transform (DWT) and the singular value decomposition (SVD) has been used to estimate a person's age using his/her fingerprint. The most robust

K nearest neighbor (KNN) used as a classifier. The evaluation of the system is carried on using internal database of 3570 fingerprints in which 1980 were male fingerprints and 1590 were female fingerprints. Tested fingerprint is grouped into any one of the following five groups: up to 12, 13-19, 20-25, 26-35 and 36 and above. By the proposed method, fingerprints were classified accurately by 96.67%, 71.75%, 86.26%, 76.39% and 53.14% in five groups respectively for male and by 66.67%, 63.64%, 76.77%, 72.41% and 16.79% for female. Finger-wise and Hand-wise results of age estimation also achieved.

Rijo Jackson Tom, et al, (2013) have proposed a method for Fingerprint Based Gender Classification through frequency domain analysis to estimate gender by analyzing fingerprints using 2D Discrete Wavelet Transforms (DWT) and Principal Component Analysis (PCA).A dataset of 400 persons of different age and gender is collected as internal database. They have used minimum distance method for classification and achieve overall success rate in age classification of around 60% [9].

Ajitha T.Abraham, Yasim Khan M,(2014) have proposed a method for Age classification from fingerprints through frequency domain analysis to classify gender by analyzing fingerprints using 2D Bi-orthogonal Wavelet Transform. A dataset of 250 persons of different age and gender is collected as internal database. They have used wavelet entropy as a classifier and formulated equation is used as threshold for easy classification and achieve overall success rate in gender classification of around 58%[5].

TABLE 1:COMPARISON

Sl. No	Author/Title	Method	Classification Method	Accuracy
1	P.Gnanasivam, Dr.S Muttan[6]2012	SDA & FDA	KNN Classifier method	70%
2	Rijo Jackson, T.Arulkumaran [11] 2013	FDA(2D DWT and PCA)	Minimum distance method	65%
3	Ajitha T.Abraham,Yasim Khan M[5]	FDA(2D BWT)	Threshold setting	58%

3. PROPOSED METHOD

3.1 PREPROCESSING

The actual size of obtained fingerprint is 310 X 420 no of pixels. We took the only center portion of a fingerprint pattern. So firstly the image was cropped by 300 X 350 no of pixels by using image cropping technique. Secondly it resized in to 256 X 256 no of pixels. Then the resized image undergoes enhancement technique like histogram equalization. Histogram equalization is a method in image processing of contrast adjustment using the image's histogram. This method usually increases the global contrast of many images, especially when the usable data of the image is represented by close contrast values. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to gain a higher contrast.

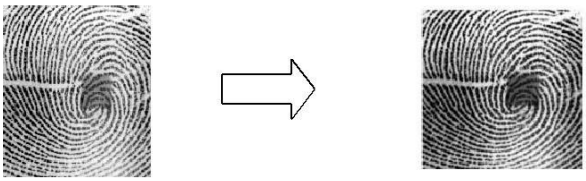


Fig. 2 Input image and Enhanced image

Enhancement techniques used on the fingerprints varies with the quality of the image and types of database used. Poor quality fingerprint image obtained is enhanced for better implementation of algorithm.

3.2 SINGULAR VALUE DECOMPOSITION(SVD)

Image synthesis is the process of creating new images from some form of image description. For synthesis here we used SVD decomposition. The Singular Value Decomposition (SVD) is an algebraic technique for factoring any rectangular matrix into the product of three other matrices[10]. The SVD is the factorization of any real matrix into three matrices, each of which has important properties. Any real m X n matrix *A* can be decomposed uniquely as

$$A = U D V^T \quad (1)$$

U is m X m and column orthogonal (its columns are eigenvectors of AA^T)ie

$$AA^T = U D V^T V D U^T = U D^2 U^T \quad (2)$$

V is n X n and orthogonal (its columns are eigenvectors of A^TA ie

$$A^T A = V D U^T U D V^T = V D^2 V^T \quad (3)$$

D is m X n diagonal (non-negative real values called singular values)

$$D=diag(\sigma_1,\sigma_2 ,\ldots\ldots,\sigma_n) \quad (4)$$

D is ordered so that $\sigma_1 \geq \sigma_2 \geq \ldots \sigma_n$ (σ is singular value of matrix *A*).

The rank of matrix *A* is equal to the number of its nonzero singular values. In many applications, In many applications, the singular values of a matrix decrease quickly with increasing rank. This property allows us to reduce the noise or compress the matrix data by eliminating the small singular values or the higher ranks.

3.3. SYNTHESIS AND RECONSTRUCTION

The generation of an image from a mathematical model rather than observation is known as image synthesis. For makes use of singular value decomposition (SVD) perturbation, which at first, applies the SVD decomposition on input image (I).

$$[U D V]=svd(I) \quad (5)$$

Here U and V are left and right odd vectors respectively, D is the diagonal matrix of particular values. SVD perturbation [13]uses these singular values to make the derived image (J).

$$J=U * D^i * V \quad (6)$$

where *i* varies between 1 and 2. Finally the derived image is combined with the original image.

$$C = \frac{I+(a * J)}{1+a} \quad (7)$$

where *a* is the combination parameter and it varies from 0 to 1.

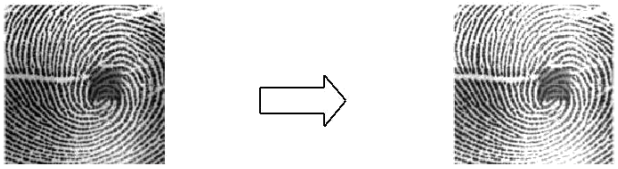


Fig. 3 Enhanced image and Reconstructed image

Individually, this step is not able to perform well under varying conditions. So finally, this paper makes use of wavelet transforms to handle those variations. Wavelet transforms decompose a face image into a number of coefficients that represent an image into different frequency sub bands.

3.4. ANALYSIS USING BI-ORTHOGONAL WAVELET TRANSFORM

Wavelets are developed from the Fourier transform to overcome the drawback of overall domain analysis for which wavelet uses a localized time and frequency analysis. Wavelet plays a vital role in image compression in the part of improving the signal strength. Hence wavelets are widely used in the field where the degradation is not tolerated. Wavelets can also effectively remove the noise in an image.

Wavelet transform is defined as the infinite set of various transforms. Which uses the function that are localized in both the real and Fourier space. Wavelet transform of any function *f* at frequency *a* and time *b* is computed by correlating *f* with wavelet atom as

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt \qquad (8)$$

Wavelet transform is always defined in terms of a " mother wavelet ψ" and a scaling function φ , along with their dilated and translated versions. Wavelet transform is defined as the infinite set of various transforms. Which uses the function that are localized in both the real and Fourier space.

A bi-orthogonal wavelet is one type of wavelet in which the associated transform is inversing but it is not necessary to be orthogonal. It gives freedom in designing bi-orthogonal wavelets than orthogonal wavelets. Additional freedom is the option to create symmetric wavelet function [10]. It compactly supports symmetric analyzing and synthesis wavelets and scaling functions. There is quite a bit of freedom in designing the bi-orthogonal wavelets, as there are no set steps in the design process [11]. It has a property of linear phase which is needed for image reconstruction. The properties can be derived by using two wavelets – Decomposition and Reconstruction instead of using a single wavelet[11].

Analysis (decomposition) and synthesis (reconstruction) filter orders for Biorthogonal filters Specify the order of the analysis and synthesis filter orders for Biorthogonal filter banks as 1.1, 1.3, 1.5, 2.2, 2.4, 2.6, 2.8, 3.1, 3.3, 3.5, 3.7, 3.9, 4.4, or 5.5, 6.8. [10] Unlike orthogonal wavelets, Biorthogonal wavelets require two different filters one for the analysis and other for synthesis of an input. The first number indicates the order of the synthesis filter while the second number indicates the order of the analysis filter. The default is 1/1. For the perfect reconstruction equation to hold, the scaling and the wavelet coefficients have to fulfil the following equations:

$$\tilde{g}(n)=(-$$
$$g(n)=(-$$

It is clear that when the analysis and the synthesis filters "orthogonality" condition in this case is defined by:

$$h(1-n) \quad (9)$$

$$\tilde{(1-n)} \quad (10)$$

are similar, the system becomes orthogonal. The

$$\sum_n \tilde{h}(n)h(n+2k)=\delta(k) \quad (11)$$

Depending upon the various performance factor, the efficiency of various wavelets is analyzed and it is instituted that bior4.4 has the greatest efficiency in compressing the fingerprint image.

3.5. SUBBAND DECOMPOSITION

There are two approaches to the subband decomposition of two – dimensional signals using two – dimensional filters, or using separable transforms that can be implemented using one – dimensional filters on the row first and then on the columns. Most approaches, use the second approach. Figure 6 shows how an image can be decomposed using subband decomposition. Of the four sub images, the one obtained by low – pass filtering the rows and columns is referred to as the LL image; the one obtained by low – pass filtering the rows and high – pass filtering the columns is referred to as the LH image; the one obtained by high – pass filtering the rows and low – pass filtering the columns is called the HL image; and the subimage obtained by high – pass filtering the rows and columns is referred to as the HH image [11]. Figure 4 demonstrate the subband decomposition of an NxM image.

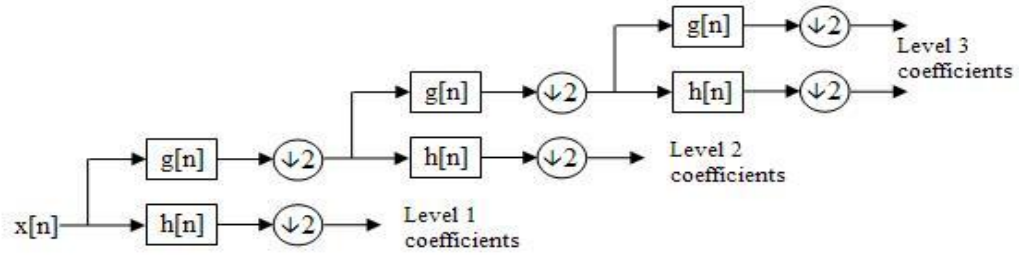


Fig.4 Subband decomposition of an N × M image

Each of these sub-bands represents different image properties. Most of the information's of the images is in the lower frequencies. So the further decomposition of sub band is repeated in LL sub band. For *k* level DWT, there *are (3*k) + 1* sub-bands available. Here we using 4 levels of decomposition[11].

3.6 Age Classification

Further processing here we using LL sub band only. By an experimental study we choose an image statistical property as a parameter. Here mode is selected as estimated parameter. The mode is the value that occurs most often. If no number is repeated, then there is no mode for the list. Mode is the most suitable statistical property used for age determination. Mode of the

2D matrices can be calculated by using following steps.

[1] Convert the given matrix to a column matrix.
[2] Calculate the minimum and maximum number of column matrix.
[3] Counting the number of times, each number is present between minimum and maximum number.
[4] Calculate which number is occurs most often.
[5] Mode = most occurred number + minimum of column matrix – 1.

By an experimental study we concluded that in each group of age classification most often number is different. By applying this method ,we can classified 7 different groups such as 6-7,8-12,13-15,16-19,20-30,31-50 and above 50.

4. RESULTS AND DISCUSSIONS

The algorithm of the proposed system is written in MATLAB R2014 and run in Intel Core 2 Duo, 2.20 GHz processor with 2.00 GB memory. Here, we proposed a new and simple method for Age Classification of fingerprints using BWT and SVD. In this section, the performance of the proposed Age estimation algorithm is verified by using the internal database. The success rate (in percentage) of age classification using the combination of both BWT and SVD are summarized and discussed. BWT level 5, 6, 7 and 8 were tried and from the results, BWT level 4 is identified as the optimum for the age estimation.

Mode of each age groups is different. Results after each steps were given in previous sections. The success rate is more than 60%. By proposing this new approach overcame all drawbacks of our previous work. 7 group of classification was obtained. Percentage of result after this study can be obtained is shown in below table.

In table 2 the success rates (in percentage) of age estimation for the fingerprints are tabulated. For the fingerprints of the persons whose age lies between 16-19 years, the success rate is achieved with 52%. The success rate in the age group of above 50 is reasonably good (60%) and thus useful for crime investigation, as this group crime rate is higher than other groups. Similarly the success rate for the remaining group is achieved as 56% for age groups 6-7 and 20-30, 53% for 8-12 and 13-15 and 55% for the age group of 30-50 respectively. Maximum success rate is achieved in the age group of ‘above 50’ for the left thumb. Low success rate is recognized for the age group of 16-19.

TABLE 2. AGE CLASSIFICATION ACCURACY

	Age Classification			
Sl. No	***Age Groups***	***Total fingerprints-60***	***Accuracy***	***Over all Accuracy***
1	6,7	34	56%	60%
2	8-12	32	53%	
3	13-15	32	53%	
4	16-19	31	52%	
5	20-30	34	56%	
6	30-50	33	55%	
7	Above 50	36	60%	

Age group-wise average success rate for fingerprints is shown in the line diagram of figure 6. Maximum success rate of 60% is achieved for the age group of ‘above 50’ years. For the age

group of ’16-19’, the success rate is low(52%).

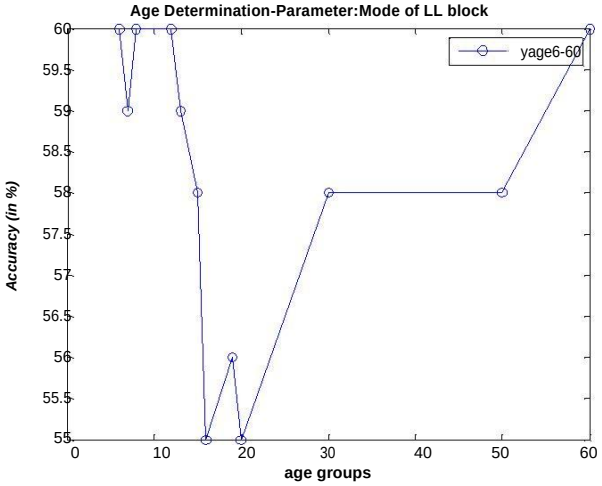


Fig.5 Age group-wise Average success rate

5. CONCLUSION

Here we proposed a new and simple method for age classification from fingerprint images based on Wavelet Transform and SVD technique. This method considered the frequency features of the wavelet domain .The LL block is selected for further processing for the age classification. Mode was choose as the parameter for age classification.

The proposed system is experimented only on the optical scanned image. Better result will be obtained for digital image. It was found that increasing the database population in each category improves the performance of the system. The further improvements hence planned to be done in conjunction with this are age determination, blood group determination and heredity checking.

ACKNOWLEDGMENTS

I would like to place my gratitude to all whose cooperation was vital for the success of this paper. I would like to acknowledge and extend my heartfelt gratitude to my guide , all staffs in ECE department and Dr.Minu K K, for their invariable suggestions and guidance that helped me to successfully complete this project. Even I would like to extend my gratitude to all for helping me in collection of fingerprints.

REFERENCES

[7] W.Babler, "Embryologic development of epidermal ridges and their configurations", Science Transition

[8] J.John,Mulvihill and David W.Smith, "The Genesis of Dermatoglyphics" ,Journal of Pediatrics.Vol.75,no.4,pp579-589

[9] Pankanti, S. Prabhakar, and A. K. Jain, "On the individuality of fingerprints," IEEE Trans. Pattern Anal.
Mach. Intell.", vol. 24, no. 8, pp. 1010–1025, Aug. 2002

[10] P. Gnanasivam & Dr. S. Muttan "Fingerprint Based Gender and Age Classification Using Wavelet Transformation and Singular Value Decomposition" IJCSI , Volume. 9: Issue 2, No 3: pp 274-282, March 2012

[11] Ajitha T.Abraham and Yasim Khan M, "Dermatoglyphics –Wavelet Approach" IJAEST, Volume 3,Number
3,pp208-216,Jan2015

[12]Acree, M. "Is there a gender difference in fin gerprint ridge density?" Forensic Science International 1999
May; 102 (1): 35 - 44.

[13]A. Badawi, M. Mahfouz, R. Tadross, and R. Jantz "Fingerprint - based gender classification" The
International Conference on Image Processing, Computer Vision, and Pattern Recognition, June 2006.

[14] Manish Verma and Suneeta Agarwal.'' Fingerprint Based Male - Female Classification. '' in Proceedings of the international workshop on computational intelligence in security for information systems (CISIS'08),
Genoa, Italy, 2008, pp.251 – 257.

[15] Rijo Jackson Tom, T.Arulkumaran , " Fingerprint Based Gender and Age Classification Using 2D Discrete Wavelet Transforms and Principal Component Analysis ". International Journal of Engineering Trends and Technology, Volume 4 Issue 2, 2013.

[10] Lijie Cao"Singular Value Decomposition Applied To Digital Image Processing", Arizona State University Polytechnic Campus

: Bouden Toufik and Nibouche Mokhtar, University of the West of England, "The Wavelet Transform for Image Processing Applications"

BIOGRAPHY

Ajitha T Abraham received Diploma in Electronics and B.Tech. Degree in Electronics and Communication Engineering from Govt. Poly Technic College, Kottayam and C S I Institute of Technology, Nagarcoil , Tamilnadu respectiveley in 2007 and 2010.Currently she had completed her M.Tech Degree in Signal Processing from College of Engineering. Poonjar,CUSAT . Her subject of interests includes Image Processing

BORDER SECURITY ROBOT

Minni Mohan[1]And Siddharth Shelly[2]

[1]Department of electronics and communication, M.A College of Engineering, Kothamangalam,A P J Abdul Kalam Technological University, Kerala, India
[2]Associate Prof, Department of electronics and communication Engineering, Mar Athanasius College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

The ordinary border patrol system suffers from intensive human involvement. Recently unmanned border patrol system consist of high tech devices, like unmanned aerial vehicles, unattended ground sensors, and surveillance towers equipped with wireless camera. However, any single technique encounters inextricable problems, such as high false alarm rate and line of sight constrains. There require a coherent system that co-ordinates various technologies to improve the system accuracy. In this project general idea of boarder security robot, wireless sensor network architecture for border patrol system, is introduced. Border security robot utilize a PIR sensor for human detection, a metal detector to detect the presence of explosives and a wireless camera for monitoring the scenario continuously at the remote station. Mechanical control of robotic vehicle along with robotic arm can be done from the remote station. This is initiated with a Bluetooth module.

KEYWORDS

PIC, PIR, Metal detector, Wireless camera

1. INTRODUCTION

Border patrol systems have recently achieved interest to address the concerns about national security. The major problem in protecting long stretches of borders is the need for large human involvement in patrolling the premises. In our border patrol system consists of security checkpoints and border troops. All vehicle traffic is need to stop in security check points which are set up on the international roads to detect and apprehend illegal aliens, drugs, and other illegal activity. The border troop watches and maintain control in a specific section of the border. The troops patrol the border according to predetermined route and time interval [1]. Under the conventional border patrol system, even modest-sized areas require large human resources if manual patrolling is considered alone. To monitor the border in real-time with accuracy and minimize the need for human support, multiple surveillance technologies, which complement each other are required. To address the challenges still facing by the existing surveillance techniques, we introduce Border security robot, a new border patrol system framework based on hybrid wireless sensor networks, which can accurately detect the border intrusion with minimum human involvements. Border security robot utilizes the PIR sensor for human detection and a metal detector for explosive detection. Also a wireless camera is used to continuously monitor the border. While the potential benefits of Border security are significant, several research challenges need to be addressed before a practical realization. In this project, a framework to deploy and operate Border Sense for border patrol is described.

DOI: 10.5121/ijci.2016.5230

2. Technology Used

2.1. Bluetooth Technology

The convergence of computing and communications has led to the development of Bluetooth technology. Taking the short-range wireless data usage to a new level, this technology is predicted to dominate both the home and business market. Bluetooth is a wireless technology standard for exchanging data over short distances (using short-wavelength radio transmissions in the ISM band 2400–2480 MHz) from fixed and mobile devices, creating personal area networks with high levels of security. It was originally conceived[2] as a wireless alternative to RS-232 data cables which is Created by telecom vendor Ericsson in 1994. It can connect several devices; problems of synchronization can be overcome.

In this project Bluetooth technology a wireless protocol that used to connects robotic side and personal computer in remote station. Configuring Bluetooth module all communications between remote station and robotic side can be done.

2.2 Infrared Technology

In order to understand what thermal imaging is, it is important to understand something about light. The total amount of energy in a light wave is related to its wavelength and shorter wavelengths have higher energy. Of visible light, the most energy for violet, and the least energy for red. Just next to the visible light spectrum is called as infrared spectrum. Shown in fig.1

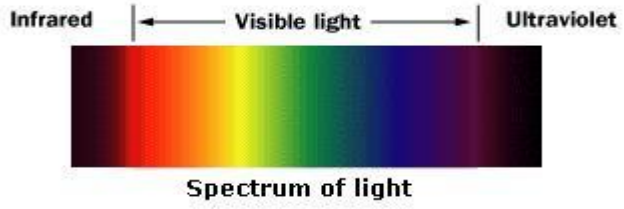


Fig.1.Spectrum of light

Infrared light can be split into three categories:

1. Near-infrared (near-IR)

It closest to visible light, near-Infrared has wavelengths that range from meters 700 billionths to 1,300 billionths of a meter.

2.2.2 Mid-infrared (mid-IR)

Mid-Infrared has wavelengths ranging from 1300 to 3000 billionths of a meter. Both near-Infrared and mid-Infrared are used in variety of electronic devices.

2.2.3Thermal-infrared (thermal-IR)

It Occupying the largest part of the infrared spectrum, thermal-IR has wavelengths ranging from 3000-30000 billionths of a meter (3 microns to over 30 microns).

In this border security robot continuous monitoring of border locations is done by using thermal imaging camera, which use infrared technology to capture images. The main difference between thermal-Infrared and the other two range is that thermal-Infrared is emitted by an object instead of reflected off it. In an object some process happening at the atomic level leads the emission of infrared lights. The signal-processing unit sends the information to the display, depending on the intensity of the infrared emission it appears as various colors. The combination of all the emissions coming from all of the elements used to creates the image

3. Proposed Model Explanation

3.1 Robot

The module consists of an Embedded System device which includes a central Microcontroller with a PIR sensor, metal detector and a Bluetooth interface. Two L293D driver ICs are used to drive the four motors. Drivers ensure the proper working voltage for DC motor and also protect the microcontroller from being harmed due to the back emf generated in the motor. Out of four DC motors, two are for the mechanical control of robotic vehicle and two are for robotic arm. The four mechanical controls that is being given to the robotic vehicle are forward, backward, right and left and that for robotic arm are up, down, expand and contract. The analog output of metal detector is connected to the controller through the inbuilt ADC in PIC and PIR sensor is connected to the external interrupt pin. Whenever an intruder or explosive is detected, PIC initiates the corresponding warning message transfer through Bluetooth module. Block diagram is shown in fig.2

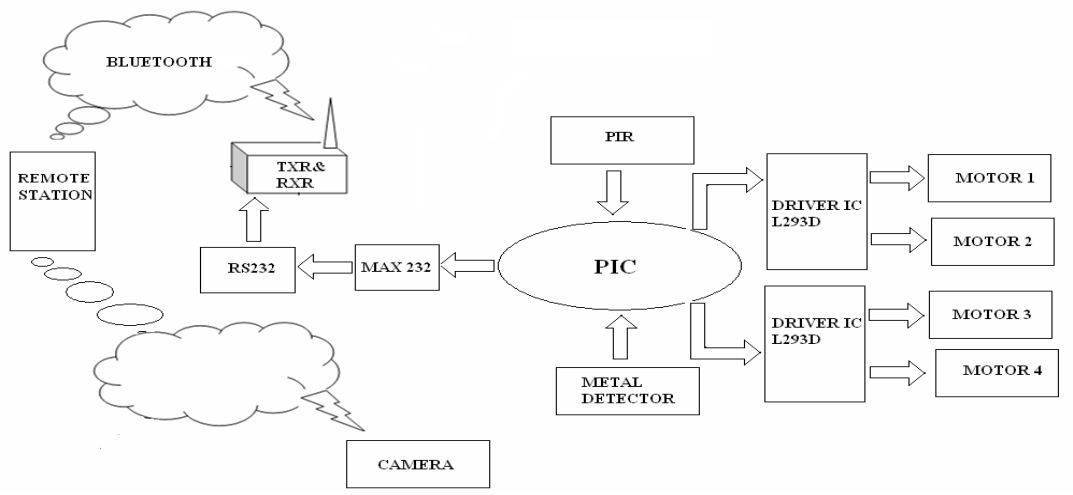


Fig.2 Block diagram

3.2 Pc Side

PC is connected to the robot through the inbuilt Bluetooth module. Whenever a warning message is received, corresponding secure measures can be taken by controlling both the robotic arm and vehicle by monitoring the border through wireless camera. With the help of wireless camera border scenario can be continuously monitor in control section

4. Flow Chart

Software part in remote station and robotic side are start by some initializations, which is provided for configuring wireless modules. In remote station configurations are set, with the use of visual basic GUI robotic part can be controlled. Both arm control and robotic motions can be controlled from remote station by using command button in GUI interface or by using configured icons in keyboard. For that icons in keyboard S, G, C, E and H, M, K, U configured. At the same time border scenario can be monitored using wireless camera, which work with infrared technology. It provides good quality images in nights, in presence of fog, mist etc. Thermal imaging camera is capturing their images by using infrared lights coming from each and every hot body in surrounding. So it is possible to find any hidden intruders in the surrounding from thermal imaging camera captured video. Remote station gets complete visualization of border scenario with high quality than conventional manual patrolling. If any intruder or explosive is found by robot the moment itself it inform to the control section of remote station. In remote station software is continuously checking the massages coming from robots and which is displayed in GUI and informs the authorized persons by alarm. Remote station flow chart is shown in fig.3.

In robotic side after initialization controller provided in robotic side continuously checking the output of PIR sensor, metal detector and remote station information's. If the PIR sensor output is high means "an intruder found", Controller send this information to remote station through Bluetooth interface. But any explosive is detected, which is sensed by controller using an interrupt (INTF) and a message "bomb detected" is send to remote station. Information's from remote stations are also received by monitoring RCIF interrupt. According to the information robot movements and its arm can be controlled. Flow chart explanation of robotic side is shown in fig.4

4.1 Remote Station

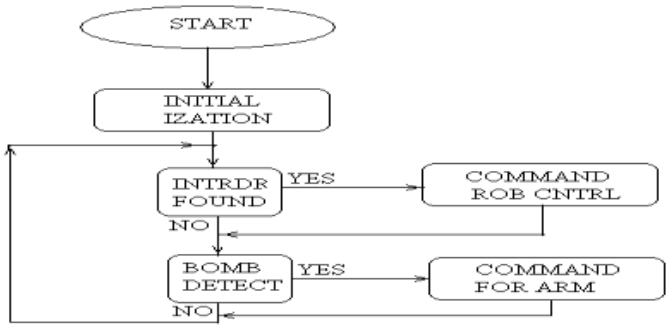


Fig3.Flow chart Remote Station

4.2 Robotic Side

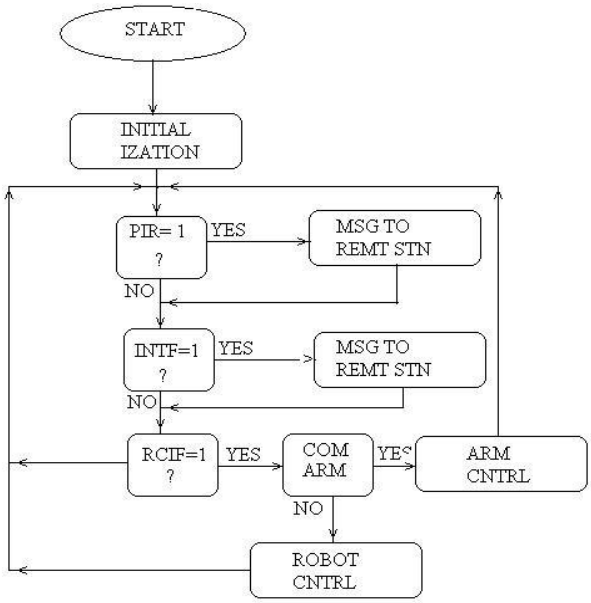


Fig.4.Flow chart –robotic side

5. Discussions And Result

5.1 Mplab Software:

It is possible to create the source files in a text editor like Notepad. Then run the Compiler on each C source file, which specifying a list of controls and run the Assembler on each Assembler source file, which specifying another list of controls, run either the Library Manager or Linker (again specifying a list of controls) and finally use the Object-HEX Converter to convert the Linker output file to an Intel Hex File. Once that completed the Hex File can download to the target hardware and can be debugged. Alternatively MPLAB can be used to create source files, which automatically compile, link and convert using options set with an easy to use user interface and finally simulate on the hardware with access to C variables and memory. Unless you have to use the tolls on the command line. MPLAB Greatly simplifies the process of creating and testing an embedded application.

5.2 VISUAL BASIC 6.0

Microsoft Visual basic 6.0 is the application software used to interface Bluetooth to PC. Robot.exe file on execution opens a window as shown in figure. Visual Basic is a third-generation event-driven programming language and (IDE) integrated development environment[3]. It from Microsoft for its COM programming model and it is first released in 1991. Visual Basic is designed to be comparatively easy to learn and use. It was derived from BASIC

and it enables the rapid application development of graphical user interface applications, which access to databases using Data Access Objects and Remote Data Objects.

Fig.5 VB application home page

In the select port combo box all available com ports of a PC will be shown in fig 5. From that select the appropriate com port and then press the connect button. In this project comport 13 is configure to connect with Bluetooth module. This will connect the Bluetooth module in PC and Bluetooth module in robot. Whenever an intruder or explosive detected, the message "an intruder detected" and "bomb detected" will be displayed in the dialogue box of the above shown form. FRONT, BACK, LEFT, RIGHT and STOP are the command buttons that has been provided for the mechanical control of robotic vehicle. At the same time some keyboard icons also configured for vehicle control which are D, S, F, C, and E. By pressing Keyboard key E or command button FRONT robotic vehicle can be move forward. By pressing Keyboard key C or command button BACK robotic vehicle can be move backward. By pressing Keyboard key S or command button LEFT robotic vehicle can be tern left. By pressing Keyboard key F or command button RIGHT robotic vehicle can be tern right. By pressing Keyboard key D or command button STOP robotic vehicle can be move stopped. UP, DOWN, CATCH, DROP and QUIT accounts for the mechanical control of robotic arm. At the same time some keyboard icons also configured for vehicle control which are J, H, M, and K, U. By using keyboard key H or by pressing command button CATCH robotic arm can move more closer, by using keyboard key K or by pressing command button DROP robotic arm can move more wider. By using keyboard key U or by pressing command button UP robotic arm can move upward. By using keyboard key M or by pressing command button DOWN robotic arm can move downward. By pressing Keyboard key J or command button QUIT robotic arm can be drop anything in their hand .All control commands sent using Bluetooth module. When the connections are to be terminated for a short while, one can use the disconnect button. On this button click, the connection between Bluetooth modules is terminated while the form will still sustain in the display. In order to re-establish a connection we can make use of connect button once more. When the application has to be terminated permanently, exit button is pressed. This will disconnect the Bluetooth module along with closing the application.

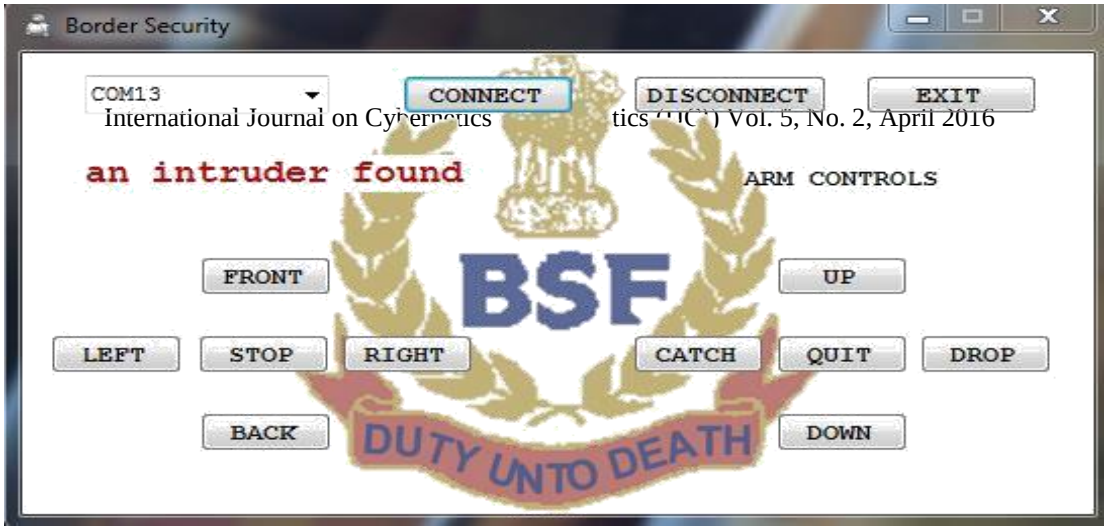


Figure 6 Notification of bomb detection

Robotic vehicle or robotic part patrolling in the borders continuously checking the presence of intruder by using PIR sensor and presence of explosive is detected by using metal detectors, which is attached in robot. Continues border scenario information is send to remote station with the help of thermal imaging camera. Infrared detector is the main part of any thermal imaging. With the help of thermal imaging camera possible to see crystal clear pictures through darkness, fog, and haze and smoke all depends on the quality of the detector. Which is a wireless camera and its receiver module is present in remote station. Receiver module receives wireless information's from thermal imaging camera and displayed in scenario monitoring PC. If any intruder is detected by robot that massage is send to remote station. "AN INTRUDER FOUND" massage is display in visual basic GUI is shown in fig.8. Similarly any explosive is detected by robot that bomb detect massage send to remote station. "BOMB DETECTED" message is display in visual basic GUI is shown in fig.7

Fig.7 Notification of intruder detection

5. ADVANTAGES

1. The multimedia sensors provide accurate detection as well as large detection range.
2. The thermal imaging camera provide additional information that cannot be detected by the multi-media sensors, e.g. in cases where the intruder is hidden behind an obstacle that cannot be detected
3. Mobile sensors provide intrusion tracking capability to track the intruders after they have been detected
4. By network processing, the heterogeneous sensors cooperatively detect the intrusion and report the results to a remote administrator. Accordingly, both the deployment and operational cost of the border patrol system can significantly be decreased.

6. CONCLUSION

At present in our country there hasn't been used a system which can automatically detect human intrusion as well as the presence of any explosive materials at the borders. Presently one or more soldiers are needed to patrol the border area in this project, we introduce Border security robot, a hybrid wireless sensor network architecture for border patrol to reduce the human involvement and used to improve the detection accuracy of current border patrol systems. The human involvement reduced with the help of PIR sensor and metal detector and thermal imaging camera. It can be concluded from the above.

ACKNOWLEDGEMENT

I take this opportunity to thank all those who have been directly or indirectly involved in making this project success. Express my honest gratitude to my project guide Prof. Aby Mathew for his constant encouragement, inspiration and guidance.

REFERENCES

[1] Zhi Sun, Pu Wang, Mehmet C,Vuran,Mznah A. Al-Rodhaan,Abdullah M.Al-Dhelaan,Ian F.Akyildiz , (2011),"BorderSense: Border patrol through advanced wireless sensor networks", Ad Hoc Networks 9 pp. 468–477

[2] K. V. S. S. S. S. Sairam, N. Gunasekaran, S. Rama Reddy, (2002), "Bluetooth in Wireless Communication", IEEE Communications Magazine ,pp.90-96

[3] Francesco Balena ,"Programming Microsoft Visual Basic 6.0", Publisher: Microsoft Press,1999

[4] I.F. Akyildiz, T. Melodia, K. Chowdhury,(2008) "Wireless multimedia sensor Networks : applications and testbeds", Proceedings of the IEEE 96 (10) pp. 1588–1605

[5] Ariponnammal S and Natarajan S (1994) "Control system for a mobile robot", Pramana Journal of Physics Vol.42, No:1, pp.421-425

AUTHOR

Minni Mohan Graduated (B-tech)in Electronics and Communication Engineering from Mahathma Gandhi University and currently pursuing M Tech in VLSI and Embedded System in APJ Abdul Kalam Technological University, Kerala, India.

Siddharth Shelly is a faculty member of Mar Athanasius College of Engineering, Kothamangalam, Kerala, India. He received his B.Tech from Mahatma Gandhi University, Kottayam and M.Tech degree from the Amrita School of Engineering, Coimbatore, India. His current research focus is in the area of vehicular ad hoc networks, embedded systems.

INTENTIONAL BLANK

COMPACT MICROWAVE PLANAR BAND PASS FILTER

Ambily K[1] and Anila P V[2]

[1]Dept of Electronics and Communication, Mar Athanasius Engineering College
A.P.J Abdul Kalam Technological University, Kerala, India
[2]Assistant Professor, Dept of Electronics and Communication Engineering
Mar Athanasius Engineering College, Kothamangalam Kerala, India

ABSTRACT

The need for the compactness of microwave filter at high frequency application has lead to the design of microwave planar band pass filter which passes certain range of frequencies. The aim of this paper is to study and implement a compact microwave planar band pass filter operates in 3.1GHz-10.6GHz which is the ultra-wide band range. It uses multiple mode resonator (MMR) and inter digital coupled lines to improve upper stop band performance and coupling degree. Simulation of the proposed microwave filter structures is done with the help of HFSS (High Frequency Structure Simulator) software.

KEYWORDS

MMR, SLFSIR ,resonator

1. INTRODUCTION

The increasing scale of modern wireless communication application and radar system in today's technology has boosted the demand for microwave filters as they are playing an essential role in Transmit–Receive system.

In circuit theory, a filter is an electrical network that alters the phase and/or amplitude characteristics of a signal with respect to frequency. A filter will not add new frequencies to the input signal, nor will it change the component frequencies of the signal, but it will change the relative amplitudes of the various frequency components and their phase relationships. Filters are used in electronic systems to emphasize signals in certain frequency ranges and reject signals in other frequency ranges. A microwave filter is a two-port network used to control the frequency response at a certain point in a microwave system by allowing transmission at frequencies within the pass band of the filter and attenuation in the stop band of the filter. Typical frequency responses include high-pass, low-pass, band reject, and band-pass characteristics. Applications can be found in virtually any type of test and measurement system, microwave communication and radar.

A band pass filter allows transmission of a limited band of frequencies and rejects all other frequencies below or above frequency band. It has parallel tuned circuit in the shunt arm and

series tuned circuit in series arm . Here a microwave planar band pass filter in ultra wide band range (3.1-10.6)GHz with circular stepped impedance stubs is implemented with a low insertion loss in pass band. For microwave range higher than 500 MHz the passive filters are mostly realized by using either planar transmission lines, or waveguides. Although waveguide components have low losses, and can handle higher power than the planar transmission lines it is not preferable in new communications systems that require the mobility, because of the their large size and heavy weight. Moreover, fabrication processes of the waveguide components are more expensive than that for planar transmission lines. Therefore most of the microwave filters are designed using the transmission lines. Variety of structures such as stepped impedance stubs, coupled lines, single mode resonator, multi mode resonators and coupled resonators can be used for implementation.

2. METHODS IN FILTER DESIGN

2.1. COMPACT UWB BAND PASS FILTER WITH MMR

The MMR used for this filter design consists of one half-wavelength $\lambda/2$ low impedance line section in the center and two identical $\lambda/4$ high-impedance line sections at the two sides. Here first three resonant frequencies of the MMR are properly adjusted to be placed quasi equally within the UWB. After that, the parallel-coupled lines at the two sides are longitudinally stretched to raise the frequency dispersive coupling degree with the coupling peak near the center of the UWB.

It was initially designed in [1] such that that the first two resonant modes of the MMR is utilized together with the input/output parallel-coupled lines to achieve a wide pass band with four transmission poles. Another modification of the filter is presented in [2],[3] by making first three resonant modes of the MMR constructed to realize five transmission poles with lowered return loss in the whole pass band range. Following the works in [1],[2] and [3] the MMR here is to be properly modified in configuration so as to reallocate its first three resonant modes close to the lower-end, center, and upper-end of the targeted UWB pass band. By forming a MMR and introducing quarter-wavelength parallel coupled lines in the input and output ports, a UWB pass band with five transmission poles is achieved and it is finalized as the result.

2.2. FILTERS WITH STUB LOADED MMR

UWB band pass filter using MMR of stepped-impedance configuration was initially reported in [4]. But these stepped-impedance MMR-based filter suffered from a high insertion loss, narrow upper stop band as well as worse selectivity.

A new approach was the design of filter based on electromagnetic bandgap (EBG) embedded MMR which is reported in [5] to widen the upper stop band . The core piece of these two filters is a stub-loaded multiple-mode resonator (MMR). The MMR is constructed by loading three open stubs in a uniform-impedance resonator that means and two uniform-impedance stubs of lengths l_s at the symmetrical side locations and one stepped-impedance stub at the center of length l_c. Five modes, includingthree even modes and two odd modes, could be designed within UWB band .The odd and even resonances of an unloaded MMR occur alternatively. All resonances move down to lower frequencies as l_s is increased and this implies circuit miniaturization. Only the

even modes move down to lower frequencies when l_c is increased**.** The two odd modes could be located within the UWB band by properly designing the horizontal impedance resonator and the side stubs and the even modes could be flexibly tuned while the odd modes are fixed. The two transmission zeros generated by the stepped-impedance stub are at the lower and upper cut-off frequencies, resulting in a sharp pass band performance. These results demonstrate that this UWB filters has much better electrical performances of UWB pass band than the filters mentioned in [4] and [5]. The filters explained in papers [1-5] had the disadvantages like narrow upper-stop owing to the appearance of parasitic pass bands, electrically large in size which is impractical in the implementation of hand-held UWB devices and the tight line spacing between parallel coupled lines for strong coupling is hard for fabrication.

2.3. FILTER WITH SLFSIR

Q X Chu etal came up with a more compact UWB band-pass filter using a stub loaded folded stepped impedance resonator (SLFSIR) and aperture-backed inter digital coupled lines .This filter comes up with a solution to all problems discussed in [1-5]. Several prototype UWB filters have been reportedly developed based on principles, such as dual-stop band features [6] and pseudo-inter digital structure [7]. The SLFSIR is designed such that the first four resonant frequencies of it are properly adjusted to be placed evenly within the UWB while the fifth resonant frequency is made above 16.0 GHz to improve upper stop band performance. The aperture-backed inter digital coupled lines can create a transmission zero to eradicate the fifth resonant mode of the SLFSIR to make a wider upper-stop band. To make a wide pass band covering the whole UWB band, the aperture width and parallel-coupled lines length should be suitably chosen to get proper coupling [8]. The inter-digital coupled lines can increase the coupling degree to relax the tight line spacing. The two side stubs with varied lengths in SLFSIR can provide an extra degree of freedom for adjusting the locations of the first four resonant frequencies in an alternative way. The filter proposed by them has some attractive features like a wide and deep upper-stop band with insertion loss higher than 20 dB in the 11.7–19.0 GHz range, relaxed tight line spacing and miniaturized size. These MMR based UWB filters are still limited by the existence of periodic and narrow pass bands in the upper-stop band.

3. PROPOSED FILTER

The compact UWB band pass filter uses MMR which is made by uniting three pairs of circular impedance stepped stubs in shunt to a high impedance micro-strip configuration, as shown in figure 1. A band-pass filter can be constructed using any units that can resonate. Filters using stubs can clearly be made band-pass. Stub is the common component of distributed filters. Over a narrow range of frequencies, a stub can be used as an inductor or a capacitor (its impedance is determined by its length) but over a wide band it act as a resonator. This compact UWB band pass filter uses stepped stubs. It can easily implement steps in characteristic impedance of the configuration, which introduces a discontinuity in the transmission.

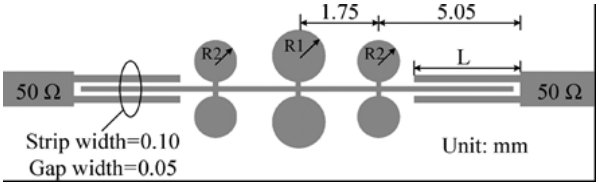


Figure 1: Filter structure

4. STRUCTURAL DESCRIPTION

Resonators are fundamental elements which usually determine the size of the filter. The effective approach for miniaturizing the filter size is to reduce the resonator size. This can be achieved by two methods. One approach is to modify the physical structures another approach is to modify the traditional resonators to generate additional modes, which makes the resonator to behave as an MMR.

Multiple mode resonator resembles to a metal box of fixed dimensions that can support a number of resonant modes at a given frequency. They combine different multiple modes which constitute narrow bands into a wide band. That is multiple mode resonator has multiple resonance behaviour in its pass band.

The compact UWB band pass filter uses MMR which is formed by attaching three pairs of circular impedance-stepped stubs in shunt to a high impedance micro-strip line. A band-pass filter can be constructed using any units that can resonate. Filters using stubs can clearly be made band-pass. Stub is the common component of distributed filters.Short-circuit, nominally quarter-wavelength stubs behave as shunt LC resonator and an open-circuit nominally quarter-wavelength stub behaves as a series LC resonator.

The compact UWB band pass filter uses stepped stubs. It can easily implement steps in characteristic impedance of the configuration, which paves the way for a discontinuity in the transmission characteristics. This is done in planar topology by changing the width of the transmission line. Here choose the dimensions of the stubs connecting to the circles as fixed 0.1 mm in width and 0.15 mm in length. These circles have different impedances and which can be changed by adjusting the radius of the circles. After adjusting the radius of the circles of the stubs, the resonant modes of the MMR can be roughly designated within the 3.1–10.6 GHz UWB band while demolishing the spurious harmonics in the upper-stop band.

The length can be reduced and the compactness successively increased by introducing inter digital filters, comb-line filters, and parallel-coupled line filters. In order to enhance the coupling degree, two inter digital coupled-lines are used. Conventional inter digital coupled line which has been widely used as a capacitive coupling element in multi-stage band pass filters. The optimized inter digital coupled lines is used to achieve design-specified coupling factor between two adjacent line resonators. For this, the common procedure is to decrease both strip and slot widths in order to achieve a lower insertion and tight coupling. But, it introduces some difficulties into the design and fabrication process because the coupling response is susceptible to the strip or slot

widths of the configuration. In this paper, the optimized values of strip and slot width are chosen as 0.1 mm and 0.05 mm respectively. Coupling length denoted as L should be optimized to enhance the coupling degree.

A new topology of MMR is utilized to replace the traditional MMR. It is formed by attaching three pairs of circular impedance-stepped stubs in shunt to a simple high impedance micro strip line.

5. OPTIMISATION OF FILTER

This structure of compact UWB band pass filter is implemented with optimized values of R_1, R_2 and coupling length L. The optimization is completed as follows .Implement the structure with weak coupling , that is consider L as .6 mm .Change the values of R_1 by keeping R_2 as constant and L as .6 mm. Check the graphical result of S_{21} for an ultra wide band , high insertion loss in the stop band and wide upper stop band. Allocate the resonating modes in the ultra wide band for a pass band in UWB by keeping R_2 constant and R_1 varying. Repeat the above steps for R_1 constant and R_2 varying with same weak coupling L as .6 mm. Finally optimize the values of R_1 and R_2 for a pass band in UWB and wide upper stop band. For enhancing the coupling, increase the L value. This is for the proper feeding of MMR by inter-digital coupling line. Optimize the coupling length for a 0 dB insertion loss in the pass band. The optimisation can be completed as follows.

1. Implement the structure with weak coupling , that is consider L as .6 mm
2. Change the values of R_1 by keeping R_2 as constant and L as .6 mm.
3. Check the graphical result of S_{21} for an ultra wide band , high insertion loss in the stop band and wide upper stop band.
4. Allocate the resonating modes in the ultra wide band for a pass band in UWB by keeping R_2 constant and R_1 varying.
5. Repeat the above steps for constant R_1 and varying R_2 with same weak coupling L as .6 mm.
6. Finally optimize the values of R_1 and R_2 for a pass band in UWB and wide upper stop band.
7. For enhancing the coupling, increase the L value. This is for the proper feeding of MMR by inter digital coupling line.
8. Optimize the coupling length for a 0 dB insertion loss in the pass band.

5.1 EFFECT OF CHANGING R_1, FIXED R_2 =.5 MM

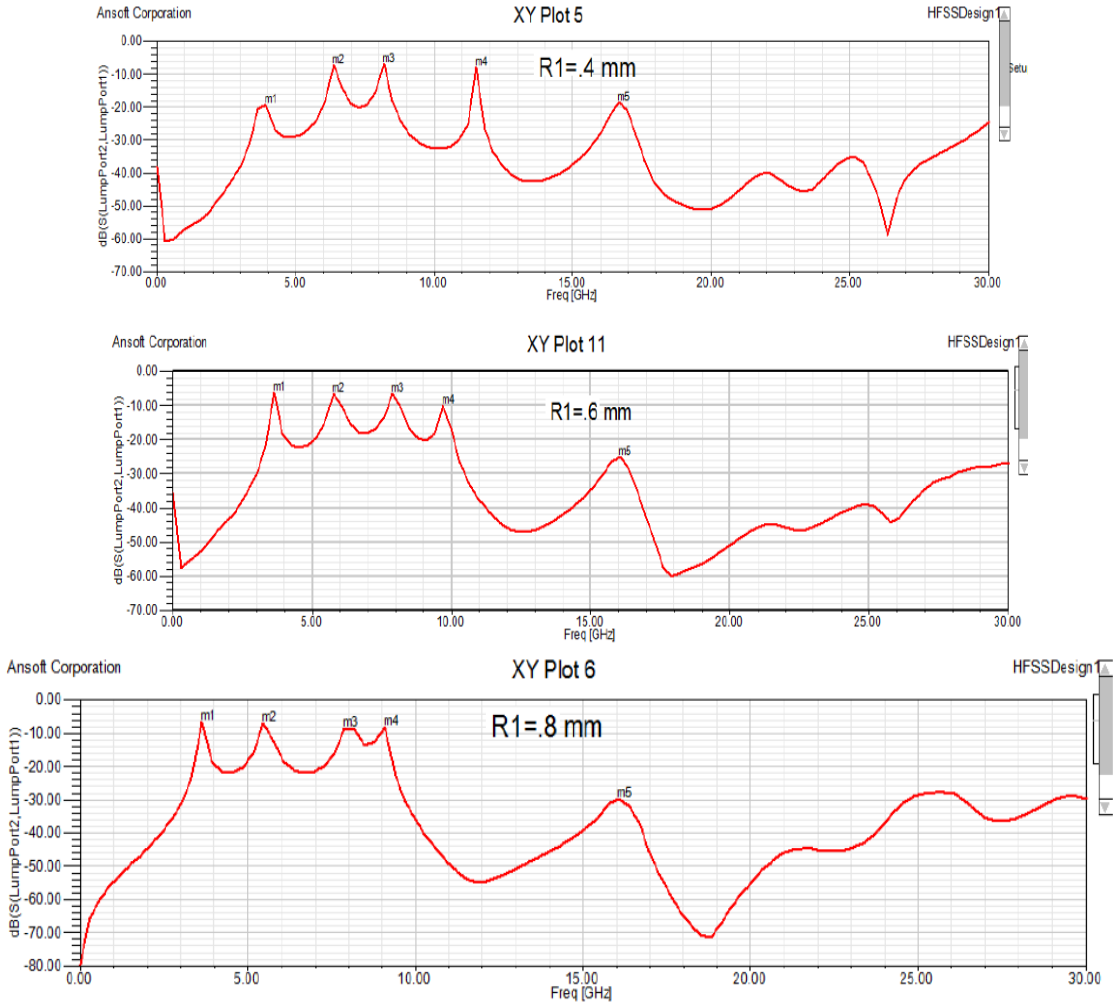


Figure 2: Simulated S_{21} -magnitude of the filter under weak coupling with fixed L=0.6 mm, R_2=0.5mm and varied R_1

As radius varies from 0.4 to 0.8 mm, the odd resonant (m1,m3,m5) remain stationary, while the even resonant (m2,m4)modes tend to slowly shift downwards. It is well valid in theory that the center location of the resonator corresponds to a short circuit or perfect electrical wall for odd modes, and its characteristics are hardly affected by the attachment of the shunt stub, whereas it indicates an open circuit or perfect magnetic wall for all the even resonant modes. Thus, the second mode can be adjusted to the middle of the pass band, while the fourth mode can be reduced and allocated within the UWB pass band.

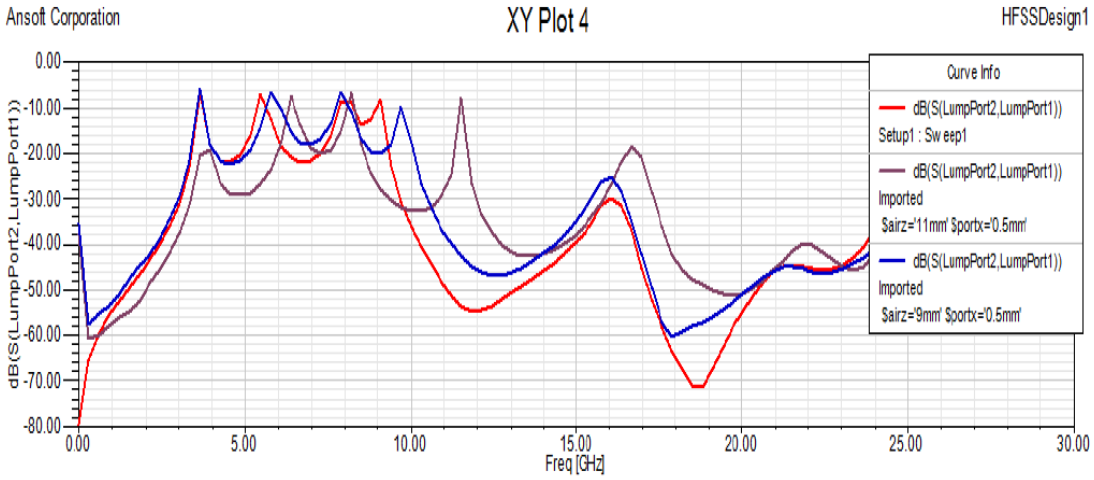
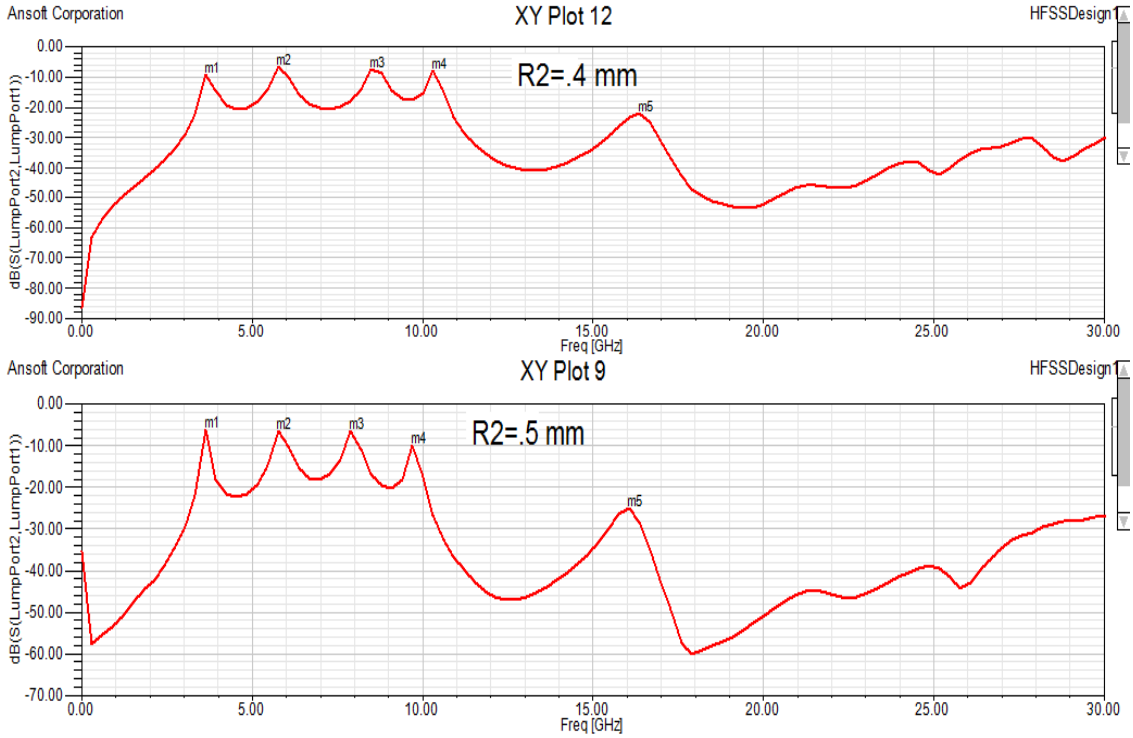


Figure 3: Comparison of S_{21} for different values of R_1.

5.2 EFFECT OF CHANGING R_2, FIXED R_1=.6 MM



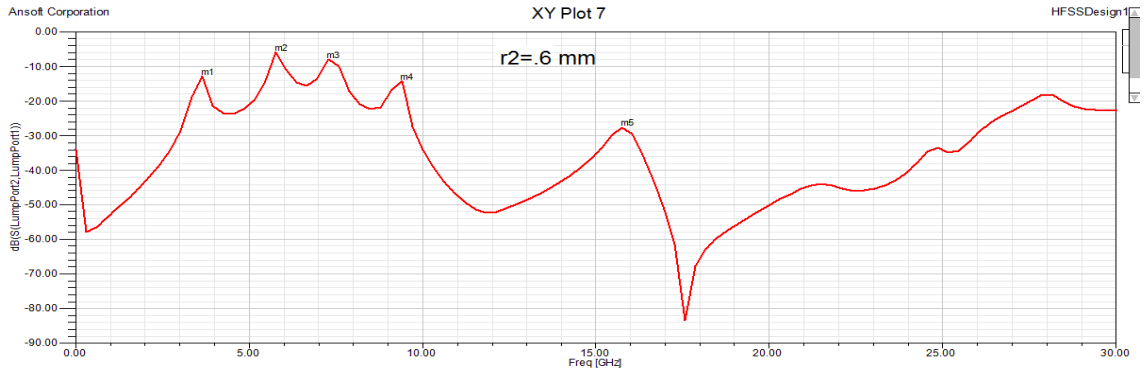


Figure4: Simulated S_{21} -magnitude of the filter under weak coupling with fixed L=0.6, R_1=0.6and varied R_2

The five resonant modes all move towards the lower frequency except the second mode which is almost unchanged, while changing the radius R_2 from .4 mm to .6mm. Thus, the two side circular impedance-stepped stubs with varied can provide an additional degree of freedom to adjust the locations of the first four resonant frequencies in an alternative way.

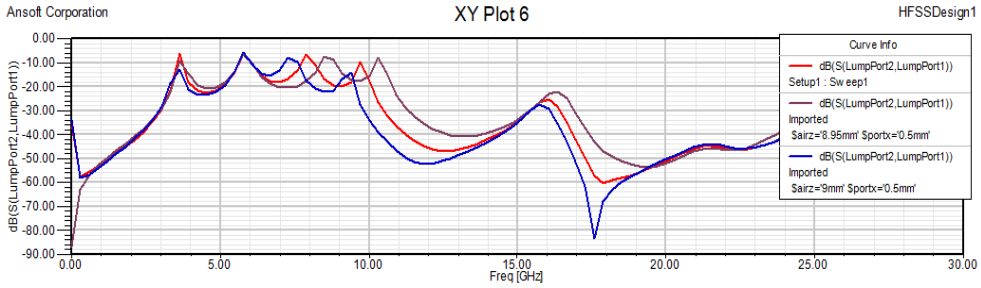
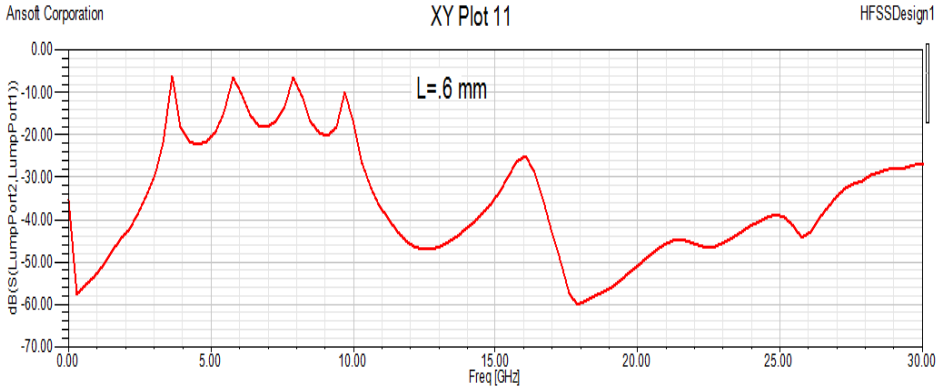


Figure 5: Comparison of S_{21} for different values of R_2.

5.3 EFFECT OF CHANGING L, FIXED R_1=.6 MM AND R_2=.5 MM



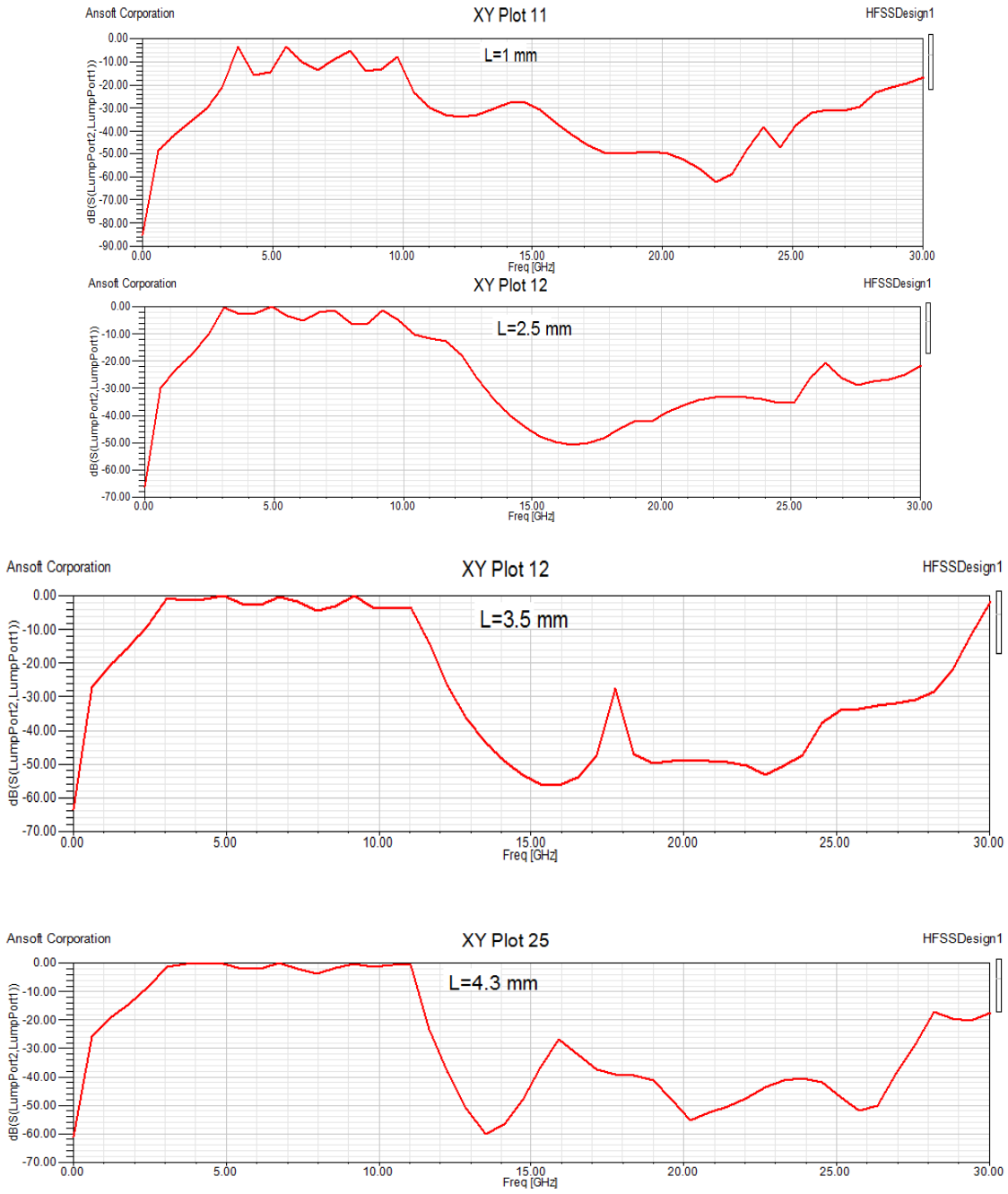


Figure6: Simulated S_{21} -magnitude of the filter under weak coupling with varied L and fixed R_1=0.6 and R_2=.5mm

From the above figures, it can be seen that on increasing L improved pass band performance can be achieved. From all these optimisation methods it can be seen that upper stop band performance of the proposed filter is improved.

6. RESULTS AND DISCUSSIONS

6.1 EQUIVALENT CIRCUIT OF UWB MICROWAVE BAND PASS FILTER AND ANALYSIS

By Richard transformation and kuroda's identity the schematic of the equivalent circuit of a UWB planar microwave band pass filter . The circuit parameters have the following values

C_1=C_3=.267 pF
L_1=L_3=2.887 nH
C_2=.7488 pF and
L_2=1.03 nH

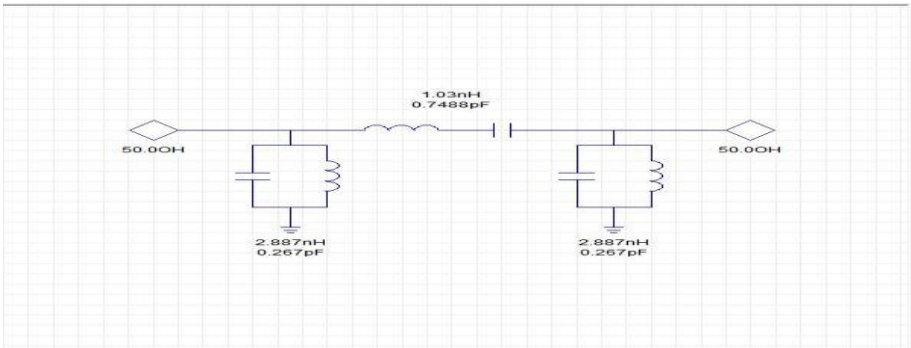


Figure 7: Equivalent circuit

On simulation in HFSS circuit design software S_{21} and S_{11} were obtained as in figure 8.

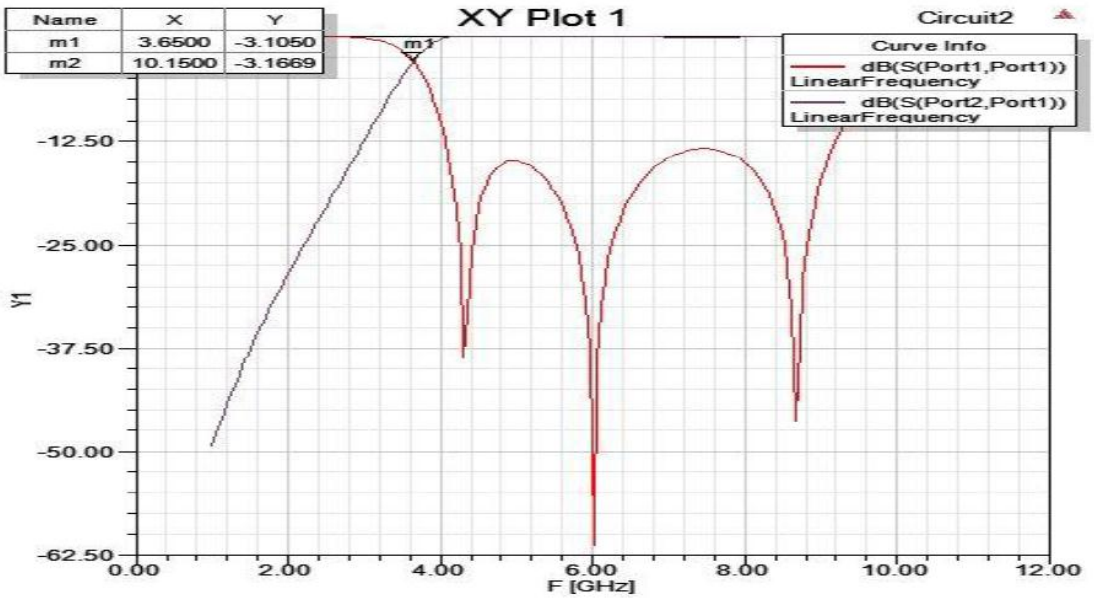


Figure 8: Simulation results

7. FEATURES OF THE DESIGN

1) Small insertion loss in the pass band.
2) The wide and deep upper-stop band with an insertion loss of 30 dB in the 12.1 to 27.8 GHz.
3) Compact size with 13.6 mm in length.

8. CONCLUSION

In this work, a compact UWB BPF with improved upper stop band performance is proposed and demonstrated using the new MMR which is formed by uniting three circular impedance-stepped stubs in shunt to a high impedance micro-strip line. The design procedure of this circular MMR is much simpler compared to the rectangular MMR. Just by simply adjusting the radius of stubs, the first four resonant modes of this MMR can be successfully allocated within the regulated UWB pass band, which makes the 3-dB bandwidth from 2.8 to 11 GHz. Meanwhile, a wide upper-stop band with the insertion loss higher than 30 dB in range of 12.1 to 27.8 GHz is achieved. In addition, it has a compact size with 13.6 mm in length.

ACKNOWLEDGEMENT

I express my deep and profound feeling of gratitude to Mr. Harikrishnan A I, Assistant Professor, NSS College of Engineering, Palakkad for providing me with the guidance for the successful completion of the project

REFERENCES

[1] L. Zhu, H. Bu, and K. Wu, "Aperture compensation technique for innovative design of ultra-broadband micro strip band pass filter," in IEEE MTT-S Int. Dig., vol. 1, 2000, pp. 315–318.
[2] W. Menzel, L. Zhu, K. Wu, and F. Bogelsack, "On the design of novel compact broad-band planar filters," IEEE Trans. Microw. Theory Tech., vol. 51, no. 2, pp. 364–370, Feb. 2003.
[3] L. Zhu, W. Menzel, K. Wu, and F. Boegelsack, "Theoretical characterization and experimental verification of a novel compact broadband micro strip band pass filter," in Proc. Asia-Pacific Microwave Conf., Dec.2001, pp. 625–628.
[4] M. Makimoto and S. Yamashita, "Band pass filters using parallel coupled strip line stepped impedance resonators," IEEE Trans. Microw. Theory Tech., vol. 28, no. 12, pp. 1413–1417, Dec. 2001.
[5] S. W. Wong and L. Zhu, "EBG-embedded multiple-mode resonator for UWB band pass filter with improved upper-stopband performance," IEEE Microw.Wireless Compon.Lett., vol. 17, no. 6, pp. 421–423, Jun. 2007.
[6] Ishida, H., and Araki, K.: 'Design and analysis of UWB bandpass filterwith ring filter,' IEEE MTT-S Int. Microw. Symp. Dig., June 2004,pp. 1307–1310.
[7] Ji, M.Z., and Chu, Q.X.: 'A compact UWB band pass filter usingpseudo-interdigital stepped impedance resonators'. Proc. ChinaMicrowave Millimetre-Wave Conf., Ningbo, China, October 2007,pp. 1096–1098.
[8] Zhu, L., and Wang, H.: 'Ultra-wideband bandpass filter on aperturebacked microstrip line,' Electron. Lett., 2005, 41, (18), pp. 1015–1016.

AUTHOR

Ambily K did her B. tech in Electronics and Communication from Calicut university at NSS College of Engineering in 2014.Currently she is pursuing her M. tech in VLSI and Embedded systems at Mar Athanasius College of Engineering, Kothamangalam

Anila P V working as Assistant Professor in Mar Athanasius College of Engineering, Kothamangalam received her M.Tech degree in Electronics with specialization of Microwave and Radar Engineering from Cochin University of Science and Technology (CUSAT) in 2011 and B.Tech in Electronics and Communication Engineering in 2009. She is also pursuing her Ph.D degree from CUSAT. Her areas of interest include Electronic Circuits, Computational Electromagnetics, Multi band antennas, Planar antennas, Metamaterials and its applications in the field of compact antennas etc

COMPRESSION AND DECOMPRESSION OF BIOMEDICAL SIGNALS

AnjalyJosephT[1] and Arun.K.L[2]

[1]Department of Electronics and Communication Engineering, Mar Athanasius College Of Engineering, A P J Abdul Kalam Technological University, Kerala, India
[2]Assistant Professor, Department of Electronics & Communication,
Mar Athanasius College of Engineering

ABSTRACT

In this work, a novel ECG data compression method is presented which employs set partitioning in hierarchical trees algorithm(SPIHT) on two dimensional electrocardiogram(2D-ECG).The 2D ECG is a two dimensioned array, in which each row of this array indicates one or more period and amplitude[7] normalized ECG beats. When SPIHT algorithm is used to compress one or two-dimensional signals separately it is observed that they achieve precise rate control, progressive quality, high compression ratio and low root mean square[7] difference(PRD).Better results are attained because of the wavelet transform eliminating effect[7] on redundancies between adjacent samples. It is also eliminated in one- dimensional ECG-and between adjacent beats by applying 2D wavelet transform. The SPIHT algorithm has achieved prominent success in signal compression. Experiments on selected records of ECG from MIT-BIH arrhythmia database revealed that the proposed algorithm is significantly more efficient for compression.

KEYWORDS

SPIHTAlgorithm, 2Dwavelet decomposition, encoding

1. INTRODUCTION

An ECG is a physiological signal for cardiac disease diagnostics. As the sampling rate, sample resolution, observation time and number of leads increase, the amount of ECG data also increases and so large storage capacity is required. Specially, when data transmission is required, the amount of transmission time also increases and it needs more bandwidth for compensation. With all of these limitations, ECG data compression has been the most active research areas in biomedical engineering and signal processing. Techniques for ECG compression can be classified into three categories:

1) Direct time-domain methods (AZTEC, CM, TP, CORTES, and SAPA, FAN)
2) Transform methods (Fourier, KLT, DCT, Wavelet) and
3) Parametric techniques (linear prediction, long-term prediction).

In technical literature, a number of time–frequency methods are available for the high resolution signal decomposition. This includes the short time Fourier transform(STFT),Wigner–Ville transform(WVT),Choi–Williams distribution (CWD) and the WT.DWT is the appropriate tool for the analysis of ECG signals as it removes the main shortcomings of the STFT[4]; it uses a single analysis window which is of fixed length in both time and frequency domains. This is a major

DOI: 10.5121/ijci.2016.5232

drawback[2]of the STFT, since what we really needed are a window of short length(in time domain) for high frequency content of a signal and a window of longer length for low frequency content of the signal. The WT is more better compared to STFT, it is because it is possible to vary the window length depending on the frequency range of analysis. This is obtained by scaling (contractions and dilations)as well as shifting the basis wavelet. The continuous wavelet transform (CWT)transforms a continuous signal into a highly redundant signal of two continuous variables—translation and scale. The resulting transformed signal is easy to interpret and for time-frequency analysis.

An ECG is a pseudo periodic signal, means that not periodic in the strict mathematical sense and not completely a random signal. By looking at the time of evolution of the signal, a concatenation of similar events which almost never reproduce themselves identically is observed. Based on these, several compression methods have been developed. These methods are average beat subtraction with residual[6] differencing, long-term prediction and vector quantization (VQ). Most of these methods are using correlation between adjacent samples in a single cycle(intra-beat dependencies) and not using correlation between adjacent beats across cycles (inter-beat dependencies).Some works have been done to utilize inter-beat dependencies. In this project a new method of ECG compression, using set partitioning in hierarchical trees algorithm (SPIHT) in wavelet domain is presented and it is applied to 2D-ECG.Simulation results, based on data from MIT/BIH arrhythmia database is provided to show the effectiveness of this approach. All ECG data used here are sampled at 360 Hz, 11 bits/sample.

The ECG signal is composed from five waves labeled using five capital letters from the alphabet : P,Q,R,S and T (Fig.1).Compressing the ECG signal while preserving the original shape of the reconstructed signal and especially the amplitudes of Q,R and S peaks, without introducing distortions in the low amplitude ST segment, P and T waves are the main objectives of our project.

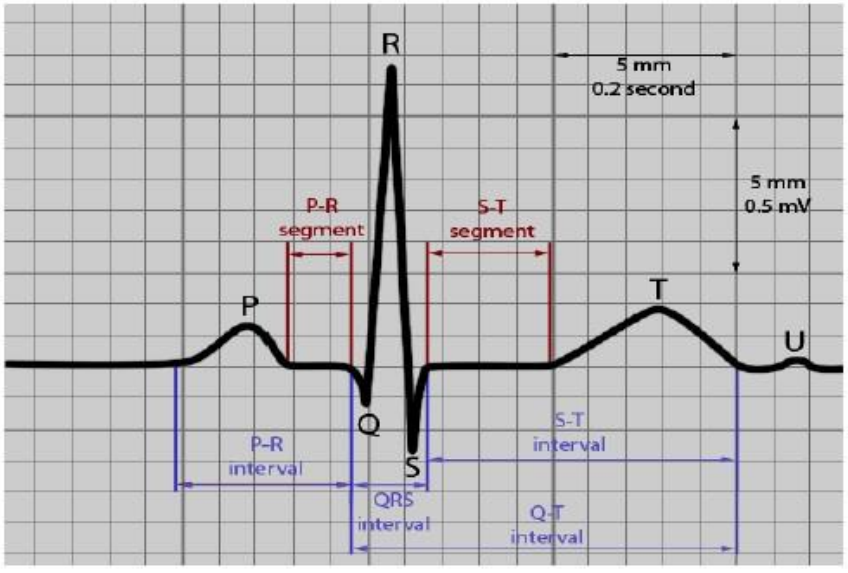


Figure1. Schematic Representation of ECG Signal in Time Domain

2. SYSTEM DESIGN

The project is based on the following block diagram. The first stepis to procure an ECG signal from an existing ECG database over the World Wide Web.A signal cannot be compressed as such and certain pre-processing is needed that is coefficients have to be extracted from the signal.Different types of transforms can beused for this purpose.But,wavelet transform is opted because of its advantages over other transformation methods in this specific compression method.Once the coefficients are obtained,SPIHT compression technique is applied and we obtain the compressed form of the ECG signal as the input.

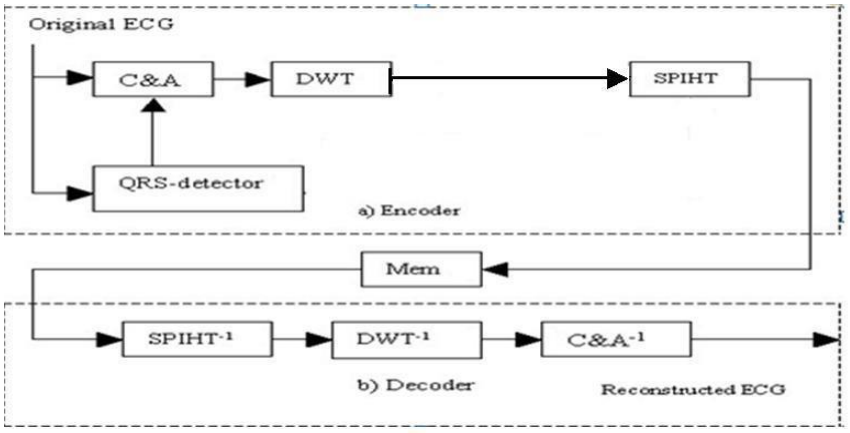


Figure2.Blockdiagram

According to the intra beat correlation and the inter beat correlation of ECG signals, a[5]2-DECG signal compression algorithm is proposed. The proposed algorithm is generally implemented in the following steps.

- QRS detection of the 1-DECGsignal.
- 2-DECGdata array construction.
- 2-D wavelet decomposition.
- Coefficients encoding.
- Reconstruction.

2.1 QRS complex detection of the 1-d ECG signal

In order to fully utilize the inter beat correlation, the input1-DECGsignalhastobefirstQRS complex detected and the original 1-Dsignalcanbe segmented and aligned according to the results of QRS detection. By comparing ever altypical QRS detection algorithms, it is found that the QRS detection scheme is based on wavelet transform, and it has robust noise performance.Itemploysthemulti-resolutionanalysisandbearsflexibilityinanalyzingthe time-varying morphology of ECG data.

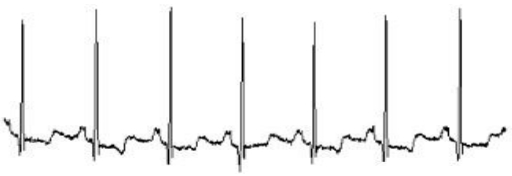


Figure 3.ECG signal

2.2 2-D ECG Data array Detection

Construction of a cut and aligned 2-D ECG data array after the QRS detection is shown in Fig3.Segment the1-D ECG signal according to the heart beat period(namely the R–R interval).A number of zeros are padded to the end of each heart beat data sequence, so that the length of each segment becomes uniform .Because the length of each heartbeat is coded, the number of padded zeros need not be preserved for decoding process. In order to improve the coding efficiency further, estimate the mean heart beat period from some initial cycles of the ECG data and preserve or send it to the decoder. The difference between the mean heart beat period and each heart beat period is coded.

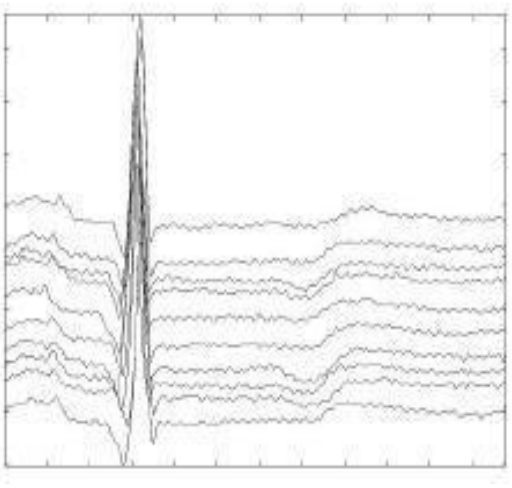


Figure 4. 2-DECGdata array after cutting, padding and aligning

Amplitude normalization brings further similarity between the ECG data. Each sample of a heart beat is divided by the magnitude of the largest sample of that beat. This makes the highest amplitude sample of each beat equal to one. Thus, the variations between the magnitudes of different segments are decreased. To improve the coding efficiency, estimate the mean R-peak value from some initial values of R peaks and preserve or send it to the decoder. The difference between the mean R-peak value and the value of each R peak is noted.

2.3 Iterpolation and decemation

The process of increasing the sampling rate is called interpolation. Interpolation is up sampling followed by filtering. The process of decreasing the sampling rate is called decimation. Decimation is down sampling with appropriate filtering. Interpolation involves inserting new samples between existing samples of a sequence with values derived from the values of the existing samples. Interpolation is up sampling followed by appropriate filtering. so y(n)obtained by interpolating x(n),is generally represented as:

$$y(n)=x(n/L)$$

The simplest method to interpolate by a factor of L is to add L-1 zeros in between the samples, multiply the amplitude by L and filter the generated signal, with aanti-imaging low pass filter at the high sampling frequency.

Decimation is the exact opposite of interpolation. To decimate or down sample a signal x(n) by a factor of M implies collecting every Mth value of x(n)to a new signal. This is given by:

$$y(n)=x(Mn)$$

Down sampling by an integer factor M indicates retaining one sample and discarding the remaining M-1 samples and this is done forever yM samples.ie,It involves the deleting of samples of a sequence with the values derived from values of existing neighbours.Instead of adding signals like in interpolation, signals carrying information about neighbours are deleted, once it is ensured that all the existing samples is enough to reconstruct the signal.

2.4 D wavelet Decomposition

In technical literature, a number of time frequency methods are available for the high resolution signal decomposition. This includes the short time Fourier transform (STFT),Wigner–Ville transform (WVT),Choi–Williams distribution (CWD) and the WT. Out of these, the wavelet transform is the most favored tool by researchers as it does not contain the cross terms in herent in the WVT and CWD methods while possessing frequency-dependent windowing which allows for arbitrarily high resolution of the high frequency signal components.DWT is the appropriate tool for the analysis of ECG signals as it removes the main shortcomings of the STFT; it uses a single analysis window which is of fixed length in both time and frequency domains. This is a major drawback of the STFT, since what are really required area window of short length(in time domain) for the high frequency content[1]of a signal and a window of longer length for the low frequency content of the signal. The WT improves upon STFT by varying the window length depending on the frequency range of analysis. This effect is obtained by scaling as well as shifting the basis wavelet. The continuous wavelet transform (CWT) transforms a continuous[3] signal into a highly redundant signal of two continuous variables translation and scale. The resulting signal is easy to interpret and valuable for time-frequency analysis. In this step we will use wavelet transform.

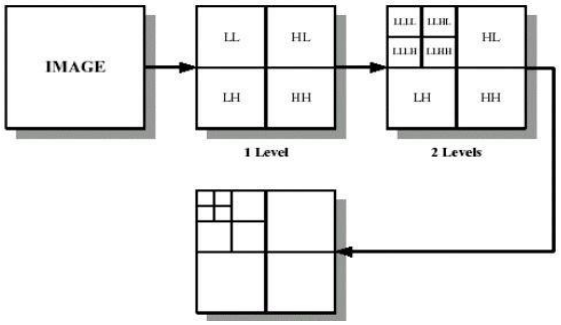


Figure 5.3level2-Ddecomposition

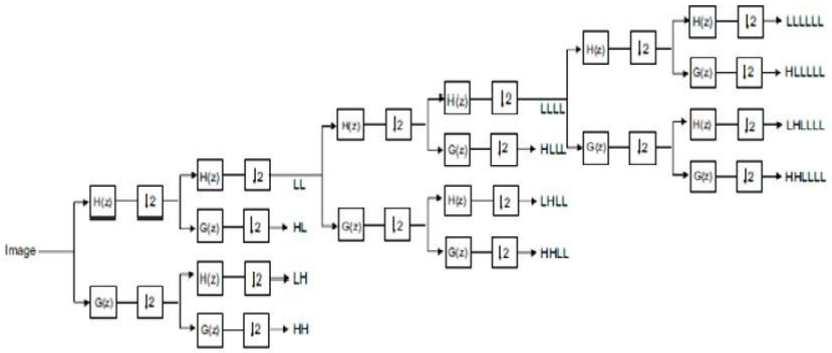


Figure 6.3 level 2-Ddecomposition

Here first, constructed 2-D ECG is decomposed by using 2D wavelet transform upto three levels.After that,the coefficients in each subband[3]of wavelet tree are thresholded. Thresholding does not create significant distortion in reconstructed signal because of the energy invariance property of orthogonal wavelet transforms.This decomposition is repeated to increase the frequency resolution and the approximation coefficients decomposed with high and low pass filters and then downsampled. This is represented as a binary tree with nodes representing a sub-space with a different time-frequency localization. The tree is known as a filter bank.

2.5 Compression algorithm-set partitioning in hierarchical trees

The SPIHT algorithm is a generalization of the EZW algorithm. The SPIHT algorithm was proposed by Amir Said and William Pearlman. In EZW we transmit a lot of information for little cost[1]when we declare an entire sub tree to be insignificant and represent all coefficients in it with a zero tree root label. The SPIHT algorithm uses a partitioning of the trees (which in SPIHT are called spatial orientation trees) in a manner that tends to keep in significant coefficients together in larger subsets. The partitioning decisions are binary decisions[1] are transmitted to the decoder, providing a significance map encoding that is more efficient than EZW.In fact, efficiency of the significance map encoding in SPIHT is such that arithmetic coding of binary decisions provides very little gain.

Thresholds used for checking significance are powers of two, so in essence the SPIHT algorithm sends the binary representation of the integer value of the wavelet coefficients. A sin EZW,the significance map encoding, or set partitioning and ordering step,is followed by a refinement step in which the representations[1]of the significant coefficients are refined.

– O(i,j):off spring of node (i,j)
– D(i,j):all descendants of node (i,j)
– L(i,j)=D(i,j)-O(i,j)
– H:tree roots

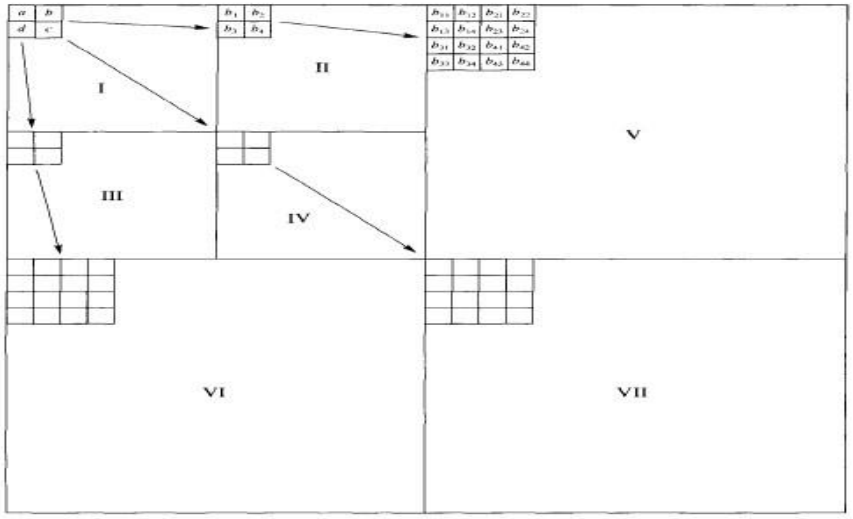


Figure 7.SPIHT wavelet transform example

The algorithm makes use of three different lists: the list of insignificant pixels(LIP),the list of significant pixels(LSP),and the list of insignificant sets(LIS).LSP and LIS lists will contain the coordinates of coefficients, while the LIS will contain the coordinates of the roots of sets of type D or L. We start by determining the initial value of the threshold. We do this by calculating

$$n = [\log 2\ C_{max}]$$

where C_{max} is the maximum magnitude of the coefficients to be encoded. The LIP list is initialized with the set H. Those elements of H that have descendants are also placed in LIS as type D entries. The LSP list is initially empty.

•LIP-list of insignificant pixels

• LIS-list of tree roots (i,j)of insignificant descendant sets D(i,j)(TypeA) or insignificant descendant of offspring sets L(i,j)=D(i,j)-O(i,j)(Type B)

•LSP-list of significant pixels.

2.6 Reconstruction

In the decoder side, the same process is running. The only difference is that the significant/ insignificant decisions found in the encoder by comparing the coefficients to a threshold are input to the decoder. Since the lists are initialized identically, they are formed in the decoder exactly as in the encoder. In there finement pass,the threshold is added to the significant coefficients, instead of subtracted.(The addition or subtraction of threshold is equivalent to adding or removing a bit in a bit plane representation of the coefficient's magnitude.)

Note that the encoding and decoding are comprised of simple operations : comparison to threshold, movement of co-ordinates to lists, and bit manipulations. There are no complex calculations needed for modeling and training prior to coding. The only search is the single search for the initial threshold. The method is completely self-adaptive, always finding the

most significant bits of the largest coefficients and sending them before those bits of smaller coefficients.The method is also extremely efficient, as it has the capability to locate large descendent sets with maximum magnitude smaller the final threshold and representing with a single

3. Experimental result

- Compression ratio(CR) : It is defined as the ratio of total number of bits used to represent the digital signal before and after the compression.

- Mean squared error(MSE) : It is defined as the mean squared error. Given a noise-free m×n monochrome image I and its noisy approximation K,MSE is defined as:

$$MSE = \frac{1}{m\,n}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2$$

- Peak signal to noise ratio(PSNR) : It is the ratio between maximum possible power of a signal and power of corresponding noise that affects the fidelity of its representation. It is usually represented in logarithmic decibel scale. The signal in this case is the original data and noise is the error introduced by compression.

$$\begin{aligned} PSNR &= 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \\ &= 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \\ &= 20 \cdot \log_{10}(MAX_I) - 10 \cdot \log_{10}(MSE) \end{aligned}$$

Here, *MAXI* is the maximum possible pixel value of the image
Our simulation results are
MSE = 5.5342e-04
PSNR= 48.2475
CR= 1.7104

4. Conclusion

In literature,numerous ECG compression methods have been developed.They may be defined either as reversible methods(offering low compression ratios but guaranteeing an exactor near- lossless signal reconstruction), irreversible methods (designed for higher compression ratios at the cost of a quality loss that must be controlled and characterized),or scalable methods(fully adapted to data transmission purposes and enabling lossy reconstruction). Choosing one method mainly depends on the use of the ECG signal. In the case of the needs of a first diagnosis, a reversible compression would be most suitable. However, ifcompresseddatahastobestored on low-capacity data supports, an irreversible compression would be necessary. Finally, scalable techniques clearly suit data transmission.

In this project, an ECG signal is compressed effectively. It is a new hybrid electrocardiogram (ECG) data compression technique. Firstly, in order to fully utilize the two correlations of heart beat signals,1-DECG data are segmented and aligned to a 2-Ddata arrays. Secondly, 2-D wavelet transform is applied to the constructed 2-D data array. Thirdly, the set partitioning

hierarchical trees (SPIHT) method is modified, according to the individual characteristic of different coefficient sub band and the similarity between the sub bands. Finally, a hybrid compression method of the modified SPIHT is employed to the wavelet coefficients. Records selected from the MIT/BIH arrhythmia database are tested.

SPIHT is a simple and efficient algorithm with many unique and desirable properties. One of the stand out features of the SPIHT algorithm is idempotency,that is,lossless recompression at the same bitrate. The algorithm is multiresolution scalable which highlights the ability of the encoder or the decoder to track resolution of bits automatically. The lack of complexity makes SPIHT a very convenient tool for compression ofsignals, especially ECG signals. The entire algorithm is based explicitly on 3 specificlists: List of significant and significant pixels and list of tree roots.Another important feature of the algorithm is there finement pass used.After the completion of the first pass, lossless compression is achieved. Low order bits are more efficiently coded by switching from bit plane transmission at some given highrate there by improving the overall efficiency over a constant value. Due to this, efficient reconstruction is also available which is managed by truncation of file.In other words it is scalable in fidelity. The color planes are coded together. Bits are not allocated in memory locations explicitly providing the algorithm flexibility. Due to this property, different bits will have different transforms suiting the irrequirement. Overall, the algorithm is robust and efficient than other algorithms used for compression of ECG signals.

Acknowledgment

The author would like to thank Project Guide Mr. Nishanth Krishnan(Asst. Prof.ECE) and Mr. Dileep P(Asst.Prof.ECE) whose constant persistence and support helped us in the completion of this report.

References

[1] Introduction to datacompression,Khalid Sayood,Morgan Kaifman Publisher.1996,ELSEVIER.

[2] Z.LU,D.Youn Kim and W.A.Pearlman"Wavelet compression of ECG Compression by SPIHT" IEEE Trans,biomed engg.Vol 147,no.7,July 2000.

[3] ECG Compression with Thresholding of 2-D Wavelet Transform Coefficients and Run Length Coding,Tahere Izack Mohammadpour Faculty of Electrical & Computer Engineering, Noushirvani University of Technology,Mohammad Reza Karami Mollaei Faculty of Electrical &Computer Engineering,Noushirvani University of Technology

[4] Short-Time Fourier Transforms and Wigner–Ville Distributions Applied to the Calibration of Power Frequency Harmonic Analyzers,PaulS.Wright

[5] Hanwoo Lee,Kevin MBuckley ECG data compression cut and align beata pproach and 2D transform"IEEE Translations on Biomed Engineering,Vol.46,Issue no.5,May1999.

[6] AliBilgan,Member, IEEE, Michael W.Marcellin,Fellow,IEEE and MariaI.Altbuch,Member IEEE"Compression of Electrocardiogram Signals using JPEG 2000",IEEE Transactions on Consumer Electronics,Vol.49,Issue no:4,November2003.

[7] Implementing of SPIHT and subband energy compression(SEC) method on two dimensional ECG compression:A noval approach,Mohammad Rezazadeh,Hassan Moradi,conf Proc IEEE Eng Med BoilSoc,2005.

Authors

AnjalyJoseph Treceived BTechgraduation from university of CUSAT in electronics and communication. Now pursuing M-Tech from A P J Abdul Kalam Technological University in VLSI and embedded systems.

Arun K L is a faculty member of Mar Athanasius College of Engineering, Kothamangalam, Kerala, India. He received his B.Tech from Mahatma Gandhi University, Kottayam and M.Tech degree from the Karunya University, Coimbatore, India. His current research focus is in the area of VLSI Signal Process.

DIGITAL RECEIVER SUBSYSTEM USING DDSFREQUENCY SYNTHESIZER

Ahalya R S[1]&Mary Joseph[2]

[1]Department of electronics, Mar Athanasius College Of Engineering,
A P J Abdul Kalam Technological University, Kerala, India
[2]Associate Professor, Department of Electronics &Communication Engineering
M.A.College of Engineering, Kothamangalam

ABSTRACT

An important aspect of Electronic Warfare is the development of Radar Warning Receivers (RWR). Radar Warning Receiver/Electronic Warfare Receiver provides superior performance to protect tactical, transport and special mission aircraft in today's modern electronic combat environment. Radar warning receiver (RWR) systems detect the radio emissions of radar systems. Their purpose is to issue a warning when a radar detects a threat. The warning can then be used, to find out the detected threat.The performance and design aspects of radar receivers will be considered in this project. For this purpose, the radar receiver will be defined as that assemblage of components within the radar system which is required to detect, amplify, and present the desired information as gathered at the radar location. This project is focused mainly on electronic warfare radar systems and describes the receiver front end design, as well as the software algorithms used to determine signal characteristics.

KEYWORDS

RWR, RF signals, pulsed radars,DFT,FFT

1. INTRODUCTION

In Electronic Warfare (EW), radar is used to detect vehicles, ships, and aircraft, and to guide weaponry. Radars can also be used to detect threats and can be called as Electronic Counter Measures. For every newly developed threat, there is the need to neutralize that threat. This is a computer controlled radar warning receiver (RWR) that provides detection and display of RF signals from the threats. The digital system usually includes time-frequency transfer. The following parameters are used to identify and locate the origin of the threat signal:

DOI: 10.5121/ijci.2016.5233

- FREQUENCY
- PULSE WIDTH
- PULSE REPETITION RATE

For the scope of this project, it determines the initial detection and classification of the radar signal. The frequency, pulse width, and pulse repetition rate are the parameters, which are to befind out. A radio frequency (RF) signal is given to the Digital Radar Receiver and processed, in order to determine the above three parameters. Future applications to this project are to further analyze and manipulate the received RF signal to jam the enemy's communication systems. Here it can be as simple as detecting the presence of energy in a specific radar band[2]. For more critical condiions, such as military , the systems are often capable of classifying the source of the radar by the signal's strength[5], phase and waveform type, like pulsed power wave or continuous wave with amplitude modulation or frequency modulation[4] . The signal's strength and waveform will be helpful to estimate the most probable type of threat the detected radar possess.

2. WORKING PRINCIPLE

The Radar warning receivers functional elements can be divided into six elements.

Reception	: interfacing the equipment to electromagnetic environment.
Conversion	: amplification, up/down converting and filtering.
Detection	: receiver specific signal detection and measurement.
Extraction	: extraction of signal characteristics from time and frequency domain.
Analysis/control	:deinterleaving signal characterization and identification.
Indication	: display generation and parameter distribution.

The RF signal from the target is initially down converted to IF signal and the analogue input signal is transferred into Time discrete amplitude samples. Maximum IF bandwidth is defined by sampling frequency and it must be greater than 2 times IF bandwidth. Then the Time domain samples are combined into frames each of N samples. These samples are used to perform a Fast Fourier Transformation (FFT). Time effective calculation can be performed if number of samples = 2^N. And then the analog signal is converted to pulse descriptor words (PDW). Then for each valid signal, the information such as pulse width, coarse and fine frequency, signal amplitude, phase etc are derived from pulse descriptor words. These parameters are derived using parallel and sequential correlation. The product of sampling rate and FFT length defines the Time

Resolution. Frequency selection and S/N can be improved and time resolution can be decreased with higher FFTs.

3. SYSTEM DESCRIPTION

3.1Block Diagram

The system block diagram is shown in figure. The user input is a software input, which is used to set the LO range. The pulse radar RF signal is mixed with the frequency synthesizer signal, which acts like a local oscillator. The internal signal, IF, is the down converted pulse radar signal which is fed into the A/D board which is connected to the PC. Then the signal is processed using DSP methods, and if a radar signal is present the signals characteristics are sent as outputs.

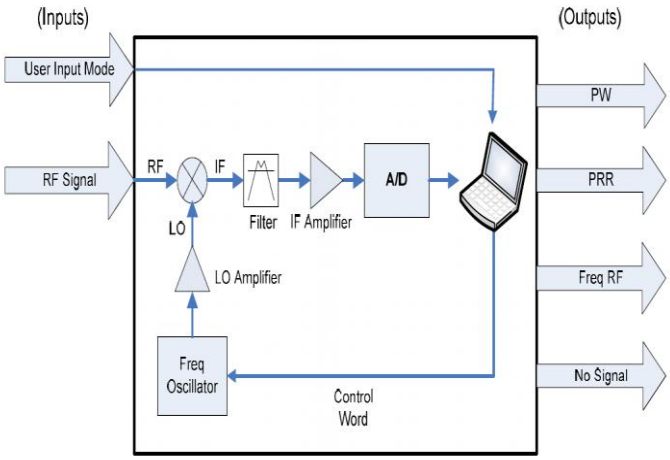


Fig:1 Block Diagram

The function of the system is to detect a pulsed radar signal over a large frequency and power range. The frequency ranges from 50 to 120 MHz and the power level can be as low as -20 dBm. The mixer multiplies the incoming RF signal of range 50-120 MHz with the LO signal and produces the down converted IF signal. A control word is sent to frequency synthesizer through the parallel port, which determines its output. In order to boost the signal from the frequency synthesizer a local amplifier is used. The amplifier is needed to get the amplitude from the frequency synthesizer up to the 7 dBm level needed for the mixer. This signal is fed to the mixer. Thus the signal output from the mixer is adapted to detector IF range. Then the down converted signal is conditioned and for that the signal is then fed to a low pass filter of cut off frequency 200 kHz. after eliminating noise the filtered signal is fed to an amplifier which make up the signal loss

encountered in various parts of the system and also compensate the losses due to filter. The amplifier is used to set the voltage level of IF signal for the best data acquisition. And then conditioned signal is then fed to analog to digital converter. The ADC converts the analog signal to Pulse Descriptor Words which the information about pulse width, amplitude, phase and frequency of the input radar signal. The PDW from ADC (Data acquisition card) is fed to processor and the signal is processed using DSP methods. The signal processing involves taking the FFT of the signal and detecting whether a radar signal is present or not. If the radar signal is absent it outputs no signal and the frequency synthesizer output is incremented and if a radar signal is present it outputs the information about pulse width, pulse repetition rate and frequency of the input radar signal.

Data acquisition subsystem consists of an analog filter, amplifier, and the data acquisition board. The IF signal is filtered using a low-pass filter with a cut off frequency 200Khz , which passes the intermediate frequencies with little attenuation, but attenuates signals with frequencies above 200 KHz. After filtering, the signal is amplified using a 741 amplifier. The amplifier can add a 20 dB of gain to compensate for the attenuation caused by the filter.

3.1.1Frequency Synthesizer

A frequency synthesizer is a system that generates a number of precise frequencies from a single reference frequency.

A frequency synthesizer can replace the array of crystal resonators in a multichannel radio receiver frequency synthesizers are relatively inexpensive and can be easily controlled by digital circuitry. Hence they are used in new communication system designs. Frequency synthesizers are found in many devices, including radio receivers, mobile telephones, radiotelephones, walkie-talkies, satellite receivers, GPS systems, etc. to obtain the required output signal it can perform frequency multiplication and frequency division. The control word is sent to the frequency synthesizer, through the parallel port, and it decides the frequency of the synthesizer output. Foe power boosting a10 dB amplifier is used. The amplifier should get a amplitude up to 7dBm from the DDS frequency synthesizer for the mixer. The DDS 9854 is used as the frequency synthesizer. It uses a technique called direct digital synthesis creating a stable sinusoidal wave form. The maximum frequency of the synthesizer is 120 MHz to prevent aliasing.

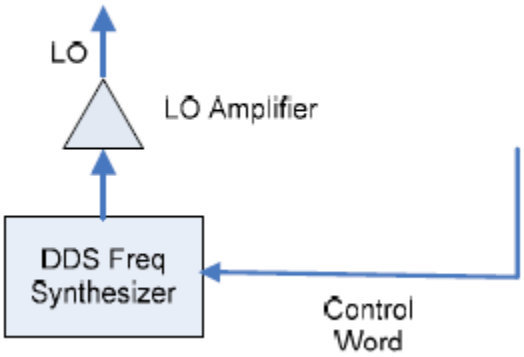


Fig 2 Frequency synthesizer subsystem

3.2DSP Flow chart

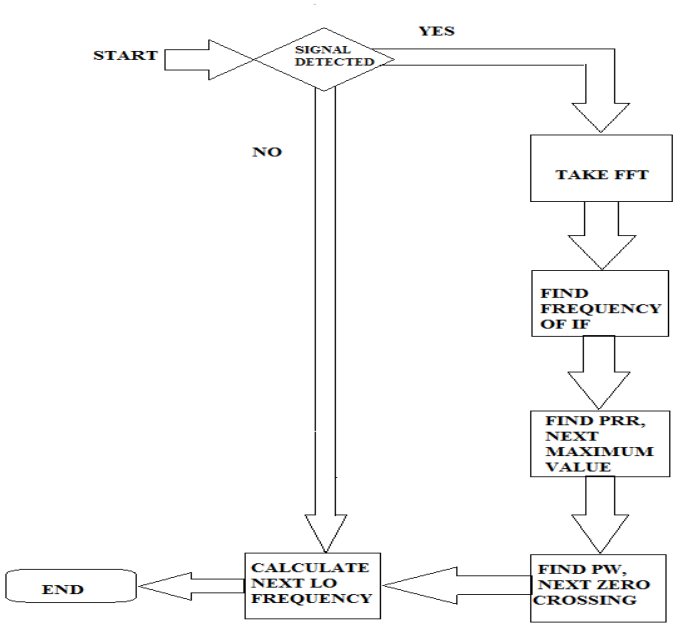


Fig: 3 DSP FLOW CHART

A fast Fourier transform (FFT) is an algorithm to find out the discrete Fourier transform (DFT) and its inverse. A fast Fourier transform converts time to frequency and vice versa; an FFT rapidly computes such transformations by factorizing the DFT matrix into a product of mostly zero factors.

The DFT is obtained by decomposing a sequence of values into components of different frequencies. An FFT is a way to compute the DFT of *N* points. FFTs are of great importance to a wide variety of applications, from digital signal processing and solving partial differential equations to algorithms for quick multiplication of large integers.

Figure 3 shows how the radar signal is processed in Mat lab.

- The process checks whether a signal is detected
- If there is no signal, the frequency synthesizer is incremented and the program exits.
- If a signal is there, then its FFT is calculated.
- The FFT will be having an array whose points correspond to various magnitudes.
- A frequency array is made based on the number of points in the FFT and the sampling rate.
- To calculate the IF frequency the point with the maximum value in FFT array is found and is correlated with same point.
- To calculate the PRR the point with largest magnitude next to previous value is found.
- PRR is obtained by subtracting this value from he IF frequency.
- Value of local minimum when subtracted from the IF frequency gives the pulse width.

4. RESULTS AND DISCUSSIONS

Before signal processing signal has two level of detection. Figure illustrates this detection process. The first level1 detection checks whether the signal's amplitude is above the preset threshold value. The threshold value is 200mV. It corresponds to about -20 Db m signal level on the input. Once Level 1 is over, Level 2 Detection starts. During this level Detection the system takes an FFT with a small number of points and then searches for the spikes on either side of the main spike. If these spikes are detected then a pulse radar signal has been detected and the signal can continue to be processed.

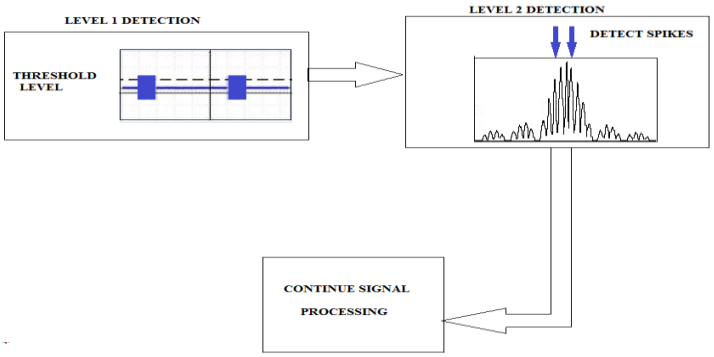


Fig:4 Signal Detection

Level 1 and Level 2 works successively to detect whether a radar signal is been detected. This is to speed up the system.

The illustration of a pulse radar signal is shown in Figure 4. By processing the signal in,time domain characteristics can be determined. The following is the list of outputs and how they relate to Figure

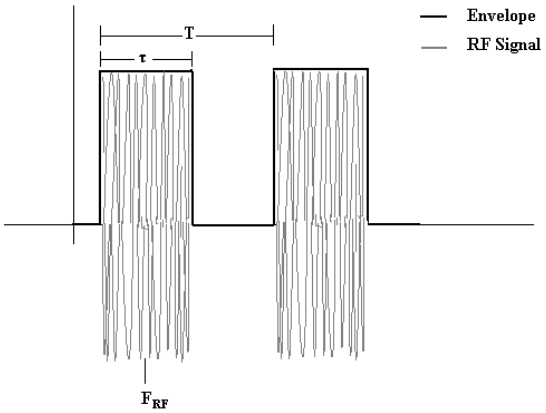


Fig 5 Time representation of the pulsed radar signal

- ζ Gives the Pulse Width
- The inverse of period T in the time domain gives the Pulse Repetition Rate (PRR)
- The frequency under pulse envelope is the frequency RF denoted as F_{RF}.

4.1 EXPERIMENTAL RESULTS

In Figure 6 the time domain of a sampled signal can be seen. The frequency of the RF signal is 100 MHz

- The Local Oscillator frequency is obtained as 99.9 MHz
- Calculated Pulse repetition rate is 500 Hz
- Calculated value of Pulse width is 400uS corresponds to frequency of 2500 Hz.

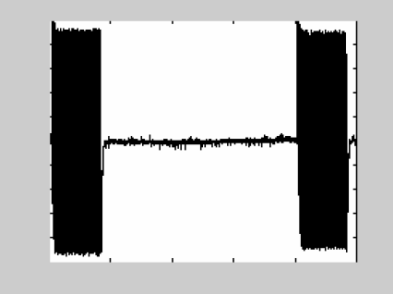


Fig 6 Sampled pulse Radar signal in Mat lab

Figure 7 is the results for the signal in the frequency domain from the FFT, which is used to determine the characteristics of the radar signal.

- The IF frequency is obtained as 100 KHz which is expected.
- Thus, LO + IF = 100 MHz which as expected.
- Also the PRR and pulse width are the expected values.

5. CONCLUSION

The Project was an overall success. The main objective was to make sure the system was able to automatically scan the entire frequency range and detect and characterize a pulse radar signal within the range. Also there was no error when determining the PRR of the signal. The system could not detect much large pulse width but it doesn't cause much problem because larger pulse width may cause wastage of power.

ACKNOWLEGMENT

The author would like to thank Mrs.Renu Agarwal(Scientist Engineer SD), Vikram Sarabhai Space Centre and Anu Mohandas(Assistant professor), Pankaja kasthuri College Of Engineering And Technology for their guidance and support.

REFERENCES

[1] Couch, Leon W., Digital and Analog Communication Systems, 3rd edition (New Jersey: Prentice Hall, Inc. 1990)
[2] David.M.Pozar, Microwave Engineering New York John Wiley & sons, second ed.. 1998
[3] David.M.Pozar, Microwave Engineering New York John Wiley & sons, second ed.. 1998
[4] Haykin, Simon, Communication Systems, 3rd edition (1994)
[5] Jon Hagen , Radio frequency Electronics
[6] Mahafza, Bassem R., Radar Systems Analysis and Design Using Matlab, 2nd edition (New York: Chapman and Hall/CRC 2005)
[7] P.E. Chadwick, High Performance IC Mixers, IEEE Conference on Radio Receivers and Associated Systems, Leeds, 1981, IERE Conference Publication No. 50.

Authors

Ahalya R S: Received her B.Tech degree in Electronics And Communication engineering from University Of Kerala in 2014. Currently pursuing her M.tech degree in VLSI And Embedded System from Dr.APJ Abdul Kalam Technological University, Kerala.

Mary Joseph: received M.Tech Degree in Microwave and Radar from Cochin University of Science and Technology (CUSAT), Kochi, India, in 1997. Currently she is working as Associate Professor in M. A. College of Engineering, Kothamangalam. She has joined in M. A. College of Engineering in 1991 as Assistant Professor. In between she worked at Birla Institute of Science & Technology-Pilani's (BITS-PILANI) Dubai Campus for 9 years as Assistant Professor during 2000-2008. Her Research interests include Microstrip Antennas and Uniplanar Antennas.

INTENTIONAL BLANK

EMG ANALYSIS AND CONTROL OF ARTIFICIAL ARM

Anjali Raghavan[1] & Prof. Sunny Joseph[2]

[1]Department of Electronics and Communication, Mar Athanasius College Of Engineering,
A.P.J Abdul Kalam Technological University,Kerala, India
[2]Associate Prof, Department of Electronics and Communication,
Mar Athanasius College Of Engineering, Kothamangalam, Kerala, India

ABSTRACT

Robotic hand for prosthetic applications is a unique structure intended to be driven by electro myographic (EMG) signals captured from human body. The main characteristic of this robotic hand is its actuation system, which is based on the behaviour of EMG signals. The direct relation between signal and actuation system lends itself well to interpreting the EMG signals from the muscles into effective task execution, with the goal of helping the user to achieve a good approximation of the full capabilities associated with the human hand, without compromising strength, dexterity, appearance, or weight; which are common issues associated with prosthetic hands. EMG signal capturing capability was added to control the DC motors. The emg signal was then filtered and scaled using MAT lab to a value representing the amplitude of the EMG signal, which was then used to control the direction of the motors. The controller is a simple feed forward system in the project but provides the appropriate framework to integrate more elaborate control schemes and EMG signal conditioning. The goal is to create a direct relation between the EMG signals.

KEYWORDS

Robotic hand, prosthetic, EMG, DC motor, MAT lab

1. INTRODUCTION

There have been many different approaches taken in the development of an effective prosthetic hand. These varying strategies often find themselves focusing on one of the following categories: implementing a new actuator type, developing a more effective kinematic structure, integrating effective compliance, generating effective control strategies, and interpreting/conditioning input signals. Advances in these areas have resulted in robotic hands that perform many tasks with a high similarity to that of the human hand, such as the DLR hand, I-Limb hand, Shadow hand, and Fluid hand to name a few. However, a prosthetic hand that is nimble, quick, strong, lightweight, quiet, and efficient has yet to be achieved. The primary reason for the current state of prosthetic hands has been the complexity associated with the human hand as a result of its multiple bones and joints. This is further compounded by the fact that the human hand as a functioning unit does not just embody the palm and its digits but also the wrist, forearm muscles, nervous system, and the body's energy generation system. As a result, the entire prosthetic hand actuation structure (inputs, power, strength, kinematics, etc.) must fit in a significantly reduced volume compared to

DOI: 10.5121/ijci.2016.5234

the human hand that it is replacing. To address some of the challenges described above, this project implements a unique perspective of the muscles signals in the human forearm and proposes a novel design and parallel actuation structure that complements this perspective. The goal is to create a direct relation between the forearm's EMG signals and the actuation system, in order to help the user achieve a good approximation of the full capabilities associated with the human hand in a compact design.

1.1 Motivation of the work

Cerebrovascular accident (CVA) also known as stroke is being the major cause of long term disabilities. The only way to aid stroke patient is restoring loss through rehabilitation training, which can be enhance using rehabilitation device. As referred, the training can be enhanced using exoskeleton systems by guiding patients in relearning motions on correct trajectories, or by giving them force support to perform certain motions. However, mathematical representation of human lower extremities is required in order to build the device. From the past studies, it is clear that EMG system provide changes of neuromuscular system while the muscles experiencing fatigue. Thus here relationship between EMG signal and human movement is studied.

2. ELECTROMYOGRAM

Electromyography (EMG) is a technique for evaluating and recording the electrical activity produced by skeletal muscles.[2] An electromyograph detects the electrical potential generated by muscle cells when these cells are electrically or neurologically activated. The signals can be analyzed to detect medical abnormalities, activation level, and recruitment order or to analyze the biomechanics of human or animal movement. Understanding EMG implies understanding muscles and the way they generate bioelectric signals. It also implies how specific mechanisms influence the signals as well as how the signals reflect certain mechanisms and phenomena and allow their identification and description. The EMG signal is a representation of the electric potential field generated by the depolarization of the outer muscle membrane. Its detection involves the use of intramuscular or surface electrodes that are placed at a certain distance from the sources. The surface EMG signal is an effective and important system input for the control of powered prosthesis. This control approach, referred to as myoelectric control, has found widespread use for individuals with amputations or congenitally weak limbs. The filtered form of obtained EMG signal can be used for these purposes.

Understanding EMG Signals implies the understanding of muscles and the way they generate bioelectric signals. A basic description of the physiological system that generates the EMG signal is presented here.

2.1 Prosthetic Hand Design

The design is dimensionally consistent with that of an average male human hand and possesses the same degrees of freedom.[3] The anthropomorphic aspect of the hand is intended to enhance the amputee's acceptance and usability. The DIP and PIP joints of the finger and the IP and MCP joints of the thumb are coupled. This is achieved by connecting a single actuator to both the PIP joint (bevel gears) and DIP joint (pulley connection on metacarpal phalange). The movement associated with a region is achieved by two DC motors. The DC motors actuating the coupled

DIP/PIP joints of the finger and IP/MCP joints of the thumb are embedded in the proximal phalange of the finger and the metacarpal phalange of the thumb. The DC motor in the metacarpal phalange of the finger actuates the horizontal degree of freedom of the MCP joint. The DC motor at the base of the thumb actuates the CMC joint to obtain an approximation of the abduction/adduction motion. The second degree of freedom of the finger's MCP joint (abduction/adduction) is only subject to compliance without actuation. The second degree of freedom in the thumb's CMC joint (flexion/extension) is actuated by another region actuation structure. The actuation structure corresponding to other region for the finger includes a light cable that passes over two restraining shafts in the MCP joint of the finger, coils in the proximal phalange, and embeds in the middle phalange. The string is kept in light tension by a tension unit at the back of the hand while that region actuation structure is active. When other region actuation is required the shape memory alloy actuates a spring loaded cam which in turn pinches the string between itself and a roller beneath it. As the shape memory alloy continues to actuate, the cam introduces the additional force required for other region tasks. At task completion the electric signal causing the shape memory alloy to heat up is stopped and the DC motors and cam spring extend the shape memory alloy back to its original state. The thumb's region 2 actuation structure is similar to that of the finger's region 2 actuation structure. However, unlike the finger, this structure actuates the degree of freedom at the CMC joint that is not actuated by the DC motor. This is based on the observation that this degree of freedom is more apposing of the fingers during tasks that would require additional force. The design shown in this section has been manufactured using a rapid prototyping machine.

2.2 Human-Machine Interface: Processing of The EMG Signal

The user can control the grasping task adopting a consolidated technique, the processing of the EMG signals[2]. Two EMG signals generated by two antagonist muscles of the forearm devoted to the wrist flexion-extension movement (the extensor carpi radialis and the flexor carpi radialis) are acquired and processed by two small boards (each 0.8" by 1.5" in dimension), in the first board, the signal is rectified and filtered (pre-processing phase); in the second one, the pre-processed EMG-signal is used to control the actuators of the RTR II hand (low level control). The function of the two boards is to interpret the user's intention and to send the appropriate commands to the actuators. Two small boards are used to preprocess the EMG signal and to control the actuators of the RTR II hand.

2.3 EMG Pre-Processing

The first board receives as inputs two EMG signal measured by 2 Delsys DE2.3 differential Ag electrodes, designed with a built-in gain of 1000 V/V and a built-in pass band filter from 20±5 Hz to 450±50 Hz. These signals are then rectified (by a full-bridge precision rectifier with adjustable gain) and low-pass filtered through a Butterworth filter (2nd order, pass-band edge frequency: 8 Hz, pass-band attenuation: 0.5 dB, fp=13.56 Hz), whose transfer function is:

$$T(s) = T0 \frac{A}{Bs2+Cs+D}.$$

Where T0 = 1, A = 7.23E3, B = 1, C = 1.2E2, D =7.2E3.

The two rectified signals are then compared in order to detect the direction of movement (i.e. extension=open, flexion=close; SGN signal); moreover, the signal is also compared with an adjustable threshold with a regenerative comparator, in order to detect if there is a movement or not (EDG signal).

2.4 Low Level Control of the Hand

The second board receives 3 signals from the first one: activation (EDG), direction (SGN) and amplitude of the movement (AMP).[1] The first signal (EDG) is used to understand if there is a movement or not; the 2nd (SGN) to 2232 choose between opening and closing, and the 3rd (AMP) to control the speed of the movement in a proportional way. The second board is composed of 2 blocks: the control block, which is composed by a PIC16F870 microcontroller, and the driver block, which contains 2 drivers capable of driving 2 motors up to 18W each. The control algorithm (in this first prototype) is a simple proportional open loop control. It remains in the stand-by state (low power consumption) until there is a positive edge of the EDG signal. The controller reads the SGN signal (SGN0), and starts to sample the AMP signal (obtained by the rectified difference of the 2 rectified EMG signals). After twait (twait=300msec in this first version of the controlling algorithm) the controller reads once again the SGN signal (SGN1). If SGN1 6= SGN0, it means that the user is changing the control from one motor to the other. Instead, if SGN1 = SGN0, the controller starts to drive the motor of the hand by using the Pulse Width Modulation (PWM) technique. The duty cycle is calculated from the average of the previous 64 samples of the EMG signal. The direction of movement is determined by the SGN signal. During twait the microcontroller samples the EMG signal, and averages the last 128 samples. If this average is greater than a pre-determined threshold for more than 70% of the time, it means that the user wants to move the hand at low speed. The next version of the control algorithm, currently under development, will exploit the force and position sensors, in order to control locally some grasping functions.

3. BLOCK DIAGRAM AND FLOWCHART

3.1Block Diagram

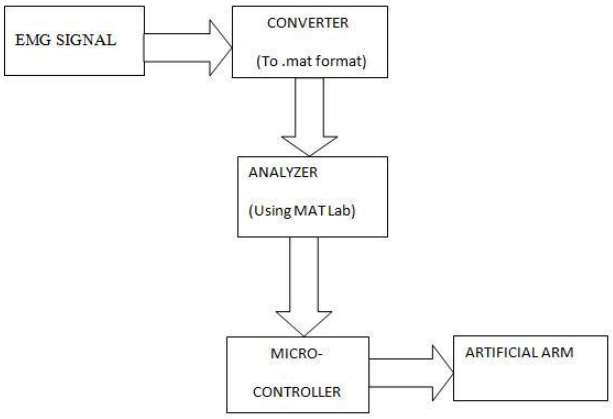


Fig:1 Block Diagram

The EMG signals obtained from human body are converted to wave format. This '.wav' format is converted to '.txt' forma in MAT Lab. This '.txt' format is analyzed using MAT Lab. It is normalised and rectified. This obtained output is given to microcontroller. The power supply given to the circuit using transformer. The regulator IC is regulated the 12v to 5v. This 5v and 12v are used in circuit. The microcontroller connected to motor which drives the artificial part through relays. Two relays are used to function one motor. Here we are using total 4 relays and two motors.

3.2Flow chart

The relays works as H-bridge with motors.Filtering and rectifications done to get a correct signal for the circuit section. If the maximum peak value of signal is less than 0.17, then mat lab output gives a value of 3 to controller and the artificial arm moves left. If the maximum peak value of signal is in between 0.17 and 1.11, then mat lab output gives a value of 1 to controller and the artificial arm picks something. If the maximum peak value of signal is in between 1.11 and 1.2, then mat lab output gives a value of 2 to controller and the artificial arm drops out. If the maximum peak value of signal is greater than 1.2, then mat lab output gives a value of 4 to controller and the artificial arm moves right.

According to the switching of relays motors moves clockwise and anti clockwise. DC motors are the main parts in the circuit section. Based on the rotation of motors, the artificial arm moves. Here the 250v A/C supply is converted to 12v AC first using a 250/12v transformer. Then to 12v DC using a rectifier. This DC voltage is regulated to 5v for the proper functioning of ICs. Most of the ICs work in 5v. There is a chance to dropout the voltage at the time of rectification. To avoid this we are using capacitors in circuit with appropriate values to minimize them.

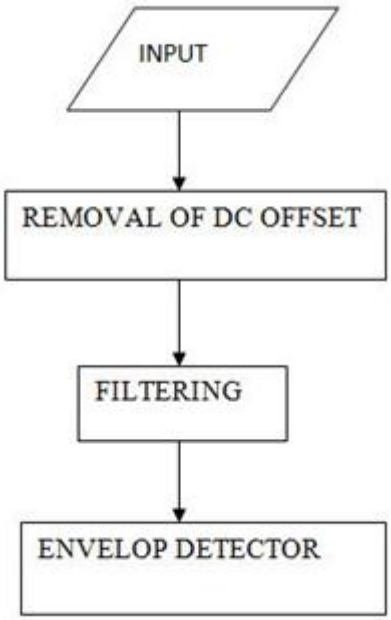


Fig:2 Flow Chart

4.RESULT

Designed and set up the signal processing circuit for "EMG ANALYSIS AND CONTROL OF ARTIFICIAL ARM".Four raw EMG signals took for the functions pick, drop, left movement and right movement. The signals are shown below. We are drawing original signal, rectified signal and the filtered signal here. Signal rectified to take the absolute value. It is easy to work out on peaks from the filtered signal after finding the peaks using function.

4.1 Drop Out

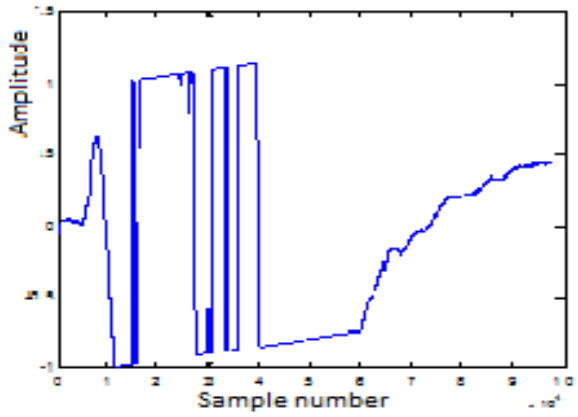


Fig:3 Input Signal for Input

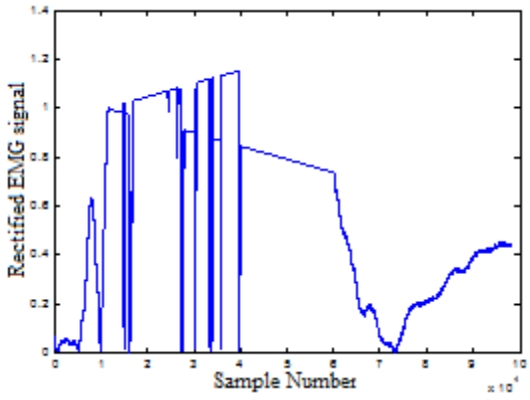


Fig: 4 Rectified Form of Signal

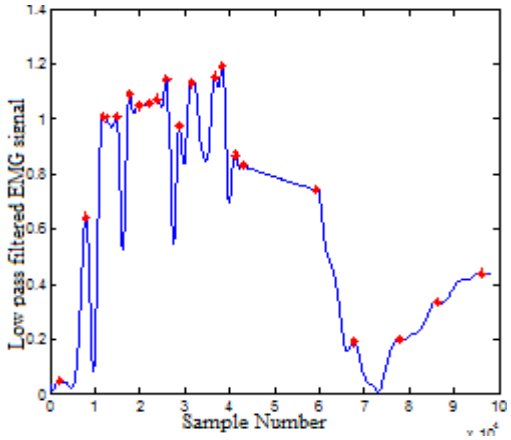


Fig:5 Filtered Form of Signal

4.2 Pick-up

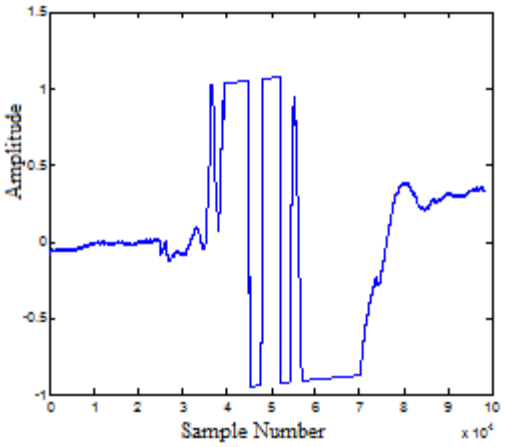


Fig:6 Input Signal To Pick-up

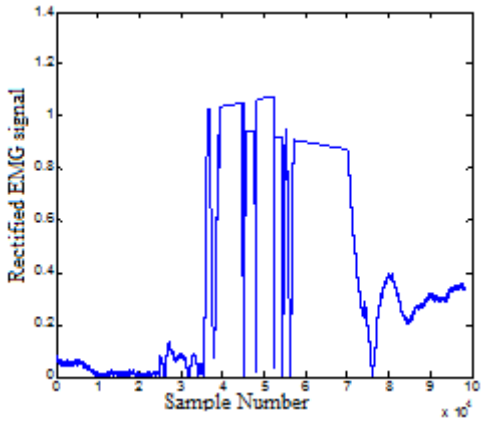


Fig:7 Rectified Form of Signal

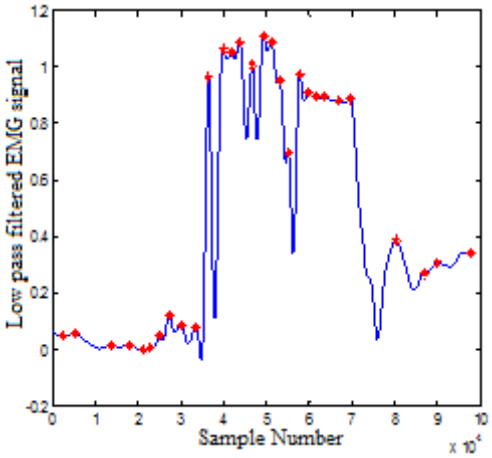


Fig:8 Filtered Form of Signal

4.3 Left Movement

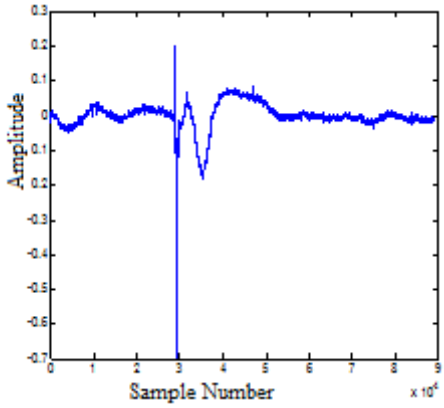


Fig:9 Input Signal For Left Movement

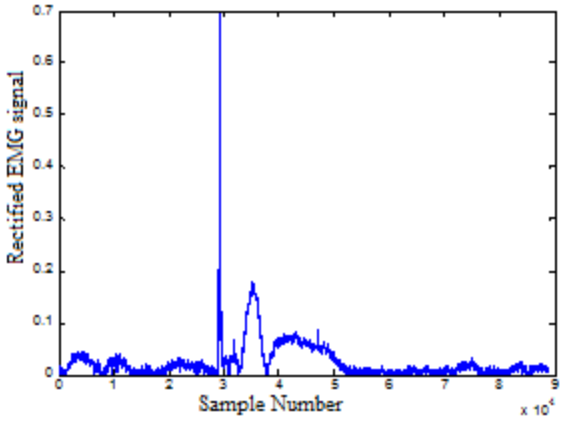


Fig:10 Rectified Form of Signal

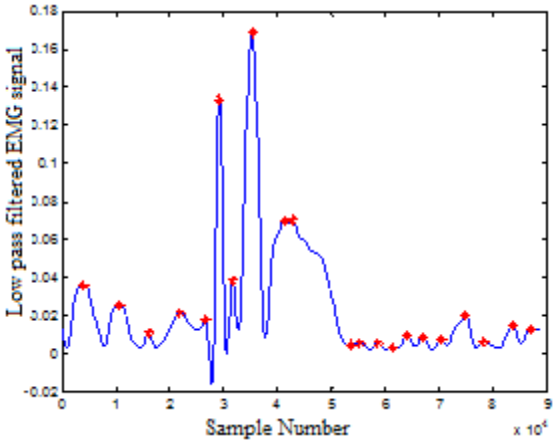


Fig:11 Filtered Form of Signal

4.4 Right Movement

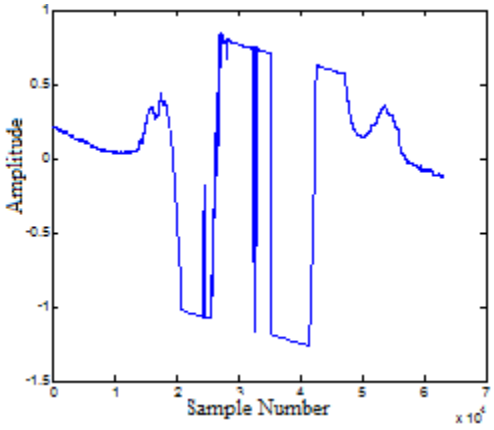


Fig:12 Input Signal For Right Movement

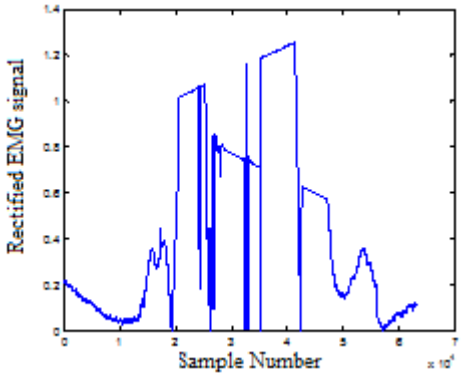


Fig:13 Rectified Form of Signal

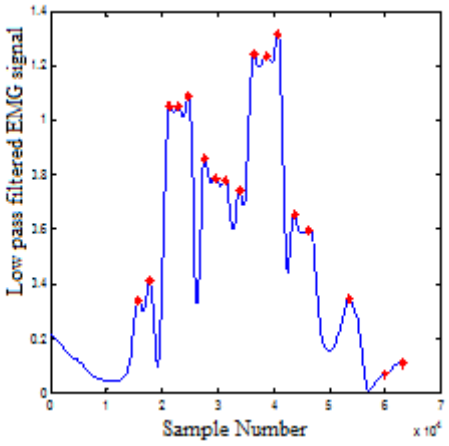


Fig:14 Filtered Form of Signal

Raw EMG signals can be obtained by using a surface electrode placed on arm for various actions such as pick, drop, left movement and right movement. These signals are shown in fig:3,fig:6,fig:9 and fig:12. Then signals are processed. Finally the four filtered signal shown in fig:5,fig:8,fig:11 and fig:14 are fed to MAT LAB. On the basis of which signal is running in MAT LAB, the artificial arm will be performing above mentioned actions.Fig:13 shows the working model of an artificial arm, which consist of a prototype of an arm.

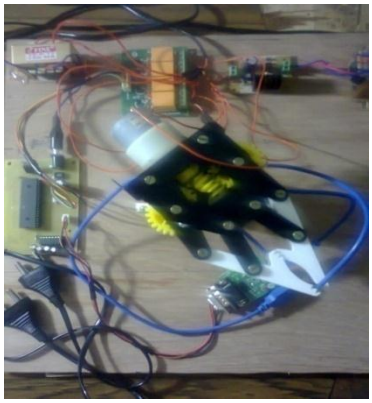


Fig:13 Working Model

5. CONCLUSION

EMG can be used to detect muscle fatigue and in treatment of motor disorder problem. Nonetheless, due to insufficient of mathematical model that can fit for general case, more research need to be done for improvement on available model. From the results, it was found that combination of EMG and GA is sufficient for determining best fit model for human lower limb although improvement is still required. The paper has provided an enhanced understanding of human movement in walking motion, which is essential for building of stroke rehabilitation device.This paper presents a novel design and actuation system for a prosthetic hand. The actuation structure was shown to effectively execute the specific actions similar to those found in

the human hand. Results also validated that the design could be effectively driven by an EMG signal .Complete testing of the actuation system's performance will require the expansion of the EMG inputs to perform different grasps and manipulations.

ACKNOWLEDGEMENT

The authors would like to thank Abdhu Rahiman ,project coordinator and Dr.Shajee Mohan,Head of the Department of applied electronics and instrumentation of Government Engineering College,Kozhikode for the assistance and support.

REFERENCES

[1] Dr. Anthony L. Crawford, Jeffrey Molitor, Dr. Alba Perez-Gracia, Dr. Steve C. Chiu, "Design Of A Robotic Hand And Simple EMG Input Controller With A Biologically-Inspired Parallel Actuation System For Prosthetic Applications".

[2] Tan Chee Weil, S.Parasuraman2 And I.Elamvazhuthi ,"Electromyography(EMG) And Human Locomotion"2012 4th International Conference On Computer Engineering And Technology.

[3] Robertoo Merlittie ,Philip A Parker "Electromyography Physiology Engineering And Non-Invasive Application",IEEE,press.

[4] M.C.Carroza,F.Vecci,F.sebastiani,G.Cappiello,S.Roccella,M.Zecca,R.Lazzarini P.Dairo"Expiremental Analysis Of An Innovative Prosthetic Hand With Proproceptive Sensors" 2003 IEEE International Conference On Robotics And Automation

AUTHORS

Anjali Raghavan recieved B-Tech degree from University of Calicut in Applied Electronics and Instrumentation.Now pursuing M-Tech degree in A.P.J Abdul Kalam Technological University in VLSI and Embedded System.

Prof. Sunny Joseph is the Head of the Department of Electronics and Communication Engineering in Mar Athanasius College of Engineering, Kothamangalam, Kerala, India. He has a teaching experience of 29 years. He received his B.E from Bangalore University and M.Tech degree from Kerala University. His current research focus is in the area of VHDL, high speed digital design and microwave engineering. He is a member of ISTE and FIE.

INTENTIONAL BLANK

HANDS FREE COMPUTER CONTROL

Pooja Antony[1] & Prof. Sunny Joseph[2]

[1]Department of Electronics, Mar Athanasius College of Engineering, Kothamangalam,

A .P.J Abdul Kalam Technological University,Kerala, India

[2]Associate Prof., Department of Electronics & Communication,
Mar Athanasius College of Engineering, Kothamangalam

ABSTRACT

For modern human computer interaction, traditional method based of using keyboard and mouse do not prove sufficient. Especially when thinking of physically handicapped persons, we must try to find the easiest and most comfortable method for human system interaction. The good method for communication process going from man to machine is speech communication. In case of control system, speech communication offers free hand manipulation, which is sometimes very important. In modern era, mouse control has become an important part human computer interaction, which is difficult for physically disabled persons. This paper 'Hands free computer control' presents two systems called as vocal mouse and virtual keyboard. The device vocal mouse will allow the users to work on continuous motion control over computer screen including the virtual keyboard created as GUI. This includes commands consists of nine mouse controlling commands, low level acoustic features are extracted in real time using MFCC. Pattern recognition is performed using minimum feature distance technique. Vocal mouse can be used by users without extensive training which is the main advantage of the system.

KEYWORDS

MFCC, GUI, k-mean clustering, Speech recognition

1. INTRODUCTION

Voice input has a number of potential benefits, especially for physically disabled people. People who are suffering from physical disabilities cannot use standard input devices like keyboard and mouse to access the system, so they have only one option for gaining access to the computer i.e. hands-free input methods[2].

The project "Hands Free Computer Control" mainly presents two systems, Vocal Mouse andVirtual Keyboard.Vocal Mouse is an enhanced voice-based interaction will give benefit to physically handicapped people who find themselves in impairing situations like driving or interacting with computer, hands-free interaction can be more suitable than traditional manual input devices.

DOI: 10.5121/ijci.2016.5235

This system was reliable and speaker independent. Operator training was not required. Any 32-bit Windows software, can access this voice control system. Standard programming languages such as Borland or Microsoft C/C+ + and Visual Basic (VB) and commercial packages such as Lotus Notes and Microsoft Word support this system. In this system, the user is allowed to enter data and to control the software flow by voice command or from the keyboard or mouse.

Main objective of designing Vocal Mouse is to make interactions with existing computer applications possible. This objective put more emphasis on the practical needs of physically challenged people who cannot use existing computer and applications, which have been designed with keyboard and mouse inputs. However, Vocal Mouse cannot not provide the ideal solution from the interaction design perspective. Even then it will give benefit to people who are physically challenged more than the alternative of not being able to access the functionality of computers at all. Another objective of the proposed system is to provide interactions with existing computer applications more effective for any users. The main aim of this system is to use voice input can as an additional input to augment the standard keyboard and mouse interaction for greater control i.e. giving sound input parallel to the keyboard and mouse.The other motivation is to design a new application environments optimized for voice-based control.

2. PROPOSED SYSTEM

The speech command or input can be determined by power of audio signal which can be taken by the help of microphones being connected to the computer itself. Using MATLAB the speech signal is sampled at a rate of 8000 samples/sec according to nyquist criteria i.e. F=2*fm Where F=Sampling Frequency, fm = maximum component of frequency being present in speech signal. Then the sampled signal is filtered off by using band pass filter lying in the range of 300 Hz-4000 Hz, which filters all the audio signals lying below 300 Hz of frequency range. Moreover, it includes the algorithm for the creation of speech parameters which can be achieved by calculating the power of each sampled signals respectively.

Initially, a database is created in which same command is taken many times and thus the average of these will represent a speech template which can be stored in a library. Now the command speech signal that are being received by microphone are then compared with the speech signal being stored in library according to Euclidean`s Distance. Thus, the command will be detected by the system and it will performs the operation accordingly & if it does not match, the system will not perform any operations. The general block diagram of the system is shown in figure2.1.[2]

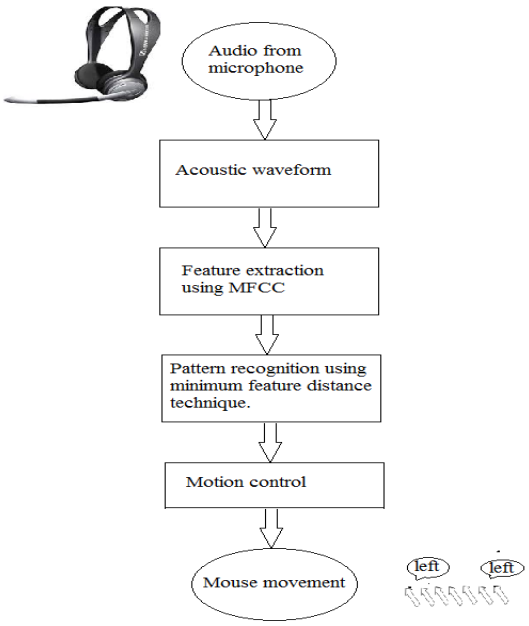


Fig 2.1 Block diagram to represent the working of hands free computer control

This system truly based on speech recognition which can be achieved by the use of Mel Frequency Cepstrum Coefficients which will process the input audio signal and further will perform the speech recognition. This above task can be performed by using MATLAB Programming. In the above diagram, when the input signal i.e. the speech signal is provided the Speech Extraction process simply extracts all the signal components such as pitch, frequency etc. which are needed at the end for the recognition of speech. Then, the Speech Matching Block will match all the properties that are extracted before with the set of audio command signals being stored as database.

The speech recognition system consists of two phases.The first one is referred to the training phase while the second one is referred to as the testing phase.

2.1. BACKGROUND ON SPEECH FEATURE EXTRACTION[1]

Methodology of constructing the proposed system will consists of different modules. Each module uses different methods and algorithms to perform its tasks. After a particular module completes its task, the output will be given as input for the next module.
1:- Acoustic signal processing
2:- Pattern recognition
3:- Motion control.

2.1.1 Signal Processing

Speech feature extraction is a fundamental requirement of speech recognition system; it is the mathematical representation of the speech signal. In a human speech recognition system, the main

aim is to classify the source files using a reliable representation that reflects the difference between utterances of audio signals.

2.1.1.1 Pre-processing

Pre-processing is the fundamental signal processing method applied before extracting features from speech signal, for enhancementof the performance of feature extraction algorithms. Commonly used pre-processing techniques contain DC component removal, pre-emphasis filtering, and amplitude normalization.

2.1.1.2 DC Component Removal

The speech signal often has a constant component, i.e. a non-zero mean known as the DC component of the signal. This is typically due to DC bias within the recording instruments. The DC component in the input signal can be easily removed by subtracting the mean value from all samples within an utterance.

2.1.1.3Pre-emphasis Filtering

A pre-emphasis filter consiststhe dynamic range of the speech signal's power spectrum by flattening the spectral tilt.The filter is in form of
$P(z) = 1 - az^{-1}$
Where$a$ ranges between 0.9 and 1.0
In speech processing, the glottal signal can be modelled by a two-pole filter with both poles close to the unit circle. However, the lip radiation characteristics models its single zero near to $z=1$, which tends to cancel the effect of one of the glottal pole. In addition, the spectral slope of a human speech spectrum is usually negative since the energy is concentrated in low frequency. Thus, a pre-emphasis filter is introduced before applying feature algorithms to increase the relative energy of the high-frequency spectrum.

2.1.1.4 Amplitude Normalization

Recorded signals often have varying energy levels due to speaker's volume and the distance ofmicrophone. Amplitude Normalization can cancel the inconsistent energy level between signals, and thus can enhance the performance of the energy-related features. There are several methods usedto normalize a signal's amplitude. One of them is achieved by the point-by-point division of the signal by its maximum absolute value, so that the dynamic range of the signal is constrained between -1.0 and +1.0.

2.1.1.5 Mel Frequency Cepstral Coefficient (MFCC) [1]

The first step in speech recognition system is to extract features i.e. identify the components of the audio signal that are good for identifying the linguistic content in the signal and discarding all the other things which carries information like background noise, emotion etc.

Mel Frequency Cepstral Coefficients (MFCCs) is one of the feature widely used in automatic speech and speaker recognition. Prior to the introduction of MFCCs, Linear Prediction

Coefficients (LPCs) and Linear Prediction Cepstral Coefficients (LPCCs) were used for automatic speech recognition (ASR).

Steps InCalculating MFCC:

1. Frame the signal into short frames.
2. Calculate the periodogram estimate of the power spectrum for each frame.
3. Apply the Mel filter bank to the power spectra; sum the energy in each filter.
4. Take the logarithm of all filter bank energies.
5. Take the DCT of the log filter bank energies.
6. Keep DCT coefficients 2-13, discard the rest.

The Mel scale relates perceived frequency, or pitch, of a pure tone to its actual measured frequency.

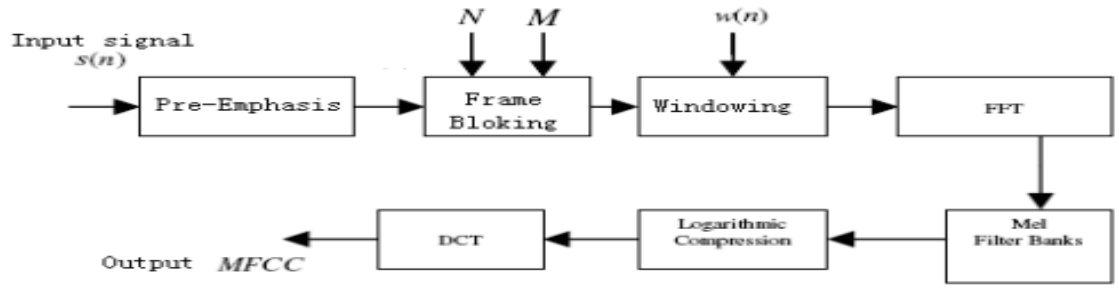


Fig 3.1 Mel Frequency Cepstral Coefficients[1]

2.1.1.6Windowing Functions

Speech is a non-stationary signal. Thus a window function is applied in the pre-processing stage to divide the speech signal into small segments to approximate a stationary signal,. There are many window functions are available. The proposed system is designed using *Hamming window*, defined as

$$w(n) = \begin{cases} (1-\alpha) - \alpha 0.46 \cos\left[2\pi n/(N-1)\right] & 0 \le n \le N-1 \\ 0 & n = else \end{cases}$$

Where$N$ is the window length andα=.46

2.1.1.7K-Mean Clustering

K-means clustering is a vector quantizationmethod, originally from signal processing, that is popular for cluster analysis. k-means clustering algorithm is used to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster.Let($\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$) be the set of observations, where each observation is a *d*-dimensional real vector, *k*-means clustering partitions the *n* observations into *k* ($\le n$) sets $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$ so that it minimize the within-cluster sum of squares (WCSS).

2.1.2 Pattern Recognition[2]

This module uses the various audio features to extract desired parameters. Approach used for pattern recognition is:-

Minimum Feature Distance Technique

This technique is based on calculating distances between the spoken word and each word in the library.

1) D=features of spoken word –features of word stored in library at training time.
2) Sum up all the corresponding differences.
3) Take the square root of the total calculated difference.
4)For each command in library perform the above step 1, 2 and 3 calculations
5) Above steps will result in nine feature distance values
6) Print the word with minimum feature distance.
7) The result will correspond to word spoken by user

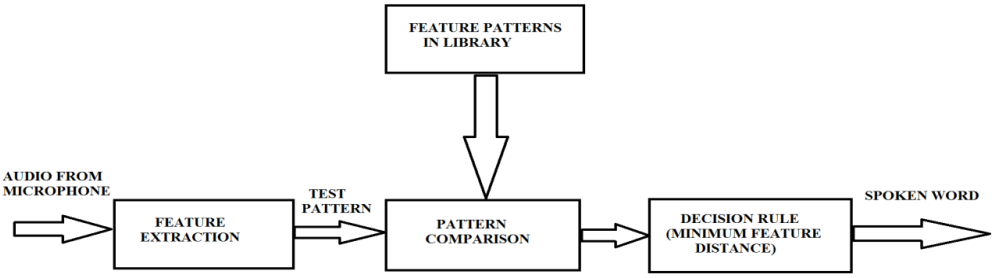


Fig. 3.3 Pattern Recognition[2]

2.2. SPEECH RECOGNITION

Speech recognition refers to the process which finds the distance or probability between the vector sequence of the unknown speech feature and each input speech sample, when matching the unknown speech features with different training patterns. The system digitized analog signals according to Nyquist sampling frequency. After pre-processing the input speech signal by using several algorithms, Mel-frequency cepstral coefficient (MFCC) is calculated. The MFCC analysis consists of four steps.

a)Perform a fast Fourier transform on the input speech signal.
b)Model the frequency axis by the Mel-scale. The Mel–frequency melf can be computed from the frequency f
c)Calculate the amplitude spectrum of the triangle filter to the filtered signal output
d) Perform the discrete cosine transform on the logarithm of the filter-bank energies and append first order differentials.
Then Cluster the features of the training speech samples using K-means clustering algorithm.Fig. 2.2 shows that when recognizing, the system calculates the feature of input speech and determines

its cluster group k at first. In the case of appropriate cluster group the system will save a considerable amount of computation. Then consider the test file, input the test command through microphone if the input is not strong enough then displays the message 'TRY AGAIN'. This is to avoid noise signals in the surroundings. Then after pre-processing perform MFCC calculation to extract MFCC features. Calculate the distances between the test speech samples and the cluster centres. For each sample, the minimum distance determines its Euclidian distance. Check whether the target group contains the training sample. If included, corresponding command displayed on the screen and the mouse pointer moves up , down or in specified direction

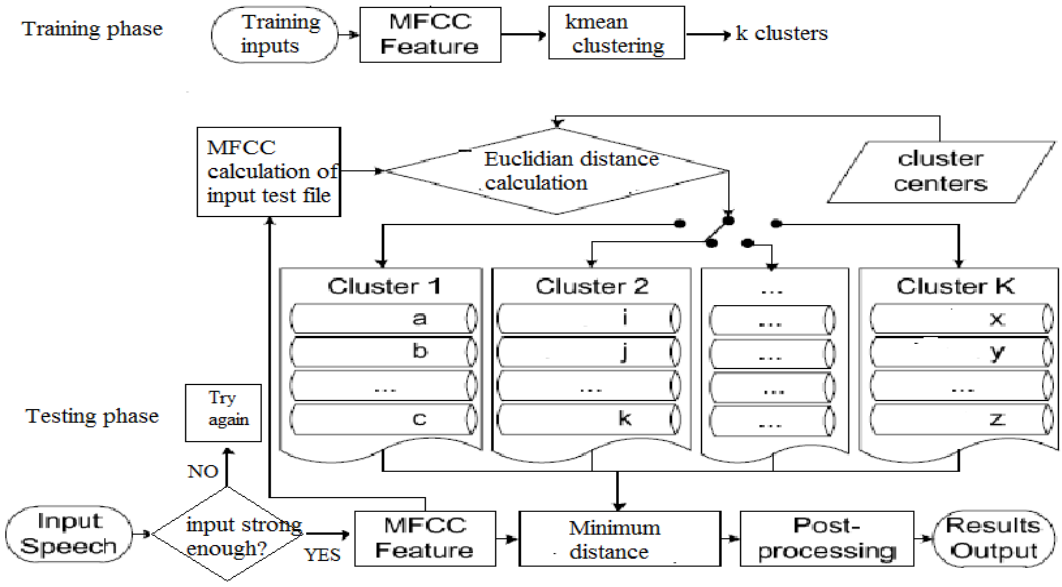


Fig.2.2 flowchart for training phase and testing phase

2.3. GRAPHICAL USER INTERFACE

"Hands free computer control" allows the user to continuously control various applications running on the system using vocal mouse and virtual keyboard. A set of operating modes which include, database recording and recognition for sound, and virtual keyboard for documentation are integrated under the same programming platform. This requires a set of graphical user interfaces (GUI).The main GUI is provided to allow the users to select the type of function:-

[1] Recording of speech commands

[2] Adding the speech commands to database

[3] Testing the input command from the microphone with the database.

[4] Virtual keyboard for various functions like creating and editing documents, and other keyboard applications

[5] Pop up menu for pointer location selection.

[6] Exit from main GUI

Fig2.3 Main GUI

2.3.1 Creation Of Database Using GUI

Fig. 2.4GUI for Database creation

If user selects ‘START’ option another GUI pops up saying to record the sound after pressingOK. And when we press OK two more dialog boxes appear on the screen saying to give a name to the .wav file and the name of the character to be stored in the database. Then select ‘DATABASE’ for adding the sound file to database.

2.3.2 Testing

The sound recognition and matching starts with ‘TEST’option. After clicking on this option the user will have a small time gap for making the sound. The system process the sound based on MFCC calculation and gesture corresponding to best match is displayed.

Fig 2.5 command window(MATLAB)

3. IMPLEMENTATION TOOL AND RESULTS

The proposed system can be implemented using MATLAB. To do work on speech recognition, users can take the help of VOICEBOX. There is no need of any other expensive, bulky hardware.

The only thing that is needed for the system is good quality microphone.

A. Training Phase

Vocal Mouse project starts with training phase. Get nine commands (up, down, left, right, click, doubleclick, right click, start and stop) from microphone and compute their features and to create database. Initially, save the calculated features in a feature matrix and then store them in some other file. User is given 1 second to say each word. User can press START button and say the specified word in 1 second. Features of all the spoken words are stored in a feature matrix. Matrix is a 2-D matrix with 9 rows (one for each word) and 17 columns (17 features are extracted). Contents of this matrix can be loaded to library data file for further reference. Here is the completion of training phase.

B. Testing Phase

After clicking 'TEST' button on the GUI, the Testing phase started. During the testing phase, User can continuously say any word. Features of the spoken words are computed and these features thus extracted are compared with feature patterns already stored during training time. For this comparison,approach is used is minimum feature distance. Vocal Mouse system will compute resulting spoken word on the basis of minimum distance. Output will be displayed on the console window of MATLAB giving spoken word by the user. After the command recognition, system will perform task as per the requirement of the user.

4. FUTURE SCOPE

One area of enhancement of the proposed system is in the user interface for supporting self-diagnosis of recognition issues. The user can be provided with a testing mode to test whether the system is responding properly or not. More information should be conveyed to the user that would allow them to troubleshoot when the system fails to recognize certain inputs. The main obstacle and sources of frustration that a user faces is false signal that occur while operating the system. False positive means that system recognize event when the user did not intend to vocalize. This will happen when the user forgets that the system is processing vocal input and begins speaking or making some sounds, typically when the users attention is away from the user interface with the system, or when the system picks some background noise and incorrectly recognizes it as some valid vocalization. In these cases, the user cannot not realize that the system has processed the false positive events later, e.g., when the user returns his attention back to the system, at which point the user become confused what had happened, and possibly quite frustrated about not knowing what to undo.

In such a situation, quick method is required to disable current operation. A possibility for future work for addressing this issue is the use of various external contexts such as the users head posture and gaze to disable voice input when the user is likely disengaged from the system interface. Current method of Hands free Computer control uses only standard spoken words such as up, left, right, down, click etc is inefficient for performing continuous tasks. Therefore the system can be improved by including commands given as words or regular sounds such as continuous speech.

5. CONCLUSION

The Hands Free Computer Control is a system that enables the user to continuously control the mouse cursor using their voice. This uses a technique minimum feature distance algorithm for pattern recognition. The system provides more interactive and easy to user interface. Hands free Computer control is mainly designed for physically handicapped persons and mainly presents two systems, Vocal Mouse(VM) and Virtual Keyboard. Now, VM will allow users to work on continuous motion control over the computer screen including the virtual keyboard created as GUI. This includes commands consist of nine mouse controlling command. The system is based on the recognition of the commands and vocal sounds which is a very robust and accurate method as compared to the recognition of words using conventional speech recognition systems.

ACKNOWLEDGEMENT

The author would like to thank Mrs. Jibi John, Associate Professor, Govt. Model Engineering College, Cochin, for the guidance and support.

REFERENCES

[1] Aseemsaxena, Amitkumarsinha,"speech recognition using matlab", International Journal of Advances In Computer Science and Cloud Computing, ISSN: 2321-4058 Volume- 1, Issue- 2, Nov-2013

[2] "Mouse Movement using Speech and Non-Speech Characteristics of Human Voice", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-5, June 2012.

[3] "To Design Voice Control Keyboard System using Speech Application Programming Interface", IJCSI, Vol. 7, Issue 6, November 2010, ISSN : 1694-0814.
SelinaChu,ShrikanthNarayana, and Jay Kuo,‟‟ Sound Recognition With Time and Frequency Audio Features‟‟ IEEE Transactions audio,speech,and language processing,VOL.17,NO.6,August2009.

[4] M Abdeen, H Moshammad and M C E Yagoub, "An Architecture for Multi-Lingual Hands Free Desktop Control System for PC Windows", Niagara Falls, Canada:IEEE , 2008

Authors

PoojaAntony, received BTech degree in Electronics and Communication Engineering from Cochin University Of Science And Technology, Kerala in 2015. Currently pursuing MTech degree in VLSI and Embedded Systems from APJ Abdul Kalam Technological University, Kerala.

Prof. Sunny Joseph is the Head of the Department of Electronics and Communication Engineering in Mar Athanasius College of Engineering, Kothamangalam, Kerala, India. He has a teaching experience of 29 years. He received his B.E from Bangalore University and M.Tech degree from Kerala University. His current research focus is in the area of VHDL, high speed digital design

INTER INTRA VEHICULAR COMMUNICATION

Neethu P P[1] and Siddharth Shelly[2]

Department of Electronics& Communication, Mar Athanasius College of Engineering, A.P.J Abdul Kalam Technological University, Kerala, India
Assistant Professor, Department of electronics & Communication, Mar Athanasius College of Engineering, Kerala, India

ABSTRACT

CAN communication based transmission system in vehicle is used to implement an anti-collision system for vehicles that includes a speed sensor for sensing the speed of the vehicle, an ultrasonic sensor for measuring the distance of the vehicle from an object, a system for computing a anti collision distance to the object, an alarm actuated by the system when the sensed distance of the object is equal to or less than the anti collision distance compared by the system, and a brake light actuated upon the actuation of said alarm, and an alert system for driver's to shift gear of vehicle when the ideal speed is reached, and cornering head light for the vehicle. The data will be transmitted over CAN bus and will be displayed on LCD.

KEYWORDS

CAN Protocol, Anti-collision system

1. INTRODUCTION

The modern vehicles require more complex control connections, thus more electronic control units (ECU) interacting through actuators and sensors. When a driver drives the car the ECU contain information about current state of the car. But in an in-vehicle communication technology uniting these many controllers has become essential. Controller area network (CAN) is introduced to solve the problem of sending this message efficiently .This focuses on a better in-vehicle communication network used to combine these ECUs: the controller area network communication protocol, also known as Controller area network[1]. It is a serial communication protocol developed by Robert Bosch for the German car industry. The CAN protocol is an ISO 11898 standard for serial data communication[1]. In advances data communications has created efficient methods for several devices to communicate using a minimum number of system wires. The Controller Area Network (CAN) is one of these methods. CAN transmits and receives messages over a two wire bus (CAN Bus). The protocol was developed for automotive applications. Today CAN have gained widespread use and is used in industrial automation as well as in automotives and mobile machines. But it is also used for economical standard in today. The

DOI: 10.5121/ijci.2016.5236

vehicular communication provide traffic management, efficient and easier maintenance, which ultimately leads to safer roads.

2. CAN PROTOCOL

2.1 CAN BUS

In 1980s, engineers at Bosch evaluated serial bus systems regarding their possible use in passenger cars and found that the available network protocols were not able to fulfill the requirements of the automotive. In February of 1986, CAN was introduced: at the SAE (Society of Automotive Engineers) congress in Detroit, the new bus system developed by Bosch was introduces 'Automotive Serial Controller Area Network'. Controller Area Network (CAN) is a serial data communications bus for real-time applications[2]. Fig.1 is the CAN network topology follows the bus network topology, which gives it the advantage that easily adding new CAN nodes to a current network. Furthermore, the standardization of the protocol means all ECUs will conform to the CAN standards when transmitting data. Note that in the Fig.1 all CAN nodes are fitted with a transceiver chip that connects it to the CAN bus. CAN protocol is a message based protocol not an address based protocol. This means that messages are not transmitted based on addresses . All nodes in the system receive message transmitted on the bus. Another feature built into the CAN protocol is the ability for a node to request information from other nodes. One benefit of this message based protocol is that additional nodes can be added to the system without the necessity to reprogram all nodes. This is called Remote Transmit Request (RTR)[2]..

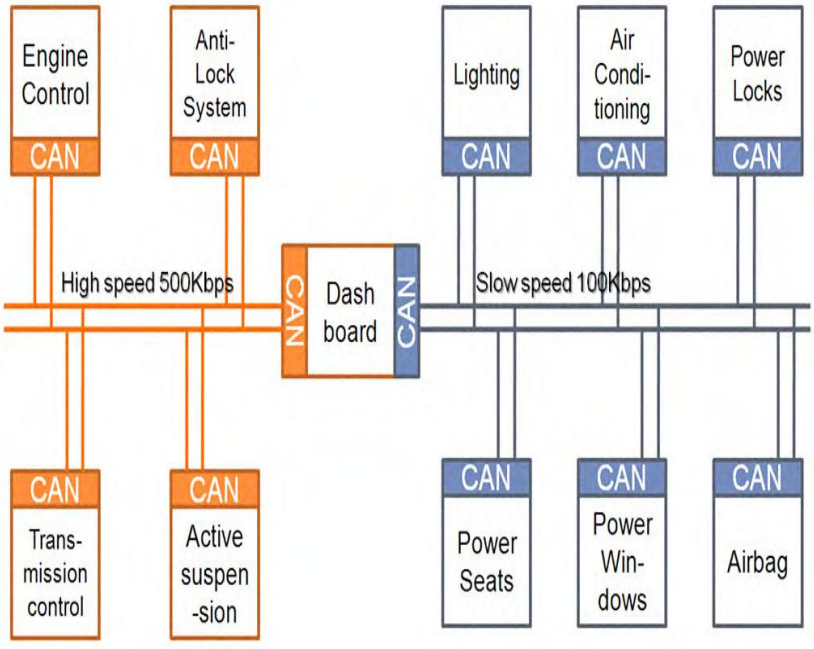


Fig .1 The CAN network topology

2.2 CAN FRAME STRUCTURE

Fig.2 CAN Frame

SOF– The start of a message, it is used to synchronize the nodes on a bus after being idle.
Identifier- The standard CAN has 11-bit identifier which establishes the priority of the message. Lower the binary value, higher its priority.

Rtr–the single remote transmission request bit is dominant when information is required from another node. All nodes receive the request, but the identifier specifies the node. The responding data is also received by all nodes and used by the node interested. So, all data that is used in a system is uniform.

IDE–a dominant single identifier extension bit means that a standard can identifier with no extension is being transmitted.

R0–reservedbit for future standard amendment.

DLC–the 4-bit data length code contains the number of bytes of data being transmitted. Data up to 64 bits of application data may be transmitted.

CRC–the 16-bit (15 bits plus delimiter) cyclic redundancy check contains the checksum (number of bits transmitted) of the preceding data for error detection.

ACK–every node receiving an accurate message overwrites this recessive bit in the original message with a dominate bit, indicating an error-free message has been sent. Should a receiving node detect an error and leave this bit recessive, it discards the message and the sending node repeats the message after re arbitration. In this way, each node acknowledges the integrity of its data. ack is 2 bits, one is the acknowledgment bit and the second is a delimiter.

EOF–The end-of-frame is 7-bit field which indicate the end of a can frame (message) and disables bit-stuffing. When 5 bits of the same logic level occur in succession during normal operation, a bit of the opposite logic level is *stuffed* into the data.

IFS–The 7-bit inter frame space contains the time required by the controller to move a correctly received .

2.3. THE CAN STANDARD

CAN is an International Standardization Organization (ISO) defined serial communications bus developed for the automotive industry. Which replace the complex wiring with a two-wire bus..

The features of CAN increases its popularity in a variety of industries like building automation, medical, manufacturing etc .The CAN communications protocol, ISO-11898: 2003, describes how information is passed between devices on a network. It conforms to the Open Systems Interconnection (OSI) model that is defined in terms of layers. Actual communication between devices connected by the physical medium is defined by the physical layer of the model.

The data link and physical layers of which are normally transparent to a system operator, are included in any controller that implements CAN protocol. The connection to physical medium is implemented using a line transceiver to form like a system node .

3. IMPLEMENTATION

In this section, discuss on the block diagram, circuit and its layout .A brief idea is given on how it is implemented. Here used PIC16F73, which manages the different sensors used such as ultra-sonic sensor, speed sensor .And the data read from the sensor is passed to the receiver part where it displays the speed, distance of collision, mileage of the vehicle for current speed, and gear shift indicator for the speed specified. The program is written in Embedded C. Embedded C is a set of language extensions for the C Programming language.

Another software used in this project is MPLAB Integrated Development Environment (IDE), is a free, integrated toolset for the development of embedded applications employing Microchip's PIC® and dsPIC® micro controllers. A set of compilers and assemblers are also available with it. To begin with download and install MPLAB IDE from the website. Don't forget to install HITECH C compiler when it asks you during the installation process. For the demonstration purpose I am using MPLAB version 8.60.

Then use Proteus it is best simulation software for various designs with microcontroller. It is mainly popular because it contains almost all microcontrollers in it. So it is a handy tool to test programs and embedded designs for electronics hobbyist. We can simulate our programming of microcontroller in Proteus 8 Simulation Software. ISIS lies at the main of the Proteus system. It combines a powerful design environment with the ability to define most aspects of the drawing appearance. Whether your requirement is the rapid entry of complex designs for simulation and PCB layout, or the creation of attractive schematics for publication, ISIS is the tool for the job.

3.1 BLOCK DIAGRAM

Fig.3 shows the intra vehicular communication, the ultrasonic sensor measures the distance of the vehicle from an object .And the speed sensor senses the speed. This information is passed over to the receiver section through a CAN bus. Then in the receiver section the buzzer is activated indicating the collision distance has reached, and the corresponding break light indicator is activated.

The distance will be displayed on LCD. If the speed is increasing, it will command for a gear shift, which in turn helps to increase the fuel efficiency of the vehicle. The LCD display displays

the speed of the vehicle. This information is given to second vehicle through RF Transmitter by encoding the data. Encoder is a parallel to serial converter. Both transmitter and receiver uses PIC 16F73.

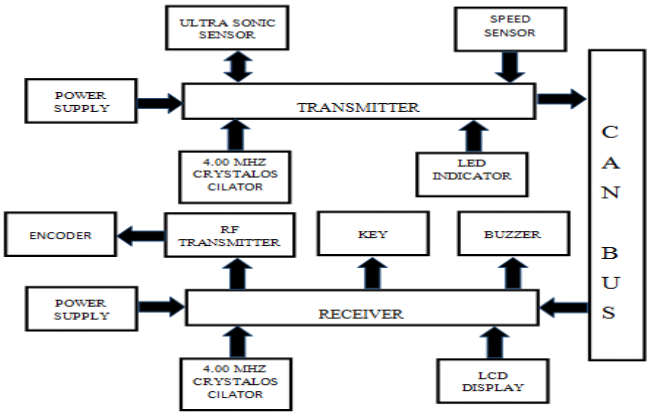


Fig .3 Intra communication

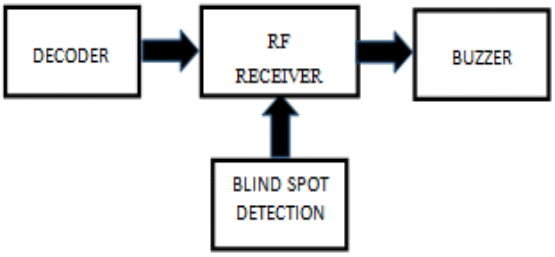


Fig.4 Inter communication

Fig.4 is used for inter vehicular communication. Data from the RF transmitter of the first vehicle is decoded by the RF receiver in the second vehicle. The decoder converts the serial input to parallel output. Then alerts the driver in the second vehicle. Using IR LED and Photodiode the blind spot of the vehicle is detected in order to inform the overtaking.

4. FLOW CHART

A flowchart is representation of the sequence of steps and decisions needed to perform a process. Each step in the sequence is noted within a diagram shape. Steps are linked by connecting lines and directional arrows. This allows anyone to view the flowchart and logically follow the process from beginning to end.

4.1TRANSMITTER

Algorithm

1. Start

2. Initialization
3. Check sensor speed
4. Measure distance
5. Send to receiver
6. Go to step 3

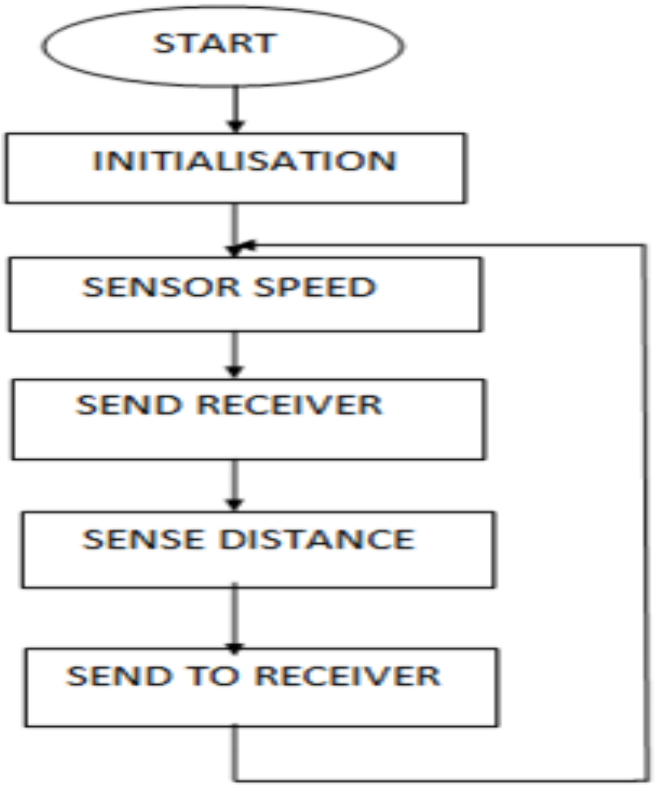


Fig 5 Flow chart of transmitter

4.2 RECEIVER

Algorithm

1. Start
2. Initialization
3. Check receiver value
4. If speed is received then display speed
5. Otherwise display distance then go to step 10
6. If the received speed is low display gear state
7. The received speed is high check whether it is increasing or decreasing
8. If the receiver speed is reducing then alarm off
9. Otherwise go to step 7
10. If distance is high go to step 14
11. Otherwise alarm on and measure distance

12. If distance is increases alarm off and go to step 14
13. Distance decreases go to step 11
14. Check the need of overtake
15. If there is a need of overtaking alarm on and displayed
16. If there is no need of overtaking check the value of s1,s2
17. If s1,s2 is equal to zero , adjust servomotor then go to step 3
18. If s1s 2 is not equal to zero then go to step 3

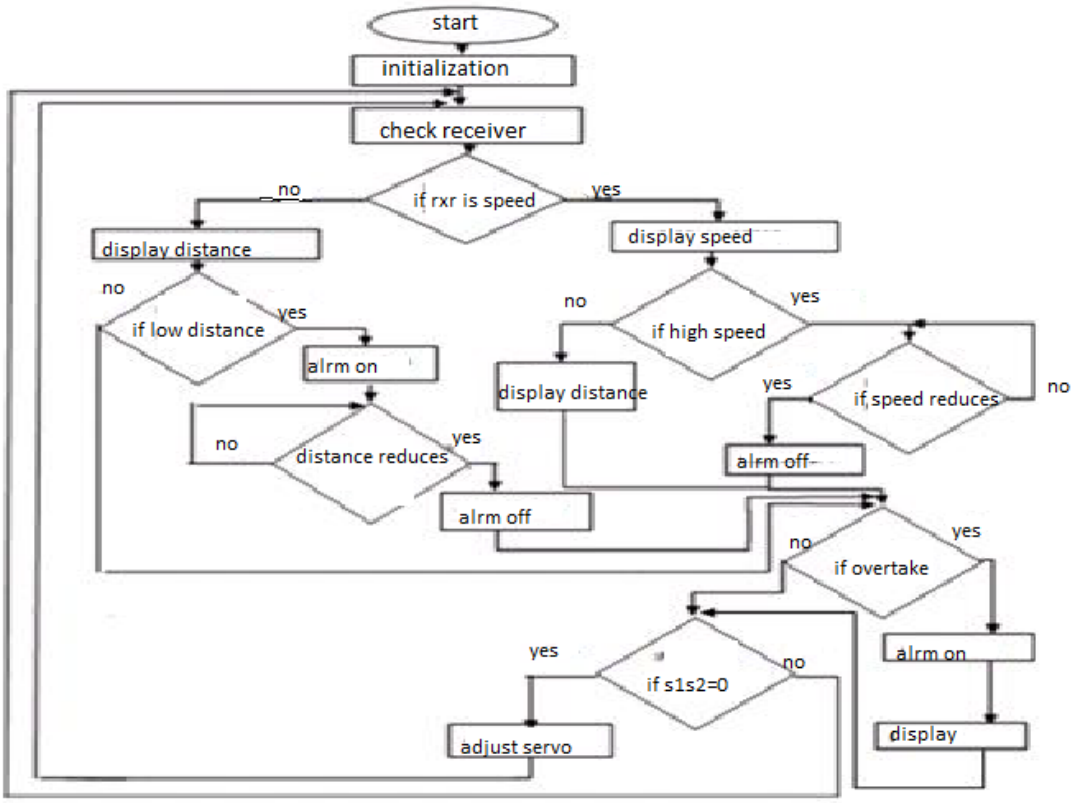


Fig .6 Flow chart of receiver

5. CIRCUIT DIAGRAM

The transmitter section consists of speed sensor, IR LED, and the photodiode. Ultrasonic sensor is used to measure the distance of collision. The ultrasonic sensor has 3 pins, GND, VCC, SIG. Port C of PIC16F73 is used to connect the ultrasonic sensor .PIN 18-RC7 is connected to the SIG of ultrasonic sensor. The port A of PIC 16F73 is used to connect the motor to sense the speed. The ultrasonic sensor used to check the distance between the two vehicles. It sends an ultrasonic signal and receives the reflected signal from other vehicle and calculate the distance,.

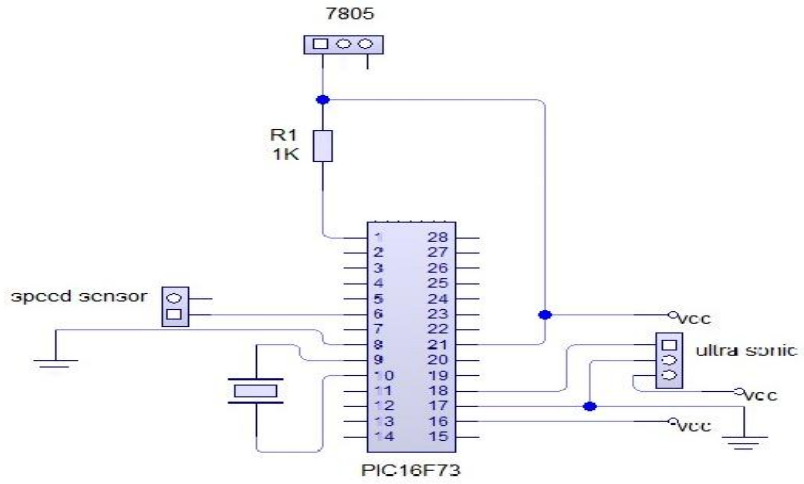


Fig 7 Main board transmitter section

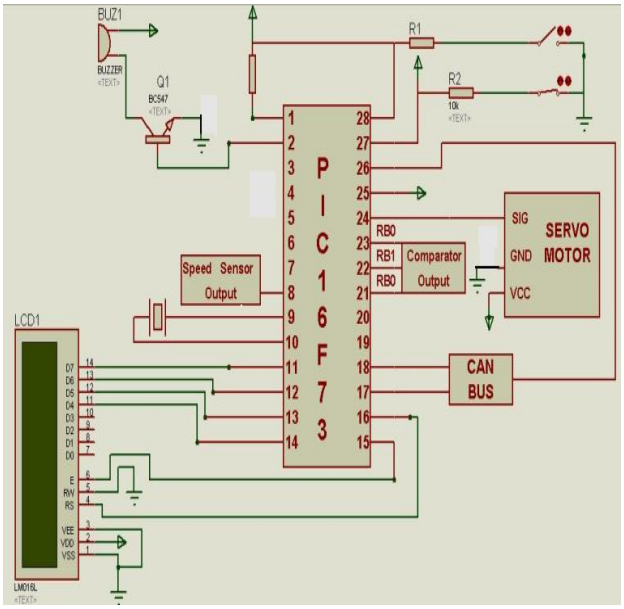


Fig 8 Receiver section

Fig 8shows the receiver section which has servo motor, comparator section, speed sensor output and LCD. The distance measurement value is obtained through CAN bus and the distance measured is displayed on the LCD display. The buzzer is actuated only when the measured distance is less than 1m. The transistor connected to buzzer amplifies the input signal to actuate the buzzer. The measure of speed, distance and gear status are displayed on LCD, the two switch connected to PIC16F73 controls the servo motor as part of cornering headlight A servomotor is a rotary actuator that allows for precise control of angular position, velocity and acceleration. It consists of a suitable motor coupled to a sensor for position feedback. It

also requires a relatively sophisticated controller, often a dedicated module designed specifically for use with servomotors

6. CONCLUSION

In the near future, automobile manufacturers are considering to use wireless ad hoc networks to improve traffic flow and safety, as they proves to be more cost effective than continually undertaking massive construction projects, which have only limited success. Consequently, future developments in automobile manufacturing will include new communication technologies that offer more effective spacing and collision avoidance systems, gas mileage (less braking),less pollution (cars are in movement), more information and entertainment, etc. In order to reduce communication costs and guarantee the low delays required to exchange data between cars, inter-vehicle communication (IVC) systems, based on wireless ad-hoc networks, represent a promising alternative for future road communication scenarios, as they permit vehicles to organize themselves locally in ad hoc networks without any pre-installed infrastructure.

ACKNOWLEDGEMENT

I would like to thank to Mr. Altaf Hamed Shajahan Assistant Professor and internal Guide, Department of Electronics and Communication Engineering in AWH Engineering college , Calicut, for his valuable suggestions and timely help.

REFFERENCES

[1] D.Sridhar1, N.Mallika2 and Chirivella Anjaneyulu "Implementation of intra inter vehicular communication" ,IJTTCS 2012 Volume 1, Issue 3, September – October 2012

[2] HuaqunGuo, lunlie Ang, and YongdongWu ," Extracting Controller Area Network Data for Reliable Car Communications " IJTTCS 2009

[3] MilindKhanapurkar, Dr. Preeti Bajaj, DakshataGharode, "A Design Approach for Intelligent VehicleBlack Box System with Intra-vehicular communication using LIN/Flex-ray Protocols" IEEE-ICIT, April- 2008

[4] Yousef Al-Ali, Ghaleb[Al-Habian, SadiqSaifi, "Automobile Black Box for Accident Simulation", published in CSIDC 2005, American university of Sharjah

AUTHOR

NEETHU P P did her B.Tech degree from AWH Engineering college Calicut, Kerala, India, university of Calicut. Now pursuing M.Tech in VLSI &EMBEDDED SYSTEM at Mar Athanasius College of Engineering, Kothamangalam, Kerala, India, APJ Abdul Kalam Technological University.

Siddharth Shelly is a faculty member of Mar Athanasius College of Engineering, Kothamangalam, Kerala, India. He received his B.Tech from Mahatma Gandhi University, Kottayam and M.Tech degree from the Amrita School of Engineering, Coimbatore, India. His current research focus is in the area of vehicular ad hoc networks, embedded systems.

INTENTIONAL BLANK

MAXIMAL MARGINAL RELEVANCE BASED MALAYALAM TEXT SUMMARIZATION WITH SUCCESSIVE THRESHOLDS

Ajmal E B[1] and Rosna P Haroon[2]

[1]Department of CSE, Ilahia College of Engineering and Technology, Muvattupuzha, India
[2]Department of CSE, Ilahia College of Engineering and Technology, Muvattupuzha, India

ABSTRACT

Automatic text summarization has prime importance in the area of Natural Language Processing. As we are aware a large quantity of information are there on web, it is very difficult to extract the needed information from the huge. Text summarization is the process of shorten the document, so that it retains only the important points of the original document. As the problem of information overload has grown, and the quantity of data has assumed a greater significance, the need for an instant summarization of the un-touched language -Malayalam- assumes vital importance. Lots of summarization systems have already been developed for various languages, there is no such well performing system for Malayalam. In this paper propose a Malayalam text summarization system which is based on MMR technique with successive threshold. Here the sentences are selected based on the concept of maximal marginal relevance. The key idea is to use a unit step function at each step to decide the maximum marginal relevance and the number of sentences present in the summary would be equal to the number of paragraphs or the average number of sentences present in the text document, which can be achieved by using successive threshold approach.

KEYWORDS

Maximum Marginal Relevance, Successive Threshold, Unit step function

1. INTRODUCTION

Automatic text summarization has prime importance in the area of Natural Language Processing. As we are aware a large quantity of information are there on web, it is very difficult to extract the needed information from the huge. Text summarization is the process of shorten the document, so that it retains only the important points of the original document. As the problem of information overload has grown, and the quantity of data has assumed a greater significance, the need for an instant summarization of the untouched language - Malayalam assumes vital importance. Lots of summarization systems have already been developed for various languages, there is no such well performing system for Malayalam. The existing systems have high computational cost, time and storage capacity. To address the issues of computational cost time and storage capacity, here

DOI: 10.5121/ijci.2016.5237

proposes a text summarization system that works on the concept of maximal marginal relevance between the sentences or the words. The key idea is to use a unit step function at each step to decide the maximum marginal relevance and the number of sentences present in the summary would be equal to the number of paragraphs or the average number of sentences present in the input text document, which can be achieved by using successive threshold approach.

Malayalam is the official language of Kerala and there are around 33 million people who speak Malayalam. There is a vast amount of online data available in Malayalam. This warrants creating a tool that can be used to explore digital information presented in Malayalam and other native languages. In this concept, we propose the MMR based Malayalam document Summarization with Successive Thresholds.
Concept is presented in five sections. Section II reviews the related works. Section III discusses the proposed scheme. In section IV, the evaluation of the proposed scheme is presented. Section V concludes the work and future scope is discussed.

2. RELATED WORK

Attempts to automatically summarize documents started early since 1958. The method based on word frequencies by Luhn is one of the oldest but still relevant method. This method measures the importance of a sentence based on the presence of keywords (most frequently occurring words in a document otherthan the stopwords) in the sentence. Text summarization method by Ed-mundson used cue words, title words, and sentence location for determining the sentence weights [4]. Text summarization for Malayalam documents by Rajina Kabeer and Sumam Mary Idicula [1].

Graph theoretical approaches for summarization represents a document as an undirected graph, in which the nodes represent the sentences in the document. Two nodes in the graph are connected if the cosine similarity of the sentences corresponding to the nodes is above some particular threshold. The sentences corresponding to the nodes with the highest cardinality or in other words the sentences which are more similar to other sentences in the document are considered important and are included in the summary [8]. Methods based on Co-reference chains and Lexical chains are based on the semantic structure of the document.

Semantic graph based approaches extracts semantic triplets (Subject-Object-Predicate triplets) from each of the sentence in the document. These triplets are used to generate a graph of the document. A sub-graph of this graph is selected using machine learning techniques and the sentences in the sub-graph are used to form the summary [2].

Machine Learning approaches to summarization models the summarization process as a classification problem. Naïve Bayes method, Neural networks and Hidden Markov Model (HMM) are some of the machine learning approaches [4] used for text summarization.

Information extraction by abstractive text summarization for Telugu language [7], summarization of tamil document using semantic graph method [14], Text extraction for an Agglutinative Language by Sankar K, VijaySundar Ram R and Sobha Lalitha Devi which was used for summarizing Tamil documents [8], Bengalitext summarization by sentence extraction by Kamal Sarkar [6] are some text summarization works done for Indian languages.

3. PROPOSED SCHEME

In the proposed method, a single-document input is summarized based on the concept of maximal marginal relevance between the sentences or the words. The key idea is to use a unit step function at each step to decide the maximum marginal relevance and each word meaning is calculated with the help of a dictionary, finally the number of sentences present in the summary would be equal to the number of paragraphs or the average number of sentences present in the input text document, which can be achieved by using successive threshold approach.

3.1. MAXIMAL MARGINAL RELEVANCE

The key idea in this technique is to use a unit step function at each step to decide the maximum marginal relevance. The automatic summarization process is explained below:

1. Input a document to be summarized
2. Now the document is traversed and eliminates the words that are not useful (stop word removal)
3. Starting with the starting position of the sentence until the document finishes
4. Identify the most important word/sentence (by meaning) with the help of a malayalam dictionary
5. Using the unit step function we can calculate the relevant information required. The unit step function used in the algorithm is given as:

$$u_k + 1 = \arg\max(Sim1(u_i, Q) - \max(sim2(u_i, u_j)))$$

   Where
   Q : User input document
   u_i : Most important word/sentence
   u_j : Remaining sentences in the document
   U : Selected list of sentences
6. The process may be terminated once an appropriate number of words or sentences are in U. Which can be achieved by using successive threshold approach

3.2. SUCCESSIVE THRESHOLD APPROACH

The concept behind this approach is that the number of sentences present in the summary would be equal to the number of paragraphs or the average number of sentences present in the input text document. That is initially count the total number of paragraphs and sentences in the given text document, if the total number of paragraphs in the input text document is meet a threshold value then take the value of ‘n’ as number of paragraphs otherwise take ‘n’ as average number of sentences in the input document. After applying all the pre-processing steps and the MMR technique to select the relevant information or the sentences from the document. Then counts the total number of sentences say it is ‘m’, if m is equal to ‘n’, then these are the sentences finally included in the summary. Else, repeat the steps of MMR technique until the ‘m’ value will be equal to ‘n’.

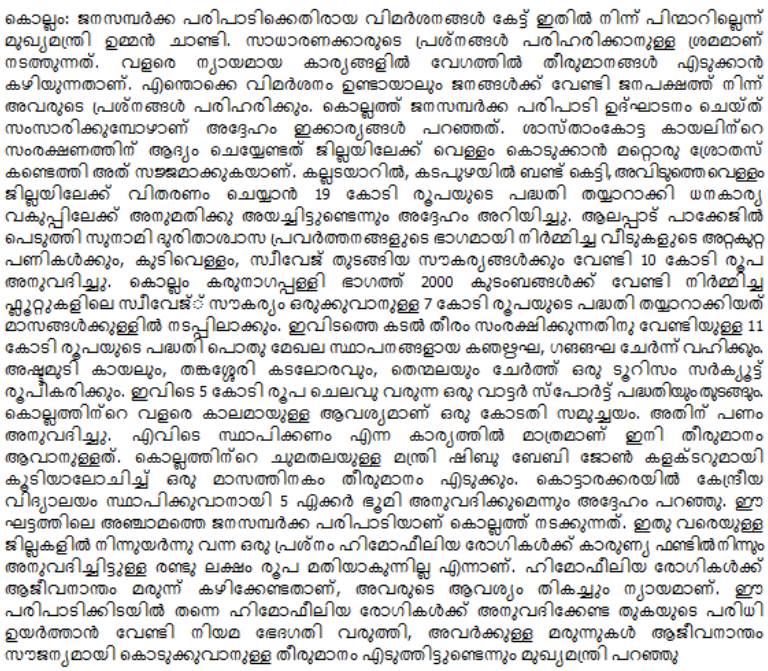
The proposed system uses the following algorithm. The algorithm consists of two sections; the first section uses a unit step function that identifies the maximum marginal relevance that is the

relevant sentences from the input document. Then the next section uses a successive threshold approach. By using this approach the total number of sentences in the final summary can be calculated. Process is explained below:

Input: Malayalam document
Output: Summarized document

1. Input a document to be summarized
2. Now the document is traversed and eliminates the words that are not useful (stop word removal).
3. Identify the most important word (by meaning) with the help of a Malayalam dictionary from the input document
4. Starting with the starting position of the sentence until the document finishes.
5. Using the unit step function and dictionary calculate the first level important sentence from the document by using the important word identified in step 3
   The unit step function used in the algorithm is given as:
   $$u_k + 1 = \arg\max(Sim1(u_i, Q) - \max(sim2(u_i, u_j)))$$
   Where
   Q : User input document
   u_i : Most important word/sentence
   u_j : Remaining sentences in the document
   U : Selected list of sentences
6. Then the second level sentence is identified using the first level sentence and the dictionary by using the unit step function
7. The next level sentence is identified using the sentence identified in step 6 and the dictionary by using the unit step function
8. Repeat the step 7 until an appropriate number of sentences is in U. Which can be achieved by using successive threshold approach
9. Stop

As an example consider the following input text shown in figure 1 and the corresponding output obtained for the text using proposed method is shown in figure 2.



കൊല്ലം: ജനസമ്പർക്ക പരിപാടിക്കെതിരായ വിമർശനങ്ങൾ കേട്ട് ഇതിൽ നിന്ന് പിന്മാറില്ലെന്ന് മുഖ്യമന്ത്രി ഉമ്മൻ ചാണ്ടി. സാധാരണക്കാരുടെ പ്രശ്നങ്ങൾ പരിഹരിക്കാനുള്ള ശ്രമമാണ് നടത്തുന്നത്. വളരെ ന്യായമായ കാര്യങ്ങളിൽ വേഗത്തിൽ തീരുമാനങ്ങൾ എടുക്കാൻ കഴിയുന്നതാണ്. എന്തൊക്കെ വിമർശനം ഉണ്ടായാലും ജനങ്ങൾക്ക് വേണ്ടി ജനപക്ഷത്ത് നിന്ന് അവരുടെ പ്രശ്നങ്ങൾ പരിഹരിക്കും. കൊല്ലത്ത് ജനസമ്പർക്ക പരിപാടി ഉദ്ഘാടനം ചെയ്ത് സംസാരിക്കുമ്പോഴാണ് അദ്ദേഹം ഇക്കാര്യങ്ങൾ പറഞ്ഞത്. ശാസ്താംകോട്ട കായലിന്റെ സംരക്ഷണത്തിന് ആദ്യം ചെയ്യേണ്ടത് ജില്ലയിലേക്ക് വെള്ളം കൊടുക്കാൻ മറ്റൊരു ശ്രോതസ് കണ്ടെത്തി അത് സജ്ജമാക്കുകയാണ്. കല്ലടയാറിൽ, കടപുഴയിൽ ബണ്ട് കെട്ടി, അവിടുത്തെവെള്ളം ജില്ലയിലേക്ക് വിതരണം ചെയ്യാൻ 19 കോടി രൂപയുടെ പദ്ധതി തയ്യാറാക്കി ധനകാര്യ വകുപ്പിലേക്ക് അനുമതിക്കു അയച്ചിട്ടുണ്ടെന്നും അദ്ദേഹം അറിയിച്ചു. ആലപ്പാട് പാക്കേജിൽ പെടുത്തി സുനാമി ദുരിതാശ്വാസ പ്രവർത്തനങ്ങളുടെ ഭാഗമായി നിർമ്മിച്ച വീടുകളുടെ അറ്റകുറ്റ പണികൾക്കും, കുടിവെള്ളം, സ്വീവേജ് തുടങ്ങിയ സൗകര്യങ്ങൾക്കും വേണ്ടി 10 കോടി രൂപ അനുവദിച്ചു. കൊല്ലം കരുനാഗപ്പള്ളി ഭാഗത്ത് 2000 കുടംബങ്ങൾക്ക് വേണ്ടി നിർമ്മിച്ച ഫ്ലൂറ്റുകളിലെ സ്വീവേജ് സൗകര്യം ഒരുക്കുവാനുള്ള 7 കോടി രൂപയുടെ പദ്ധതി തയ്യാറാക്കിയത് മാസങ്ങൾക്കുള്ളിൽ നടപ്പിലാക്കും. ഇവിടത്തെ കടൽ തീരം സംരക്ഷിക്കുന്നതിനു വേണ്ടിയുള്ള 11 കോടി രൂപയുടെ പദ്ധതി പൊതു മേഖല സ്ഥാപനങ്ങളായ കഞ്ഞിഘ, ഗങ്ങഘ ചേർന്ന് വഹിക്കും. അഷ്ടമുടി കായലും, തങ്കശ്ശേരി കടലോരവും, തെന്മലയും ചേർത്ത് ഒരു ടൂറിസം സർക്യൂട്ട് രൂപീകരിക്കും. ഇവിടെ 5 കോടി രൂപ ചെലവു വരുന്ന ഒരു വാട്ടർ സ്പോർട്ട് പദ്ധതിയും തുടങ്ങും. കൊല്ലത്തിന്റെ വളരെ കാലമായുള്ള ആവശ്യമാണ് ഒരു കോടതി സമുച്ചയം. അതിന് പണം അനുവദിച്ചു. എവിടെ സ്ഥാപിക്കണം എന്ന കാര്യത്തിൽ മാത്രമാണ് ഇനി തീരുമാനം ആവാനുള്ളത്. കൊല്ലത്തിന്റെ ചുമതലയുള്ള മന്ത്രി ഷിബു ബേബി ജോൺ കളക്ടറുമായി കൂടിയാലോചിച്ച് ഒരു മാസത്തിനകം തീരുമാനം എടുക്കും. കൊട്ടാരക്കരയിൽ കേന്ദ്രിയ വിദ്യാലയം സ്ഥാപിക്കുവാനായി 5 ഏക്കർ ഭൂമി അനുവദിക്കുമെന്നും അദ്ദേഹം പറഞ്ഞു. ഈ ഘട്ടത്തിലെ അഞ്ചാമത്തെ ജനസമ്പർക്ക പരിപാടിയാണ് കൊല്ലത്ത് നടക്കുന്നത്. ഇതു വരെയുള്ള ജില്ലകളിൽ നിന്നുയർന്നു വന്ന ഒരു പ്രശ്നം ഹിമോഫീലിയ രോഗികൾക്ക് കാരുണ്യ ഫണ്ടിൽനിന്നും അനുവദിച്ചിട്ടുള്ള രണ്ടു ലക്ഷം രൂപ മതിയാകുന്നില്ല എന്നാണ്. ഹിമോഫീലിയ രോഗികൾക്ക് ആജീവനാന്തം മരുന്ന് കഴിക്കേണ്ടതാണ്, അവരുടെ ആവശ്യം തികച്ചും ന്യായമാണ്. ഈ പരിപാടിക്കിടയിൽ തന്നെ ഹിമോഫീലിയ രോഗികൾക്ക് അനുവദിക്കേണ്ട തുകയുടെ പരിധി ഉയർത്താൻ വേണ്ടി നിയമ ഭേദഗതി വരുത്തി, അവർക്കുള്ള മരുന്നുകൾ ആജീവനാന്തം സൗജന്യമായി കൊടുക്കുവാനുള്ള തീരുമാനം എടുത്തിട്ടുണ്ടെന്നും മുഖ്യമന്ത്രി പറഞ്ഞു

Figure 1. Input Text

കല്ലടയാറിൽ, കടപുഴയിൽ ബണ്ട് കെട്ടി, അവിടുത്തെ വെള്ളം ജില്ലയിലേക്ക് വിതരണംചെയ്യാൻ 19 കോടി രൂപയുടെ പദ്ധതി തയ്യാറാക്കി ധനകാര്യ വകുപ്പിലേക്ക് അനുമതിക്കു അയച്ചിട്ടുണ്ടെന്നും അദ്ദേഹം അറിയിച്ചു. ആലപ്പാട് പാക്കേജിൽ പെടുത്തി സുനാമി ദുരിതാശ്വാസപ്രവർത്തനങ്ങളുടെ ഭാഗമായി നിർമ്മിച്ച വീടുകളുടെ അറ്റകുറ്റ പണികൾക്കും, കുടിവെള്ളം, സ്വീവേജ് തുടങ്ങിയ സൗകര്യങ്ങൾക്കും വേണ്ടി 10 കോടി രൂപ അനുവദിച്ചു. കൊല്ലം കരുനാഗപ്പള്ളി ഭാഗത്ത് 2000 കുടംബങ്ങൾക്ക് വേണ്ടി നിർമ്മിച്ച ഫ്ലൂറ്റുകളിലെ സ്വീവേജ് സൗകര്യം ഒരുക്കുവാനുള്ള 7 കോടി രൂപയുടെ പദ്ധതി തയ്യാറാക്കിയത് മാസങ്ങൾക്കുള്ളിൽ നടപ്പിലാക്കും. ഇവിടത്തെ കടൽ തീരം സംരക്ഷിക്കുന്നതിനു വേണ്ടിയുള്ള 11 കോടി രൂപയുടെ പദ്ധതി പൊതു മേഖല സ്ഥാപനങ്ങളായ കഞ്ഞിഘ, ഗങ്ങഘ ചേർന്ന് വഹിക്കും.

Figure 2. Output Text

2. EVALUATION

As shown in the below table is the different parameter evaluation of the existing method. The method is implemented on 6 different dataset of different sizes and various parameters such as precision, recall and F-measure is calculated.

Table 1. Parameter evaluation – Existing method

Dataset	**Precision**	**Recall**	**F- Measure**
Dataset1	0.485	0.525	0.530
Dataset2	0.5309	0.5765	0.5665
Dataset3	0.5807	0.6635	0.5978
Dataset4	0.6679	0.7814	0.7449
Dataset5	0.656	0.7756	0.7645
Dataset6	0.772	0.7901	0.7801

Table 2 shows parameter evaluation of the proposed method. The method is implemented on 6 different dataset of different sizes and various parameters such as precision, recall and F-measure is calculated.

Table 2. Parameter evaluation – Proposed method

Dataset	**Precision**	**Recall**	**F- Measure**
Dataset1	0.535	0.565	0.543
Dataset2	0.5407	0.5805	0.5785
Dataset3	0.5917	0.6743	0.6537
Dataset4	0.6779	0.7896	0.7549
Dataset5	0.673	0.7826	0.7775
Dataset6	0.852	0.8910	0.8018

The following chart shows the comparison of proposed MMR method with existing Sentence scoring method.

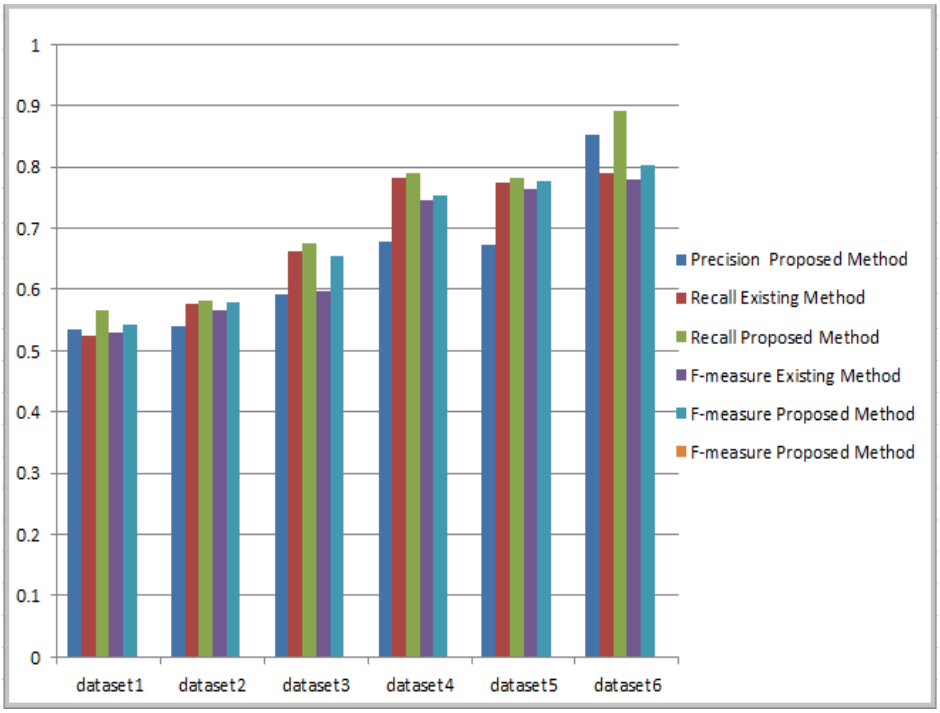


Figure 3. Graphical evaluation

3. CONCLUSIONS AND FUTURE WORK

The text summarization provides the summary of the input document. Here in this concept an efficient technique of document summarization is proposed. The proposed method is works on the concept of maximal marginal relevance between the sentences or the words. The key idea is to use a unit step function at each step to decide the maximum marginal relevance, and the number of sentences present in the summary would be equal to the number of paragraphs

or the average number of sentences in the input text document, which can be achieved by using successive threshold approach. Analysis shows that proposed method is more accurate. More quality parameters are generated by incorporate another methods is future work

ACKNOWLEDGEMENTS

The authors would like to thank, Prof. Rosna P Haroon, Head of the Department, Department of Computer Science and Engineering, Ilahia College of Engineering And Technology,Muvattupuzha, kerala, for her timely advices and suggestions.

REFERENCES

[1] Kabeer, Rajina, and Sumam Mary Idicula (2014) "Text summarization for Malayalam documents—An experience." *2014 IEEE International Conference On Data Science & Engineering (ICDSE).*

[2] Leskovec, J., Milic-Frayling, N., Grobelnik, M. and Leskovec, J., (2005) "Extracting summary sentences based on the document semantic graph", *Microsoft Research, Microsoft Corporation*.

[3] Bijalwan, Vishwanath, Pinki Kumari, Jordan Pascual, and Vijay Bhaskar Semwal (2014) "Machine learning approach for text and document mining." *arXiv preprint arXiv:1406.1580*

[4] Gupta, Vishal, and Gurpreet Singh Lehal. (2010) "A survey of text summarization extractive techniques." *Journal of Emerging Technologies in Web Intelligence*2, no. 3: 258-268..

[5] Dipanjan Das, Andre F.T. Martins, (2007) "A Survey on Automatic Text Summarization, Language Technologies Institute", Carnegie Mellon University.

[6] Sarkar, K., August. (2012) "An approach to summarizing Bengali news documents." In *proceedings of the International Conference on Advances in Computing, Communications and Informatics* (pp. 857-862). ACM..

[7] Jagadish S Kallimani, Srinivasa KG, Eswara Reddy B, (2011) "Information Extraction by an Abstractive Text Summarization for an Indian Rgional Language Natural Language Processing and Knowledge Engineering (NLP-KE)", *2011 7th International Conference on* Natural Language Processing and Knowledge Engineering.

[8] Sankar K, VijaySundar Ram R and Sobha Lalitha Devi (2011) "Text Extraction for an Agglutinative Language"

[9] Martin Hassel, (2004) "Evaluation of AutomaticText Summarization: A practical implementation".

[10] Bindu.M.S, Sumam Mary Idicula, (2011) "A Hybrid Model For Phrase Chunking Employing Artificial Immunity System And Rule Based Methods", *International Journal of Artificial Intelligence Applications(IJAIA)*, Vol.2, No.4.

[11] Rajeev RR, Rajendran N, Elizabeth Sherly, (2005) "A Suffix Stripping Based Morph Analyser For Malayalam Language", *Proceedings of 20th Kerala Science Congress*.

[12] Jayashree.R, SrikantaMurthy.K,Sunny, (2011) , " Keyword extraction based summarizationof categorised Kannada text documents*", International Journal on Soft Computing (IJSC)* , Vol.2, No.4.

[13] Aysun Guran, Eren Bekar, Selim Akyokus, (2010) "A Comparison of Feature and Semantic-Based Summarization Algorithms for Turkish", *International Symposium on Innovations in Intelligent Systems and Applications*, Kayseri Cappadocia,TURKEY.

[14] Banu, M., Karthika, C., Sudarmani, P. and Geetha, T.V., (2007) "Tamil Document Summarization Using Semantic Graph Method", *IEEEConference on Computational Intelligence and Multimedia Applications, 2007.* (Vol. 2, pp. 128-134).

[15] Lin, Hai, Lusheng Wang, and Ruoshan Kong. (2015) "Energy Efficient Clustering Protocol for Large-Scale Sensor Networks." *Sensors Journal, IEEE* 15, no. 12 7150-7160.

[16] Liu, Dawei, Saifeng Cai, and Xiaohong Guo. "Incremental sequential pattern mining algorithms of Web site access in grid structure database." *Neural Computing and Applications*: 1-9.

[17] Demertzis, Kostantinos, Lazaros Iliadis, Stavros Avramidis, and Yousry A. El-Kassaby. "Machine learning use in predicting interior spruce wood density utilizing progeny test information." *Neural Computing and Applications*: 1-15.

[18] Sahoo, G. "A two-step artificial bee colony algorithm for clustering." *Neural Computing and Applications*: 1-15.

[19] Stein, Procópio, Anne Spalanzani, Vitor Santos, and Christian Laugier (2014) "Leader following: A study on classification and selection." *Robotics and Autonomous Systems*

[20] Acampora, Giovanni, Matteo Gaeta, and Vincenzo Loia (2011), "Hierarchical optimization of personalized experiences for e-Learning systems through evolutionary models", *Neural Computing and Applications*,Vol.20, No.5, PP 641-657.

Authors

Ajmal E B received his **B Tech degree** in Computer Science and Engineering from Ilahia College of Engineering and Technology, Muvattupuzha, Kerala, India in the year 2013 and received **M Tech degree** in Computer science and engineering from the same college in the year 2015. His research interests include Natural Language Processing, Data mining and Evolutionary Algorithms, and Digital Image Processing.

MULTIRESONATOR CIRCUIT USING Λ/4 SIR FOR CHIPLESS RFID TAGS

Sajitha V R, Nijas C M, Roshna T K and Mohanan

Department of Electronics, Cochin University of Science and Technology, Kerala, India, 682022

ABSTRACT

a compact multi resonator circuit for Chipless rfid tag is presented in this paper. The basic resonator is a shorted λ/4 Stepped Impedance Resonator (SIR). The tag is designed and simulated using High Frequency Structure Simulator and a 6 bit prototype is fabricated on rtduroid(εr=2.2 and tanδ=.0008) Substrate in area of 4.5x2cm2. Data is coded as amplitude variations in the frequency domain. Harmonic separation in the desired band by proper tuning of design parameters and ease of coding by simply adding or removing the shorts are achieved with λ/4 SIR.

KEYWORDS

Chipless RFID, Multiresonator, spectral signature, stepped impedance resonator

1. INTRODUCTION

Automated detection, tracking, identification etc., are ever interested topics in the field of wireless communication and hence radio frequency identification, in which communication relies on Radio Frequency waves between the reader and tag. Conventional RFID tags contain integrated circuits to carry data. Chipless RFID is an emerging technology in this scenario. Though in its infancy, a lot of research is carrying out on this topic. Main objectives are reducing the cost of tags by eliminating silicon chips, identifying good resonators and developing better reading techniques.

Chipless RFID tags can be classified into three: frequency domain (spectral signature based), Amplitude/phase backscatterer and time domain reflectometry (TDR) based tags[1-2]. Spectral signature based chipless tag which uses multiple resonators is a multi-stop band filter that encodes data in the frequency spectrum. In Amplitude/Phase backscatter modulation based chipless RFID tags, data encoding is performed by varying the amplitude or the phase of backscattered signal based on the loading of the chipless tag. TDR based chipless RFID tags are interrogated by sending a signal from the reader in the form of a pulse and listening to the echoes of the pulse sent by the tag. In this paper frequency domain tags with reception-retransmission antennas are concentrated which are rather simple in design and coding. A multiple spiral resonator based Chipless RFID tag is presented in [3]. The spiral resonators are compact and provide ease of coding after fabrication of the tags by simply shorting the resonators. But the usable bandwidth is limited due to harmonics. In [4] C M Nijaset .al presented a λ/4 open stub

DOI: 10.5121/ijci.2016.5238

resonator based Chipless RFID tag. It also has limited usable bandwidth due to harmonics. In [5] a number of modified complementary split ring resonators (MCSRR)placed along the transmission line as data bit encoding element. A novel data detuning technique is also presented. In [6] a dual multi-resonant dipole antenna performs as multiresonator.

In this paper a multiresonator circuit for chipless RFID tag using l/4 stepped impedance resonator is proposed. The basic resonator is compact, has harmonic separation capability, has flexible design and has ease of coding. The tag is designed and fabricated on RTDuroid (εr=2.2, tanδ=0.0008) substrate. A 6 bit multi resonator is prototyped and the tag size is 4.5 x 2 cm2. Either frequency shift coding or presence/absence technique is applicable depending on the number of items.

2. MULTIRESONATOR CIRCUIT DESIGN

Stepped impedance resonators are very interesting structure due to their flexible design. Size and harmonic frequency separation can be independently controlled by choosing impedance ratio (K) and length ratio (α) properly [7]. The structure of the basic resonator is shown in figure 1.

Fig 1: Basic stepped impedance resonator in fundamental mode

The λ/4 stepped impedance resonator shown in Figure 1 gets excited in fundamental mode when it is grounded and in first harmonic frequency when the short is removed. In this work only the fundamental mode is used. Equation 1 represents the fundamental mode of λ/4 SIR.

$$K\cot(\alpha.\theta_t)=\tan[(1-\alpha).\theta_t] \quad (1)$$

where $\alpha = \theta_2/(\theta_1+\theta_2)$, $\theta\_t=\theta_1+\theta_2$ and θ_1,θ_2 are electrical length corresponding to L_1 and L_2, respectively. Figure 2 shows how the separation between fundamental and first harmonic frequency can be controlled by properly setting the impedance and length ratio. This property is effectively utilized here to remove the harmonic frequencies from the desired band.

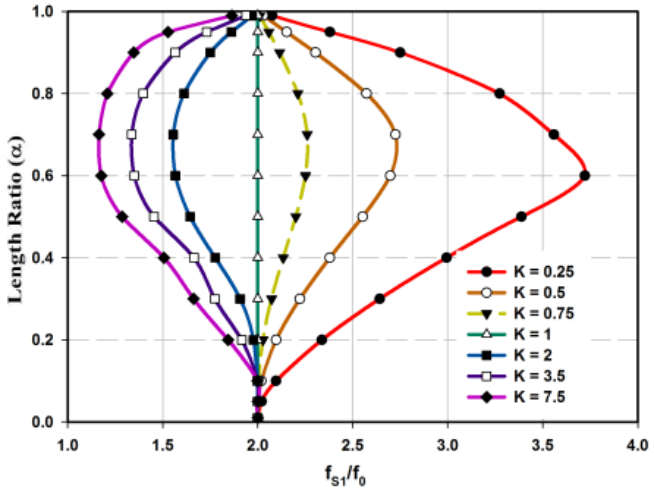


Fig 2: Dependence of fundamental and first harmonic separation on K and α

The resonators are arranged on both sides of 50 microstrip transmission line. When the received signal travel from one port to the other the present resonators absorbs energy at their resonance frequencies and the signal at the output port has a unique signature of the tag. The presence of dips in the frequency spectrum used indicates the presence of the particular resonator and can be considered as '1' and correspondingly absence as '0'. For tagging large number of items frequency shift coding can be used as described in [2].

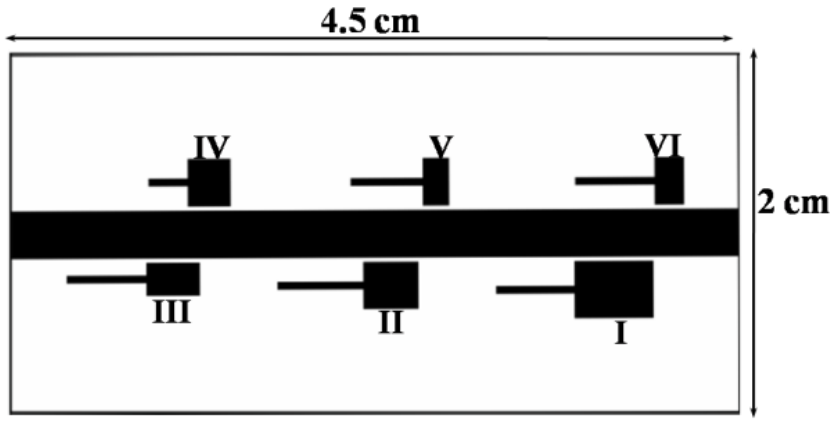


Fig 3: Geometry of the Multiresonator, height=1mm

Design of a six bit multiresonator section is presented in this section. The substrate used for fabrication is RTduroid (εr=2.2 and tanδ=.0008) with a thickness of 1mm. The geometry of the multiresonator circuit is shown in Fig 3 and the parameters of the resonators are given in Table I. Note the high impedance end of each resonator is shorted to excite the fundamental mode.

Table I: parameters of the resonators in the multiresonator circuit shown in Fig 3

Resonators	I	II	III	IV	V	VI
k	0.35	0.4	0.5	0.4	0.4	0.5
α	0.5	0.4	0.4	0.3	0.3	0.3
L1	5.05	5.53	5.14	5.12	4.65	4.59
L2	3.56	3.56	3.45	1.9	1.7	1.8
W1	0.5	0.5	0.5	0.5	0.5	0.5
W2	3.03	3.56	2.13	3.07	3.06	2.16

2. MULTIRESONATOR CIRCUIT DESIGN

A single SIR coupled to a microstrip transmission line and the response of the system in frequency domain is simulated in HFSS and the S12 is shown in Fig 4.

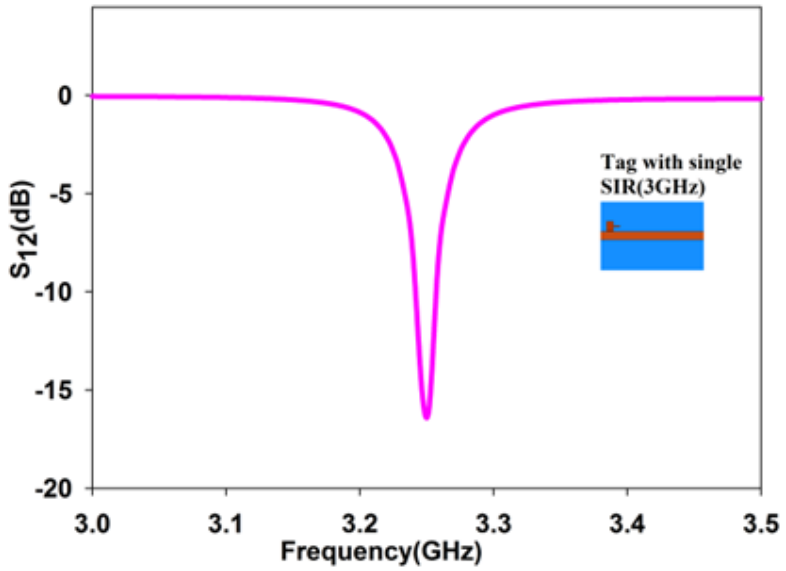


Fig 4: S_{12}(dB) of Single SIR coupled microstrip line

A 6 bit multi resonator circuit is simulated using HFSS. Simulation studies of coupling effects between microstrip line, between adjacent resonators, and different codes are conducted. Finally a prototype is also fabricated on RTduroid (εr=2.2 and tanδ=.0008). The simulation and measurement results are discussed in this section.

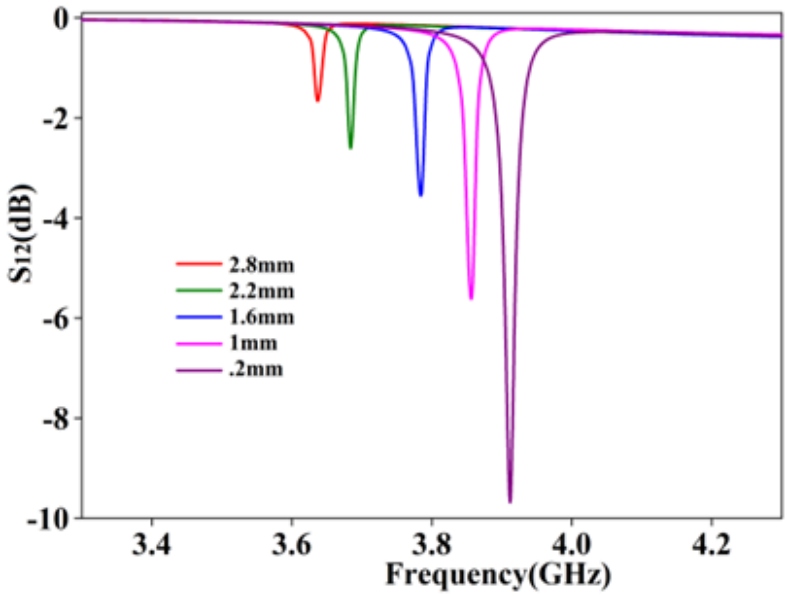


Fig 5: Coupling effect between a single resonator and microstripline

Coupling effect between the resonator and microstrip line for different distances is shown in Fig 5. Of course there is frequency shift and magnitude variation due to the capacitive effect due to the gap. For RFID applications narrow, deep resonances are desirable. With a minimum gap maximum identifiable dip is obtained. The optimum case of 0.3mm is selected.

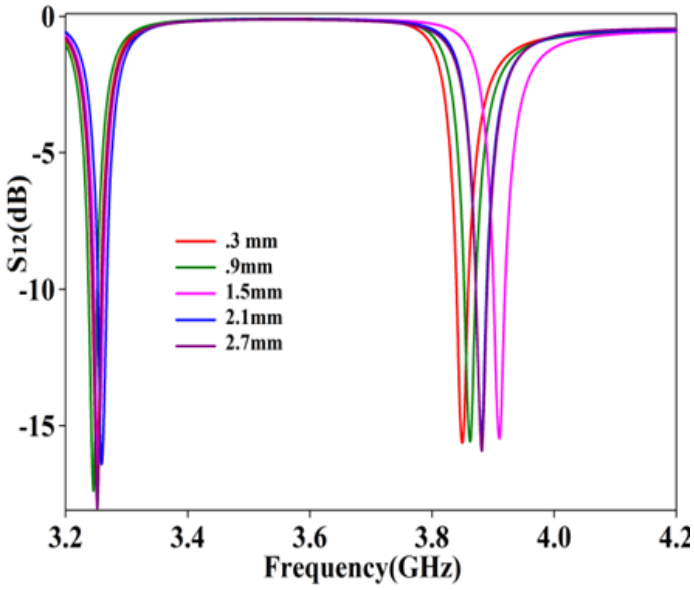


Fig 6: Coupling effect between two adjacent resonators

Effect of coupling between two adjacent resonators for different distances is shown in Fig 6. For distances greater than 0.3mm between the resonators shows good performance. Larger separation always provides good isolation but distance between the resonators must be accountable for the compactness of the tag.

Performance of multiresonator circuit with six resonators when signal is given between its two ports and observing S12 is shown in Fig 7. Results for two different codes are shown. The frequency dips are clearly identifiable so that the data can be easily decoded.

Finally the measurement result of fabricated multiresonator circuit of Fig 8 is shown in Fig 9. All the six resonances can be distinguished clearly.

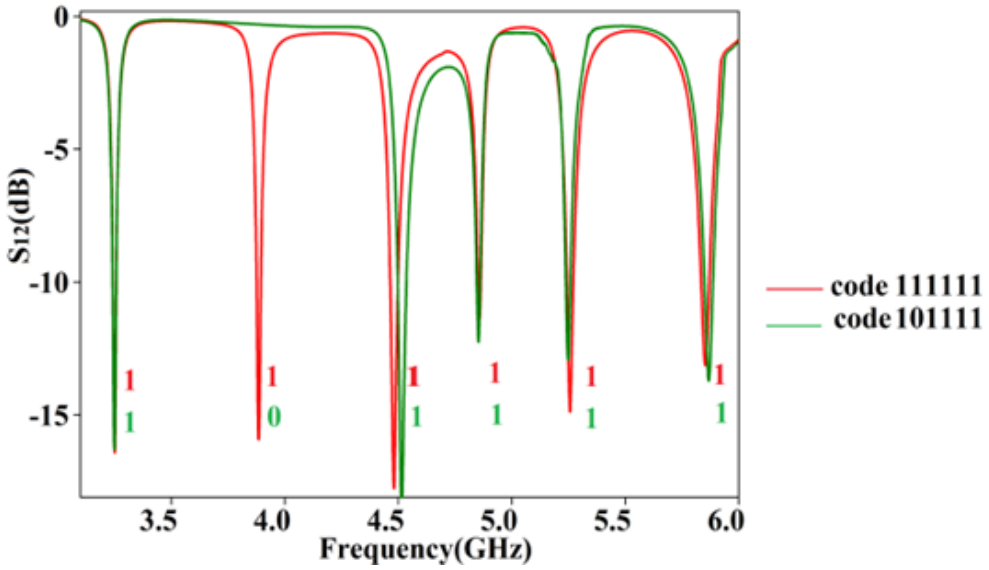


Fig 7: Simulated S12(dB) of the 6 bit multi resonator circuit for two different codes

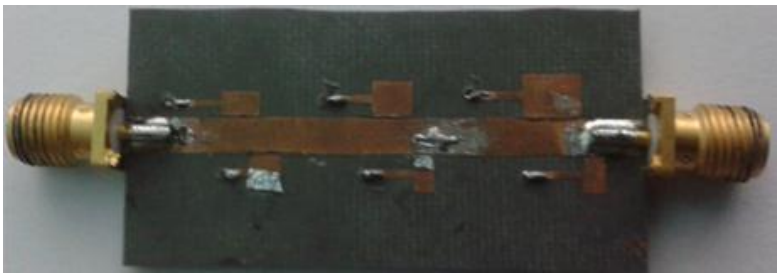


Fig 8: Fabricated Multiresonator circuit with all the resonators shorted

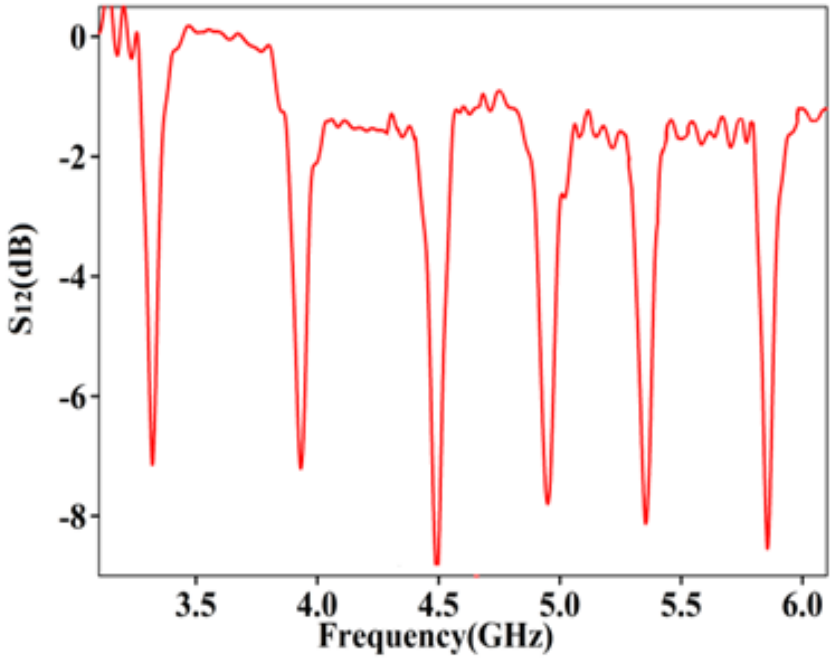


Fig 9: Measured S12 of 6 bit multi resonator circuit for the code 111111

More efficient bandwidth utilization can be achieved by using frequency shift coding technique and is very reliable particularly for this resonator due to its flexible design. More closely spaced and specially arranged resonators can enhance bandwidth as well as surface utilization. These two points and the complete tag measurement will be included as future works

3. CONCLUSIONS

6 bit Multiresonator circuit for Chipless RFID application is realized in an area of 4.5x2.5 cm2. Harmonic separation in the desired band is achieved. Presence/absence technique for Coding is done by simply removing or adding the short circuit. Frequency shift coding is also reliable for large number of items due to the flexible design of the basic resonator.

ACKNOWLEDGEMENTS

The Authors acknowledge Kerala State Council for Science Technology and environment, University Grants Commission Government of India and Department of Science and Technology Govt. of India for financial support.

REFERENCES

[1] S. Preradovic, N. C Karmakar, "Chipless RFID: bar code of the future", Microwave Magazine, IEEE,vol.11, No.7, December 2010, pp.87–97.

[2] A. Vena, E. Perret, S. Tedjini, "Design rules for chipless RFID tags based on multiple scatteres" Antenna telecommunication, 68, 2013:pp.361-374.

[3] S. Preradovic, Nemai Chandra Karmakar, Gerhard F. Swiegers "Multiresonator-Based Chipless RFID System for Low-Cost Item Tracking," IEEE Transactions On Microwave Theory And Techniques, VOL. 57, NO. 5, MAY 2009, 1411-1419

[4] C. M. Nijas, R. Dinesh, U. Deepak, Abdul Rasheed, S. Mridula, K. Vasudevan and P. Mohanan, "Chipless RFID Tag Using Multiple Microstrip Open Stub Resonators"IEEE Transactions On Antennas And Propagation, VOL. 60, NO. 9, SEPTEMBER 2012 4429-4432.

[5] Md. ShakilBhuiyan, AKM Azad, N Karmakar, "Dual band modified complementary split ring resonator (MCSRR) based multi resonator circuit for chiplessrfid tag.", IEEE Eightth international conference on Intelligent sensors, Sensor networks and information processing, APRIL2013,Melbourne,VIC

[6] I.Balbin,N.C.Karmakar, "Novel Chipless rfid tag for conveyor belt tracking using multi resonant dipole antenna.", proceedings of 39th European microwave conference,october2009,1109-1112, Rome Italy.

[7] M. Makimoto and S. Yamashita, "Microwave Resonators and Filters for Wireless Communication: Theory, Design and Application", Springer Series in Advanced Microelectronics. New York, NY, USA: Springer,1994,vol.4,pp.19–10

AUTHORS

V.R.Sajitha received the B.Sc. degree in electronics from the University of Calicut, Kozhikode, India, and the M.Sc. degree in electronics from Cochin University of Science and Technology (CUSAT), Kochi, India, in 2009 and 2011, respectively. She is currently working toward the Ph.D. degree at the same university. Her research interests include designing of MIMO antennas multiband antennas, ZOR antenna, electrically small antennas, inductive tuned antennas, chip less RFIDs, and UWB antennas.

C. M. Nijas received the B.Sc. degree in electronics from Mahatma Gandhi University, Kottayam, India, M.Sc and Ph.D. degrees in electronic science from Cochin University of Science and Technology (CUSAT), Kochi, India, in 2007 and 2009, respectively. His research interests include designing of chip less RFIDs, dielectric diplexer, multiband antennas, ZOR antenna, and UWB antennas.

K. Roshna received the B.Sc. degree in electronics from the University of Calicut, Kozhikode, India, and the M.Sc. degree in electronics from Cochin University of Science and Technology (CUSAT), Kochi, India, in 2009 and 2011, respectively. She is currently working toward the Ph.D. degree at the same university. Her research interests include designing of MIMO antennas multiband antennas, ZOR antenna, electrically small antennas, inductive tuned antennas, chip less RFIDs, and UWB antennas.

P. Mohanan (SM'05) received the Ph.D. degree in microwave antennas from Cochin University of Science and Technology (CUSAT), Kochi, India, in 1985. He worked as an Engineer with the AntennaResearch and Development Laboratory, BharatElectronics, Ghaziabad, India. Currently, he is aProfessor with the Department of Electronics, CochinUniversity of Science and Technology (CUSAT). He has authored more than 250 referred journal articlesand numerous conference articles. He also holdsseveral patents in the areas of antennas and material science. His research interests include microstrip antennas, uniplanar antennas, ultra wideband antennas dielectric resonator antennas, superconducting microwave antennas, reductionof radar cross sections, chipless RFID, dielectric diplexer, and polarization agile antennas.Dr.Mohanan was the recipient of Dr. S. Vasudev Award 2011 from Kerala Sate Council for Science, Technology and Environment Government ofKerala, in 2012, and Career Award from the University Grants CommissioninEngineering and Technology, Government of India, in 1994.

SIMULATION OF BASK,BPSK,BFSK MODULATORS USING VERILOG

Lakshmi S Nair[1] and Arun.K.L[2]

Department of Electronics & Communication, Mar Athanasius College Of Engineering,
A.P.J Abdul Kalam Technological University, Kerala, India
Assistant Professor, Department of Electronics& Communication,
Mar Athanasius College Of Engineering

ABSTRACT

This project presents the simulation results of digital modulation schemes BASK,BPSK and BFSK. In long distance transmission digital communication is more efficient and secure. In digital communication part noise detection and correction is very simple than analog communication. Digital modulation represents the transfer of digital bit stream from the transmitter to the receiver via the analog channels in an easy way. During the modulation the information signal modifies one or more carrier signal parameters leading to shift keying techniques. So it has more importance in modern communication systems. These three digital modulation schemes can be implemented using FPGA (Field Programmable Gate Array). This project employs the advantages of reliability, concurrent operation and minimum cost .It uses minimum number of blocks necessary for achieving BASK,BPSK and BFSK modulation techniques. In this project BASK,BPSK and BFSK modulation techniques have been implemented on FPGA using Verilog Hardware Description Language on Xilinx ISE 10.1 and simulated with Modelsim SE 6.5.

KEYWORDS

FPGA,BASK,BPSK,BFSK

1.INTRODUCTION

The objective of this paper is to design BASK,BPSK,BFSK digital modulators on FPGA.This paper employs the minimum number of blocks for achieving these three modulation techniques in an easy way. This is mainly used in software defined radio. Software Defined Radio (SDR) is defined as a a"Radio in which some or all of the physical layer functions are software defined". SDR defines a collection of hardware and software where some or all of the radio's operating functions (also referred to as physical layer processing) are implemented through software operating on programmable processing technologies. It has the ability to be transformed through the use of software or programmable logic.As we are implementing these modulators in FPGA we can reprogramme the values used by it that is its parameters can be varied.

In this project we are developing and compiling the modulators using verilog language.The verilog is a simple language for starters because it is similar in syntax with C programming .In this project these digital modulators are implemented on FPGA.The advantage of implementing these modulators are flexibility,low cost,low power consumption and time delay is very less compared to microcontrollers.

The choice of digital modulation scheme will affect the characteristics, performance and resulting physical realization of a communication system.while designing the system we have to consider the required data rate, predicted level of latency, available bandwidth, anticipated link budget and target hardware cost, size and power consumption. The objective of a digital communication system is to transmit digital data between two or more nodes. In BASK (binary amplitude-shift keying or OOK) modulation, the amplitude of the sinusoidal carrier signal is varied according to the information level , while keeping the frequency and phase of carrier signal constant. If information is 1, BASK modulated signal is carrier signal that is carrier signal is transmitted with out any change . But when information is 0, BASK modulated signal is 0. In a BPSK (binary phase-shift keying) modulation process, the phase of the sinusoidal carrier signal is changed according to the information level (“0” or “1”) while maintaining the amplitude and frequency of carrier signal constant.

The BPSK modulated signal is of positive values, if transmitting symbol is 1. But if transmitting signal is 0, starting of BPSK modulated signal is of negative values. In a BFSK (binary frequency-shift keying) modulation process, the frequency of the sinusoidal carrier signal is changed according to the information level (“0” or “1”) while keeping the amplitude and phase constant.

The methodology adopted in doing the project is that at first BASK,BPSK,BFSK modulators are stimulated in MATLAB SIMULINK one by one .In step 2 ,All the digital modulation schemes were coded in Verilog .In step 3 These codes are stimulated in ISE & MODEL SIM software In step these were then load in to XILINIX FPGA board.

2.LITERATURE SURVEY

Our project is FPGA implementation of BASK-BPSK-BFSK digital modulators .In this project each of these digital modulation technique was done using a multiplexer using coding. Here we are using minimum number of blocks[1].This section presents a broad overview of digital modulators, applications and commonly used hardware platforms for modulators.

Traditional modulators are using large number of building blocks .They are quite inflexible because it is difficult to change the parameters of modulating and carrier signal ,The most commonly used methods for modulator implementation are matlab implementation ,FPGA implementation, generation with self starting optoelectronic oscillator, DSP(Digital Signal Processor), general purpose microprocessors ,graphic processing units (GPU),ASICS(Application Specific Integrated Circuit) and through hardware circuits consiting of resistors ,CD 4016 IC etc.

General purpose microprocessors, such as the Intel and AMD devices usually found in personal computers, are not specialized for any particular application[2].Therefore, they are very flexible. However, SDR systems using general purpose processors are often wasteful since these processors are designed for speed and generality rather than power efficiency or mathematical

operations[3].
Graphics processing units uses massively parallel architectures that are optimized for vector manipulations and other graphical operations. Such parallel designs are very well suited for signal processing, but general purpose processors are relatively difficult to program and they consume high power .

A digital signal processor solve these two problems by fetching instructions and data from memory, does operations, and stores the results back to memory, just like a regular CPU[4]. The difference between a DSP chip and a CPU chip is that a DSP chip usually has a block that does high-speed signa lprocessing, especially a block called MAC (Multiply and Accumulate). By calling different routines in memory, a DSP chip can be reconfigured to perform functions. On the other hand, their narrow focus makes them slow for other applications.
ASIC (Application-specific Integrated Circuit) is an integrated circuit that is used to perform a fixed specific task[5]. Examples of signal-processing specific ASIC's are DDC (digital down converter) chip, and digital filter chips. The disadvantage of ASIC is that its functionalities are fixed and thus cannot be changed by the user according to his interest

MATLAB is a language and provides an interactive environment for numerical computation, visualization, and programming. Using MATLAB, we can easily analyze data, develop algorithms, and create models and applications[6]. The language offers tools, and built-in math functions enable us to explore various approaches and reach a solution faster than with spreadsheets or traditional programming languages, such as C/C++ or Java™. we can use MATLAB for a wide range of applications, including signal processing techniques and communications, image and video processing, control systems, testing and measuring, computational finance, and computational biology.

The disadvantage with hardware circuits is that whenever we are trying to vary the parametrers of modulating and carrier signal we have to vary the design of the circuit resulting in change of values of capacitors, resistors etc. Also they will result in large time delay compared to FPGA.

FPGA (Field Programmable Gate Array) is capable of performing any task by mapping the task to the hardware. One of the advantages of FPGA is its re-configurability capability that ASIC does not have. Re-configurability is a feature, which enables FPGA to realize any user hardware by changing the configuration data on a chip as many times as needed by the user, they are often programmed with a hardware description language ,like as Verilog or VHDL .

In summary, with its many advantages, FPGA has become key component in realizing high performance digital signal processing systems and digital communication systems. In this project, we will be using FPGA as the hardware platform for implementing BASK, BPSK,BFSK digital modulators.

3. SYSTEM DESIGN

We have an analog carrier signal & binary modulating signal .In modulation circuit the corresponding modulation will be performed. In a BASK modulation process as shown in figure1, the amplitude of the sinusoidal carrier signal is changed according to the message level (0/1) while keeping the frequency and phase of carrier signal constant .The transmitted signal for

OOK signal can be represented as: s(t)=m(t)c(t), where,m(t) is the modulating signal &c(t) is the carrier signal. The logic "1" and "0" are represented during any bit interval, Tb by the following signal set :

$$s(t) = \begin{cases} 0 & if\ "0" \\ A\cos 2\pi f_c t & if\ "1" \end{cases}$$

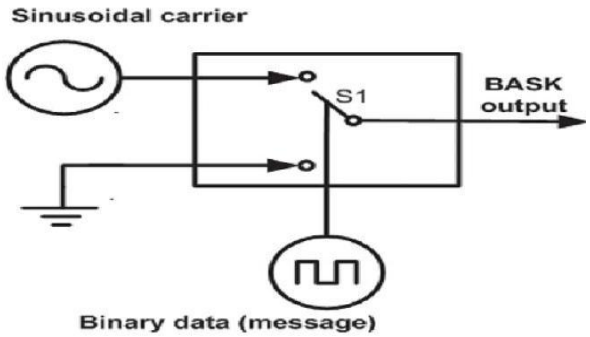


Figure1. Modulation circuit of BASK

In a BFSK modulation process as shown in figure2, the frequency of the sinusoidal carrier signal is changed according to the information level(0/1) while keeping amplitude and phase constant..A BFSK signal is represented as:

$$S_{BFSK}(t) = A\sin\left\{2\pi\left[f_c + m(t)f_m\right]t + \Phi_0\right\}, \quad 0 \le t \le T,$$

where m(t)=0 or m(t)=1 (the binary message), T is the bit duration, and A, f_c, and Φ_0 are the amplitude, frequency, and phase of the sinusoidal carrier signal.

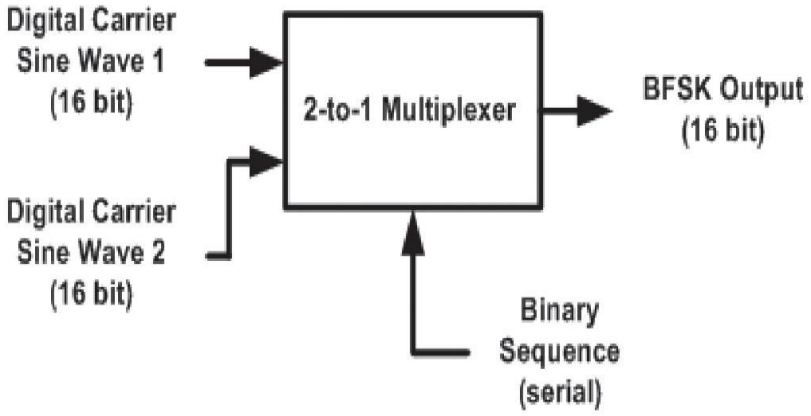


Figure2. Modulation circuit of BFSK

In a BPSK modulation process as shown in figure3, the sinusoidal carrier signals phase is changed according to the message level(0/1) while keeping the amplitude and frequency constant.A BPSK signal of frequency fc and amplitude Ac is represented as:

$$s(t)= \begin{cases} +Ac\ Sin(2\pi fct)\ ,if\ message=1 \\ -Ac\ sin(2\pi fct)\ ,if\ message=0 \end{cases}$$

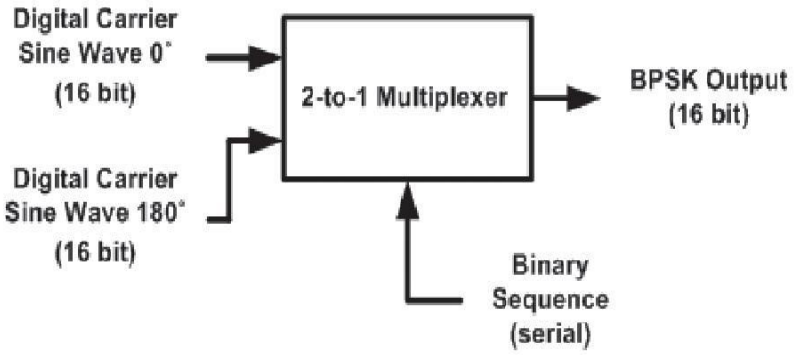


Figure3. Modulation circuit of BPSK

4.SOFTWARE REQUIREMENTS

4.1 MATLAB

MATLAB is a software tool providing an interactive enviornment for students.Matlab allows data to be plotted in graphical format,implementing data algorithmsvery easily,It will also provide the creation of user interfaces,it will also provide the easy interfacing of programs written in other software languages.

4.2VERILOG

Verilog, standardized as IEEE 1364, is a hardware description language (HDL) used to model electronic systems in an simplest way. It is mostly used in the designing and verification of digital circuits at the register-transfer with high level of abstraction. The portion of a hardware design isdescribed in Verilog as a Module. The module defines both the interface to the block or the

hardware and its internal structure or behavior. A number of primitives, or logic Gates, are built into the Verilog language. They represent basic logic gates (example and, xor). In addition User Defined Primitives(UDPs) may be defined according to our need .

4.3 MODELSIM

Mentor Graphics was the first to combine single kernel simulator (SKS) technology with a unified debug environment for Verilog language, VHDL, and SystemC. The combination of industry leading, native SKS performance with the integrated debug and analysis environment make ModelSim the simulator of choice for both ASIC and FPGA design technology. The best standards and platform in the industry make it easy to adopt in the majority of process. ModelSim SE 6.2 b combines high performance and high capacity with the code coverage and debugging capabilities required to simulate larger blocks and systems and attain ASIC gate-level sign-off. Comprehensive support of Verilog, VHDL, and SystemC provide a foundation for single and multilanguage design verification environments. ModelSim is a simulation tool. It doesnot create any hardware, even on the monitor. ModelSim just compiles the code, checks syntax and provides the waveform of the design behaviour according to the inputs values defined at the Test Bench file. Therefore, ModelSim is a tool for the functional checking of the designed one.

4.4 XILINX

Xilinx ISE 10.1 is a software tool manufactured by Xilinx for synthesis and analysis of HDLdesigns, which helps the developer to synthesize ("compile") their designs, performing timing analysis, examine RTL diagrams, simulate a design's reaction to different stimuli,and configure the target device with the programmer. The standard design flow consists of the following three major steps as shown in figure 4

4.4.1 DESIGN ENTRY AND SYNTHESIS

In the first step of the design flow, the design is created using a Xilinx-supported schematic editor, a Hardware Description Language (HDL) for text-based entry, or both. If HDL is used for text-based entry, the HDL file must be synthesized into an industry-standard Electronic Data Interchange Format (EDIF) file. If Xilinx Synthesis Technology (XST) is used a Xilinx-specific NGC netlist file is created, which can be converted to an EDIF file.

4.4.2 DESIGN IMPLEMENTATION

By implementing the specific Xilinx FGPA generation architecture, the logical design is converted into a file format, such as EDIF, that has been created in the design entry or synthesis stage into a physical file format. The physical data is contained in the Native Circuit Description (NCD) file. Then a bit stream file is created from these files and optionally programs a PROM for subsequent programming of the FPGA device

4.3 DESIGN VERIFICATION

Using a gate-level simulator, it is promising that the design meets timing requirements and

functions properly. In-circuit testing can be done by downloading the design to the device using Xilinx IMPACT Programming Software. Design verification can start immediately after design entry and can be repeated after various steps of design implementation

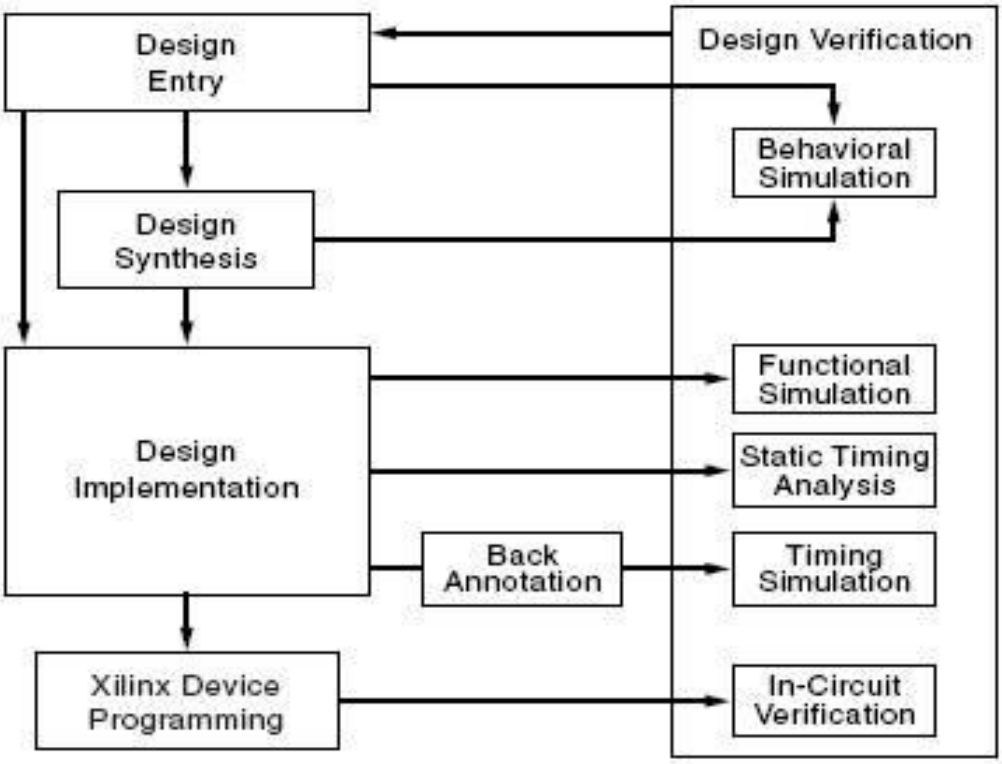


Figure4.Xilinix Design Flow

5. SIMULATION RESULTS

5.1 Modulations in Matlab

In the first phase of this project implementation of these digital modulators in Matlab Simulink has been done for studying the characteristics of each of these modulation technique. In Matlab we are generating sinewave with the help of sine function and we are performing modulation schemes by different arithmetic operations such as multiplication and logical operations by coding in a similar way to a mux.

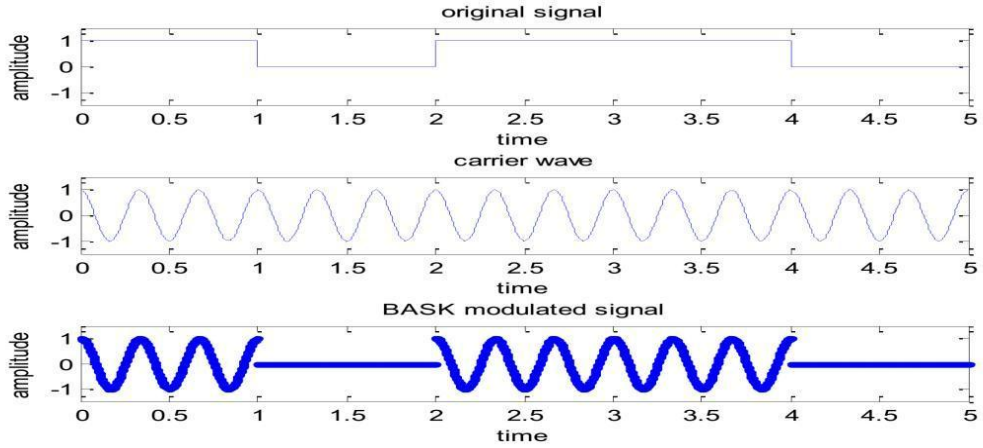


Figure5.BASK Modulation In Matlab

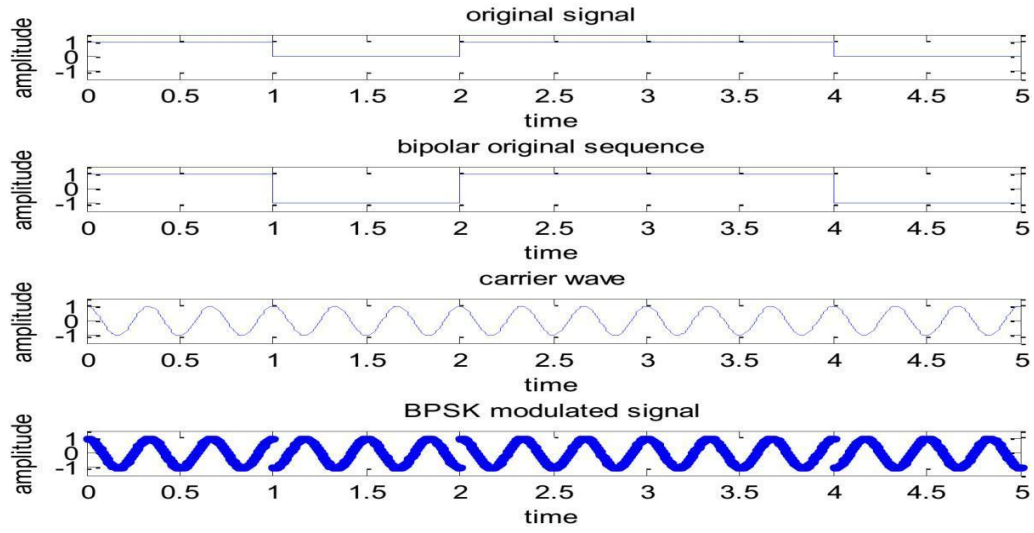


Figure6.BPSK Modulation In Matlab

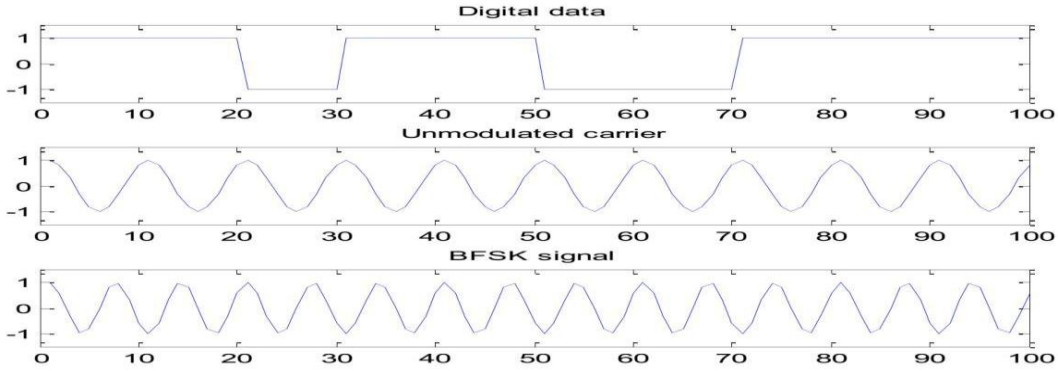


Figure7.BFSK Modulation In Matlab

In the BPSK signal for simplicity in doing modulation the binary input signal is converted to bipolar signal.

5.2 Modulations in Modelsim

Modelsim is a simulation tool. It doesn't create any hardware even on the monitor.Modelsim just compiles the code,check syntax of the code,and provides the waveform of the design behaviour according to the inputs values defined at the Test Bench file. ModelSim is a tool for the functional checking of the design.In modelsim the code is written in Verilog.The sine wave is generated using look up table approach.The values for the look up table are generated using sine function in MATLAB. This project we are using behavioural level approach.Actually in this project the carrier wave reaches a mux module,where the information signal acts like an selection input.The second phase of the project had been done in modelsim for functional verification.

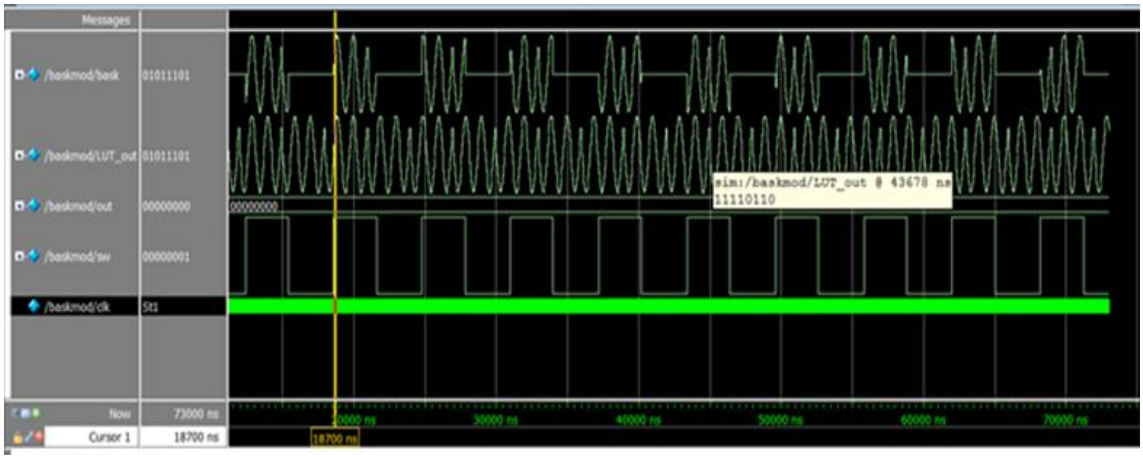


Figure8.BASK Modulation In Modelsim

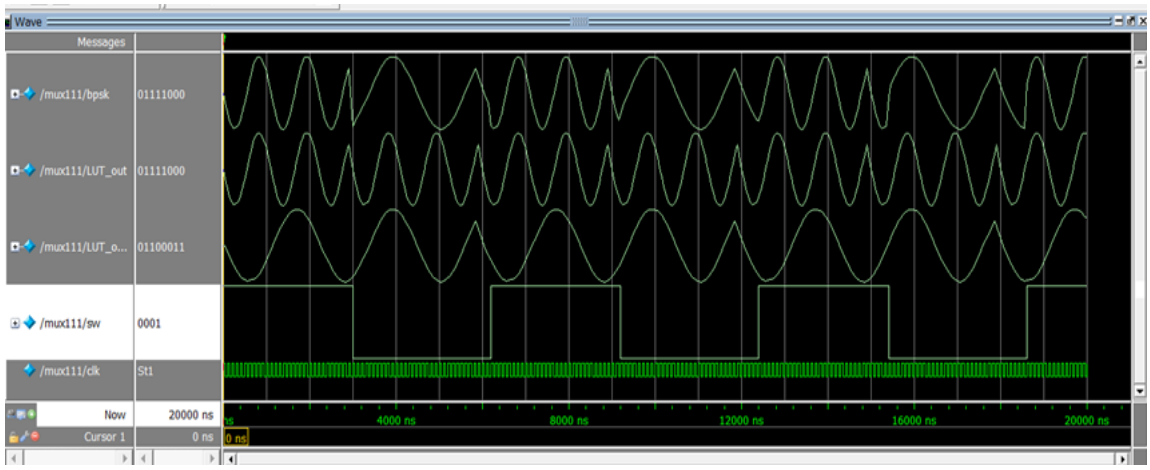


Figure9.BFSK Modulation In Modelsim

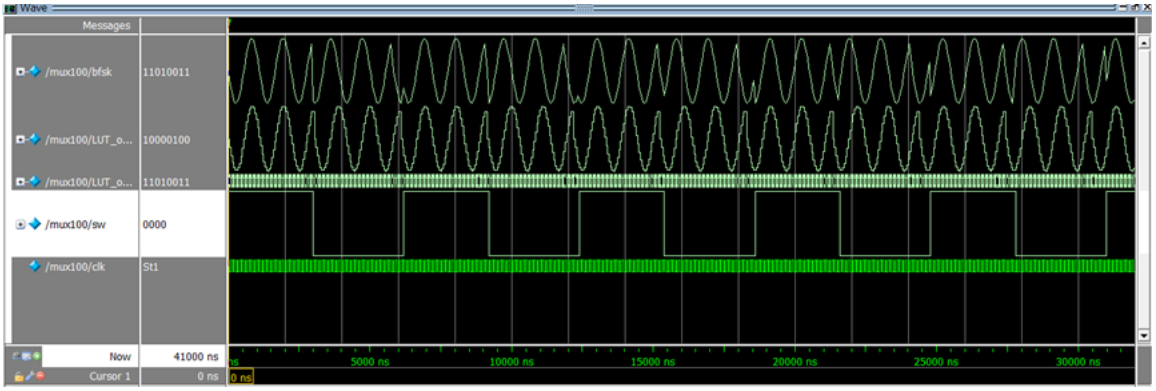


Figure10.BPSK Modulation In Modelsim

5.3 Modulations in Xilinx

Xilinx is a software tool which help the designer to synthesize their design.By doing simulation in this software designer will get a clear cut idea of how the software will be converted in to hardware in the FPGA.

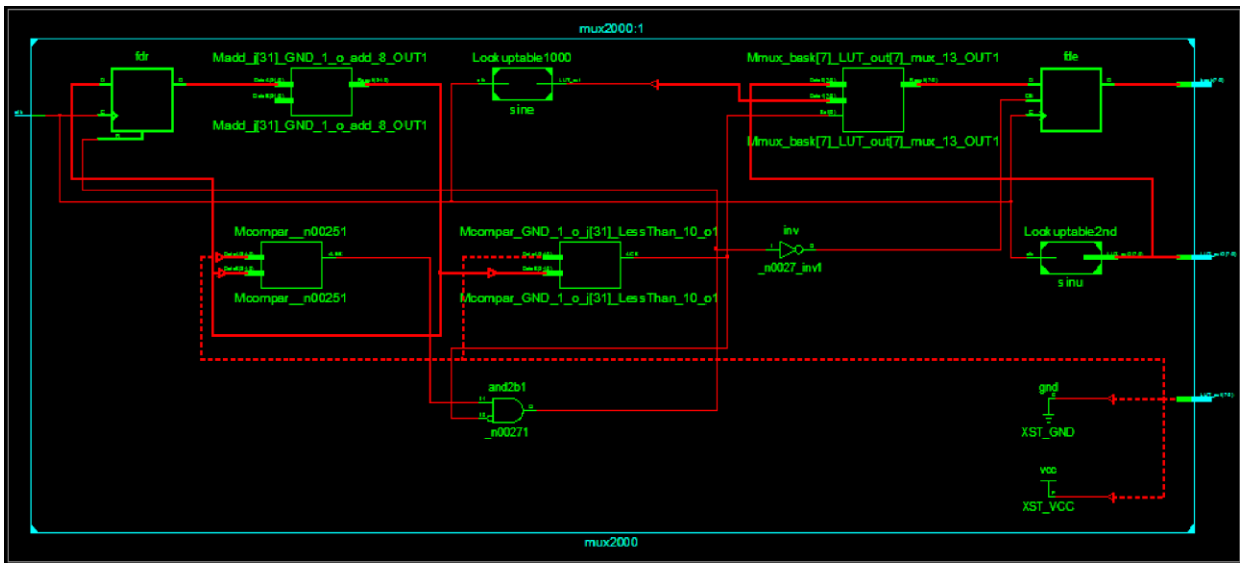


Figure 11:RTL View

The RTL view of modulation schemes is shown in figure11,here for BASK modulation the Lookup table 2nd will be filled with zeroes, Look uptable1000 should be filled with normal sine wave generating values, because for BASK modulation the output of the mux is switching between carrier signal and zero depending on the information signal. For BPSK modulation scheme both look up table are filled with same values for generating sine wave, but the BFSK modulation is made possible switching the Look up table 2nd with a different time delay. In BPSK, modulation both look up table have the same values, but the Look uptable2nd is switched in such a way that the look up table value will be negative values.

5.CONCLUSION

The project has been designed and developed successfully. Three modulations BASK,BPSK and BFSK has been coded in MATLAB successfully. Programs for each building block of modulation systems were written in Verilog. Mentor Graphics simulation tool ModelSim was used for writing the code, simulating the programs and to test its behaviour. Xilinx Synthesis tool was used to synthesize the modulation module by choosing Spartan 6 as the FPGA target device. The future scope is creating a single IC which can perform BASK,BPSK,BFSK modulations.

ACKNOWLEDGEMENT

The satisfaction and euphoria of successful completion of any task would be incomplete without the mention of the people who made it possible through their constant guidance and encouragement. I would like to extend my heartfelt thanks to Asso. Prof. Deepa Elizabeth George& Mr. George M Jacob of ECE DEPARTMENT TOCH Institute Of Science And Technology, ARAKKUNNAM for the inspiration inculcated in us and for the apt guidance.

REFERENCE

[1] Ms. Neha P. Shrirao & Prof. Ajay P. Thakare" Design of Digital Modulators: BASK, BPSK and BFSK usingVHDL "International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 1, January 2013

[2] Mehmet Sonmez,Ayhan Akbal"FPGA Based BASK,BPSK,BFSK Modulators Using VHDL:Design,Applications and Performance Comparison for Different Modulator Algorithm ,International Journal of Computer Applications ,Volume 42 ,March 2012

[3] C. Erdoğan, I. Myderrizi, and S. Minaei "FPGA Implementation of BASK-BFSK-BPSK DigitalModulators"IEEE Antennas and Propagation Magazine, Vol. 54, No. 2, April 2012

[4] www.dspg.com [5]www.asicworld.com [6]www.math.mtu.edu

About author

Lakshmi. S. Nair did her B.TECH in ELECTRONICS AND COMMUNICATION from TOCH Institute Of Science And Technology .Currently she is pursuing M.TECH in VLSI and Embedded System at Mar Athanasius College Of Engineering, Kothamangalam**.**

Arun K L is a faculty member of Mar Athanasius College of Engineering, Kothamangalam, Kerala, India. He received his B.Tech from Mahatma Gandhi University, Kottayam and M.Tech degree from the Karunya University, Coimbatore, India. His current research focus is in the area of VLSI Signal Processing.

SVD AUDIO WATERMARKING

Veena Gopan[1] and Mary Joseph[2]

[1]Department of Electronics and Communication, Mar Athanasius College Of Engineering, A.P.J Abdul Kalam Technological University, Kerala, India
[2]Associate Professor, Department of Electronics &Communication Engineering M.A.College of Engineering, Kothamangalam

ABSTRACT

It proposes an approach for audio watermarking using the singular value decomposition (SVD) mathematical technique. After transforming it into a 2-D format the encrypted image is embedded in the singular values of the audio signal. After watermark embedding, the audio signal is transformed again into a 1-D format. For encrypt the image chaotic encryption is used. It improves the quality of extracted images as proved experimentally, where it resists the noise and different attacks.

KEYWORDS

SVD, watermarking, baker mapping, chaotic encryption, decryption, SVs (Singular Values)

1. INTRODUCTION

Digital Watermarking has found many applications in image, video and audio transmission. In audio watermarking algorithms most of them are designed to achieve an efficient detection of the watermark without extracting meaningful information from the watermarked audio signal designed to achieve an efficient detection of the watermark without extracting meaningful information from the watermarked audio signal [8-9]. There is a need for a robust audio watermarking approach with a higher degree of security, which can be achieved by embedding encrypted images in audio signals [8-9]. In this project, the chaotic Baker map is used for the encryption of the watermark image [1-3]. Then, the watermark is embedded in the audio signal using the SVD mathematical technique [1]. The audio signal is transformed into a 2-D format .The singular values (SVs) of the resulting 2-D matrix are used for watermark embedding. From the speech signal at the receiver end the water mark is extracted and then the watermark is decrypted to get the message.

2. SYSTEM IMPLEMENTATION

In this project the cover speech signal is converted to 2-D for SVD water marking, then the message to be transmitted is encrypted by using baker mapping version A [1-3], and watermarked the message into the 2-D changed cover signal using SVD technique [1]. Then for transmission it is again converted to 1D.

DOI: 10.5121/ijci.2016.5240

At the receiver end the 1-D matrix is changed to 2-D for extracting watermark. The extracted watermark is decrypted to get the message, and then the result is compared with the original message. The block diagram is given in fig:2.1

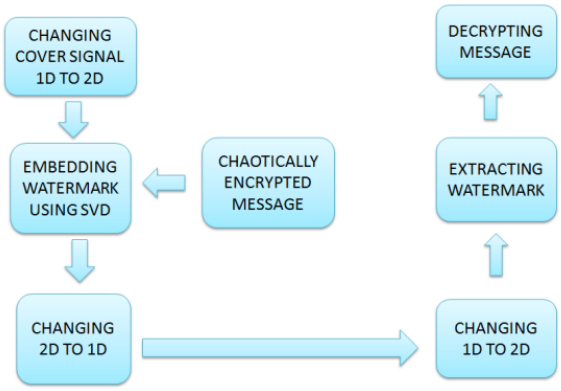


Fig 2.1 block diagram of system

To find out its response with noise, noise is added with the transmitted signal and the received message is compared with the original message using different signal to noise ratio.

2.1. SVD (SINGULAR VALUE DECOMPOSITION)

The Singular Value Decomposition decomposes the cover signal into 3 matrices each of size same as that of cover 2-D matrix [1].

SVD (Cover signal) = U S V'

Its singular value diagonal matrix is S of the cover signal, the encrypted message matrix is watermarked into the matrix S using a constant k=0.01, which makes the signal undistorted. And the watermarked matrix is again converted to 1-D for transmission.

2.2. BAKER MAPPING (ENCRYPTION)

Baker mapping encrypt the message block by block, which will be easy for decryption [2,3]. Here we uses version A baker mapping of 8X8 matrix. Version A baker mapping can be done by using 5 different keys. The key elements are selected according the following considerations

- The elements sum must be 8
- 1 is not taken as element
- 8/each element must be perfectly divisible
  The steps for baker mapping are given below
- Take 8*8 message matrix
- Encrypt the message using the equation
  $B(x', y') = A(\frac{N}{n_i}(x - N_i) + y \bmod \frac{N}{n_i}, \frac{n_i}{N}(y - y \bmod \frac{N}{n_i}) + N_i)$, Where N=8.
- B will be the baker mapped message

2.3. EMBEDDING WATERMARK USING SVD

The flowchart for embedding watermark is given in the figure 2.2. The audio signal is transformed into a 2-D format and the singular values (SVs) of the resulting 2-D matrix are used for watermark embedding. The chaotically encrypted message is embedded into the cover matrix using SVD method. And the 2-D matrix is transformed to 1-D

The steps for embedding a watermark using SVD method are given below

- The 1- D audio signal is transformed into a 2-D matrix (A matrix).
  The SVD is performed on the A matrix.
  $$A = U\,S\,V^{T}$$
- The watermark (W matrix) is added to the SVs of the original matrix.
  $$D = S + K\,W$$
- A small value of K of about 0.01 is required to keep the audio signal undistorted. The SVD is performed on the new modified matrix (D matrix
  $$D = U_W S_W V_W^T$$
- The watermarked signal in 2-D format (Aw matrix) is obtained using the modified matrix of SVs (Sw matrix).
  $$A_W = U\,S_W V^{T}$$

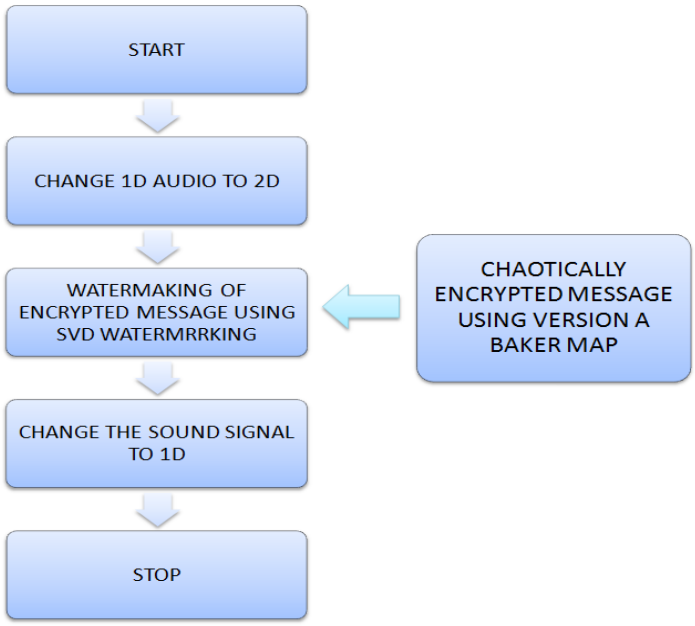


Fig:2.2 flowchart for embedding watermark

2.4. EXTRACTING WATERMARK

The flowchart for extracting watermark is given in the figure 2.3.The received 1-D is transformed to 2-D and the watermark is extracted. The watermark is decrypted to get the original message [1-3].

The steps for extracting watermark are

- The 2-D Aw matrix is transformed again into a 1-D audio signal.
- The SVD is performed on the possibly distorted watermarked image (A*w matrix).

$$A_W^* = U^* S_W^* V^{*T}$$

- The matrix that includes the watermark is computed.

$$D^* = U_W S_W^* V_W^T$$

- The possibly corrupted encrypted watermark is obtained.

$$W^* = (D^* - S)/K$$

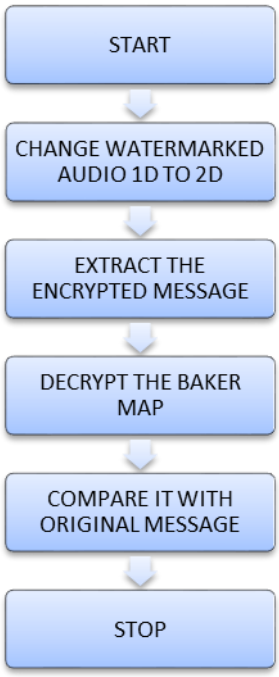


Fig.2.3 flowchart for extracting watermark

2.5. DECRYPTION OF BAKER MAP

- Take L=B'

- Use the following equation for decryption
  $K(x', y') = L(\frac{N}{n_i}(x - N_i)+y \bmod \frac{N}{n_i}, \frac{n_i}{N}(y-y \bmod \frac{N}{n_i})+N_i))$, Where N=8,

- Take E=K', then E will be the decrypted message

The message before baker mapping is given in the fig:2.4 and after baker mapping is given in the fig :2.5.Then change the message matrix with the baker mapped matrix for embedding.

1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

Fig: 2.4 message before baker mapping

1	9	17	25	2	10	18	26
33	41	49	57	34	42	50	58
3	11	19	27	4	12	20	28
35	43	51	59	36	44	52	60
5	13	6	14	7	15	8	16
21	29	22	30	23	31	24	32
37	45	38	46	39	47	40	48
53	61	54	62	55	63	56	64

Fig: 2.5 message after baker mapping using a key

3. RESULTS

Matlab software is used for simulation of the system which is explained in chapter 2. The cover signal required for system is 3 x 256 x 256 elements for transmission. For decryption we need U matrix and V matrix.256 x 256 U and 256 x 256 V is the watermarked matrix after SVD decomposition. They are embedded in to the cover signal. The encrypted matrix is embedded using SVD technique. The U and V matrix elements multiplied with 0.1 and added respectively with the elements of the cover signal. The cover signal is a sound signal and waveform of original sound is given in fig:3.1.

The message is of 8 x 8 size which is to be encrypted using baker mapping and watermarked using SVD to the above cover signal. The message example used is given in Fig: 3.2.The message is encrypted using baker mapping technique and the result will be like in figure fig:3.3 Then the encrypted message is watermarked into the cover signal using SVD method. And the U and V matrices are also added to the cover signal for decryption purpose. The resultant waveform of the water marked speech signal is given Fig: 3.4. The speech after watermarking makes no difference when played. And the waveforms also differ a little, which is hardly detectable. The extracted encrypted message from the watermarked sound wave is given in fig:3.5.

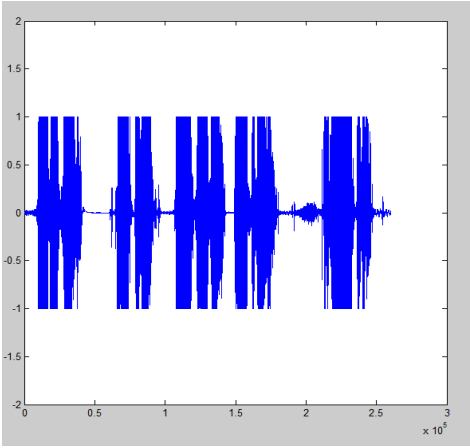


Fig: 3.1 waveform of original sound signal

123	114	90	89	98	102	119	139
132	127	101	97	111	116	132	145
116	125	156	146	157	162	143	141
104	98	169	148	161	169	124	115
134	59	147	146	161	143	93	94
157	109	113	134	138	116	106	105
166	149	141	132	132	130	147	131
141	122	150	140	139	138	117	103

Fig: 3.2 message

123	132	116	104	114	127	125	98
134	157	166	141	59	109	149	122
90	101	89	97	98	111	102	116
156	169	146	148	157	161	162	169
147	113	146	134	161	138	143	116
141	150	132	140	132	139	130	138
119	132	143	124	139	145	141	115
93	106	147	117	94	105	131	103

Fig: 3.3 encrypted message

The U and V matrices are extracted from the received speech signal and by using the reverse operation in baker mapping the message is decrypted. The decrypted message is given Fig: 3.6 the decrypted watermark is same as the original message. Then for studying the performance of the system, watermarked speech signal is subjected to noise. And by varying the SNR from 0 to 50, the performance is evaluated for 64 elements. A graph is plotted verses SNR and error. SNR greater than 45 dB indicates no error and hence the signal is extracted. The error increases with decreasing SNR. The figure is given in fig:3.7

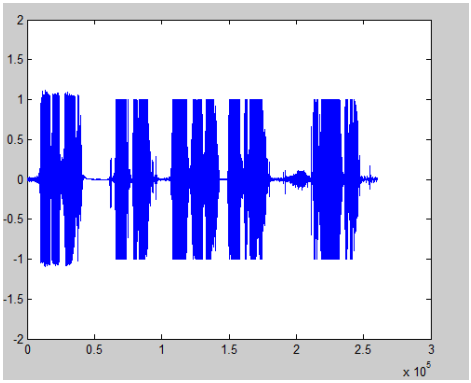


Fig: 3.4 waveform of watermarked audio signal

123	132	116	104	114	127	125	98
134	157	166	141	59	109	149	122
90	101	89	97	98	111	102	116
156	169	146	148	157	161	162	169
147	113	146	134	161	138	143	116
141	150	132	140	132	139	130	138
119	132	143	124	139	145	141	115
93	106	147	117	94	105	131	103

Fig: 3.5 extracted watermark from audio

123	114	90	89	98	102	119	139
132	127	101	97	111	116	132	145
116	125	156	146	157	162	143	141
104	98	169	148	161	169	124	115
134	59	147	146	161	143	93	94
157	109	113	134	138	116	106	105
166	149	141	132	132	130	147	131
141	122	150	140	139	138	117	103

Fig: 3.6 decrypted message

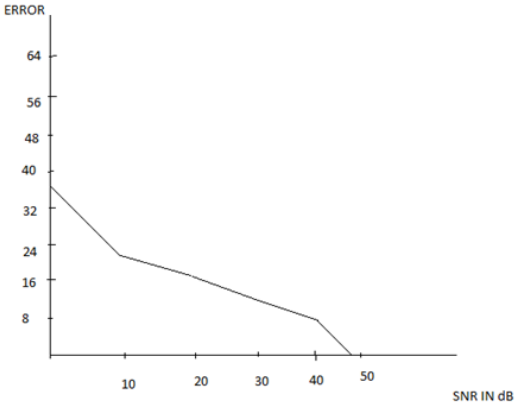


Fig 3.7 graph of SNR in dB against the error

3.1. ADVANTAGES AND DISADVANTAGES

The advantages and disadvantages of the SVD audio watermarking method are listed below
Advantages

- When SNR> 45 dB the received message has no errors.
- The encryption scheme used here is baker mapping which is more secure.
- The keys used for encryption are more secure which reduces Bruce-force guessing attacks.

Disadvantages

- Only 8x8 messages can be watermarked at a time, because encryption using baker mapping can be done for only 8X8 matrix at a time.
- The message length greater than 64 need to be cut into blocks of 64 elements.

4. CONCLUSION

This paper has presented an efficient security algorithm for Bluetooth network through SVD audio watermarking approach. In this algorithm, encrypted image is embedded as watermarks in audio signals to achieve a high degree of security. Experimental results have proved that watermark embedding in the proposed approach does not deteriorate the audio signals. It has been clear through experiments that the chaotic Baker map encryption algorithm is an efficient algorithm for watermark encryption.

ACKNOWLEDGEMENTS

The author takes this opportunity to thank all those who have been directly or indirectly involved in making this project a success. The author express her honest gratitude to the project guide Prof. DAVID SOLOMON GEORGE(Associate Professor, Dept. of ECE, RIT Kottayam) who is also the project co-coordinator, for his constant encouragement, inspiration and enthusiastic guidance and without whose continuous support and patience, it would have been extremely difficult to complete this work. The author expresses her sincere gratitude to all the teachers and staff of Rajiv Gandhi Institute of Technology.

REFERENCE

[1] M. A. M. El-Bendary, A. Haggag, F. Shawki, and F. E. Abd-El-Samie, "Proposed Approach for Improving Bluetooth Networks Security through SVD Audio Watermarking "2012 IEEE ,pp-594-598
[2] Feng HUANG, Yong FENG," Security analysis of image encryption based on two dimensional chaotic maps and improved algorithm"2009, pp-5-9
[3] Jiri fridrich, "Symmetric ciphers based on two-dimensional chaotic maps"1998, pp-1259-1284
[4] B. Macq, J. Dittmann, and E. J. Delp, "Benchmarking of Image Watermarking Algorithms for Digital Rights Management", Proceedings of The IEEE, Vol. 92, No. 6, pp. 971-984, 2004.
[5] Z. M. Lu, D. G. Xu, and S. H. Sun, "Multipurpose Image Watermarking Algorithm Based on Multistage Vector Quantization" IEEE Transactions on Image Processing, Vol. 14, No. 6, pp. 822-831, 2005.
[6] H. S. Kim and H. K. Lee, "Invariant Image Watermark Using Zernike Moments", IEEE Transactions on Circuits And Systems For Video Technology, Vol. 13, No. 8, pp. 766-775, 2003.
[7] W. C. Chu, " DCT-Based Image Watermarking Using Subsampling", IEEE Transactions on Multimedia, Vol. 5, No. 1, pp.34-38, 2003.
[8] L. Ghouti, A. Bouridane, M. K. Ibrahim, and Said Boussakta, "Digital Image Watermarking Using Balanced Multiwavelets", IEEE Transactions on Signal Processing, Vol. 54, No. 4,pp. 1519-1536, 2006.
[9] S. Xiang and J. Huang, "Histogram-Based Audio Watermarking Against Time-Scale Modification and Cropping Attacks", IEEE Transactions on Multimedia, Vol. 9, No. 7, pp. 1357-1372, 2007.
[10] Z. Liu and A. Inoue, "Audio Watermarking Techniques Using Sinusoidal Patterns Based on Pseudorandom Sequences", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 13, No. 8, pp. 801-812, 2003

AUTHOR

VeenaGopan received B-tech graduation from M G University in Electronics and Communication engineering. Now pursuing M-tech from A P J Abdul Kalam Technological University in VLSI and Embedded System.

Mary Joseph received M.Tech Degree in Microwave and Radar from Cochin University of Science and Technology (CUSAT), Kochi, India, in 1997. Currently she is working as Associate Professor in M. A. College of Engineering, Kothamangalam. She has joined in M. A. College of Engineering in 1991 as Assistant Professor. In between she worked at Birla Institute of Science & Technology-Pilani's (BITS-PILANI) Dubai Campus for 9 years as Assistant Professor during 2000-2008. Her Research interests include Microstrip Antennas and Uniplanar Antennas.

PERFORMANCE EVALUATION OF MYSQL AND MONGODB DATABASES

Dipina Damodaran B[1], Shirin Salim[2] and Surekha Marium Vargese [3]

Department of Computer Engineering, M A College of Engineering, Kothamangalam, Kerala, India

ABSTRACT

A database is a collection of information that is organized so that it can easily be accessed, managed, and updated. There are many databases commonly, relational and non relational databases. Relational databases usually work with structured data and non relational databases are work with semi structured data. In this paper, the performance evaluation of MySQL and MongoDB is performed where MySQL is an example of relational database and MongoDB is an example of non relational databases. A relational database (the concept) is a data structure that allows you to link information from different 'tables', or different types of data buckets. A non-relational database just stores data without explicit and structured mechanisms to link data from different buckets to one another.

KEYWORDS

Relational database,MySQL, MongoDB.

1. INTRODUCTION

The relational database has been the foundation of enterprise applications for decades, and since MySQL's release in 1995 it has been a popular and inexpensive option. Due the explosion of large volume and variety of datas in recent years, non-relational database technologies like MongoDB become useful to address the problems faced by traditional databases. MongoDB is very useful for new applications as well as to augment or replace existing relational infrastructure.

MySQL is a popular open-source relational database management system (RDBMS) that is distributed, developed, and supported by Oracle Corporation. The relational systems like, MySQL stores data in tabular form and uses structured query language (SQL) for accessing of data. In MySQL, we should pre-define the schema based on requirements and set up rules to control the relationships between fields in the record. In MySQL, related informations may be stored in different tables, but they are associated by the use of joins. Thus, data duplication can be minimized.

MongoDB is an open-source database developed by MongoDB, Inc. MongoDB stores data in JSON-like documents that can vary in structure. Related information can be stored together for fast query access through the MongoDB query language. MongoDB uses dynamic schemas,

DOI: 10.5121/ijci.2016.5241

which helps to create records without first defining the structure, such as the attributes or the data types. It is possible to change the structure of records by simply adding new attributes or deleting existing fields. This model helps to represent hierarchical relationships, to store arrays, and other more complex structures very easily. Documents in a record need not have an identical set of fields. MongoDB is designed with high availability and scalability includes replication and auto-sharding. In this paper, we perform a comparison on both MySQL and MongoDB on the platform of supermarket application.

The "Supermarket Management System "which manages the sales activity in a supermarket, maintaining the records of stock details, maintaining the records of the sales done for a particular month/year etc. Thus users will consume less time for calculation and the sales activity can be completed within a fraction of seconds whereas manual system will make the user to write it down which is a long procedure and it also needs a lot of time. The data can be stored in the database. Because of this software, paper work can be reduced and the user can spend extra time for monitoring the supermarket. MongoDB is more applicable to large databases but for the simplicity we take supermarket data.

2. PROBLEM DEFINITION

This section gives a brief definition on MySQL and MongoDB. Then evaluate the performance of both the databases on the application of hypermarket. When compared to MySQL it is observed that Mongodb is much better in query processing [9][12]. The MongoDB database consists of a set of databases in which each database contains multiple collections. Because MongoDB operates with dynamic schemas, every collection might contain different types of datas. Every object also called as documents is represented by a JSON structure: a list of key value pairs. The value can be of mainly three types: a primitive value, an array of documents or a list of key-value-pairs. For to query these objects, the client can set the collections expressed as a list of key value pairs. It is also possible to query nested fields. The queries are also JSON like structured; hence a complex query can take much more space than the same query for the relational databases. If the built-in queries are too limited, it is possible to send JavaScript logic to the server for more complex queries.

MongoDB supports mainly two types of replication: master-slave and replica sets. In the master-slave replication, the master has control of full data access and which writes every change to its slaves. The slaves can only possible to read data. Replica sets works same as master-slave replications, but it is possible to select a new master if the original master become down. Another important feature that supported by MongoDB is automatic sharding. Using this feature data can be partitioned to different nodes. The administrator has to verify a sharding key for each collection which defines how to partition the documents. In such an environment, the clients connect to a special master node called mongos process which analyses and redirects the query to the appropriate node or nodes. To eliminate data losses, every logical node contain physical servers which act as a replica set. Using this infrastructure it is also possible to use Map/Reduce having a very good performance.

2.1 Architecture

MongoDB supports standalone or single instance operations. The replica sets provide high performance of replication with automated failure handling, while sharded clusters make it possible to divide large data sets over different machines which are transparent to the users. MongoDB users combine replica sets and sharded clusters to provide high levels of redundancy of data sets which are transparent for applications [7]

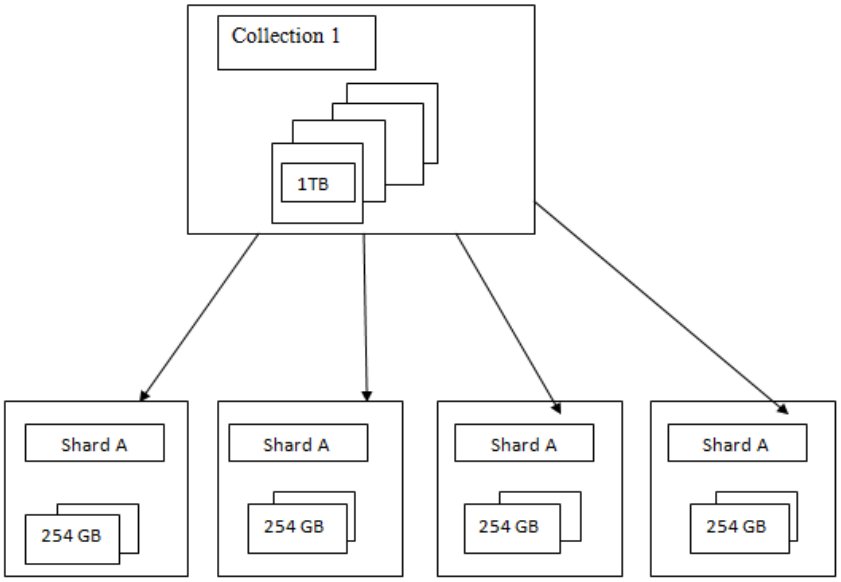


Figure 1.Deployment Architecture

MongoDB supports sharding through the configuration of a *sharded clusters*.

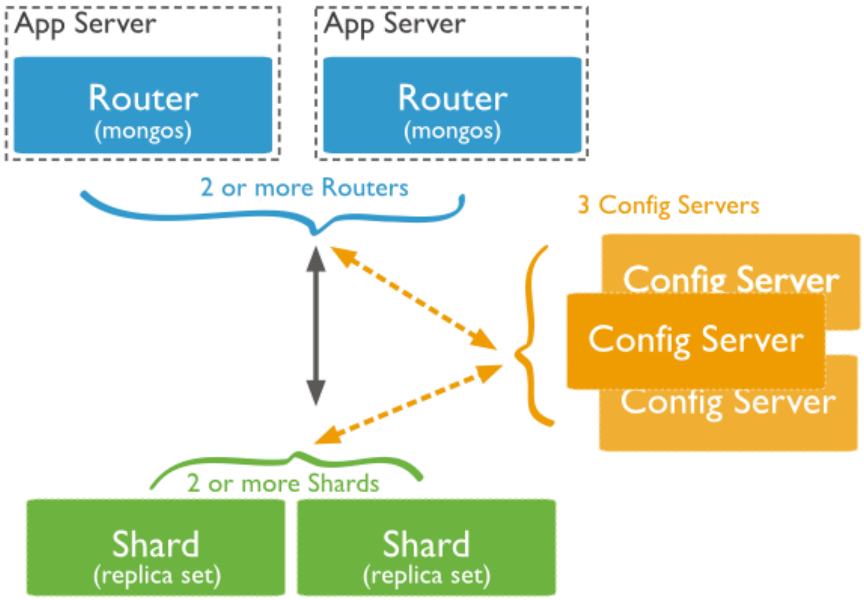


Figure 2.Sharding in Mongodb

Sharded cluster has the following components: shards, query routers and config servers.

- ***Shards*** store the data. To provide high availability and data consistency, in a production sharded cluster, each shard is a replica set For more information on replica sets, see Replica Sets.
- ***Query Routers*** interface with client applications and direct operations to the appropriate shard or shards. The query router processes and targets operations to shards and then returns results to the clients. A sharded cluster can contain more than one query router to divide the client request load. A client sends requests to one query router. Most sharded clusters have many query routers.
- ***Config servers*** store the cluster's metadata. This data contains a mapping of the cluster's data set to the shards. The query router uses this metadata to target operations to specific shards. Production sharded clusters have *exactly* 3 config servers.

3. METHODOLOGY

Organizations of all sizes commonly adopting MongoDB because it enables them to build applications which are faster, handle highly diverse types of datas, and manage applications more efficiently at scale. MongoDB documents map naturally to modern, object-oriented programming languages. MongoDB removes the complex object-relational mapping (ORM) layer which translates the objects in code to relational tables. MongoDB's flexible data model helps the database schema can evolve with business requirements. For example, the ALTER TABLE command required to add a single, new field to Craiglist's MySQL database would take months to execute. The Craigslist team migrated to MongoDB because it helps to accommodate changes to the data model without such costly schema migrations.

MongoDB can scale within and across multiple distributed data centers, providing new levels of scalability and availability which are unachievable with relational databases like MySQL. As your deployments grow in terms of data volume and throughput, MongoDB scales easily without much downtime, and without changing the application. but, to achieve scale with MySQL, it often requires significant, custom engineering work. While modern applications require a flexible and scalable system like MongoDB, there are use cases for which a relational database like MySQL are better suited. MongoDB is not a drop-in replacement for legacy applications built around the relational data model and SQL.

A concrete example would be the booking of tickets behind a travel reservation system, which also involves complex transactions. While the core booking system might run on MySQL, those parts of the app that system with users – serving booking, integrating with social networks, managing sessions – would be better when placed in MongoDB. MongoDB came with the aim of giving the new way of data storage. Therefore database provide storage of document for the World Wide Web. Began in 2007, MongoDB is built to store data in a dynamic schema, instead of a tabular representation like SQL. The data in MongoDB is stored in the form of object notation based on the format of JSON (Java Script Object Notation). JSON is a standard for the data transfer over the network between the server and web application which use human readable format. Prior to JSON, the XML was used for that purpose. MongoDB modified the JSON format into its own BSON, which store the object as a binary format. Hence the BSON stands for Binary JSON. BSON, due to its binary format provide more reliable and efficient in the area of storage space and speed.

4. EXPERIMENTAL RESULTS

The results of experiments performed to test various aspects of the implementation employed in hypermarket are provided in this section i.e., using the insertion and search operations on databases for auditing purposes .The various operations are performed on the two databases and we obtain the below results.

Table 1 for insertion and searching operations:

Operations	**No.of Records**	**Execution Time (in ms)**	
		MongoDB	**MySQL**
INSERTION	100	0.01	0.01
	1000	0.5	1.25
	10000	1.2	2.2
	25000	2.25	3
SEARCH	100	0.05	0.152
	1000	0.12	1.52
	10000	0.55	4.47
	25000	1.25	5.21

We study the performance of Mongodb while comparing with SQL by performing two operations, Insertion and Searching. A large no of records were taken and performed the operations in both databases. The graph plotted based on the performance is shown below.

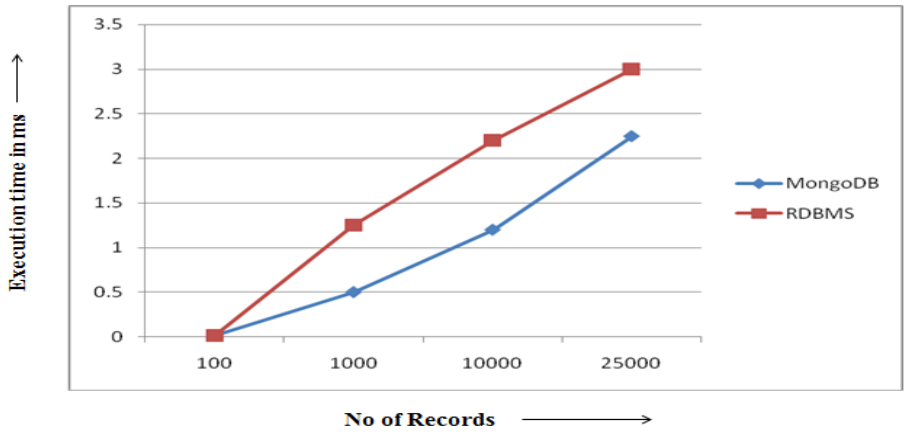


Figure 3. Insertion operation

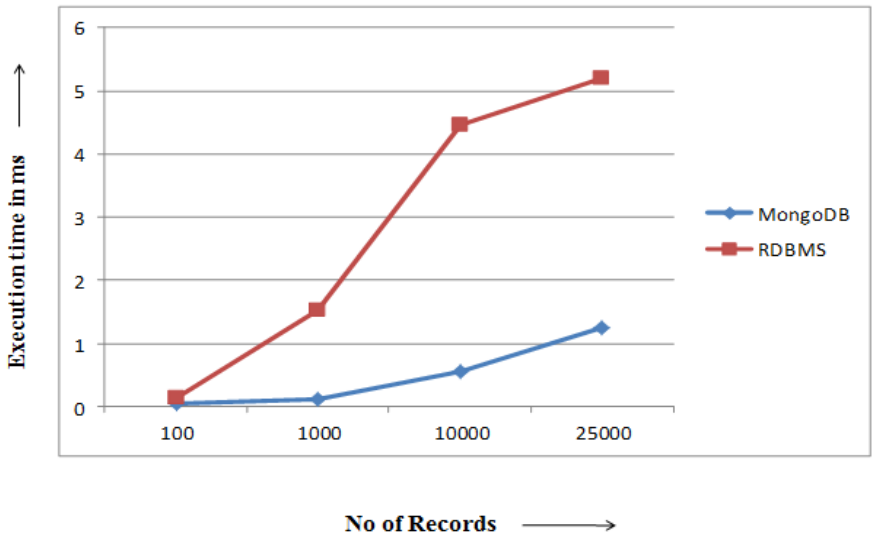


Figure 4.Searching operation

5. PERFORMANCE EVALUATION

On analysing the performance of MySQL and MongoDB databases on hypermarket application, the performance of MongoDB is more when compared to that of MySQL. Organizations of all sizes are commonly adopting MongoDB because it enables them to build applications faster, handle highly diverse types of data, and manage applications more efficiently at large scale.

Development is simplified because MongoDB documents map naturally to modern, object-oriented programming languages. Using MongoDB, it removes the complex object-relational mapping (ORM) layer that translates the objects in code to relational tables. MongoDB's flexible data model helps that the database schema can evolve with business requirements.

One of the most important drawbacks of relational databases is that each item can only contain single attribute. Consider a bank example; a customer's relationship with a bank is stored as different row items in separate tables. So each customer's master details are stored in one table, the account details of those customers are in another table, the loan details in yet another table, investment details are in a different table, and so on. But these tables are connected to each other by use of relations like primary keys and foreign keys. Non-relational databases, use key-value stores or key-value pairs, are different from this model. Key-value pairs provide possibility to store several related items in one "row" of data in the same table. For instance, in a non-relational table for the same bank example, each row can store the customer's details as well as their account details, loan and investment details. All data relating to one customer can conveniently stored as one record. This implies an obviously superior method for storing of data, but it has a major limitation: key-value pairs, unlike relational databases, it cannot use relationships between data items. In key-value databases, the customer details like (name, social security, address, account number, etc.) are stored in one data record (instead of stored in several tables, as in the relational model). The customer's transaction details (account withdrawals, account deposits, loan repayments, etc.) would also be stored as another single data record.

6. DISCUSSION

MongoDB is widely used in the field of large databases. One of the most important advantages is its scalability. MongoDB follows BASE transaction, Basically Available Soft State and Eventual consistency .Another important feature is handling of failures.For the simplicity we conduct an analysis based on super market. But mongaDB is more suitable for other applications having large volume of data where data need high security. Since it is schemaless , it supports different types of data.

7. CONCLUSION

In this paper, we undergo performance evaluation between MySQL and MongoDB on hypermarket application. For evaluating its performance execution time is considered. We came to a conclusion that when number of records inserted or searched is smaller, there is no difference in the execution time taken for each of these operations to complete for both MongoDB and MySQL databases. However, when number of records is increased, MongoDB shows significant reduction in the time taken for execution compared to MySQL. Thus, when the number of records is higher, MongoDB takes less time compared to MySQL. MongoDB can be preferred for better performance.

So in summary, RDBMS's suffer from no horizontal scaling for high transaction loads (millions of read-writes), while NoSQL databases solve high transaction loads but at the cost of data integrity and joins.

REFERENCES

[1] K. Sanobar, M. Vanita, "SQL Support over MongoDB using Metadata", International Journal of Scientific and Research Publications, Volume 3, Issue 10, October 2013

[2] https://www.mongodb.org/about/introduction/

[3] S. Hoberman, "Data Modeling for MongoDB", Publisher by Technics Publications, LLC 2 Lindsley Road Basking Ridge, NJ 07920, USA, ISBN 978-1-935504-70-2, 2014.

[4] http://dwhlaureate.blogspot.in/2013/10/features-of-mongo-db.html

[5] R. P Padhy, M. R. Patra, S. C. Satapathy, "RDBMS to NoSQL: Reviewing Some Next-Generation Non-Relational Database's", International Journal of Advance Engineering Sciences and Technologies, Vol. 11, Issue No. 1, 015-030, 2011.

[6] https://en.wikipedia.org/wiki/MongoDB

[7] https://en.wikipedia.org/wiki/shard-(database-architecture)

[8] Z. Wei-Ping, LI Ming-Xin, H. Chen, "Using MongoDB to Implement Textbook Management System instead of MySQL", IEEE 3rd International Conference on Communication Software and Networks (ICCSN), ISSN 978-1-61284-486, 2011.

[9] J. Clarence, M. Tauro, S. Aravindh, A. B. Shreeharsha, "Comparative Study of the New Generation, Agile, Scalable, High Performance NOSQL Database", International Journal of Computer Applications, ISSN 0975 – 888, Volume 48– No.20, June 2012.

[10] http://docs.mongodb.org/manual/core/sharding-introduction/

[11] https://en.wikipedia.org/wiki/JavaServer_Pages

[12] https://www.mongodb.com/compare/mongodb-mysql

AUTHORS

Dipina Damodaran B completed her B.Tech degree from Malabar College of Engineering and Technology,Trissur in 2013 which is affiliated to Calicut University. She presented paper on National Level Conference. She is currently pursuing M.Tech in Computer Science in Computer Science and Engineering in Mar Athanasius College of Engineering. Her areas of research are Modern Databases, Data Structure and Data Mining.

Shirin Salim currently pursuing M.Tech in Computer Science and Engineering in Mar Athanasius College of Engineering. She completed her B.Tech degree from Ilahia College of Engineering in 2014 which is affiliated to Mahatma Gandhi University. She presented paper in National Conference. Her areas of research are Modern Database, Data Mining and Machine Learning.

Surekha Mariam Varghese is currently heading the Department of Computer Science and Engineering, M.A. College of Engineering, Kothamangalam, Kerala, India. She received her B-Tech Degree in Computer Science and Engineering in 1990 from CET affiliated to Kerala University and M-Tech in Computer and Information Sciences from CUSAT, Kochi in 1996. She obtained Ph.D in Computer Security from CUSAT, Kochi in 2009. Her research interests include Network Security, Database Management, Data Structures and Algorithms, Operating Systems and Distributed Computing. She has published 17 papers in international journals and international conference proceedings. She has been in the chair for many international conferences and journals.

FPGABASEDACQUISITIONANDTRANSMISSIONOFDATA INSONAR

Anagha A V[1] &Mary Joseph[2]

[1]Department of Electronics and communication, Mar Athanasius College of Engineering , A P J Abdul Kalam Technological University, Kerala, India
[2]Associate Professor, Department of Electronics &Communication Engineering M.A.College of Engineering, Kothamangalam

ABSTRACT

System development in FPG As allows considerable flexibility,during development and in production use. However, this flexibility comesat the cost of increased complexity. I have designed a modular development framework to help to overcome these issues of increased complexity. The development of the framework hasbeen divided into two phases. The first phase Is to interface preamplifier and ADC using SPI communication protocols while the second phase is to design the Ethernet interface. In this studies I developed the system in Virtex5 FPGA on ML505 evaluation platform, here IDEtool XilinxISE13.3i and Wireshark is used for control and communication.

Keywords

SONAR, Field Programmable Gate Array, Analog to digital convertor, Ethernet

1. INTRODUCTION

Underwater acoustics a key underpinning technology in off shoreoil and gas activities, is increasingly use dinoceanographic and environmental studies and continue stop lay a crucial role in defence. [1]Sound Navigation And Ranging (SONAR) uses sound propagation (usually under water, as in submarine navigation) to navigate, communicate or detect objects on or under the surface of the water, such as other vessels.

The challenge of SONAR signal processing is to detect/classify targets in anoisy environment. In signal processing literature, SONAR signal can be classified to wide range of applications according to the SONAR mode. The class of marine vessels is defined according to the application of surveillance system for example surface and sub.[2]For decades, the trained people classified and recognized the class of marine vessels by listening to the irradiated noise. Substituting these people with intelligent systems for classifying marine vessels based on their acoustic radiated noise is one of the hot topics in signal processing and artificial intelligence. Although intelligent methods outperformed classic methods in the signal processing using the same data[2] they need large amount of real data for learning, thus a considerable number of research papers are based on simulations of SONAR signals and environments. In addition, new

DOI: 10.5121/ijci.2016.5242

Trends on mixing of classic and modern methods have been introduced. Infact, mixed methods strengthen advantages of each approach and reduce the irdis advantages. FPGA-based acquisition and transmission control systems are especially interesting because they allow parts of the system to be easily migrated between hard ware and software implementations[3]for optimal performance and resource use[4].

This project introduces a highly reliable system for realizing programmable gain amplifier(PGA)[3] based on digital potentiometers as opposed to multiple DAC.The choice of multiple DAC oradigipot involves a trade-off between resolution and speed. Although DACh as a higherre solution, digipot is better suited for multiple a coustic channel Sonar where speed is the most contributing factor.

2. SYSTEM DESIGN AND SPECIFICATION

An automatic gain circuit is provided which modifies the beam of signals based upon a history of the signal level present in previous returns. Design make suse of digitally controlled potentiometers (Digipot) programme dusing an automatic gain control algorithm to condition the transducer in put in the Programmable Gain Amplifier (PGA)as shown in the block diagram given in Fig.1

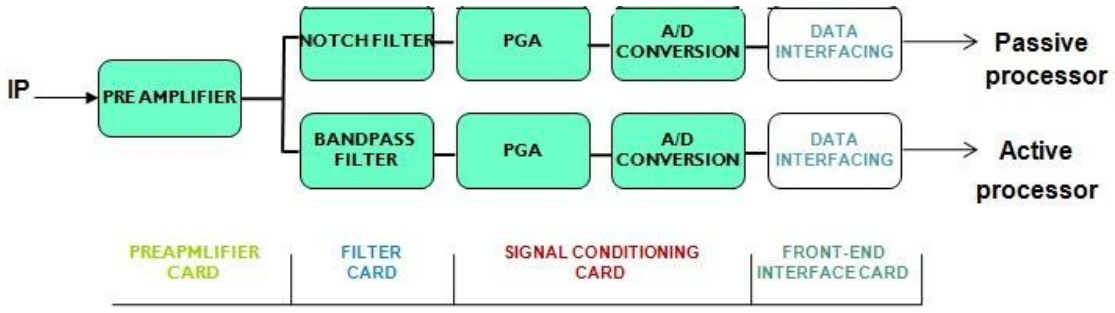


Fig.1. BlockDiagram

The basic functions of the Signal Conditioning Card are: Line Receiving the pre amplified signal from the signal conditioning PCB, High pass filter, Whitening, PGA, Anti- aliasing Filter and Analog to Digital Conversion of acoustic channels. The PCB is designed to cater for16 acoustic channels as shown in the input output diagram of signal conditioningcardinFig.2.

Specifications of the Signal Conditioning Card are Number of analog input channels: 16 Maximum, Maximum Signal Input:125mVpeak at preamplifier input2.5Vpeakat Line Receiver input, Programmable Gain : -10dB to +80dB in steps of 1dB , Signal Bandwidth:100Hzto12KHz, Sampling Frequency: 31.25K sample sperse cond, ADC Resolution:12bits.

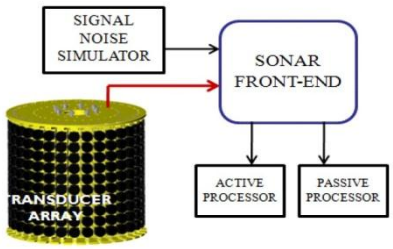


Fig.2. SignalConditioningCard

The platform is chosen to be FPGA over the conventional embedded environment. FPGAs are hard wired and the random attack of alpha ray scan not destroy/corrupt the memory are as hence collapse the device functionality. Life time of the FPGA based development is longer when they areopted in critical conditions. Micro controllers change too often. Hencelots of re-workre quired to do in order to keep pace with changing technology.

FPGAs can be easily adopted for advanced chip when situation demands. This is necessary to save the design from being obsolete. Hence the reconfigurable platform of FPGAs is preferred in real time military applications.

2.1 CIRCUITDESCRIPTION

A programmable-gain amplifier (PGA)is an amplifier (typically an op-amp) whose gain can be controlled by external digital or analog signals. The Programmable Gain Amplifier section of the PCB is realized using two independent gain stages both of them, programmable .The Programm able Gain Amplifier is realized as per the configuration shown in Fig.3.

Each of the Programmable gain stage is realized using digipot.The digipot is connected in the feed back path of the op-amp in order to vary the gain.U(x)8(AD5262) is a dual digital potentiometer with two independent 8bit register for storing the gain parameter.

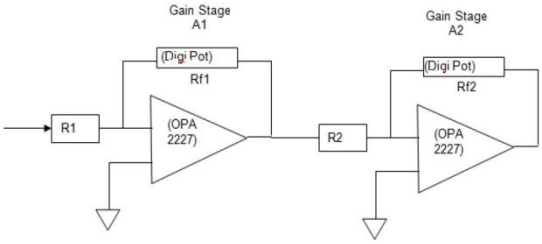


Fig.3. Programmable Gain Amplifier Configuration

The digital data interface for digipot is provided through the SPI interfacing. All the 16 devices required for 16 programmable gain amplifiers are cascaded so that only three serial lines are required for loading the gain parameter for all the Programmable gain amplifiers. The scheme for cascading the digital potentiometers is shown in Fig.4

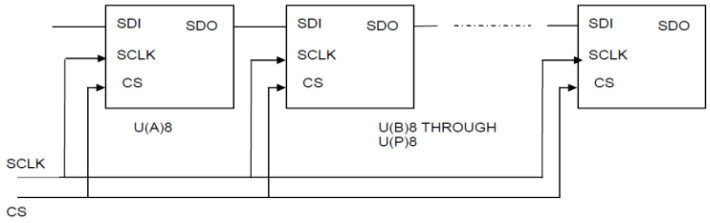


Fig.4 The scheme for cascading the digital potentiometers

3. A FRAMEWORK FOR PROTOTYPING

3.1 Interfacing of Digipot (AD5262)

The programmable gain amplifier section of the PCB is realised using two independent gain stages both of them are programmable. Each of the programmable gain stage is realised using digitally controlled potentiometers (digipot). The digital data interface for digipot is provided through the SPI interfacing. TheAD5262providesadual-channel,256-positiondigitally controlledvariableresistor(VR)deviceandoperateupto15Vmaximumvoltage[4]. Each VR has its own VR latch, which holds its programmed resistance value. These VR latches are updated from an internal serial-to-parallel shift register, which is loaded from a standard 3-wire serial-input digital interface. The AD5262 contains a 9-bit serial register. Each bit is clocked into the register on the positive edge of the CLK. The AD5262 address bit determines the corresponding VR latch to be loaded with the last 8 bits of the data word during the positive edging of CS strobe. Changing the programmed VR settings is accomplished by clocking a9-bit serial data word in to the SDI(SerialDataInput)p in.For this ADC, the format of this data word is one address bit.A represents the first bit of serial data B_8,then followed by eight data bits B_7-B_0 with MSB first. VR outputs can be changed one at a time in random sequence. TheAD5262 presets to amid-scale,simplifying fault condition recovery at power-up[3]. Mid-scale can also be achieved at anyt ime by asserting the PRp in.Both parts have an internal power ON preset[3] that places the wiper in mid-scale preset condition at power ON.Operation of the power ON preset function depends only on the state of the VLpin.TheAD5262 contains a power shutdown SHDNp in,which places the RDAC in an almost zero power consumption state,where terminals Axare opencircuited. And the wiper W is connected toB,result in g in only leakage currents being consume d in the VRstructure. In shutdown mode ,the V R l atch setting sare maintained so that,returning to operational mode from power shutdown. And the VR settings return to the irprevious resistance values. The functional block diagram of AD5262 is shown in the Fig.5.

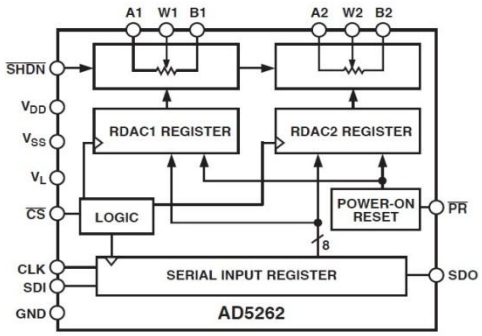


Fig.5. AD5262BlockDiagram

The AD5262 containsa4-wire SPI-compatible digital interface(SDI,SDO,CS,andCLK).For the AD5262,the 9-bit serial word loaded with address bit A0 first,then MSB of the data. The positive-edgesensitiveCLKinputrequirescleantransitionstoavoidclockingofincorrectdataintothe serial input register.When CS is low,the clock loads data in to the serial register one ach positive clock edge.For the AD5262 the last 9 bits of the serial data word entered into the serial data register are held when CS returns to high. Any extrabits are ignored. At the same time CS goe shigh,it gates the address decoder by enabling AD5262 one of two positive edge-triggered AD5262 RDA Clatches.The target RDA Clatch should be loaded with the last 8 bits of the serial data word completing one RDAC update.For theAD5262,two separate 9-bit serial data word smust be clocked into change both VR settings. During shutdown (SHDN)the SDO output pin is forced to the off to disable power dissipation in the pull-upresistor. Fig.6 shows the timing diagram of AD5262.

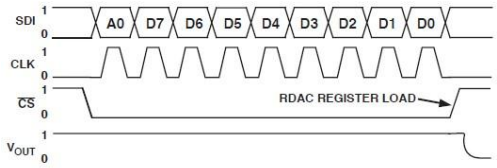


Fig.6. AD5262TimingDiagram

3.2 Interfacing of ADC (ADS1278)

This paper opted ADS1278for ADC interfacing.Based on the single-channel ADS1271,ADS1278(octal)is24-bit, delta-sigma analog-to-digital converter(ADCs)with data rates upto128ksamples per second(SPS),allowing simultaneous sampling of eight channels. The device is offered in identical packages,permitting drop-inexpandability. Industrial delta-sigma ADCs offers good drift performance use digital filters with large pass band droop. Hence they have limited bandwidth and are mostly suited for dc measurements.In audio applications high-resolution ADC sofferslargerus able band widths,but the offset and drift specifications are significantly weaker than respective industrial counter parts. This ADC combines these types of

converters, allowing high-precision industrial measurement with excellent dc and ac specifications.

Data ready for retrieval are indicated by the falling DRDY output and are shifted out on the falling edge of SCLK, MSB first. The device shifts in data on the falling edge and the user normally shifts this data in on the rising edge. On the application of the control signals, serial data bits will be available at the output pin of ADC along with DRDY signal. Timing diagram of ADS1278 is shown in Fig.7.

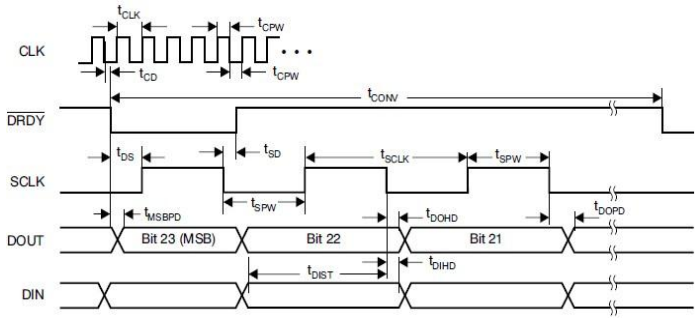


Fig.7. Timingdiagram of ADS1278

3.3 Design of Ethernet Interface

Ethernet interfacing is implemented VHDL making use of Virtex-5FPGA.Embedded Tri-Mode Ethernet MAC wrapper automates the generation of the HDL wrapper files for the Embedded Tri-ModeEthernet MAC(Ethernet MAC)inVirtex-5LXT,FXT,SXT and TXTFPGA susing the software Xilinx COREGenerator[5].Although, Verilog and VHDL instantiation templates are available in Libraries Guide for the FPGA Ethernet MAC primitive; how ever,due to the complexity and large number of ports CORE Generator software is preferred. It simplifies integration of the Ethernet MACby providing HDL examples based on user-selectable configurations.

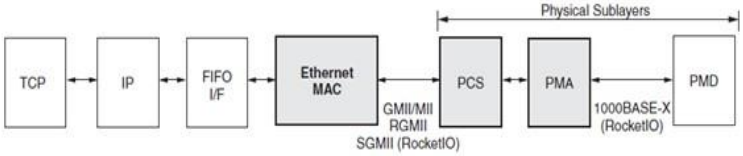


Fig.8. Typical Ethernet Architecture

Fig.8 displays Ethernet MAC architecture from the MAC to the right,as defined in the IEEE802.3specification, and it also illustrates where the supported physical interfaces fit into this architecture.TheSerial-GMII(SGMII) interface is an alternative to GMII/MII.It converts

the parallel interface of the GMII/MII in to a serial format using aRocket IOGTP or GTXtransceiver. It radically reduces the I/Ocount.For this reason,it is often the preferred interface of PCBdesigners. SGMIIcancarryEthernet traffic at10Mbps,100Mbps,and 1Gbps[6].The combination of the Physical Coding Sublayer (PCS),the Physical Medium Attachment (PMA),and the Physical Medium Dependent(PMD) sublayer comprise the physicallayersof the Ethernet protocol.

The Ethernet MAC wrapper file instantiates the full Ethernet MAC primitive. All unused input ports are tied to the appropriate logic level and all unused outputports are left unconnected. The Ethernet MAC attributes are set based on the options selected in the CORE Generator. Only usedports are connected to the ports of the wrapper file.This simplified wrapper is used as the instantiation template for the Ethernet MAC in the design.The designis downloaded on to the FPGA. All the clock management logic required to operate the Ethernet MAC and the designis provided. BUFGs, DCMs and so forth are instantiated as required.The design includes an Addresss wapping module and a FIFO[7]. Frames received by the Ethernet MAC are passed through the Receiveside of the FIFO. Data from the Receiveside of the FIFO is passed into the Address Swap Module and then onto the Transmitside of the FIFO using a Local Link interface.The Transmit FIFO queues frames for transmission and connects directly to the clientside Transmit interface of the Ethernet MAC.

The Address Swap Module switches the Destination Address and Source Address . The10Mbps,100Mbps,1Gbps Ethernet FIFO is a wrapper file around the Receive and Transmit FIFO components. The Receive(Rx) Client FIFO and theTransmit(Tx) Client FIFO,both4k-byte FIFO simplemented in block RAMS, is connected directlyt to the Rx Client Interface and the TxClient Interface of the Ethernet MAC respectively. The transmit and receive FIFO component simple menta Local Link user interface, through which the frame data is read and written. The FIFO is designed to work with the client clocks running at speeds in the range of 125MHz to 1.25MHz.

The data transferred from source to destination on the Local Link interface, with the flow governed by the four active low control signals of -n, src-rdy-n, e of-n, and dst-rdy-n. The flow of data is controlled by the src-rdy-nanddst-rdy-n signals. When these signals are asserted simultaneously,data is transferred from source to destination.The individual packet boundaries are marked by thes of-nande of-nsignals.

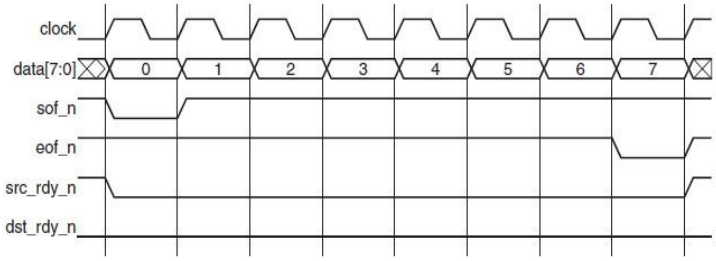


Fig.9 FrameTransferacrossLocalLinkInterface

The selected GMII/MII interface connects the physical interface of the Ethernet MAC block to the I/O of the FPGA. This component contains IOB flip-flops Input/Outputblock(IOB) buffers.

4. RESULT

Isim is a Xilinx simulation provides a complete, full featured hdl simulator which is integrated within the ISE tool. Xilinx Isim is a Hardware Description Language (HDL) simulator that helps to perform behavioral and timing simulations for VHDL. A test bench was created for the user logic module to initially simulate and verify its working.

The csbar signal is generated based on the clkout and timing at which control word is given to the digipot. The csbar signal is set low when the control word is available at the SDI pin of the digipot and made high when the control word is ready to be written on to the RDAC register. The sync pulse which is necessary to activate the ADC is set as an active low pulse of small duration. The simulation result for generating digipot and ADC control signals is shown in the figure fig10.

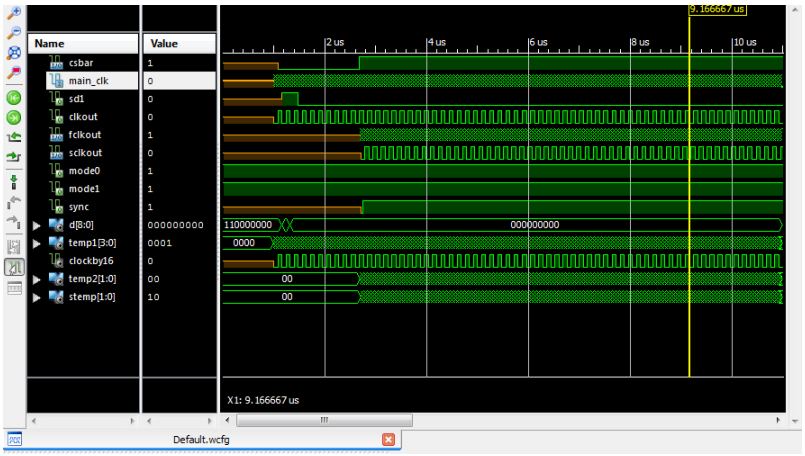


Fig 10: Generated Digipot and ADC control signals

On the application of the control signals, serial data bits were available at the output pin of ADC along with DRDY signal. The experimental results are shown in Fig11 and Fig12.

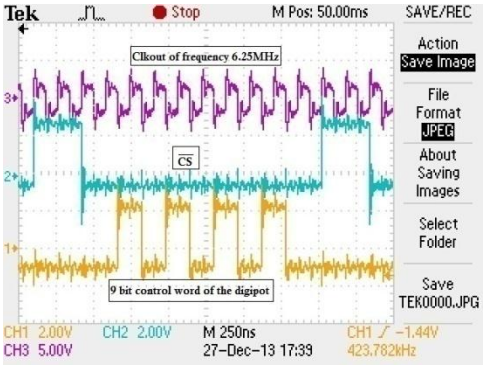


Fig.11: Generated control signals along with the required serial control word of AD5262

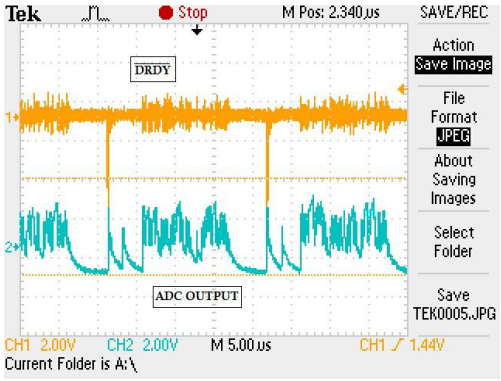


Fig.12: The output signals from ADS1278

It was observed that all the control signals namely sof_out_n, eof_out_n, src_rdy_out_n and dst_rdy_n were generated accurately, with initially the header information being transmitted through data[7:0] followed by the data field. Sof_out_n signals the beginning of frame transmission, eof_ot_n signals the end of a frame transmission, and src_rdy_out_n signals presence of valid data on data_out[7:0]. The simulation results are shown in Fig13.

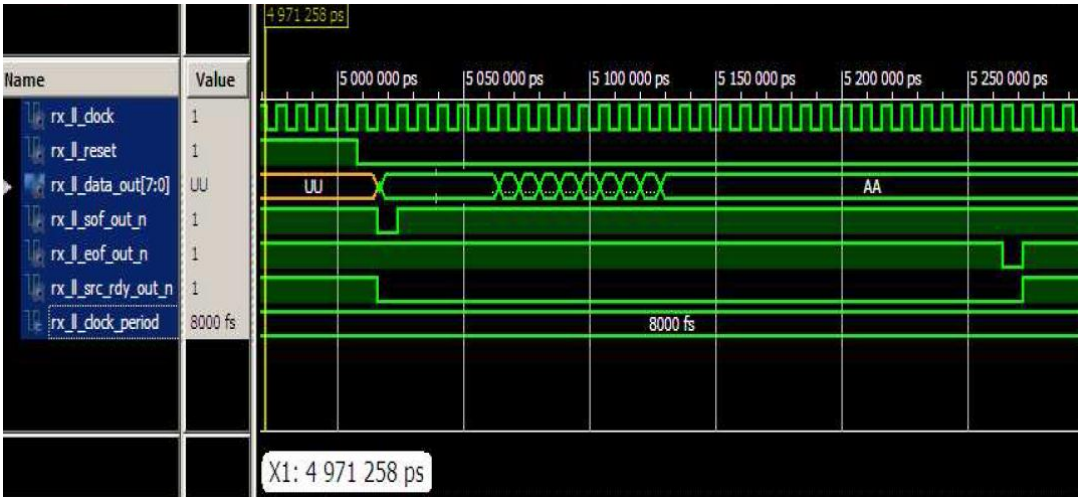


Fig13: User logic simulation results Expanded waveform

The serial data transmitted through Ethernet cable is captured using Wireshark software and the resultant data shown in Fig 14

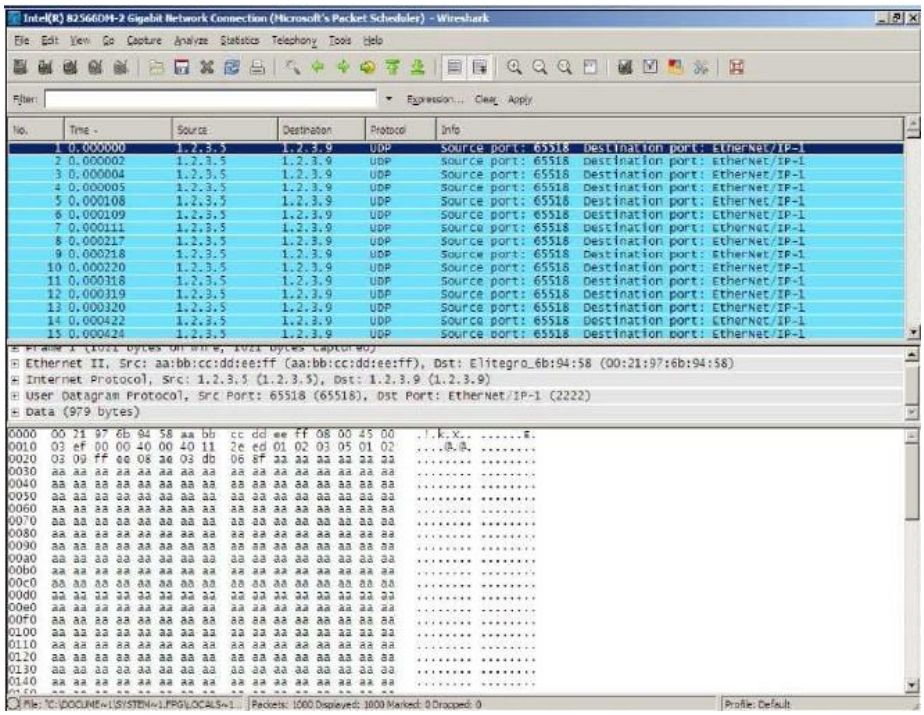


Fig14: Transmitted packets captured in Wireshark

5. CONCLUSION

The development of alow-costreal-time Data Acquisition System on FPGA and interfacing with Ethernet was studied. This paper presents an attractive combination of low cost and high performance,along with an apparent flexibility. A framework has been developed suitable for the development of an FPGA based DAQ using SPI and Ethernet protocols suitable for Sonar.This framework deceives many of the difficult issues associated with the existing DAQs, making it usable even by less-experienced developers. The potential interms off lexibility is of particular interest, asitis the key to developing large,scalable,independent data acquisition systems.

ACKNOWLEDGEMENT

The author would like to thank RejaniKrishna(ScientistD) from Naval Physical and Oceanographic Laboratory, NPOL Thrikakkara and Dr.Sindhu.R, Head of the department of electronics and communication of NSS college of engineering fortheassistanceandsupport.

REFERENCES

[1] HosseinPeyvandi, "SONAR SystemsandUnderwaterSignalProcessing: ClassicandModernApproaches", Scientific AppliedCollegeofTelecom- munication,Tehran,2010.
[2] P.Branchini "AnFGPABasedGeneralPurposeDAQModuleforthe KLOE-2Experiment", Rome,Italy,IEEETRANSACTIONSON NU- CLEARSCIENCE, VOL.58,NO.4,AUGUST2011.
[3] Swamy TN, Rashmi KM Data Acquisition system based on FPGA ,IJERA, April 2013.
[4] "http://www.alldatasheets.com" - AD52624] Datasheet.
[5] ML505 Evaluation Platform User Guide UG347.
[6] Virtex 5 FPGA Embedded Tri Mode Ethernet MAC User Guide UG194.
[7] Ethernet MAC Core Getting Started Guide. Pdf.
[8] Isim User Guide UG660, v14.1..

AUTHORS

|**Anagha A V**: Received BTech degree from university of Calicut in electronics and Communication. Nowpursuing MTech degree from A P J Abdul Kalam Technological University in VLSI and embedded systems

Mary Joseph: received M.Tech Degree in Microwave and Radar from Cochin University of Science and Technology (CUSAT), Kochi, India, in 1997. Currently she is working as Associate Professor in M. A. College of Engineering, Kothamangalam. She has joined in M. A. College of Engineering in 1991 as Assistant Professor. In between she worked at Birla Institute of Science & Technology-Pilani's (BITS-PILANI) Dubai Campus for 9 years as Assistant Professor during 2000-2008. Her Research interests include Microstrip Antennas and Uniplanar Antennas.

INTENTIONAL BLANK

A SURVEY ON WIND DATA PRE-PROCESSING IN ELECTRICITY GENERATION

Mahima Susan Abraham[1] and Jiby J Puthiyidam[2]

[1] Department of Computer Science and Engineering, College of Engineering, Poonjar
[2] Department of Computer Science and Engineering, College of Engineering, Poonjar

ABSTRACT

Wind energy integration research generally relies on complex sensors located at remote sites. The procedure for generating high-level synthetic information from databases containing large amounts of low-level data must therefore account for possible sensor failures and imperfect input data. The data input is highly sensitive to data quality. To address this problem, this paper presents an empirical methodology that can efficiently preprocess and filter the raw wind data using only aggregated active power output and the corresponding wind speed value sat the wind farm. First, raw wind data properties are analyzed, and all the data are divided into six categories according to their attribute magnitudes from a statistical perspective. Next, the weighted distance, a novel concept of the degree of similarity between the individual objects in the wind database and the local outlier factor (LOF) algorithm is incorporated to compute the outlier factor of every individual object, and this outlier factor is then used to assess which category an object belongs to.

KEYWORDS

Data mining, data preprocessing, local outlier factor (LOF), unsupervised learning

1. INTRODUCTION

All The objective of data mining is to identify valid novel, potentially useful, and understandable correlations and patterns in existing data .Finding useful patterns in data is known by different names (including data mining) in different communities (e.g., knowledge extraction, information discovery, information harvesting, data archaeology, and data pattern processing).The term "data mining" is primarily used by statisticians, database researchers, and the MIS and business communities. The term Knowledge Discovery in Databases (KDD)[7] is generally used to refer to the overall process of discovering useful knowledge from data, where data mining is a particular step in this process. The additional steps in the KDD process, such as data preparation, data selection, data cleaning, and proper interpretation of the results of the data mining process, ensure that useful knowledge is derived from the data.

Data mining, almost by definition, is primarily concerned with the operational. The second type of data mining approach, pattern detection, seeks to identify small (but nonetheless possibly important) departures from the norm, to detect unusual patterns of behaviour. Examples include unusual spending patterns in credit card usage (for fraud detection), sporadic waveforms in EEG traces, and objects with patterns of characteristics unlike others. It is this class of strategies that led to the notion of data mining as seeking "nuggets" of information among the mass of data. In

DOI: 10.5121/ijci.2016.5243

general, business databases pose a unique problem for pattern extraction because of their complexity. Complexity arises from anomalies such as discontinuity, noise, ambiguity, and incompleteness. And while most data mining algorithms are able to separate the effects of such irrelevant attributes in determining the actual pattern, the predictive power of the mining algorithms may decrease as the number of these anomalies increase.

Data pre-processing [8] is an often neglected but important step in the data mining process. The phrase "Garbage In, Garbage Out" is particularly applicable to data mining and machine learning. Data gathering methods are often loosely controlled, resulting in out-of-range values (e.g., Income: -100), impossible data combinations (e.g., Gender: Male, Pregnant: Yes), missing values, etc. Analyzing data that has not been carefully screened for such problems can produce misleading results. Thus, the representation and quality of data is first and foremost before running an analysis. If there is much irrelevant and redundant information present or noisy and unreliable data, then knowledge discovery during the training phase is more difficult. Data preparation and filtering steps can take considerable amount of processing time. Data pre-processing includes cleaning, normalization, transformation, feature extraction and selection, etc. The product of data pre-processing is the final training set.

Raw data is highly susceptible to noise, missing values, and inconsistency. The quality of data affects the data mining results. In order to help improve the quality of the data and, consequently, of the mining results raw data is pre-processed so as to improve the efficiency and ease of the mining process. Data pre-processing is one of the most critical steps in a data mining process which deals with the preparation and transformation of the initial dataset. Data pre-processing methods are divided into following categories:

- Data Cleaning
- Data Integration
- Data Transformation
- Data Reduction

Nowadays, more and more attention is paid on wind energy—a kind of clean and renewable energy. While developing the wind power, we should also keep a watchful eye on the ability of real-time data processing, in order to make the wind power develop healthily and rapidly, taking the road of sustainable development is the ultimate goal. Now wind power data is usually applied in wind power prediction, which is benefit for reducing the shock of wind power on the grid, and improving the economy of the grid operations. As we know, effective wind data pre-processing is the key to wind power forecasts. Provide clean, accurate data for, data mining, thus reduce the amount of data processing, and then deduce the valuable information. Although there are a lot of methods of data pre-processing, few of which apply in the wind power data. Wind data reprocessing existed mainly study on attribute reduction, missing values, isolated points, but offer few reference value for prediction.

Wind energy integration research generally relies on complex sensors located at remote sites. The procedure for generating high-level synthetic information from databases containing large amounts of low-level data must therefore account for possible sensor failures and imperfect input data. The data input is highly sensitive to data quality. To address this problem, this paper presents an empirical methodology that can efficiently pre-process and filter the raw wind data

using only aggregated active power output and the corresponding wind speed values at the wind farm.

RELATED WORKS

There are many methods in analysing wind data. Some of those methods are discussed. Most of these methods don't consider irregular data's or values. One of these methods identifies the irregular data's or values and removes them. Some methods improve accuracy and some improves performance speed. . It's discussed below in details

1. RAW WIND DATA PREPROCESSING

[1] presented a wind data pre-processing method including four steps: 1) validity check; 2) data scaling; 3) missing data processing; and 4) lag removal. The validity check involves a data range check that detects data values exceeding the physical limits. Data scaling normalizes data with the ratings. Missing data processing involves either neglecting or approximating the missing values. Lag removal uses the cross correlation function to identify the lag between input and output, which is useful when dealing with time-series analysis. In real-world applications, artificial judgment is limited and inconvenient when the size of the database is large, and the wind farm operation state records are often unavailable. Thus, these data classification procedures are infeasible or unreliable, which causes difficulties in applying supervised learning algorithms. Therefore, the alternative solution is to use unsupervised algorithms. To use unsupervised algorithms, we adopted an unsupervised learning approach based on the local outlier factor (LOF)-identifying algorithm. The LOF of every data point is computed using a novel concept of the degree of similarity among the individual data points, and hence invalid data are detected as abnormal outlier factors.

The contribution of this paper is to develop an empirical methodology for raw wind data pre-processing. The only information required for this methodology is the aggregated wind power output of the wind farm collected from the Supervisory Control And Data Acquisition (SCADA) system, which is available at the dispatch center, and the wind speed magnitude data at the corresponding wind farm site. The availability of wind farm operation state records or wind turbine fault logs (which are not recorded or stored by most wind farm operators) will help improve the accuracy of the methodology. If these data are unavailable, this is often the case, the methodology proposed in this paper has nonetheless been proved to be adequate for the situation.

Advantage: One of the greatest advantages of the proposed methodology is that it is a type of unsupervised learning algorithm. Therefore, it can detect and classify the raw data using solely the attributes of the data themselves. It is easier and more convenient to perform in practice, especially when the operation records are not available.

Disadvantage:. First, the total number of the data points should not be too small. An empirical minimum value is approximately 1000. Second, if most of the data are invalid, the accuracy cannot be guaranteed. This situation indicates that either the data acquisition and transmission system is broken down or manual actions are frequent.

In short, the wind farm is faulty, and the data acquired from it should not be used for research. The data pre-processing method proposed in this paper can be used for many purposes, not only wind-related applications. The idea of weighted distance can also be used in other outlier or cluster-detection algorithms to develop individual detection algorithms dedicated to specific applications.

2. WIND SPEED FORECASTING STRATEGY

[2] presented a new wind speed forecasting approach based on based on the chaotic time series modelling technique and the Apriori algorithm has been developed. The new approach consists of four procedures: Clustering by using the k-means clustering approach; Employing the Apriori algorithm to discover the association rules; Forecasting the wind speed according to the chaotic time series forecasting model; and Correcting the forecasted wind speed data using the associated rules discovered previously. This procedure has been verified by 31-day-ahead daily average wind speed forecasting case studies, which employed the wind speed and other meteorological data collected from four meteorological stations located in the Hexi Corridor area of China. The results of these case studies reveal that the chaotic forecasting model can efficiently improve the accuracy of the wind speed forecasting, and the Apriori algorithm can effectively discover the association rules between the wind speed and other meteorological factors. In addition, the correction results demonstrate that the association rules discovered by the Apriori algorithm have powerful capacities in handling the forecasted wind speed values correction when the forecasted values do not match the classification discovered by the association rules.

This paper firstly analyses the historical wind speed data for a given wind farm by applying the nonlinear time series modelling techniques. The numerical simulations results indicate that chaotic characteristics obviously exist in the wind speed time series. This finding inspires us to model the wind speed as a complex non-linear dynamic system that often exhibits chaotic behaviour. If an irregular movement characterized by the time series can be regarded as a type of chaos phenomenon, a prediction with higher precision is available with the chaos theory, which is used to address the inner uncertainties of the system. As an important method for studying the characteristics of complex systems, the chaotic time series predictions have attracted significant research interests over the past few years. Some chaotic prediction methods have been developed such as the local-region method, Lyapunov Exponents method, and artificial neural network method. Among these methods, the local-region method seems more promising for wind speed forecasting. This paper employs a weighted local-region method to forecast the wind speed series.

Advantage: This can be very useful in two occasions: one is to check the predicted wind speed values for abnormal cases based on the association rules, and the other is to estimate the value ranges of the other meteorological factors, including the air pressure, air temperature and humidity.

3. WIND SPEED AND POWER FORECASTING

[3] presents an overview of existing research on wind speed and power forecasting. It first discusses state-or-the-art wind speed and power forecasting approaches. Then, forecasting accuracy is presented based on variable factors. Finally, potential techniques to improve the accuracy of forecasting models are reviewed. A full survey on all existing models is not

presented, but attempts to highlight the most promising body of knowledge concerning wind speed and power forecasting. Wind power is one of the most rapidly growing renewable energy sources, and is regarded as an appealing alternative to conventional power generated from fossil fuel. This led to a collaborative effort to achieve 20% of U.S. electricity supplied from wind power by 2030 . Although the integration of wind power brings many advantages, high penetration of wind power provides a number of challenges in power system operations and planning, mainly due to its uncertain and intermittent nature.

In the electricity system the power supply must be equal to the power demand at all times. However, the variation of wind power output makes it difficult to maintain this balance. One of the possible solutions to the balance challenge is to improve the wind speed and power forecasting. Research in the area of forecasting wind speed or the power produced by wind farms has been devoted to the development of effective and reliable tools and many different approaches have been proposed and reviewed in . Accurate forecasting tools reduce operating costs and improve reliability associated with the integration of wind power into the existing electricity supply system. There are different users of wind speed and power forecasts. These users not only need point forecasts but also the uncertainty of the forecast is essential for determining the size of the operating reserves necessary to balance the generation with load.

The main objectives of wind speed and power forecasting is to estimate the wind speed and power as quickly and accurately as possible. Accurate forecasting tools reduce the financial risk and lead to improved scheduling and unit commitment plans. The statistical approaches provide good results in the majority of cases, including short-term, medium-term, and long-term forecasting. However, in the very short-term and short-term horizon, the influence of atmospheric dynamics becomes more important, so that the use of the physical approaches becomes necessary.

Advantage: This approach is able to not only improve the forecast accuracy, but also reduces the risk from extreme events. Ensemble forecasting models for probabilistic forecasting are used in order to obtain the expected spread of weather conditions and assess the probability of particular weather events.

4.PROBABILISTIC WIND SPEED FORECASTING

[4] presents a probabilistic forecasts of wind speed are becoming critical as interest grows in wind as a clean and renewable source of energy, in addition to a wide range of other uses, from aviation to recreational boating. Statistical approaches to wind forecasting offer two particular challenges: the distribution of wind speeds is highly skewed, and wind observations are reported to the nearest whole knot, a much coarser discretization than is seen in other weather quantities. The prevailing paradigm in weather forecasting is to issue deterministic forecasts based on numerical weather prediction models.

Uncertainty can then be accessed through ensemble forecasts, where multiple estimates of the current state of the atmosphere are used to generate a collection of deterministic predictions. Ensemble forecasts are often uncalibrated, however, and Bayesian model averaging (BMA) is a statistical way of post processing these forecast ensembles to create calibrated predictive probability density functions (PDFs).It represents the predictive PDF as a weighted average of PDFs centered on the individual bias-corrected forecasts, where the weights reflect the forecasts' relative contributions to predictive skill over a training period. In this paper we extend BMA to

provide probabilistic forecasts of wind speed, taking account of the skewness of the predictive distributions and the discreteness of the observations.

Advantage: better calibration

5. MARKOV-SWITCHING AUTOREGRESSIVE MODELS

[5] presentsWind power production data at temporal resolutions of a few minutes exhibits successive periods with fluctuations of various dynamic nature and magnitude, which cannot be explained (so far) by the evolution of some explanatory variable. Our proposal is to capture this regime-switching behaviour with an approach relying on Markov-Switching Autoregressive (MSAR) models. An appropriate parameterization of the model coefficients is introduced, along with an adaptive estimation method allowing to accommodate long-term variations in the process characteristics. The objective criterion to be recursively optimized is based on penalized maximum-likelihood, with exponential forgetting of past observations. MSAR models are then employed for 1-step-ahead point forecasting of 10-minute resolution time-series of wind power at two large offshore wind farms. They are favourably compared against persistence and Autoregressive (AR) models.

The main objective of the present paper is to introduce a MSAR model whose coefficients are adaptively and recursively estimated, with application to the modelling and forecasting of offshore wind power fluctuations. The parameterization of the model coefficients employed here is inspired by those initially proposed .Adaptively in time is achieved with exponential forgetting of past observations. In addition, the formulation of the objective function to be minimized at each time-step includes a regularization term that permits to increase the generalization ability of estimated models, in addition to improving numerical stability of the recursive estimation procedure.

Advantage: Characterizing and modelling the power fluctuations for the specific case of offshore wind farms is a current challenge

6. WIND POWER FORECASTING

[6]presents ARMA (q, p) model of time series to forecast wind speed and atmospheric pressure, and using the RBF neural network based on this to forecast wind power. Taking the data of measured wind speed and atmospheric pressure from a wind farm as example, to validate the method described above, and the result show that the method has a certain practicality. In this paper, to forecast wind speed by using the method of time series, which the data requirement is low and the cost used to forecast is also low, so it is suitable for the actual operation of businesses. Considering the high degree non-linear relationship showed between the wind speed data and the corresponded generation power, RBF neural network to be used to forecast generation power. And to verify the feasibility and effectiveness of the method in this paper through the experimental data from a wind farm.

The technical requirements (try out) of national grid wind farms accessing grid clearly pointed out the need for forecasting the wind farms power. The analysis of time series and the RBF neural network will be introduced into the wind power forecasting in this article. The forecasting of wind power has a great significance to the construction and operation of wind farm. The study

Conclusions of the above mentioned wind power forecasting system have: Time series model is a dynamic model, which has a very good extension to dynamic data, thereby it could avoid the impact of the directly adding "Window" when we strike the statistical properties of dynamic data. For the random and dynamic of wind speed, the method of time series ARMA reflects a larger advantage.

Advantage: Very good non-linear learning ability; Advantage in resolving wind power forecasting

METHOD	MODEL	ALGORITHM	ADVANTAGE/DISADVANTAGE
Raw wind data reprocessing	Similarity Measurement	LOF Algorithm	Unsupervised learning used; Easier and more convenient to perform in practice.
Wind speed forecasting strategy	chaotic time series	Apriori Algorithm	To check the predicted wind speed values for abnormal cases; To estimate the value ranges of meteorological factors.
Wind Speed and Power Forecasting	ARMA	Artificial Neural Network Approach	Reduce the financial risk; Improve the forecast accuracy.
Probabilistic Wind Speed Forecasting	Bayesian	Fully discretized method	maximum wind speed over a particular time interval.
Adaptive modelling and forecasting	Markov-switching autoregressive	Adaptive estimation Method	Characterizing and modeling the power fluctuations for the specific case of offshore wind farms is a current challenge.
Wind Power Forecasting	RBF neural network	Data Flow Algorithm	Very good non-linear learning ability ; Advantage in resolving wind power forecasting.

RESULT ANALYSIS

Raw wind data pre-processing, it is a type of unsupervised learning algorithm. Therefore, it can detect and classify the raw data using solely the attributes of the data themselves. It is easier and more convenient to perform in practice. Wind speed forecasting strategy is to check the predicted wind speed values for abnormal cases and also to estimate the value ranges of meteorological factors. Wind Speed and Power Forecasting reduce the financial risk and improve the forecast

accuracy. Probabilistic Wind Speed Forecasting gives maximum wind speed over a particular time interval. Adaptive modelling and forecasting is used for characterizing and modelling the power fluctuations for the specific case of offshore wind farms is a current challenge. Wind Power Forecasting has very good non-linear learning ability. Here it is understood that raw wind data pre-processing is more better than other papers as it is unsupervised learning .And also it identifies and elimates the outliers.

CONCLUSION

In this paper, raw wind data properties were analyzed. Invalid data can be categorized into five types. A wind data pre-processing methodology has been proposed. Because identifying the unnatural and the irrational data is challenging, this paper treats them as outliers and uses the LOF algorithm to detect and remove these outliers. To incorporate prior knowledge regarding the wind data, a new type of similarity measurement is designed and applied in the algorithm. Numerical experiments have verified the effectiveness of the algorithm and the similarity measurement. The performance evaluation of the algorithm has also been discussed. One of the greatest advantages of the proposed methodology is that it is a type of unsupervised learning algorithm. Therefore, it can detect and classify the raw data using solely the attributes of the data themselves. It is easier and more convenient to perform in practice, especially when the operation records are not available. However, as there is no universal data-mining algorithm that can handle all problems, this methodology has its limitations. First, the total number of the data points should not be too small. An empirical minimum value is approximately 1000. Second, if most of the data are invalid, the accuracy cannot be guaranteed.

REFERENCES

1. Raw Wind Data Preprocessing: A Data-Mining Approach -Le Zheng, Wei Hu, and Yong Minieee TRANSACTIONS ON SUSTAINABLE ENERGY, VOL. 6, NO. 1, JANUARY 2015
2. A new wind speed forecasting strategy based on the chaotic time series Modelling technique and the Apriori algorithm-Zhenhai Guo a,⇑, Dezhong Chi , Jie Wu, Wenyu Zhang cEnergy Conversion and Management 84 (2014) 140–151
3. Current Status and Future Advances for Wind Speed and Power Forecasting -Jaesung Junga*, Robert P. Broadwatera
4. Probabilistic Wind Speed Forecasting using Ensembles and Bayesian Model Averaging-J. McLean Sloughter, Tilmann Gneiting, and Adrian E. Raftery
5. Adaptive modelling and forecasting of offshore wind power fluctuations with Markov-switching autoregressive models -Pierre Pinson, Henrik Madsen
6. Wind Power Forecasting Based on Time Series and Neural Network -Lingling Li1,2 , Minghui Wang2 , Fenfen Zhu2, and Chengshan Wang*
7. From Data Mining to Knowledge Discovery in Databases -Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth
8. Using Neural Networks to Estimate Wind Turbine Power Generation -Shuhui Li, Member, IEEE, Donald C. Wunsch, Senior Member, IEEE, Edgar A. O'Hair, and Michael G. Giesselmann, Senior Member, IEEE
9. Z. Q. Liu,W. Z. Gao, Y. H.Wan, and E. Muljadi, "Wind power plant prediction by using neural networks," in Proc. IEEE Energy Convers. Congr.Expo., 2012, pp. 3154–3160.
10. M. Ali, I. Ilie, J. V. Milanovic, and G. Chicco, "Wind farm model aggregation using probabilistic clustering," IEEE Trans. Power Syst., vol. 28,no. 1, pp. 309–316, Feb. 2013.

11. M. Schlechtingen, I. F. Santos, and S. Achiche, "Using data-mining approaches for wind turbine power curve monitoring: A comparative study," IEEE Trans. Sustain. Energy, vol. 4, no. 3, pp. 671–679, Jul.2013
12. A. Kusiak, H. Y. Zheng, and Z. Song, "Short-term prediction of wind farm power: A data mining approach," IEEE Trans. Energy Convers., vol. 24,no. 1, pp. 125–136, Mar. 2009.

AUTHORS

Mahima susan abraham received her Bachelor of Engineering in Computer Science and Engineering from Anna University, Chennai in 2012 .She worked as an Android Developer for 1 year. She is currently doing her Master of Technology in Computer and information Science at Cochin University of Science and Technology. Her area of interest includes Data Mining. E-Mail: mahimasusan1990@gmail.com

Jiby j.puthiyidam received his Bachelor of Engineering in Computer Science and Engineering from Madras University in 1998 and Master of Technology in Computer and information Science from Cochin University of Science and Technology in 2008. He is currently working as Assistant Professor, Department of Computer Science and Engineering, College of Engineering Poonjar, Kerala. He is a life member of Indian Society of Technical Education (ISTE). He has presented many papers in National and International conferences. His area of interest includes Wireless Sensor Networks and Data Mining. E- mail id: jibyjp@gmail.com

INTENTIONAL BLANK

HIERARCHICAL PARTITION-BASED ANONYMOUS ROUTING PROTOCOL (HPAR) IN MANET FOR EFFICIENT AND SECURE TRANSMISSION

Fahmida Aseez[1] and Dr.Sheena mathew[2]

[1]Mtech Student Division of Computer Engineering, SOE, CUSAT, Cochin, India
[2]Dr.Sheena Mathew, Professor, Division of Computer Engineering, SOE, CUSAT, Cochin, India

ABSTRACT

Anonymous routing protocols are used in MANET's to hide the nodes from outsiders in order to protect from various attacks. HPAR partitions the network area dynamically into zones and chooses nodes in zones randomly as intermediate relay nodes .This relay nodes help in secure routing. In HPAR anonymity protection is given to source, destination and route. HPAR have low cost and provide high level of protection. It has techniques to counter various attacks.

KEYWORDS

Anonymous routing, Mobile ad hoc networks, Anonymity

1. INTRODUCTION

MANET Comprises of wireless mobile nodes that are freely and self-organize into a temporary network topology with-out any infrastructural support. Open nature, dynamic changing topology, no central management are the key features of MANET's. They have a variety of applications in military, banking, education, commerce etc. Security in MANET is a major issue. Data get lost or stolen by tampering and analyzing data and traffic analysis by eavesdropping method or attacking routing protocol. Solution to this is anonymous routing.

In MANET, the term Anonymity means hiding identity of the source node, receiver node and the chosen path. Anonymous routing protocols provide secure communications by hiding node identities and preventing traffic analysis attacks from outside observers. They are used in Military, Banking like application, where security of communication is a major concern. Anonymity is critical in military applications for example soldier communication. MANET deployed in a battlefield can be vulnerable to traffic analysis; enemies may intercept transmitted packets, track our soldiers (i.e., nodes), attack the commander nodes, and block the data transmission etc.

Limited resource is an inherent problem in MANETs. MANETs' complex routing and strict channel resource constraints impose strict limits on the system capacity. Nowadays multimedia applications (e.g., video transmission) require high routing efficiency. Our existing anonymous

DOI: 10.5121/ijci.2016.5244

routing protocols [1] generate a significantly high cost, which badly affect the resource constraint problem in MANETs. A MANET employed in a battlefield, with a high-cost anonymous routing and low quality of service in voice and video data transmission due to depleted resources may lead to disastrous delay in military operations.

HPAR provide high anonymity protection (for sources, destination, and route) with low cost. HPAR dynamically partitions a network field into zones and randomly chooses nodes in zones as intermediate relay nodes, which form a non-traceable anonymous route. In each routing step, a data sender or forwarder partitions the network field in order to separate itself and the destination into two zones. It then randomly chooses a node in the other zone as the next relay node and use GPSR algorithm to send the data to the relay node. At last, the data is broadcasted to k nodes in the destination zone, providing k-anonymity to the destination. Also it has strategy to hide the data initiator among a number of initiators to strengthen the anonymity protection of the source. HPAR is also resilient to intersection attacks and timing attacks.

2. LITERATURE SURVEY

Anonymity of a subject [2] means that the subject is not identifiable within a set of subjects, the anonymity set. Anonymity of a subject from an attacker's perspective means that the attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set. In simple terms Anonymity provides the privacy protection in the communication.

An existing protocol ALARM [3] is a table driven protocol, with location based routing. ALARM provides Security against active and passive attacks by advanced cryptographic techniques such as group signature. Group signature ensures that only valid members who have registered with the group manager can decrypt and read the packets. This protocol sends out Location Announcement Messages (LAM) to inform all the nodes of the network topology from time to time. Problem with ALARM is cannot protect location anonymity of source and destination node.

To preserve privacy PRISM [4] protocol suggested the use of Location bases routing along with Group signatures. It is an on-demand routing scheme. A source node will initiate a route discovery phase when it has data to transmit. PRISM is Based on the concept of location aided routing it located the destination, encrypts he packet, insert the source group signature and send the packet. Receiving packets can verify the group signature and destination is identified with the coordinates. The Route reply consists of a session key which will be used for further communication for that particular session. These Routes are discarded after communication. This protocol achieves privacy and security against active as well as passive attacks. As the nodes identity is not revealed and the destination node location is encrypted by key known only to valid group members.ALARM is a link-state protocol and exposes the entire topology to all insiders While PRISM prevents inside attacks.

Many anonymity routing algorithms are based on the geographic routing protocol for e.g., Greedy Perimeter Stateless Routing (GPSR) [5]. GPSR, packets are routed geographically. GPSR can route a packet to any connected destination. There are two distinct algorithms GPSR uses for routing first a greedy forwarding algorithm that moves packets progressively closer to the destination at each hop, and a perimeter forwarding algorithm that forwards packets where greedy forwarding is impossible. The greedy forwarding rule is simple: a node x forwards a packet to its neighbor y that is closest to the destination D as shown in Figure 1.Greedy forwarding fails when

no neighbor is closer than *x* to the destination. GPSR recovers from greedy forwarding failure using perimeter mode, which amounts to forwarding packets using the right-hand rule shown in Figure 2.

AO2P*(*ad hoc on-demand position-based private routing Algorithm) Protocol [6] is mainly proposed for communication anonymity. Route discovery is done by using only the position of the destination. Other information such as forwarding nodes positions are hiding from the network. [7] Provides an insight about Traffic Analysis. If the different routes that can be taken require different amounts of time, the system could be vulnerable to timing attacks. Intersection attacks mainly occurs by An attacker having information about what users are active at any given time can, through repeated observations, determine what users communicate with each other.

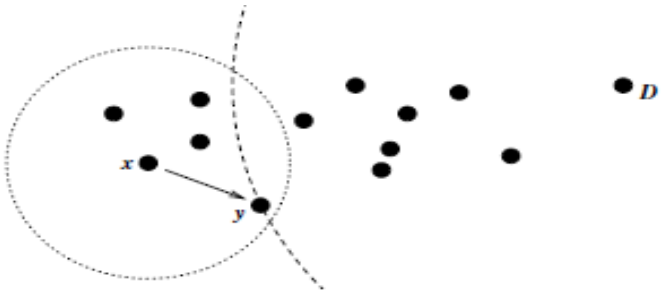


Figure 1. Greedy forwarding example x forwards to y which is closest to D.

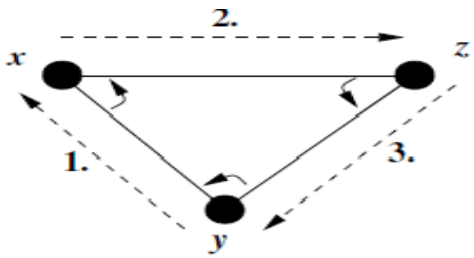


Figure 2. Right hand rule example packets travel along clockwise around the enclosed region.

Various security attacks include passive and active attacks [8].A passive attack does not alter the data transmitted within the network. Active attacks are very severe attacks on the network that prevent message flow between the nodes. Active attacks are classified into three groups: 1)Dropping Attacks Compromised nodes or selfish nodes can drop all packets that are not destined for them. Dropping attacks can prevent end-to-end communications between nodes.

2) Modification Attacks modify packets and disrupt the overall communication between network nodes. Sinkhole attacks are the example of modification attacks. 3) Fabrication Attacks the attacker send fake message to the neighboring nodes without receiving any related message.

3. PROPOSED SYSTEM

HPAR partitions given network area into two zones as horizontally (or vertically). Then again split every partition into two zones as vertically (or horizontally). This process called as hierarchical zone partition Figure 3 [1]**.** After partitioning HPAR randomly select a node in each zone at each step as an intermediate relay node. While this partitioning each data source of forwarder node checks whether itself and destination nodes are not in same zone. If it is not then partitioning continues. While in routing first source node randomly chooses a node in other zone known as temporary destination (TD) .Then uses GPSR routing algorithm to send the data to node close to TD. A node closer to TD known as Random Forwarder (RF). This repeats until destination zone is reached. But in destination zone data is broadcasted in ZD to k nodes which makes attacker or observer does not know the destination node. For successful completion of data transmission destination node send a confirmation to source node. If source node not receives to confirm during predefined time period, it will resend packets.

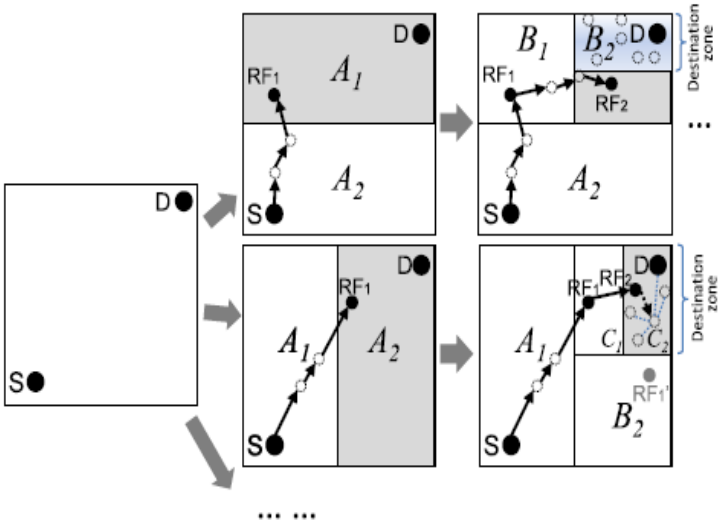


Figure 3. Zone partitioning

Different modules of the system includes Node construction, Zone partition, Source anonymity, Routing protocol , Destination anonymity Figure 4 shows the proposed system architecture.

Routing steps:

Step1: Assume rectangle network area, nodes are disseminated.
Step2: Each data source or forwarder executes the hierarchical zone partition
Step3: First check whether itself and D are in same zone.
Step4: If so, then divides the zone partition as Hierarchical zone partition.
Step5: Repeat step 4 process until itself and ZD are not in zone.
Step6: If source and ZD are not in the same zone then it randomly chooses a position in the other zone is called TD (Temporary Destination).

Step7: Using GPSR to send the data to the node closest to TD. This node is defined as a RF (Random Forwarder).
Step8: Repeat step 6 and step 7 until a data receiver finds itself residing in ZD having k node
Step9: In the last step, the data is broadcasted to k nodes in the destination zone, providing k-anonymity to the D.

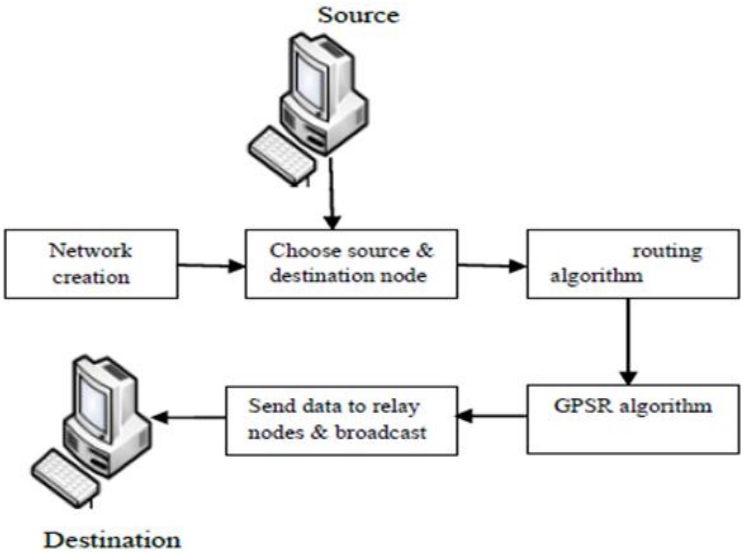


Figure 4. Proposed system architecture

A source node S sends a request to a destination node D and the destination responds with data. Each node uses a dynamic pseudonym as its node identifier rather than using its real MAC address, which can be used to trace nodes existence in the network. To avoid pseudonym collision, we use a collision Resistant hash function, such as SHA-1, to hash a node's MAC address and current time stamp. Each node periodically piggybacks its updated position and pseudonym to "hello" messages, and sends the messages to its neighbors. Also, every node maintains a routing table that keeps its neighbors pseudonyms associated with their locations.

Destination zone position is calculated by using certain equations. Zone position refers to the upper left and bottom-right coordinates of a zone.

$$H = log_2\left(\frac{\rho \cdot G}{k}\right),$$

- Let H denote the total number of partitions in order to produce ZD. Using the number of nodes in ZD (i.e., k), and node density.
- k = number of nodes in ZD
- P = node density
- G = the size of the entire network area.
- Using the calculated H, the size G, the positions (0,0) and (Xg , Yg) of the entire network area, and the position of D, the source S can calculate the zone position of ZD

Therefore, the size of the destination zone is given as:

$$\frac{G}{2^H}.$$

3.1 Anonymity Protection

HPAR makes the route between a S-D pair difficult to discover by randomly and dynamically selecting the relay nodes. The resultant different routes for transmissions between a given S-D pair make it difficult for an intruder to observe a statistical pattern of transmission. HPAR incorporates the "notify and go" mechanism to prevent an intruder from identifying which node within the source neighbourhood has initiated packets. HPAR also provides k-anonymity to destinations by hiding D among k receivers in Z d. Thus, an eavesdropper can only obtain information on Z d, rather than the destination position, from the packets and nodes en route.
In HPAR the nodes entire network is grouped to form clusters as in Figure 5. The clustering is based on the position or the coordinates of the nodes. Distance between the nodes and the source is calculated. Based on the distance the nodes are grouped. The nodes that are in the specified distance are forming a cluster. Then the communication is in the name of these clusters or groups. The packet transmission is by the communication between the groups. Inter and intra group communication is by random forwarders and relay nodes. So the communication is multi hop clustering. Each cluster has a specified range. The nodes belonging to that range are determined by the basics of their distance or coordinates. Each cluster maintains a cluster identifier or group identifier. The communication is carried out in this group id. The nodes form a cluster if they belong to particular range or distance. The nodes within the group can communicate with each other. This is known as intra group routing .They are mostly neighbors or one hop nodes. The nodes outside the group are communicated as multi hop fashion. The routing between the groups are known as inter group routing. This is by the means of relay nodes and random forwarders.

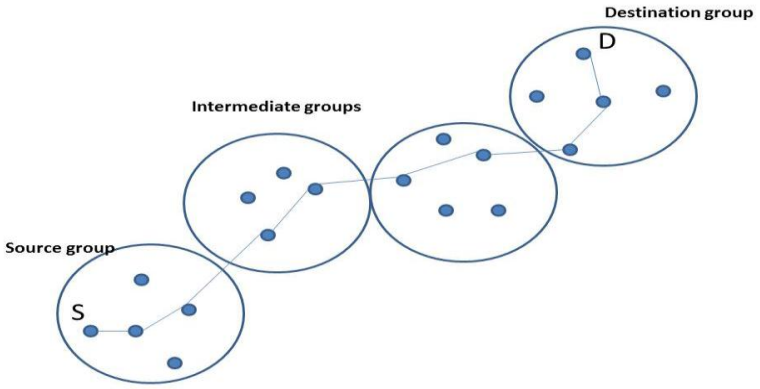


Figure 5. Clustering in HPAR

The HPAR uses the hierarchical clustering scheme and randomly chooses a node in the cluster or group in each step as an intermediate relay node as random forwarder. The source group consists

of the sender. It transmits the packets to next random forwarder from that group or next group. The random forwarder in the next group can understand that the packet is from that group that node can't get the idea of real sender. After passing through the intermediate nodes it finally reaches the destination cluster node. Then it forwards to exact destination. In the clustering the communication between the clusters are by the name of cluster group identifier. So the real identity of each node inside the cluster is maintained. The communications between the clusters are by the group or cluster identifier. Each group maintained and identifier. So the outside communication hides the real identity of the node by this group identity.

3.2 Strategies against attacks

3.2.1 Resilience to Timing Attacks

Two nodes A and B communicate with each other at an interval of 5 seconds. After a long observation time, the intruder finds that A's packet sending time and B's packet receiving time have a fixed five second difference. Then, the intruder would suspect that A and B are communicating with each other. Avoiding the exhibition of interaction between communication nodes is a way to counter timing attacks.

3.2.2 Strategy to Counter Intersection Attacks

In intersection attack an attacker with information about active users at a given time can determine the sources and destinations that communicate with each other through repeated observations. rather than using direct local broadcasting in the zone, the last RF multicasts packet pkt 1 to a partial set of nodes m.The m nodes hold the packets until the arrival of the next packet pkt 2 .Upon the arrival of the next packet, the m nodes conduct one-hop broadcasting to enable other nodes in the zone to also receive the packet in order to hide D.

Comparison of HPAR, ALARM and AO2P protocols based on some parameters:

1) Number of actual participating nodes

ALARM and AO2P is based on the GPSR method. GPSR always proceeds through the shortest paths. So the number of actual participating nodes is less compared to HPAR
.

2) Latency in packet transmission

Latency is defined as the time difference between the packet transmissions and receiving. Latency in HPAR is significantly lower than the other two. This is because of the time needed for the public key encryption of ALARM and AO2P. HPAR follows symmetric key encryption only once which reduces the latency.

3) Packet delivery rate

Fraction of successfully delivered packets to a destination is called the delivery rate .HPAR has higher delivery rates compared to AO2P and ALARM, as a result of final local broadcast process HPAR achieves enhanced route anonymity than ALARM and AO2P. HPAR has more number of actual participating nodes and its random relay node selection boost the anonymity.

4) Transmission cost

Transmission cost and latency in packet transmission are lower in HPAR compared with the other two. HPAR contributes better data delivery rate than ALARM and AO2P.

4. CONCLUSIONS

Anonymous routing protocols are crucial in MANETs to provide secure communications by hiding node identities and routes from outside observers. Anonymity in MANETs includes identity and location anonymity of senders and destinations as well as route anonymity. The aim is to make the communication between different nodes anonymous in MANET. By anonymity we mean that intermediate nodes are unaware of the sender and destination. Only the sender will know the receiver and only the receiver will know the sender.HPAR can offer high anonymity protection at a low cost when compared to other anonymity algorithms.It can also achieve comparable routing efficiency also Provide Resilience to Timing Attacks and Strategy to Counter Intersection Attacks.

ACKNOWLEDGEMENTS

The authors would like to thank everyone.

REFERENCES

1) L. Zhao and H. Shen, "ALERT: An Anonymous Location-Based Efficient Routing Protocol in MANETs," IEEE TRANSACTIONS ON MOBILE COMPUTING, JUNE 2013.
2) A. Pfitzmann, M. Hansen, T. Dresden, and U. Kiel, "Anonymity, Unlinkability, Unobservability, Pseudonymity, and Identity Management a Consolidated Proposal for Terminology, Version 0.31,"technical report, 2005.
3) K.E. Defrawy and G. Tsudik, "ALARM: Anonymous Location-Aided Routing in Suspicious MANETs," IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 10, NO. 9, SEPTEMBER 2011
4) K.E.Defrawy and G. Tsudik, "PRISM: Privacy-Friendly Routing in Suspicious MANETs (and VANETs)," IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 29, NO. 10, DECEMBER 2011.
5) Z. Zhi and Y.K. Choong, "Anonymizing Geographic Ad Hoc Routing for Preserving Location Privacy," Proc. Third Int'l Workshop Mobile Distributed Computing (ICDCSW), 2005.
6) Anupriya Augustine, Jubin Sebastian E , " A Study of Efficient Anonymous Routing Protocols in MANET"International Journal of Computer Applications (0975 – 8887) Volume 91 – No.8, April 2014.
7) J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems," Proc. Int'l Workshop Designing Privacy Enhancing Technologies: Design Issues in Anonymity and Unobservability(WDIAU), pp. 10-29, 2001.
8) Aarti ,Dr. S. S. Tyagi "Study of MANET: Characteristics, Challenges, Application and Security Attacks"International Journal of Advanced Research in Computer Science and Software Engineering ,Volume 3, Issue 5, May 2013

Authors

FahmidaAseez Post-graduate student at school of CUSAT. Had graduated from AdiSankara Institute of Engineering and Technology Kalady. Area of interest is in Network Computing.

Dr.Sheena Mathew Professor in Division of Computer Science, SOE, and Cochin University has 23 years of teaching experience in Computer Science. She had her graduation from Madurai Kamaraj University, post-graduation from Indian Institute of Science, Bangalore and doctorate from CUSAT. She was the head of Department of Division of Computer Science and engineering, school of engineering, CUSAT for the period of four years. Her areas of interest being Cryptography and Network Security. She has more than 30 publications in various international journals and conferences to her credit.

Author Index