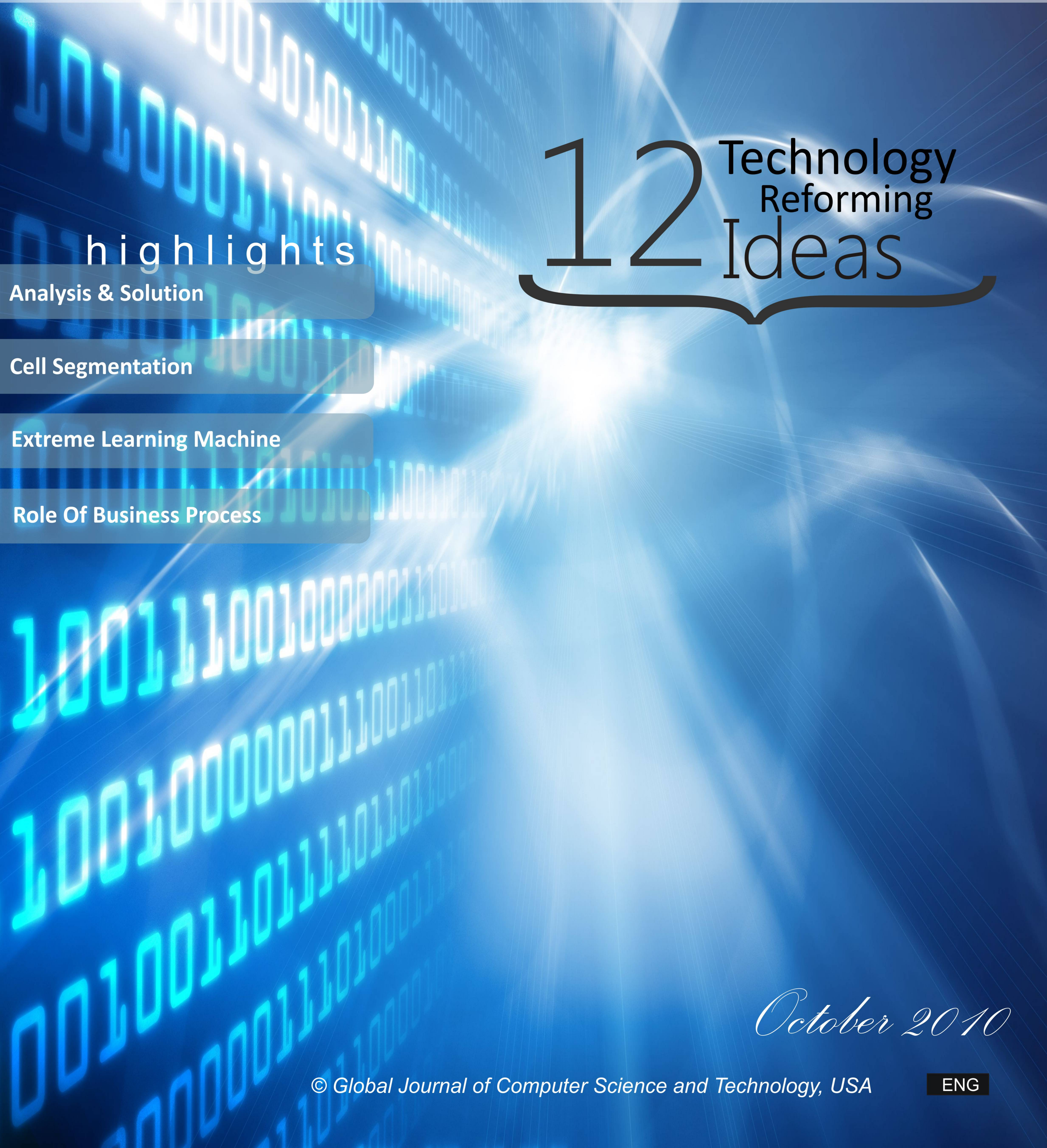
Global Journal of Computer Science and Technology

discovering thoughts and inventing future

Volume 10 Issue 13 Version 1.0 Online ISSN: 0975-4172 Print ISSN: 0975-4350



12 Technology Reforming Ideas

highlights

Analysis & Solution

Cell Segmentation

Extreme Learning Machine

Role Of Business Process

October 2010

ENG

Global Journal of Computer Science and Technology

Global Journal of Computer Science and Technology

Volume 10 Issue 13 (Ver. 1.0)

Global Association of Research

This is a special issue published in version 1.0 of "Global Journal of Computer Science and Technology." By Global Journals Inc.
All articles are open access articles distributed under "Global Journal of Computer Science and Technology"

The opinions and statements made in this book are those of the authors concerned. Ultraculture has not verified and neither confirms nor denies any of the foregoing and no warranty or fitness is implied.

Engage with the contents herein at your own risk.

The use of this journal, and the terms and conditions for our providing information, is governed by our Disclaimer, Terms and Conditions and Privacy Policy given on our website http://www.globaljournals.org/global-journals-research-portal/guideline/terms-and-conditions/menu-id-260/.

By referring / using / reading / any type of association / referencing this journal, this signifies and you acknowledge that you have read them and that you accept and will be bound by the terms thereof.

All information, journals, this journal, activities undertaken, materials, services and our website, terms and conditions, privacy policy, and this journal is subject to change anytime without any prior notice.

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374 Import-Export Code: 1109007027

USA Tax ID: 98-0673427

Global Journals Inc.

*(A Delaware USA Incorporation with "Good Standing"; **Reg. Number: 0423089**)*
Sponsors: *Global Association of Research*
Open Scientific Standards

Publisher's Headquarters office

Global Journals Inc., Headquarters Corporate Office, Cambridge Office Center, II Canal Park, Floor No. 5th, ***Cambridge (Massachusetts)***, Pin: MA 02141
United States
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Inc., City Center Office, 25200 Carlos Bee Blvd. #495, Hayward Pin: CA 94542
United States

Packaging & Continental Dispatching

Global Journals, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org* Investor Inquiries: *investers@globaljournals.org* Technical Support: *technology@globaljournals.org* Media & Releases: *media@globaljournals.org*

Pricing (Including by Air Parcel Charges):

For Authors:
22 USD (B/W) & 50 USD (Color)
Yearly Subscription (Personal & Institutional):
200 USD (B/W) & 500 USD (Color)

Editorial Board Members (HON.)

Contents of the Volume

From the Chief Author's Desk

We see a drastic momentum everywhere in all fields now a day. Which in turns, say a lot to everyone to excel with all possible way. The need of the hour is to pick the right key at the right time with all extras. Citing the computer versions, any automobile models, infrastructures, etc. It is not the result of any preplanning but the implementations of planning.

With these, we are constantly seeking to establish more formal links with researchers, scientists, engineers, specialists, technical experts, etc., associations, or other entities, particularly those who are active in the field of research, articles, research paper, etc. by inviting them to become affiliated with the Global Journals.

This Global Journal is like a banyan tree whose branches are many and each branch acts like a strong root itself.

Intentions are very clear to do best in all possible way with all care.

Dr. R. K. Dixit
Chief Author
chiefauthor@globaljournals.org

Aes Algorithm Using 512 Bit Key Implemented for Secure Communication

S.Radhika[1], A.Chandra Sekar[2]

GJCST Classification E.3

***Abstract*-The main aim of this paper is to provide stronger security for communication over the Internet by enhancing the strength of the AES algorithm. Rijndael's algorithm was selected as the Advanced Encryption Standard. The AES algorithm was believed to provide much more security without any limitations. But, recently some breaking methods on the AES have been found by cryptanalysts. In AES algorithm, the number of rounds involved in the encryption and decryption depends on the length of the key and the number of block columns. So, the number of rounds is increased to improve the strength of the AES. The strength of the AES algorithm is enhanced by increasing the key length to 512 bit and thereby the number of rounds is increased in order to provide a stronger encryption method for secure communication. Code optimization is done in order to improve the speed of encryption and decryption using the 512 bit AES.**

Keywords-Cryptography, Encryption, Java implementation, Decryption, AES algorithm

I. Introduction

Network security is becoming more and more important as people spend more and more time connected in a network. It involves all activities that organizations, enterprises, and institutions undertake to protect the value and ongoing usability of assets and the integrity and continuity of operations. Security attacks include unauthorized reading of a message of file, traffic analysis, modification of messages or files and denial of service. One of the most publicized types of attack on information systems is the computer virus. An effective network security strategy requires identifying threats and then choosing the most effective set of tools to combat them. Security involving communications and networks is not as simple as it might first appear to the novice. The expansion of the connectivity of computers makes ways of protecting data and messages from tampering or reading important. Intruders may reveal the information to others, modify it to misrepresent an individual or organization, or use it to launch an attack. One of the primary reasons that intruders can be successful is that most of the information they acquire from a system is in a form that they can read and comprehend. One solution to this problem is, through the use of cryptography. Cryptography ensures that the messages cannot be intercepted or read by anyone other than the authorized recipient. It prevents intruders from being able to use the information that they capture. Cryptography secures information by protecting its confidentiality and can also be used to protect information about the integrity and authenticity of data.

1) *Related Work*

The first open encryption algorithm, Data Encryption Standard (DES) was adopted by the National Institute of Standards and Technology(NIST) to protect the sensitive information as Federal Information Processing Standard 46 (FIPS PUB 46) in 1977 [1]. However, the shorter length of key, the complementary property and existence of weak and semi-weak keys reduce the security of DES. Differential cryptanalysis attack is capable of breaking DES in less than 2^{55} complexities. The linear cryptanalysis method can find a DES key given 2^{43} known plaintexts, as compared to 2^{47} chosen plaintexts for differential cryptanalysis. So, it was more essential to find a stronger encryption algorithm to substitute the DES.In spite of the vulnerability of DES to a brute-force attack, there has been considerable interest in finding an alternative. One approach is to design a completely new algorithm and another alternative would be the one that preserves the existing one by using multiple encryption with DES and multiple keys. Three other algorithms were found to solve the problems of DES. They are Double DES, Triple DES with two keys and Triple DES with three keys. The principal drawback of Triple DES is that it has three times as many rounds as DES and hence it is much slower. Triple DES uses a 64 bit block size which is another drawback because for both efficiency and security, a larger block size is desirable. Because of these drawbacks, Triple DES is not favorable for long term use.The Rijndael algorithm was adopted as an encryption standard, the Advanced Encryption System (AES) by the NIST as FIPS PUB 197 (FIPS 197) on November 2001 [2]. The AES algorithm was believed to provide more security than the DES [3]. The AES algorithm was designed to have resistance against all known attacks, speed and code compactness on a wide range of platforms and design simplicity [7]. AES has three variable key lengths but block length is fixed to 128 bits [2]. The three key sizes of AES are 128, 192 and 256 bits. Their number of possible keys is 3.4 x 10^{38}, 6.2 x 10^{57} and 1.1 x 10^{77} respectively [2]. There are on the order of 10^{21} times more AES 128-bit keys than DES 56-bit keys. AES with 128-bit keys has stronger resistance to an exhaustive key search than DES.

Abou[1]t- Lecturer Dept. of Electrical and Electronics Engineering Sathyabama University Chennai-600 119 radhikachandru79@gmail.com
About[2]- Professor Dept. of Computer Science and Engineering St. Joseph's College of Engineering Chennai-600 119 drchandrucse@gmail.com

2) *Drawbacks of AES 256*

Rijndael has very strong resistance against the differential cryptanalysis and linear cryptanalysis attacks since it used Wide Trail Strategy in its design [8]. Although these linear attacks are invalid for the AES, they have been extended in several ways for recent years and new attacks have been published that are relative to them [4-6, 9-11]. The newest attack combined boomerang and the rectangle attack with related-key differentials was introduced by E. Biham, et al. in 2005 [9]. It uses the weaknesses of few nonlinear transformations in the key schedule algorithm of ciphers, and can break some reduced-round versions of AES. It can break 192-bit 9-round AES by using 256 different related keys. Rijndael inherits many properties from Square algorithm. So, the Square attack is also valid for Rijndael which can break round-reduced variants of Rijndael up to 6 or 7 rounds (i.e.AES-128 and AES-192) faster than an exhaustive key search [6]. N. Ferguson et al. proposed some optimizations that reduce the work factor of the attack [5]. So, this attack breaks a 256-bit 9-round AES with 2^{77} plaintexts under 256 related keys, and 2^{224} encryptions.

II. AES USING 512 BIT KEY

AES is a block cipher and the most popular algorithm used in symmetric key cryptography. It is a substitution-permutation network and not a Fiestel network like DES. When the number of rounds is increased in AES, the complexity of AES encryption and decryption also increases. The number of rounds (Nr) in the
AES algorithm depends on the length of main keys (Nk) and the number of block columns (Nb), i.e. Nr = Nk+Nb+abs(Nk-Nb). So, the length of the key is increased to 512 bits in order to increase the number of rounds. The structure of AES is quite simple. The input to the encryption and decryption algorithms is a single 128-bit block and the key is 512 bits. It requires the same key to be used for encryption and decryption.

1) *Implementation of AES encryption*

AES operates on a 16×32 array of bytes, termed the state. The input key for encryption is 512 bits. To represent the 512 values 9 bits are required. So each entry in S-box of AES 512 is 9 bits long. The cipher is specified in terms of repetitions of processing steps that are applied to make up rounds of keyed transformations between the input plain-text and the final output of cipher-text. The encryption procedure of AES 512 has been illustrated in figure 1.Each round in AES 512 encryption includes four different round transformations namely Substitute Bytes, Shift Rows, Mix Columns and Add Round Key. The last round of AES 512 encryption alone does not include the Mix Columns transformation.

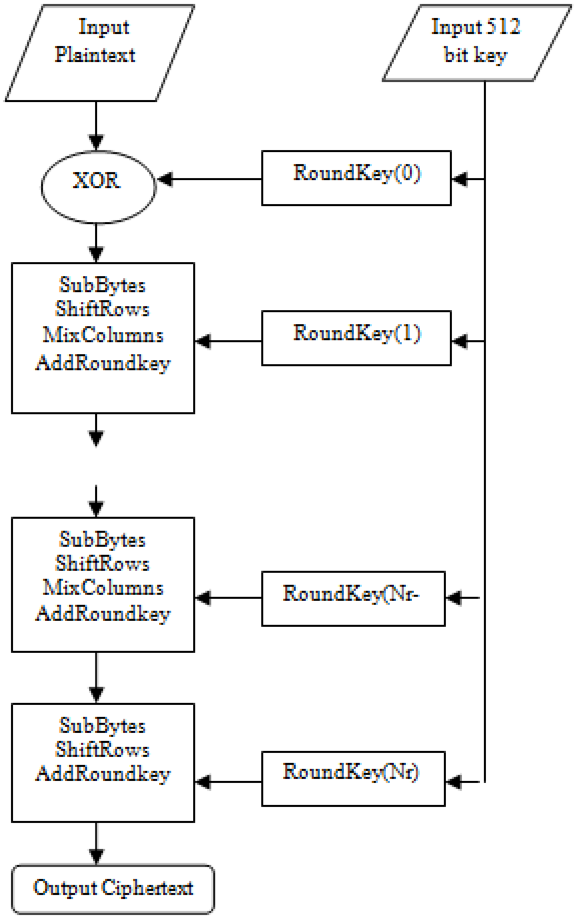


Figure: 1 Encryption procedure of AES 512

2) *Implementation of AES decryption*

A set of reverse rounds are applied to transform cipher-text back into the original plain-text using the same encryption key. The four reverse transformations used are Add Round Key, Inverse Mix Columns, Inverse Shift Rows and Inverse Substitute Bytes. The inverse S-box contains 512 values in its 16x32 array of bytes. Each round in decryption of AES 512 includes all the four reverse transformations except in the first round. The Inverse Mix Column transformation is violated in the first round of decryption since it does not occur in the last round of encryption. The decryption procedure of AES 512 is illustrated in figure 2.

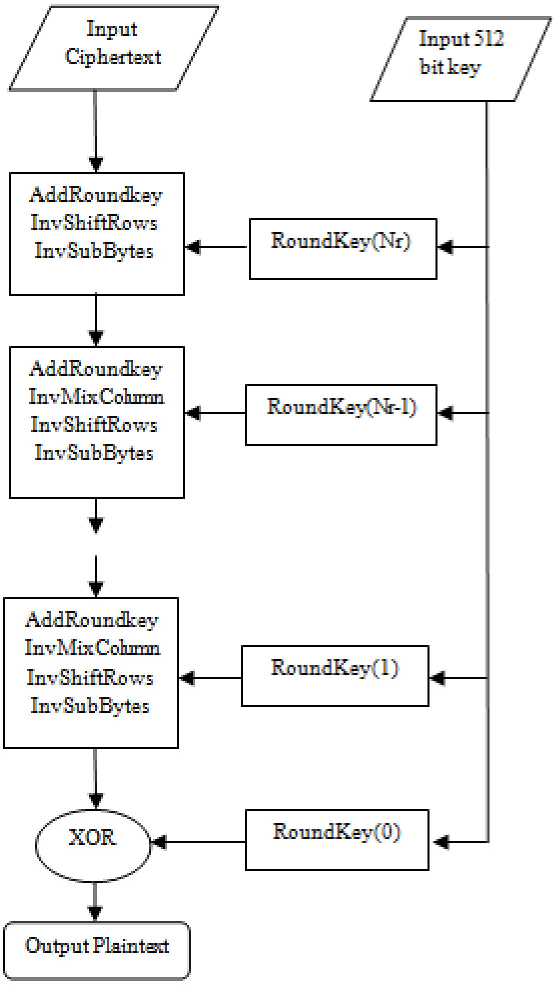


Figure: 1 Decryption procedure of AES 512

3) *Comparison of AES 256 and AES 512*

The performance of 256 bit AES algorithm is compared with the performance of AES 512 algorithm. Encryption and decryption of AES 256 is implemented to compare it with AES 512. In terms of security the 256 bit AES algorithm is weaker than the 512 bit AES algorithm. This is because the length of the key used in 512 bit AES increases the number of rounds for both encryption and decryption. But when the number of rounds increases, the encryption and decryption procedures become more complex thereby degrading the speed of the 512 bit AES algorithm. Thus there is a tradeoff between speed and security. The performance is compared in terms of time taken. The time taken for encryption and decryption of AES 256 bit and AES 512 bit are noted to measure their speed. The system time is noted at the start of encryption process and the end time, after encryption completes is also noted. The same process is repeated for decryption also to calculate the time taken for decryption.

4) *Code Optimization*

Speed and security are the most important factors that influence the transmission of any data over the network. In AES 512, a higher level of security is achieved due to the increased length of the key. But as mentioned above there is degradation in speed. It is very essential to maintain a balance between the speed and security parameters of the AES algorithm. Java implementation of AES 512 encryption and decryption is done. Code optimization techniques reduce the amount of time that a program takes to perform some task. Hence code optimization techniques are employed to improve the speed of the 512 bit AES algorithm. The source code level optimization technique is applied here. That is because avoiding bad quality coding can also improve performance, by avoiding obvious slowdowns. Creating local variables dynamic, restricting new operator, if statements are replaced with switch cases and removal of unnecessary declarations are some of the code optimizations employed.

III. Performance Evaluation

The performance of AES 512 is evaluated on the basis of two major parameters: security and speed. AES 512 has better security than the AES 256 algorithm since the number of rounds is increased. Optimization is done on AES 512 to improve its speed. We found that, the optimized AES 512 has an acceptable speed of encryption and decryption when compared to the AES 512 that was not optimized.

IV. Conclusion

AES is a new cryptographic algorithm that can be used to protect electronic data. Its security has attracted cryptographist's attentions. The methods of new attacks welled show the weaknesses of AES algorithm. When the number of rounds is increased, it improves the complexity of the algorithm making it strong against the cryptographic attacks. The length of the key is increased in order to increase the number of rounds involved as number of rounds depend on the length of the key used. Thus the increase in length of the key gives the AEs algorithm strong resistance against the new attacks and has an acceptable speed of data encryption and decryption.

V. Reference

1) U.S. Department of Commerce/NIST, "Data Encryption Standard," FIPS PUB 46-3, pp. 1-26, October 1999.
2) NIST, "Advanced Encryption Standard," FIPS PUB 197, pp. 1-51, November 2001.
3) J. Daemen and V. Rijmen, The Design of Rijndael: AES - The Advanced Encryption Standard, Springer-Verlag, Berlin Heidelberg, 2002.
4) H. Gilbert and M. Minier, A collision attack on seven rounds of Rijndael, Proceedings of the 3rd AES Candidate Conference, pp.230-241, April 2000.

5) N. Ferguson, J. Kelsey, S. Lucks, et al. "Improved cryptanalysis of Rijndael," Lecture Notes in Computer 1-4244-1035-5/07/$25.00 .2007 IEEE. 221 Science,vol. 1978, pp.213-230, Berlin: Springer-Verlag, 2001.

6) S. Lucks, "Attacking seven rounds of Rijndael under 192-bit and 256-bit keys," Proceedings of the 3rd AES Candidate Conference, pp. 215-229, April 2000.

7) J. Daemen and V. Rijmen, "The Block Cipher Rijndael," Lecture Notes in Computer Science, vol.1820, pp.277-284, Berlin: Springer-Verlag, 2000.

8) J. Daemen, and V. Rijmen, "The Wide Trail Design Strategy," Lecture Notes in Computer Science, vol. 2260, pp.222 - 238, Berlin: Springer-Verlag, 2001.

9) E. Biham, O. Dunkelman, and N. Keller, "Related-Key Boomerang and Rectangle Attacks," Lecture Notes in Computer Science, vol. 3494, pp. 507-525, Berlin: Springer-Verlag, 2005.

10) G. Jakimoski and Y. Desmedt, "Related-Key Differential Cryptanalysis of 192-bit Key AES Variants," Lecture Notes in Computer Science, vol. 3006, pp. 208-221, Berlin: Springer-Verlag, 2004.

11) J. H. Cheon, M. J. Kim and K. Kim,et al., "Improved Impossible Differential Cryptanalysis of Rijndael and Crypton," Lecture. Notes in Computer Science, vol. 2288, pp. 39-49, Berlin: Springer-Verlag, 2002.

Automatic Load Balancing Strategies for A Distributed Computer System

K.Kungumaraj, M.Sc., M.Phil., B.L.I.S.

GJCST Classification C.1.4

***Abstract*- A distributed system consists of independent workstations connected usually by a local area network. Static load balancing don't fulfill the requirements for load balancing. As in static load balancing, number of jobs at a station is fixed. Automatic load balancing does the process while job are in execution. Jobs are allocated to host or node. Load at each post is calculated (as number of process, structure of node, network bandwidth etc.) automatically. As sender initiated or receiver initiated approaches are available to find the desired load for transferring the load. Every distributed system consists of a number of resources interconnected by a network Besides providing communication facilities, it facilitates sharing computation power by migrating a local process and executing it at a remote node of the network. A process may need to be migrated because of the shortage of required resources at the local node or the local node has to be shut down for some reason. A process may also be executed remotely if the expected turnaround time needs to be better. From a user's point of view, the set of available resources in a distributed system acts like a single virtual system. Hence when a user submits a process for execution, it becomes the responsibility of the resource manager of the distributed operating system to control the assignment of resources to processes and to route processes to suitable nodes of the system according to these assignments.**

I. Introduction

The entire system is divided into number of subsets equal to the number of nodes. Each node in the system is assigned a subset. A node queries only the member nodes of its set to collect system state information and thus takes scheduling decision. The amount of state information maintained at each node of the DCS and how it is used, is an important parameter for developing distributed load balancing algorithms. The state information about the total system can be used to reduce the message traffic in the network and to recover from failures. Message traffic should be minimized in order to decrease the overhead in the communications network.

The major contributions of the thesis include:

(i) a novel workload migration technique proposed by considering the homogeneous and dynamic features of distributed computing environments

(ii) dynamic and stable technique that satisfies the quick decision making capability and provides a balanced system performance with respect to scheduling overhead.

II. The System Model

The effectiveness of a new technique depends on the suitability of the model as well as the validity of the assumptions made about the computing environment. The technique presented here is based on the following assumptions and conditions for the distributed environment:

- All nodes in the system are assigned unique identification numbers from 1 to N.
- All the nodes in the system are fully connected.
- The method used as the Load Estimation policy would be the measure of CPU utilization of the nodes. CPU utilization is defined as the number of CPU cycles actually executed per unit time.
- Process transfer policy that determines whether to execute a process locally or remotely.
- The node sends its state information to other nodes only when its state switches from normal load region to either overload or under load region.

The following aspects about the reliability of the underlying communications network should be considered.

- Message delivery is guaranteed.
- Message-order is preserved.
- Message transfer delays are finite.
- The topology of the network is known.

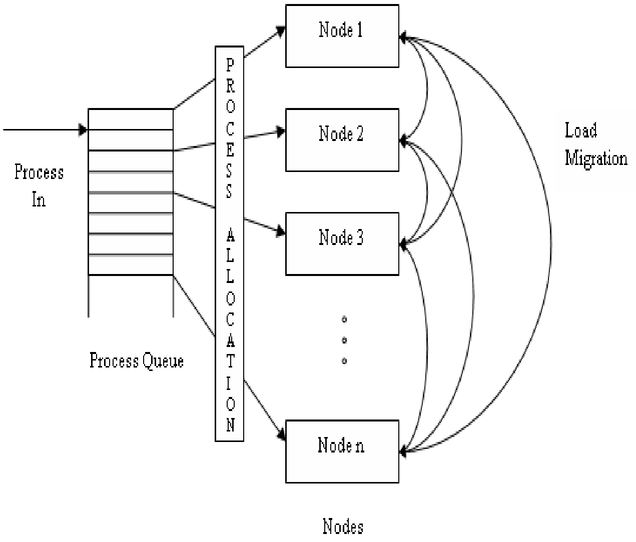


Figure – 1 Simple dynamic Load balancing to avoid overload on heavily loaded node by transferring process to light weighted node.

About- Research Scholar, Karpagam University, Coimbatore. E-Mail : kungumaraj@yahoo.co.in Mobile : 09380920577

As shown in diagram, initially processes are stored in queue or process can be allotted as they arrive. If these are placed in queue, processes are allotted one by one to primary nodes. Processes are migrated from heavily loaded node to light weighted node. Process migration is greatly affected by the network bandwidth and work load. In order to reduce the traffic, nodes are grouped into clusters. First a light weighted node is checked in the same cluster, if suitable node not found then after nearby cluster is searched and after getting a required node transfer takes place if a protocol is satisfied for load transfer.

III. SENDER-SIDE RULES (S)

Possible-destinations = { site: Load(site) -Reference(s) < d(s) }
Destination = Random(Possible-destinations)

IF Load(s) -Reference(s) > q1(s) THEN Send

RECEIVER-SIDE RULES (R)

IF Load(r) < q2(r) THEN Receive

Figure – 2 The load-balancing policy considered in this thesis

The sender-side rules are applied by the load-balancing software at the site of arrival (s) of a task. Reference can be either 0 or MinLoad; the other parameters — d, q1, and q2 — take non-negative floating-point values. A remote destination (r) is chosen randomly from Destinations, a set of sites whose load index falls within a small neighborhood of Reference. If Destinations is the empty set, or if the rule for sending fails, then the task is executed locally at s, its site of arrival; otherwise, the chosen site (r) is requested to receive the task. Upon receiving that request, the remote site applies its receiver-side rule. If the rule succeeds, the request is accepted, and the task is migrated; otherwise, the task is executed locally at s, its site of arrival.

IV. ESTIMATION OF LOAD

Before distribution change of the load can be realized the load data must be derived by special methods. Here we must separate,

- Computing load
- Communication load.

The estimation uses a simple counting method. Here within a given measuring interval the difference of in and out going communication demands and the number of calls of the measuring process are counted. The measuring process is added to the APL (Active Process List) of the processes of low priority. The resulting load data are used to calculate process related values regarding the work and data storage of the processes and then to leave it to the control instance of the processor node.

V. CONCLUSION

The motivation towards the development of this paper was to develop a automatic distributed scheduling technique that reduces the communication overhead involved in decision making. As main aim in distributed system is to execute the process at minimum cost i.e. time is most important factor can be considered in cost calculation. This is attributable to the fact that New dynamic load balancing policy achieves a higher success, in comparison to the previously used load balancing techniques, in reducing the likelihood of nodes being idle while there are tasks in the system, comprising tasks in the queues.

Cluster Reformation and Scheduling for Interference Mitigation in Coexistence Heterogeneous Wireless Packet Networks

G.M.Tamilselvan[1], Dr.A.Shanmugam[2]

GJCST Classification
C.2.1

Abstract* - The emerging IEEE 802.15.4 (Zigbee) standard is designed for low data rate, low power consumption and low cost wireless personal area networks (WPANs).In ubiquitous networking environments; we generally need two or more heterogeneous communication systems coexisting in a single place. Especially, wireless local area networks (WLANs) based on IEEE 802.11b specifications and wireless personal area networks (WPANs) based on IEEE 802.15.4 specifications need to coexist in the same Industrial, Science and Medial (ISM) band. If the WPAN communication coverage is expanded using a cluster-tree network topology, then the 802.15.4 network is more susceptible to interference from neighboring WLANs. In this paper, we propose an adaptive transmission power aware cluster reformation and scheduling algorithm using multiple channels in a WPAN in the presence of WLAN interference. The algorithm includes node identification, channel allocation, cluster reformation and time scheduling. To evaluate the performance of the proposed algorithm, the performance metrics such as Packet Error Rate (PER), Throughput, Average End-End Delay and Average Jitter is measured through Qualnet simulation. PER is calculated from bit error rate. The simulation results are compared with the conventional TDMA scheme. The measurement result shows that the proposed algorithm is effective in an IEEE 802.15.4 cluster-tree network in the presence of multiple IEEE 802.11 interferers.

Keywords-Clustering, Coexistence, Heterogeneous, Packet Error Rate (PER), WLAN and WPAN (Zigbee).

I. INTRODUCTION

As a low-power and low-cost technology, IEEE 802.15.4 is establishing its place on the market as an enabler for the emerging wireless sensor networks (WSNs) [1]. Like IEEE 802.11b and IEEE 802.11g, IEEE 802.15.4 is also used in the 2.4 GHz ISM band. Due to supporting complimentary applications, they are very likely to be collocated within the interfering range of each other and therefore their ability to coexist needs to be evaluated. In this paper we focus on the coexistence between these two major wireless standards that operate in the 2.4*GHz* ISM band. Their overlapping frequency channels are shown in energy and is designed for low rate, low cost applications over a short range of 30 to 100 meters. The IEEE 802.15.4 Fig. 1. IEEE 802.15.4 defines the physical layer and the MAC sub layer of the OSI Zigbee stack. It supports devices that consume minimum defines three physical layers; the 2.4 GHz, 868 MHz and 915 MHz frequency bands. The unlicensed industrial scientific medical (ISM) 2.4 GHz band is available worldwide, while the 868 MHz and 915 MHz bands are available in Europe and North America respectively. A total of 27 channels with three different data rates are defined for the IEEE 802.15.4: 16 channels with a data rate of 250 kbps at the 2.4 GHz band, 10 channels with a data rate of 40 kbps at the 15 MHz band, and 1 channel with a data rate of 20 kbps at the 868 MHz band. The relationship between the IEEE 802.11b (non-overlapping sets) and the IEEE 802.15.4 channels at the 2.4 GHz is illustrated in Fig.1

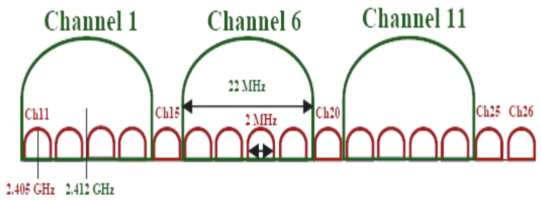


Fig.1.802.11 and 802.15.4 channels in the 2.4GHz ISM band

Figure 1 shows the operation frequency spectrum of both IEEE 802.11 and IEEE 802.15.4 networks in the 2.4 GHz ISM band. The IEEE 802.11 standard has 11 channels each of which occupies 22 MHz and up to 3 channels can be used simultaneously without mutual interference. As illustrated in the figure, channels 1, 6 and 11 can be used by the IEEE 802.11 devices to eliminate the mutual interference. On the other hand, the IEEE 802.15.4 standard defines 16 channels (2 MHz), channels 11 through 27, in the 2.4 GHz ISM band all of which can be used simultaneously without mutual interference.The IEEE 802.15.4 standard recommends using the channels that fall in the guard bands between two of the three adjacent non-overlapping IEEE 802.11 channels or above these channels to prevent interference between the IEEE 802.15.4 and the IEEE 802.11. From the figure, it is

About[1]-Senior Lecturer, Department of ECE,
About[2]-Principal BannaiAmman Institute of Technology, Sathyamangalam-638401, India
tamiltamil@rediffmail.com,dras@yahoo.com

shown that 4 of the 11 channels will have the minimal interference which in most cases is enough to cover a big region unless more IEEE 802.15.4 networks are added.There have been some studies about coexistence between the IEEE 802.11b and IEEE 802.15.4. According to [1] [2] [4] IEEE 802.15.4 has a little impact on the IEEE 802.11 performance. However, IEEE 802.11 can have a serious impact on the IEEE 802.15.4 performance if the channel allocation is not carefully taken into account [1] [3]. While the conclusion is true in general, we believe the studies so far have dealt with only limited cases of coexistence scenarios. In [3], the Packet Error Rate (PER) of IEEE 802.15.4 under the IEEE 802.11b interference is analyzed from an assumption of blind transmissions, i.e. both IEEE 802.11b and IEEE 802.15.4 transmit packets regardless of whether the channel state is busy or not. In [4], measurements are performed to quantify coexistence issues. Channel Conflict Probabilities between IEEE 802.15 based Wireless Personal Area Networks is modeled in [5]. Packet Error Rate of IEEE 802.15.4 under IEEE 802.11b interference is analyzed in [6].In [7] Packet Error Rate of IEEE 802.11b under IEEE 802.15.4 interference is analyzed. In [8] channel conflict probabilities between IEEE 802.11b and IEEE 802.15.4 have been modeled. In [9] channel collision between IEEE 802.15.4 and IEEE 802.11b for circular and grid topology is analyzed with the mobility model. The effect of inter packet delay is analyzed in [10]. The author concluded that despite its low transmit power and simple modulation technique, IEEE 802.15.4 shows a robust behavior against interference of other 2.4 GHz systems and even in the worst case conditions for frequency overlap, local distance and high traffic load for interference, some time slots remain for a successful transmission of IEEE 802.15.4.In above said related works only two WPAN nodes which are collocated with multiple WLAN nodes are considered. But today the sensor networks play a vital role in any automation; we have to consider the multiple WPAN nodes. When multiple sensor nodes are used, time slot mechanism is not helpful in WPAN network because ZigBee is a mesh networking technology.The remainder of the paper is organized as follows: Section II gives an overview of the IEEE 802.11b and IEEE 802.15.4 standard. Section III presents a heterogeneous wireless network with conventional TDMA scheme for packet transmission and cluster tree network with two different time and frequency scheduling schemes. Simulation results are shown in Section IV. Our conclusion is drawn in Section V.

II. OVERVIEW OF IEEE 802.11B AND IEEE 802.15.4

1) *IEEE 802.11b*

IEEE 802.11b standard defines the Medium Access Control (MAC) sub layer and the Physical (PHY) layer for wireless LANs. The standard operates at 13 overlapping channels in the 2.4 GHz ISM band and the bandwidth of each channel is 22 MHz. IEEE 802.11b MAC employs the Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) mechanism. Before initiating a transmission, an IEEE 802.11b node senses the channel to determine whether another node is transmitting. If the medium is sensed idle for a Distributed coordination function Inter-Frame Space (DIFS) time interval the transmission will be preceded. If the medium is busy the node defers its transmission. When the medium becomes idle for a DIFS interval, the node will generate a random back off delay uniformly chosen in an interval. This interval [0, W] is called Contention Window, where W is the size of the contention window. The initial W is set to CWmin. The back off timer is decreased by one as long as the medium is sensed idle for a back off time slot. The back off counter will become frozen when a transmission is detected on the medium, and resumed when the channel is sensed idle again for a DIFS interval. When the back off timer reaches zero, the node transmits a DATA packet. Immediately after receiving a packet correctly, the destination node waits for a Short Inter Frame Spacing (SIFS) interval and then transmits an ACK back to the source node.

2) *IEEE 802.15.4*

The IEEE 802.15.4 MAC sub layer is based on CSMA/CA (channel sense multiple access) with two modes of operation: the unslotted-CSMA (beaconless mode) and the slotted-CSMA (beacon enabled mode). The basic responsibilities for the MAC sub layer is transmitting beacon frames, synchronization and providing a reliable transmission between Zigbee devices. Link layer acknowledgments are optional in IEEE 802.15.4 which can provide extra link level reliability. For our simulations, the unslotted-CSMA is used as all sources will be continuously contending for the channel. Link layer acknowledgments are used in order to make the transmission more reliable. To minimize the energy consumption of the Zigbee nodes, the slotted CSMA/CA should be taken into consideration since it uses beacon frames that contain information about when nodes can go into sleep mode. However, this is beyond the scope of this paper.

III. .PROPOSED SCHEME

In this paper, we propose power aware time slot and frequency based spectrum access analysis for the performance metrics such as bit error Rate, PER, throughput, average End-End delay and average jitter of IEEE 802.15.4. In this proposed scheme the WPAN devices are clustered. Each cluster will have one PAN coordinator and four end devices. We consider a heterogeneous network with random topology.IEEE 802.15.4 topologies are shown in Fig.2 Here the performance of IEEE 802.15.4 under the interference of IEEE 802.11b and the interference among IEEE 802.15.4 nodes because of multiple transmissions is analyzed using Qualnet 4.5 simulation. For simulation, the unslotted CSMA/CA of the IEEE 802.15.4 model is developed using Qualnet 4.5.The random topology scenario of coexistence heterogeneous network with 20 WPAN and 20 WLAN nodes for heterogeneous wireless network with conventional TDMA scheme and two different scheduling schemes are shown in Fig.4 ,5&6 respectively.

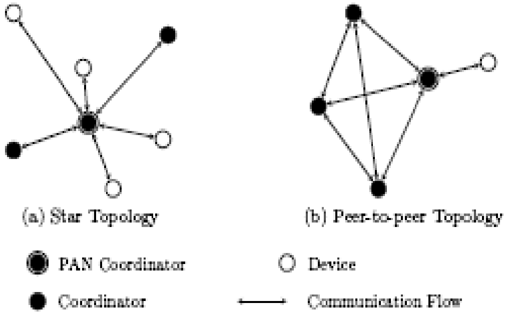


Fig.2. IEEE 802.15.4 Topologies

The PHY of the IEEE 802.15.4 at 2.4 GHz uses offset quadrature phase shift keying (OQPSK) modulation. Denote that the E_b / N_0 is the ratio of the average energy per information bit to the noise power spectral density at the receiver input, in the case of an additive white Gaussian noise (AWGN) channel. Then the bit error rate (BER), P_B, can be expressed as

$$P_B = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) \qquad (1)$$

Where Q(x) is

$$Q(x) = \frac{1}{\sqrt{x}}\int_x^{\infty} \exp\left(\frac{-u^2}{2}\right) du \qquad (2)$$

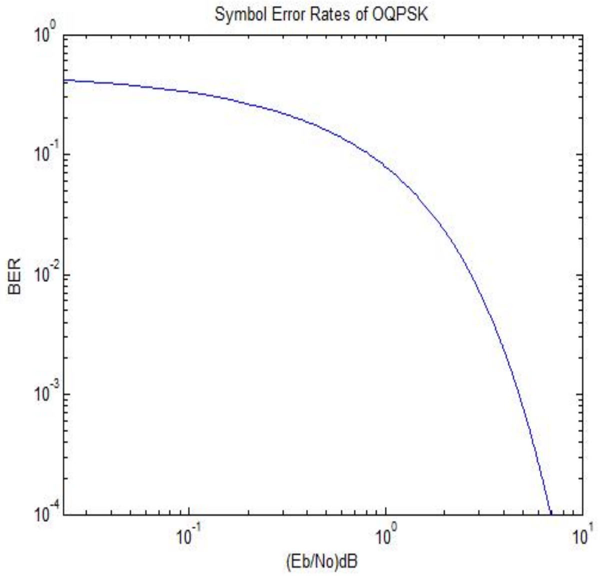


Fig.3: Theoretic Bit Error Rate of OQPSK

Fig.3 shows the relationship between the bit error rate and E_b / N_o simulated in Matlab. The bit error rate decreases when E_b / N_o increases. The noise power spectral density increases when collision increases. As the number of WLAN sources increases, the BER of IEEE 802.15.4 increases because contentions among multiple WLANs increase the channel usage and cause collisions, which is more powerful interference, source to.

The PER is calculated as a function of the BER, i.e., P_b. The probability of not having a bit error is the probability that all the bits are received correctly. Therefore the conditional probability of PER is one minus the probability of no bit errors and is computed as follows:

$$PER = 1 - \left(1 - \overline{P_b}\right)^N \qquad (3)$$

where N represents the number of bits in a packet. For the experimental setting each packet is composed of 105 bytes in the case of WPAN node and 1500 bytes in the case of WLAN node. If there is an error correction mechanism, then the PER utilizing the BER should be computed differently. However, the experimental platform does not provide an error correction mechanism and Equation 3 is the final form of the PER.

In this paper conventional TDMA scheme and two clusters based scheduling schemes are proposed and the results are compared. In cluster based scheduling the first scheme is called inter cluster scheduling. In this scheme the nodes are separated based on their transmission power. The output power of 802.15.4 devices is typically as low as 0 *dBm* , whereas the output power of 802.11b devices is 15 *dBm* or above. Then WLAN nodes are grouped under one operating frequency and WPAN nodes are clustered with cluster size 5.Each cluster will have one PAN coordinator and four end devices. Each cluster is allotted unique channel frequency for error free transmission. After frequency scheduling, in each channel specific time slot is allotted for packet transmission.In second scheme the new cluster is formed from the existing clusters and scheduling is done. The cluster members from different clusters are grouped under one channel and specific time slots are allotted for packet transmission. The figure 4, 5 & 6 shows the scenario for heterogeneous network with conventional TDMA scheme and two different clusters based scheduling schemes respectively.

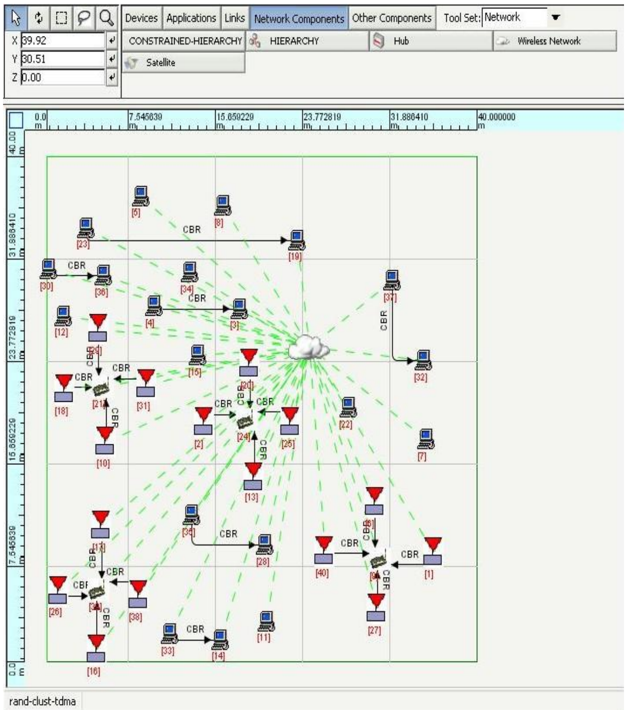


Fig.4.Random topology scenario with Conventional TDMA

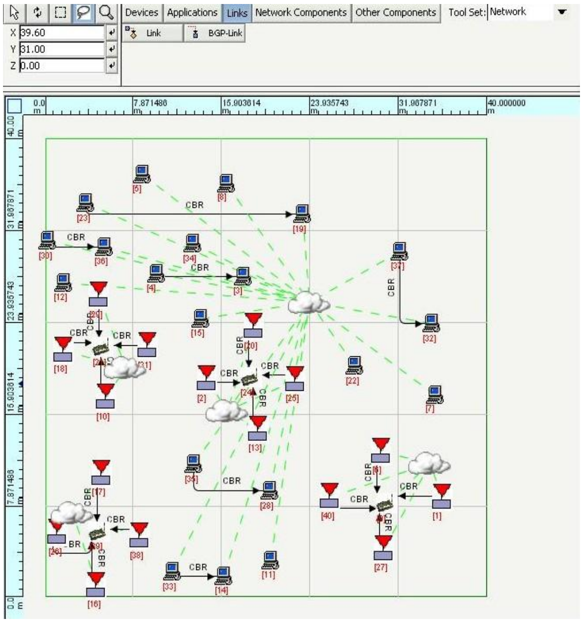


Fig.5.Random topology scenario with Inter cluster Scheduling

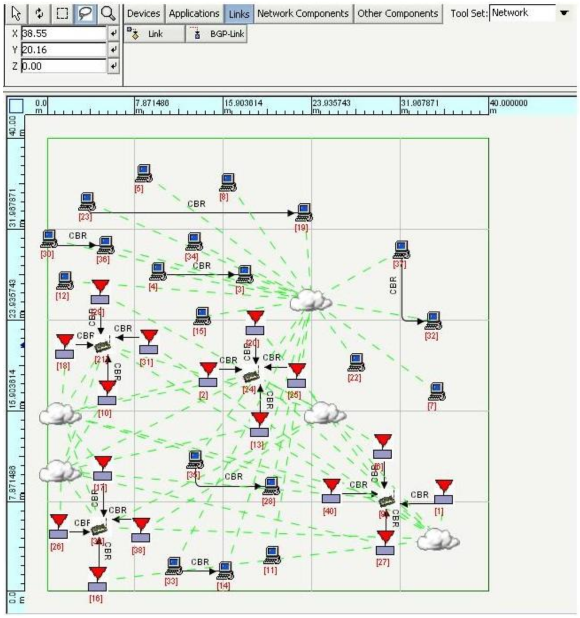


Fig.6.Random topology scenario with Reformed cluster Scheduling

IV. Simulation Results and Discussion

To evaluate the effectiveness of the proposed scheme in a coexistence heterogeneous wireless network, a simulation study was conducted using Qualnet 4.5 simulator. The simulation is conducted for three different schemes and the results are compared. The bit error is measured from the simulation. The bit error rate is calculated to find packet error rate (PER).For the conventional TDMA scheme all the nodes are linked with single channel and time slots are allotted for transmission. For this scheme the simulation time is fixed as 53s. The simulation configuration and parameters used in this paper is shown in Table 1.

TABLE 1.
SIMULATION CONFIGURATION AND PARAMETERS

Parameter	IEEE 802.11b	IEEE 802.15.4
Number of Nodes	20	20
Transmission Power	15dbm	3dbm
Modulation	CCK	OQPSK
MAC Protocol	802.11	802.15.4
Routing Protocol	Bellman ford	AODV
No of Packets	100	100
Payload Size	1500bytes	105bytes
Simulation Time	35s	
Packet Interval	100ms	1ms
Packet Transmission Time	5s	1s
Test bed size	40m × 40m	
Topology	Random	

The effectiveness of the proposed scheme was measured with different metrics such as Bit error rate, Packet Error rate, Throughput, Average End-End delay and Average jitter. The figure 7-11 shows the performance analysis of random topology with two different proposed schemes.The fig.7 shows the bit error rate analysis for random topology. In this figure bit error rate for conventional TDMA and two different schemes namely inter cluster scheduling and Reformed cluster scheduling is shown. When the reformed cluster scheduling is adopted the bit error rate becomes zero. When the conventional TDMA and inter cluster scheduling is adopted ,time slot mechanism is not helpful in WPAN network because ZigBee is a mesh networking technology, which means that devices can automatically route messages on each other's behalf (often called multi-hopping). This allows deploying larger networks without immoderately increasing the transmission power since direct communications occur only in a geographically-restricted area.

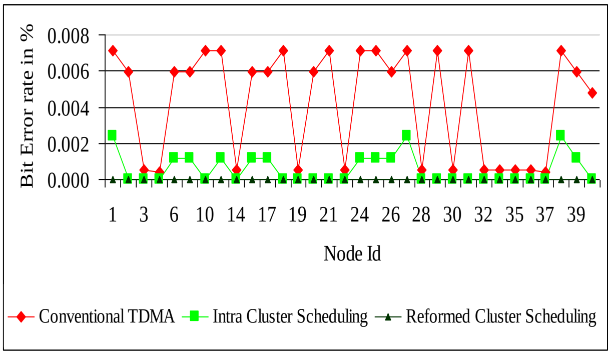


Fig .7.Bit Error Rate Analysis for Random Topology

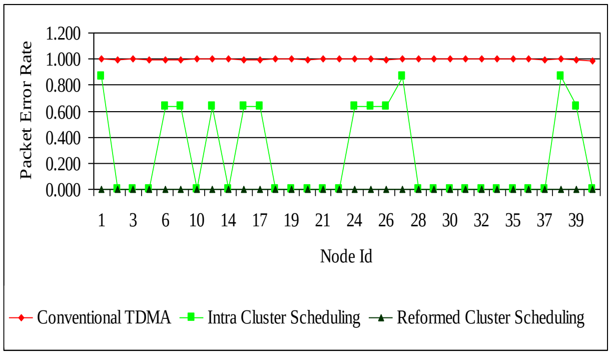


Fig.8.Packet Error Rate Analysis for Random Topology

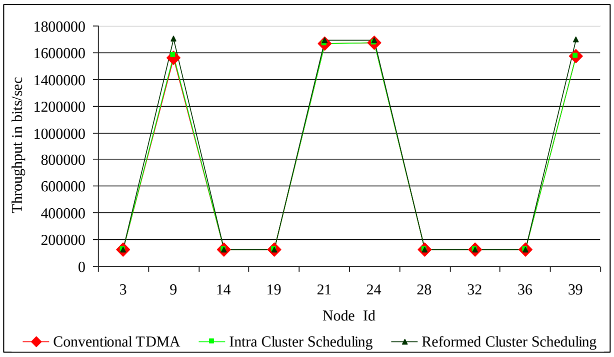


Fig .9.Throughput Analysis for Random Topology

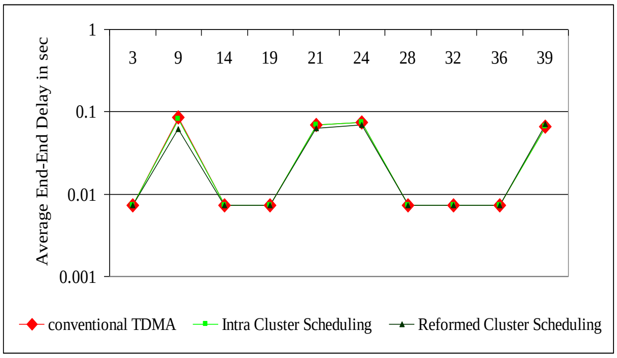


Fig .10.Average End-End Delay Analysis for Random Topology

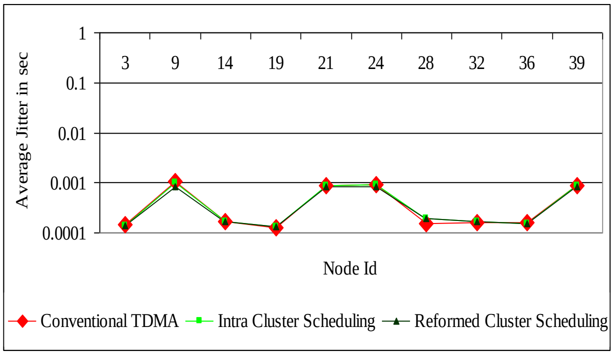


Fig .11.Average Jitter Analysis for Random Topology

V. Conclusion

We in this paper present analysis on performance of coexistence heterogeneous networks. In this paper, we propose a new power based scheme using inter & reformed cluster scheduling mechanism for the coexistence of multiple IEEE 802.15.4 LRWPAN and IEEE 802.11b WLAN.The performance metrics of IEEE 802.15.4 network such as bit error rate ,throughput, average end-end delay and average jitter is analyzed when the nodes are static. The simulation results show that the proposed scheme is effective in performance improvement for coexistence network of IEEE 802.15.4 for random topology. In future the analysis can be extended with mobility model and the same proposed scheme can be implemented with Exata emulator.

VI. References

1) M. Petrova, *et al*, “IEEE 802.15.4 Low Rate - Wireless Personal Area Network Coexistence Issues,” *Proc. IEEE WCNC’06*, Las Vegas, USA
2) I. Howitt and J. A. Gutierrez, “Low-Rate Wireless Personal Area Networks - Enabling Wireless Sensors with IEEE 802.15.4,” *Proc. IEEE WCNC’03*, vol.3, pp. 1481-1486
3) S. Shin, *et al*, “Packet error rate analysis of IEEE IEEE 802.15.4 under IEEE 802.11b interference,” *Proc. WWIC’05*, pp. 279-288
4) A. Sikora, “Coexistence of IEEE 802.15.4 (ZigBee) with IEEE 802.11 (WLAN), Bluetooth, and Microwave Ovens in 2.4 GHz ISM-Band,” *web document*, http://www.ba-loerrach.de/stzedn/
5) Ling-Jyh Chen, Tony Sun, Mario Gerla,” Modeling Channel Conflict Probabilities between IEEE 802.15 based Wireless Personal Area Networks”, Communications, 2006. ICC apos; 06. IEEE International Conference on Vol.1, Issue, June 2006 Page(s):343 - 348
6) Soo Young Shin, Hong Seong Parky, Sunghyun Choi, Wook Hyun Kwon,” Packet Error Rate Analysis of IEEE 802.15.4 under IEEE 802.11b Interference”, IEICE Transactions on Communications 2007
7) Dae Gil Yoon, Soo Young Shin, Wook Hyun Kwon and Hong Seong Park,” Packet Error Rate Analysis of IEEE 802.11b under IEEE 802.15.4 Interference” **Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd** Publication Date: 7-10 May 2006 Volume: 3, On page(s): 1186-1190
8) G.M.Tamilselvan, Dr.A.Shanmugam, “Modeling Channel Conflict Probabilities and Interference analysis of Coexistent Heterogeneous Networks” International Journal of Computer science and Knowledge Engineering (IJCSKE), Vol 3.Jan-June 2009, on page(s): 113-118
9) G.M.Tamilselvan,Dr.A.Shanmugam, “Probability Analysis of Channel Collision between IEEE” International Journal of Computer Theory and Engineering(IJCTE), Vol. 1, No.1, April 2009,On page(s):59-64
10) G.M.Tamilselvan, Dr.A.Shanmugam, “Effect of Inter Packet Delay in Performance Analysis of Coexistence Heterogeneous Wireless Packet Networks”, International Journal of Network Security and Applications (IJNSA), Vol 1, No 2, July 2009, On Page(s):40-49

.

A Survey of MANET Routing Protocols in Large-Scale and Ordinary Networks

Hossein Ashtiani, Hamed Moradi Pour, Mohsen Nikpour

GJCST Classification
C.2.2, C.2.1

***Abstract*-An ad hoc network (MANET) consists of mobile nodes that communicate with each other. Routing in ad hoc network is a challenging task because nodes are mobile. Efficient routing protocols have better performance in such networks. Many protocols have been proposed for ad hoc networks such as: Ad hoc on-demand Distance Vector (AODV), Optimized Link State Routing (OLSR), Dynamic Source Routing (DSR), and Geographic routing protocol (GRP). these approaches have not been evaluated for the same conditions in pervious researches. But in this study, the performance of these protocols is evaluated in various network conditions and with different packet size patterns. Also, different MAC layers like 802.11b, 802.11g in ordinary and large-scale networks are considered. For the evaluation, Different metrics like packet delivery ratio, end-to-end delay, Mac delay and Routing traffic received/sent, are applied. All simulations have been done using OPNET.**

I. Introduction

Manets consist of mobile nodes that communicate with each other without any infrastructure and are named as infrastructure-less networks [1]. Nodes in these networks carry out both network control and routing duties; they generate user and application traffics. Routing in ad hoc networks is much difficult because topology of such networks is dynamic. Normal routing protocols which are used in wired networks are not efficient, so, in the past years, many protocols have been designed for ad hoc networks.Routing protocols are divided into four categories: proactive, reactive, hierarchical and geographic routing protocols. The most popular ones are AODV, DSR (reactive), OLSR (proactive) and GRP (geographic). Reactive protocols like DSR, and AODV find the routes only when requested and data need to be transmitted by the source host; These protocols generate low traffic and routing overhead but this will increase delay and are suitable for energy-constrained conditions. They use distance-vector routing algorithms. Proactive protocols like OLSR are table driven protocols and use link state routing algorithms. Proactive protocols generate high traffic and routing overhead but have less delay and can be used when bandwidth and energy resources are enough [2]. Geographic routing protocols use the node position (i.e., geographic coordinates) for data forwarding. A node forwards a packet with considering its neighbors and the destination physical positions. In these protocols packets are sent to the known geographic coordinates of the destination nodes [2]. Some studies considered the evaluation of these protocols, but a little attention have been focused on evaluation and comparison of geographic routing protocols with those three protocols in ordinary and large-scale networks. Data packet size is assumed to be constant in most of the papers, e.g. 512 bytes, [3],[4]. Also, the performance of a network with two constant packet size is considered in others [5]. Evaluation and comparison the performance of ad hoc network simultaneously has not been experienced yet in the two following cases: 1- various size of data packets (uniform distribution) vs. constant size of data packets 2-different MAC layers. An ad hoc network may apply data packet with various size. In this paper, we use OPNET to evaluate these protocols with different packet size and different CBR source-destination pairs. We also evaluate them in large-scale networks; large-scale networks show different behavior in comparison with ordinary networks due to large number of connections and long paths. This paper also compares the use of different MAC layer technologies in large-scale networks. The 802.11 standard was not designed for the multi-hop ad hoc networks but because of widespread availability of 802.11 cards, this technology is the most used one in the MANETs[2]. But using this technology cause several limitation in ad hoc networks. Enhancement in the MAC layer technology (like, 802.11g, the use of OFDM[6], multi-antenna platforms, etc.) can cause these networks to perform better [2]. In this paper, a comprehensive comparison using 802.11b, 802.11g [7] in large-scale ad hoc networks is considered with different scenarios.The rest of the paper is organized as follows. Section 2 is a description of common routing algorithm using for performance comparison. In Section 3 a review of previous literature carried out in this field is provided. In Section 4 we present the scenarios used for comparison. Section 5 describes metrics used in this paper. Section 6 and 7 present the simulation results for ordinary and large-scale network respectively. And finally we provide conclusion in section 8.

II. Description of Routing Protocols

1) *OLSR*

Optimized Link State Protocol (OLSR) is a proactive protocol, so due to it’s proactive nature the routes are always available when they are needed [8]. OLSR uses hop by hop routing. It uses MPR (Multi Point Relays) flooding mechanism to broadcast and flood Topology Control (TC) messages in the network. This mechanism takes advantage of controlled flooding by allowing only selected nodes

About- Electrical Department Islamic Azad University- Babol Branch, Babol, Iran hashtiani79@yahoo.com h_moradi25@yahoo.com , mhsnnikpour@yahoo.com

(MPR nodes) to flood the TC message. Each node selects an MPR to reach its two-hop neighbors. OLSR uses topology discovery/diffusion mechanism by periodic and triggered Topology Control (TC) messages. TC messages are generated by MPR nodes and carry information about MPR selectors nodes. Neighbor sensing is done by using periodic broadcast of Hello messages. These messages are one-hop broadcasts (never forwarded) that carry neighbor type and neighbor quality information.

2) *AODV*

AODV is a reactive protocol that reduces number of broadcasts by establishing routes on demand basis. This protocol does not maintain the whole routing information of all nodes in the network [9]. For Route Discovery a route request packet (RREQ) is broadcasted whenever a node have a packet to transmit to the destination. It continues forwarding till an intermediate node which has recent route information about destination or the destination itself receive this packet. Then the intermediate node or the destination will send a Route Reply (RREP) message to the source by reverse path of RREQ, therefore AODV uses symmetric link. During forwarding a packet a node records in its tables from which the first copy of the request came. It is needed for establishing reverse path for RREP message. The intermediate nodes are allowed to inform the effected sources from link breakage. Link failure can be due to node’s movement or exhausting it’s energy. When source node receive the Route Error packet (RERR) packet, it can initiate route again if still needed. To prevent route loops, AODV uses sequence number maintained at each destination to determine how much fresh the routing information is [9]. The sequence numbers are carried by all routing packets[10]. Hello messages are responsible for the route maintenance.

3) *DSR*

DSR is another reactive protocol. The main feature of DSR is source routing. DSR is specially designed for multi-hop ad hoc networks and reduces bandwidth usage by eliminating periodic messages. In this protocol the packet includes a complete list of the all nodes which it should be forwarded towards them. DSR has two main mechanisms: “Route Discovery”, “Route Maintenance”[11]. During Route Discovery, a source node broadcasts a RREQ message; and each intermediate node that receives this packet will rebroadcast it, unless it is the destination or it has route to the destination in its route cache. Such a node will send a RREP message to the source [10]. If link failure occur then a route error packet (RERR) will be sent to the source to notify it. The source node then removes that routes consisting failed link from its cache and if there is a new route to that destination in its cache, it will replace it instead of previous one; otherwise it will reinitiate route discovery. Both Route discovery and Route maintenance are on demand. Unidirectional link and asymmetric routes are supported by DSR.[12].

4) *GRP*

A node maintains its list of neighbor nodes by periodically broadcasting Hello messages. If a node does not receive a Hello message from a neighboring node for a period exceeding the specified "Neighbor Expiry Time," it assumes the link to the neighbor is lost. Each node can determine its own position using a GPS. The position of other nodes determined through flooding. When a node moves more than a specified distance, it sends out a flooding message with its newposition. To bootstrap the network, all nodes initiate a full flooding throughout the network. To reduce the overhead caused by flooding updates, the scope of the flooding is limited. This is known as fuzzy routing. In fuzzy routing, when a node sends a position update, only nodes that “need to know” about the change receive the flood.

III. Previous Work

In [13] four different routing protocols like AODV, DSR, DSDV and TORA were compared with each other. It was shown that DSR has better performance due to aggressive use of source cache and maintain multiple routes to the destination. Moreover ,AODV suffers end-to-end delay and TORA generates high routing traffic. In [14] the authors shown that in normal cases AODV has better performance than DSR. But in constrained situation of several CBR traffic sources leading to same destination, DSR outperforms AODV, where the degradation is as severe as (30%) in AODV whereas DSR degrades marginally as 10%. Perkins et all [10] shown that DSR outperforms AODV when smaller number of nodes and lower load/mobility are used in network, however in high mobility or more load and more nodes, AODV has better performance than DSR. It was also illustrated that DSR has poor delay and throughput performances due to aggressive use of caching. Authors in [15] evaluated three routing algorithms like DSR, AODV and FSR. They represented that in city traffic scenarios AODV has a better performance than DSR and proactive protocol FSR. A limited study with considering QOS was conducted in [16] and illustrated that DSR outperforms in packet delivery fraction and routing overhead whereas OLSR shows the lowest end-to-end delay at lower network loads. [17] also discussed proactive and reactive protocols in more realistic environments and illustrated that AODV has better performance than DSR and DSDV. Mbarashimana et all [18] by using OPNET simulator shown that OLSR gets better performance than DSR and AODV. Which is different from what authors shown in [19] and [20]. G.Jayakumar et all [3] compared DSR and AODV with different CBR sources. They shown both DSR and AODV perform better under high mobility simulations. Authors of [12] compared AODV, DSR, OLSR and DSDV for variable bit rate (VBR) and shown that reactive protocols have better performance than proactive protocols. They also illustrated that DSR performs well for the performance parameters namely delivery ratio and routing overhead while AODV perform better in terms of average delay. Authors of [6] and [21] shown that AODV performs better in the networks with static traffic, with the number of source and destination pairs is relatively small for each host and OLSR protocol is more efficient in networks with high density and highly sporadic traffic. Authors of [22] evaluated AODV, DSR, LAR and TORA protocols and compared AODV, DSR and Location-Aided routing (LAR) over a large geographic area. They shown, AODV suffers in terms of packet delivery fraction but scales very well in terms of end-to-end delay; also DSR scales well in terms of packet delivery fraction but suffers an important increase of end-to-end delay. Alexander Klein [19] compared AODV and OLSR and statistic-based routing protocol (SBR) with different traffic patterns and compared them with respect to reliability and routing overhead.

IV. Simulation Environment

1) *Simulation Setup*

The ad hoc networks are implemented using OPNET simulator [23]. For having a comprehensive evaluation of these four protocols in ordinary and also large-scale networks, we use three scenarios that each of them has different geographic size. In each scenario we examined and compared two cases of application layer packets for two MAC layer protocols: 802.11b, 802.11g. The used parameters for simulation are listed in tables 1.

V. Metric

For evaluating these routing protocols we use four different metrics such as End-to-End delay,Packet delivery fraction, Media access delay and Link layer retransmission.

1) *End-to-End Delay (second)*

The End-to-End delay is the time between when the source generates the data packet to when the destination receives it.

Scenario Num.	1, 2	3
Simulation time	600 seconds	150 seconds
Area	1500×500 m²(scenario1)	-- -- -- -- --
Area	1800×800 m²(scenario2)	3000×3000 m²
Node placement	Random	Random
Mobility pattern	Random way point	Random way point
Speed	Uniform(0-10) m/s	Uniform(0-10) m/s
Pause time	0, 50, 100, 200, 300,400,500, 600 (scenario1)	0, 50
Pause time	0,50,100,200, 300,400 (scenario2)	-- -- - -- - -
Application	CBR	CBR
Packet size	1024 bytes, uniform distribution (256,512) bytes	1024 bytes, uniform distribution (256,512) bytes
Packet transmission rate	3 packet/sec	3 packet/sec
Data rate	2 Mbps	2 Mbps
MAC	802.11b, 802.11g	802.11b, 802.11g
#of connections	10, 30	10
Num of nodes	50(scenario1)	100, 150, 200, 250, 300, 350, 400, 450, 500
Num of nodes	100(scenario2)	-- -- --- ----

Table 1: Simulation parameters scenario 1, 2 and 3

2) *Packet Delivery Fraction*

The Packet delivery fraction is the ratio of number of packets successfully delivers to/received by destination to those originated by the sources

3) *Media Access Delay (MAC delay)(second)*

It is providing the result for a received packet with a routing Address Resolution Protocol (ARP) and control packet reply transmitted by MAC layer. For each frame this delay is calculated as the duration from the time it is inserted into the transmission queue, which is arrival time for higher layer data packets and creation time for all other frames types, until the time when the frame is sent to the physical layer for the first time. MAC delay is very useful metric to identify congestion hot spots and measure link interference in ad hoc network [20]. It can be used to improve network throughput in multi-rate networks.

4) *Link layer Retransmission*

Total number of transmission attempts by link layer in the network until either packet is successfully transmitted or discarded as a result of reaching retry limit. Because of large-scale fading [24], signal power attenuation and path loss are observed due to radio propagation over long distance, which itself can cause link layer retransmissions. Number of retransmission in MAC layer can affect loss rate. Retransmission can be due to congestion or internal collisions in the network. So this metric can be a useful metric for these phenomenons.

VI. Simulation Result (Ordinary Network)

1) *End-to-End Delay*

The end-to-end delay parameters are simulated here for 50 mobile nodes with different packet sizes and two different CBR sources as shown in figures (1,2 , 3, 4):

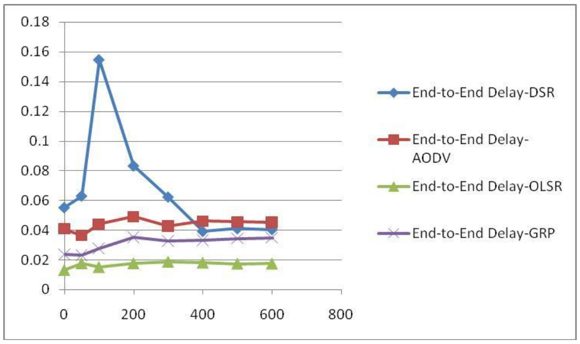
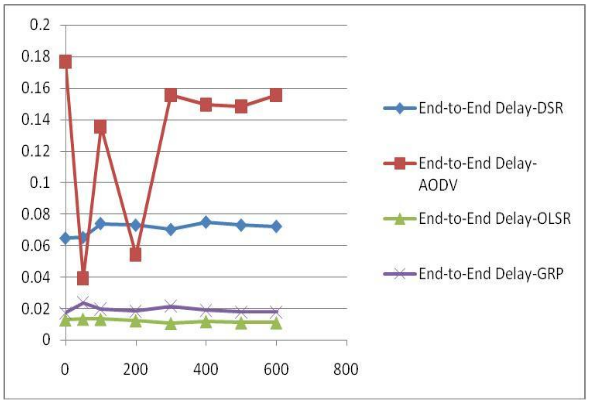


Fig. 1 packet size_1024-10



source **Fig. 2** packet size_uniform-10 source

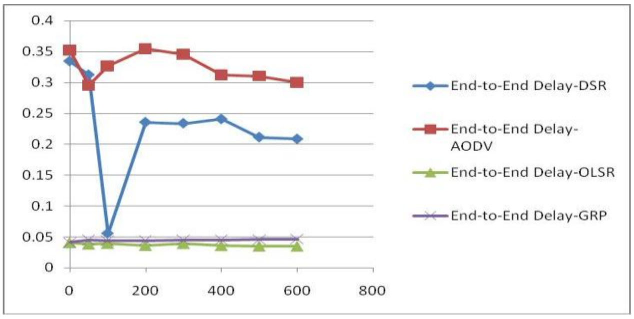


Fig. 3 packet size_1024-30 sources

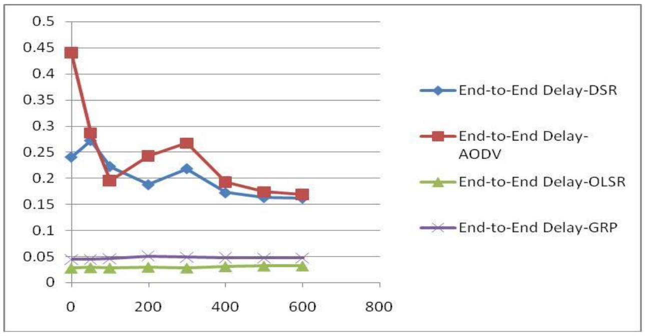


Fig. 4 packet size_uniform-30 sources

comparing figures 1 and 2, we can see that for 10 sources, if the packet size be in uniform pattern, DSR becomes better with respect to average end-to-end delay for pause time 0 to 200 and has less changes with increasing the pause time and it is more linear, however, its delay becomes slightly worse after pause time 400, but it is not significant ; But delay in AODV increases and loses its linearity. Also in GRP and OLSR delays become better slightly.For 30 sources in uniform packets, DSR and AODV delays become better for almost pause time 0 to pause time 600(except pause time 100). But for AODV in pause time 0 the delay increases about 0.1 seconds. Unlike the previous one in 10 sources, here, the delay for GRP increases slightly but for OLSR it decreases. For 100 nodes with 10 and 30 sources, delays are shown in figures (5, 6, 7, 8):

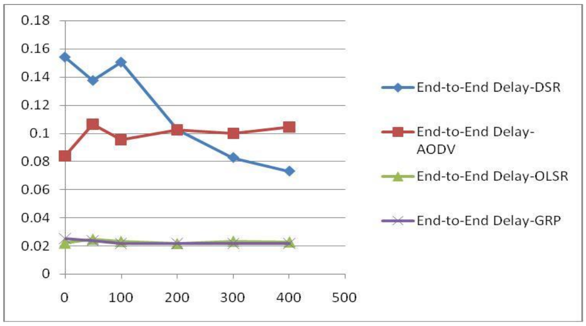


Fig. 5 packet size_1024-10 sources

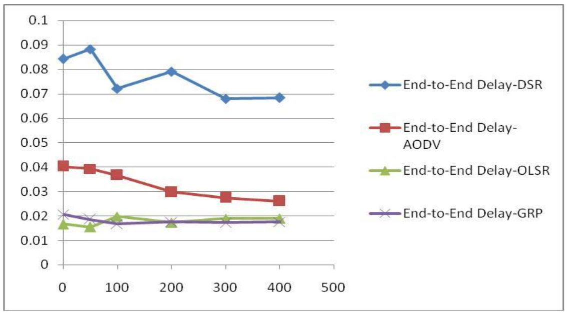


Fig. 6 packet size_uniform-10 sources

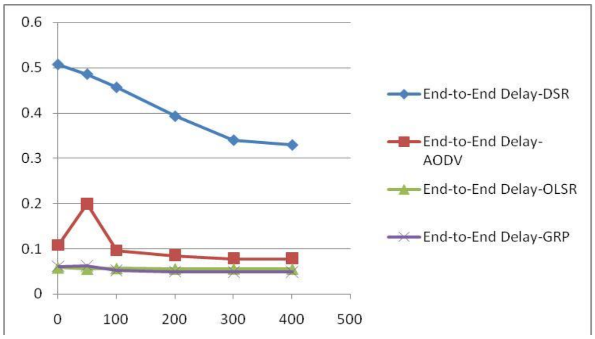


Fig. 7 packet size_1024-30 sources

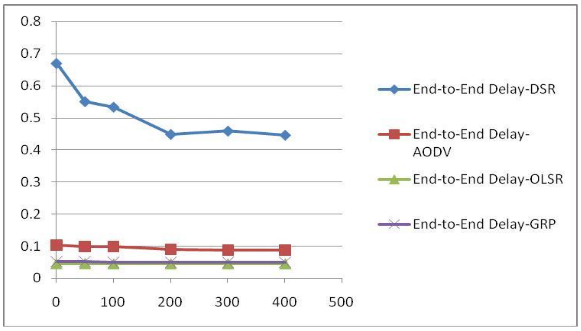


Fig. 8 packet size_uniform-30 sources

It can be seen from figure 6 that for 10 sources DSR delay becomes slightly better and it is more linear, it means that the delay in high mobility (lower pause time) and low mobility (higher pause time) is less than when the packet size is 1024 bytes. With increasing pause time, delay also becomes better in AODV about 0.04 to 0.08 sec. GRP and OLSR delays becomes slightly better but GRP unlike the previous graph has lesser delay than OLSR for pause time 100 to 400.For 100 nodes with 30 sources, for DSR, if we compare each delay in each pause time with the one when the packet size is 1024 bytes, we can see that it increases about 0.1 for every pause time. AODV delay almost does not changes. For OLSR and GRP it does not changes too.From all of these graphs it can be understood that end-to-end delays in OLSR and GRP is much better than the two reactive ones (DSR, AODV).

2) *Packet Delivery Fraction*

When 10 sources exist, in each cases (uniform packet size or 1024 bytes) reactive protocols have better performances than GRP and OLSR. This difference is about 2-3% between reactive and proactive protocols. For 30 source, we can also see that reactive protocols outperforms the others but difference between OLSR and ADV becomes lesser than previous condition (10 sources) and in some pause times their graphs overlapped. But we can see that GRP has the worst performance between these three protocols. However, it should be remind that as we said in previous section GRP has better delay than reactive protocols.

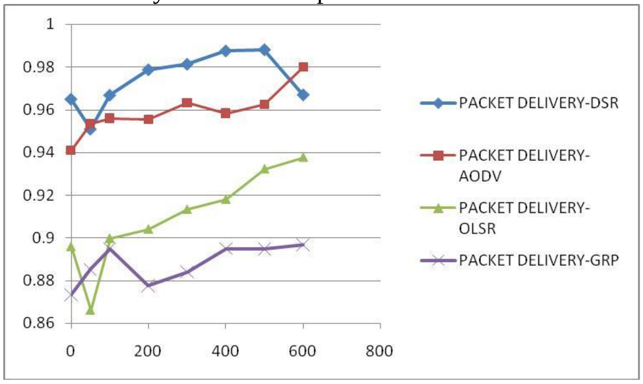


Fig. 9 packet size_1024-10 sources

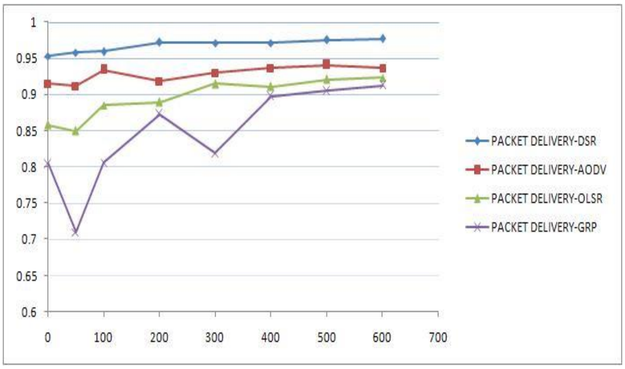


Fig. 10 packet size_uniform-10 sources

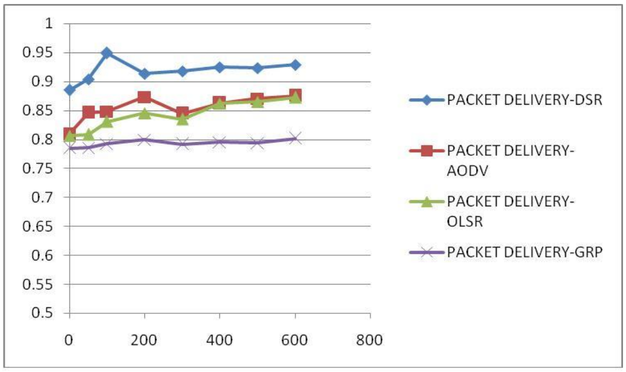


Fig. 11 packet size_1024-30 sources

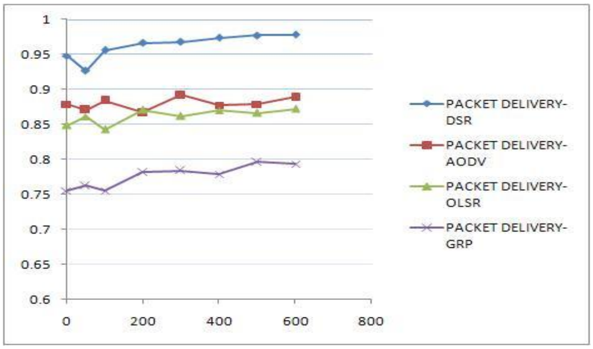


Fig. 12 packet size_uniform-30 sources

For 100 nodes:

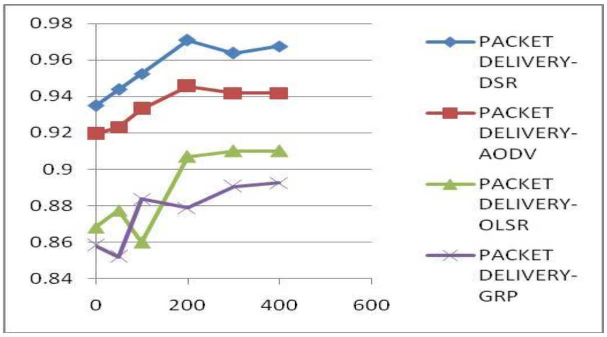


Fig. 13 packet size_1024-10 sources

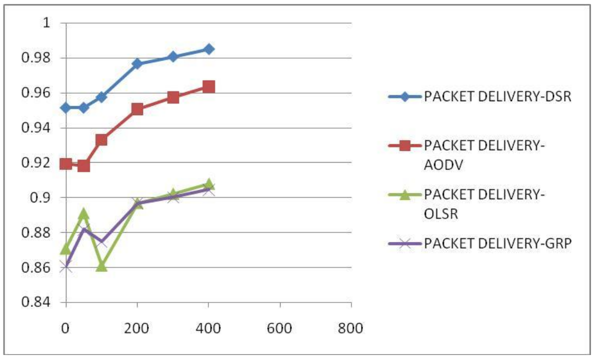


Fig. 14 packet size_uniform-10 sources

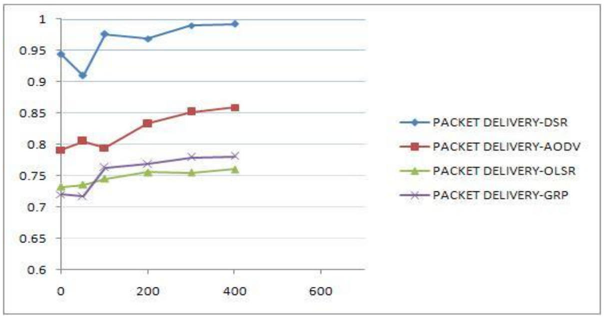


Fig. 15 packet size_1024-30 sources

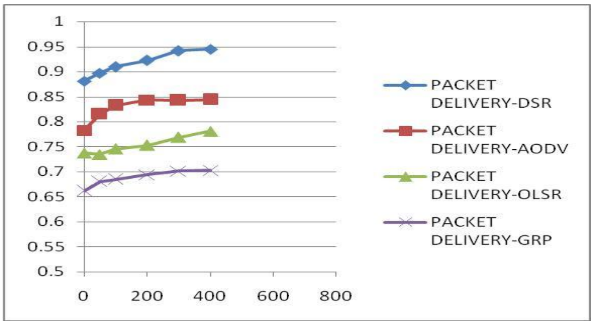


Fig. 16 packet size_uniform-30 sources

From the figures 13 to 16, the packet delivery ratio in reactive protocols (AODV, DSR) is better than OLSR and GRP. There is a little difference between the performances of OLSR and GRP; when the packet size is 1024 bytes, the difference between the performance of OLSR and GRP is about 2% but in uniform packet size(10 sources) these two graphs almost overlapped after pause time 200. Pay attention to the delivery ratios of OLSR and GRP. We can see that GRP performance is worse than OLSR performance in uniform packet sizes but in figure 15, GRP outperforms than OLSR after pause time 100. We can also notify this point that OLSR and GRP performances decreases more than the performances of reactive protocols with respect to increase of CBR connections. As the packet delivery ratio of OLSR protocol decreases about 15-20% and also the performance of GRP decreases about 10-20% for two different packet sizes. The packet delivery ratio in reactive protocols are more stable with respect to number of connections than OLSR and GRP.

3) *Media Access Delay*

MAC delay is an efficient and useful metric for measuring link interference in ad hoc networks and it can be used to improve network throughput in multi-rate networks. So evaluating this metric becomes more important. As it is shown with changes in packet size there is no changes in sequence of graphs in figure (17 to 20). As an other result from this figures, it is clearly shown that DSR has the worst MAC delay between these protocols, vise versa, OLSR has the best MAC delay or indeed, delay resulting from accessing the media during the data communication in OLSR is much lower than other three ones.

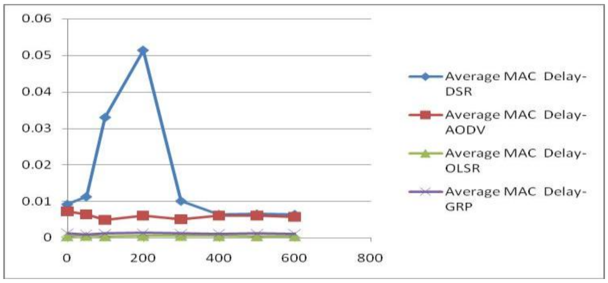


Fig. 17 packet size_1024-10 sources

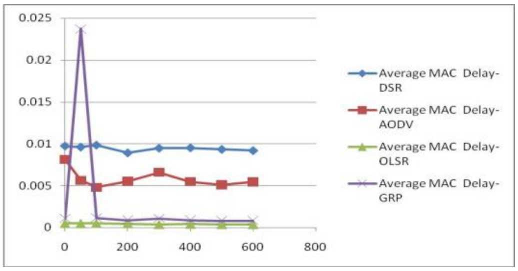


Fig. 18 packet size_uniform-10 sources

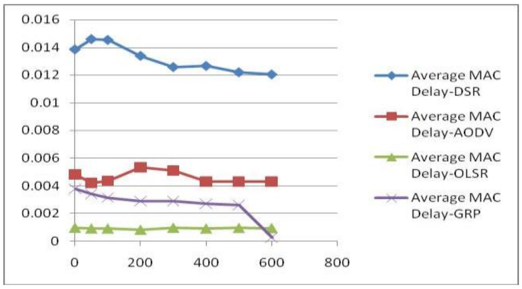


Fig. 19 packet size_1024-30 sources

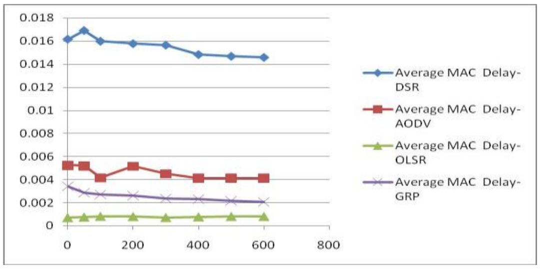


Fig. 20 packet size_uniform-30 sources

For 100 nodes (figures 21 to 24) MAC delay of GRP decreases with increasing in pause time (lower mobility) and its graphs has less changes than OLSR, DSR and AODV. Here, DSR also has the worse and OLSR has the best MAC delay. So we can say that proactive protocols like OLSR has better MAC delay than reactive protocols.

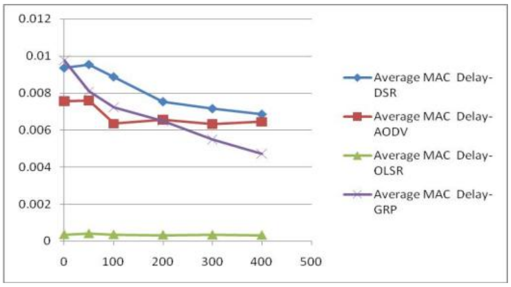


Fig. 21 packet size_1024-10 sources

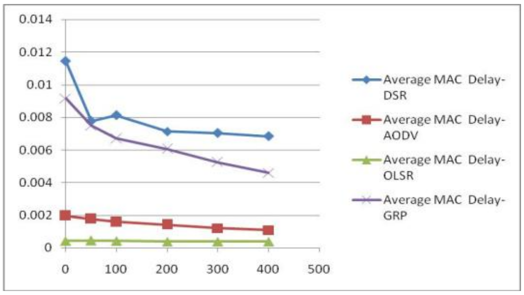


Fig. 22 packet size_uniform-10 sources

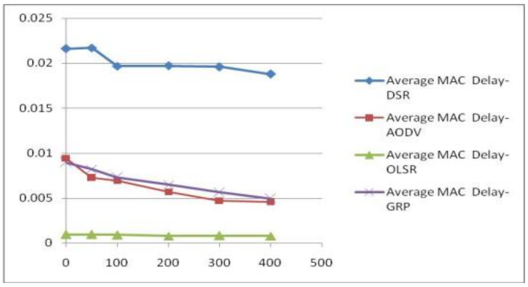


Fig. 23 packet size_1024-30 sources

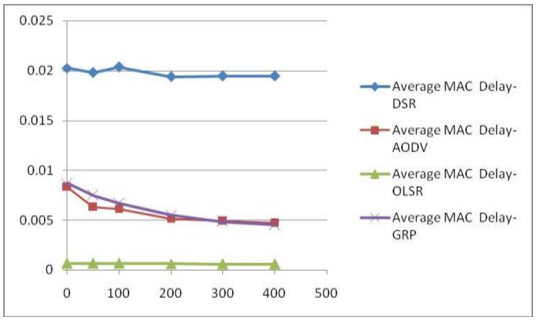


Fig. 24 packet size_uniform-30 sources

4) *Link layer Retransmission*

Total number of transmission attempts by link layer in the network until either packet is successfully transmitted or discarded as a result of reaching retry limit. Retransmission can be due to congestion in the network. Figures (25 to 28) show this metric for 50 nodes.

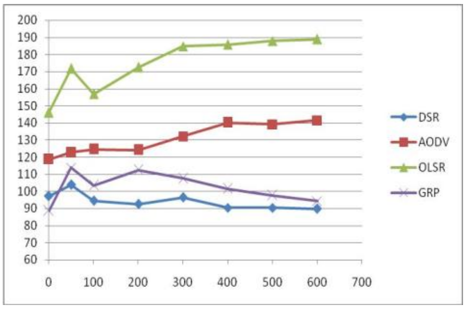


Fig. 25 packet size_1024-10 sources

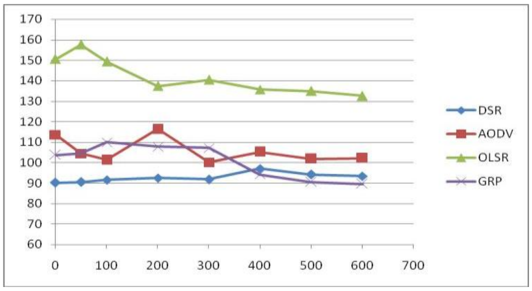


Fig. 26 packet size_uniform-10 sources

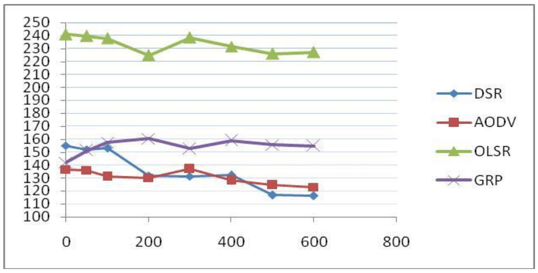


Fig.27 packet size_1024-30 sources

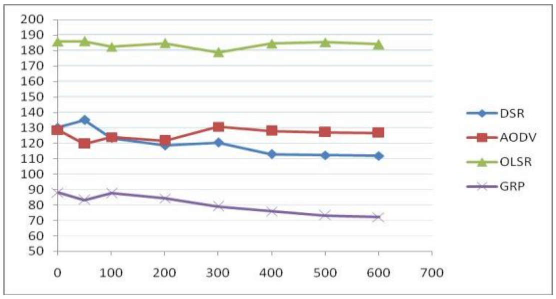


Fig. 28 packet size_uniform-30 sources

Form figures(29 to 32) we can see that OLSR has the most retransmission amount between these four protocols. It may indicate that by using OLSR, congestion increases in network due to high load. Regardless of OLSR, Notify to figures 39 and 40, GRP has the most and the least retransmission amount. So the use of different packet size is more clear in figures 39 and 40.

For 100 nodes:

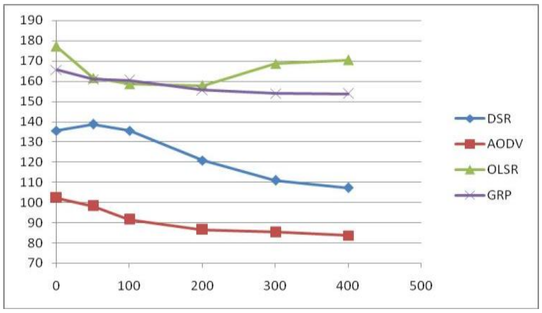


Fig. 29 packet size_1024-10 sources

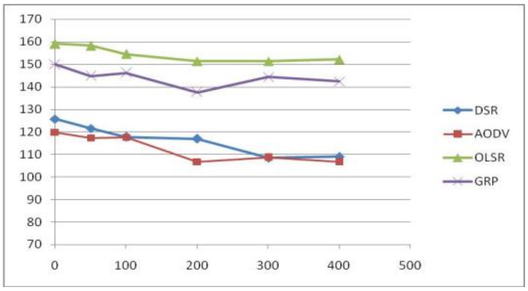


Fig. 30 packet size_uniform-10 sources

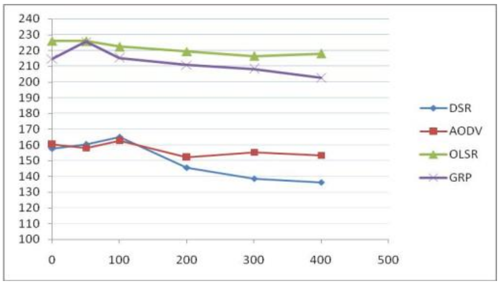


Fig. 31 packet size_1024-30 sources

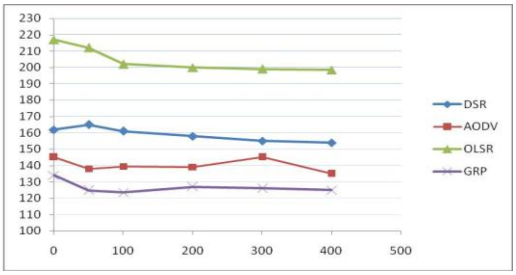


Fig. 32 packet size_uniform-30 sources

VII. Large Scale Network

It is essential to evaluate Large scale ad hoc networks, due to having large number of nodes and much more complexity, with a view to scalability and performance in different conditions like mobility, number of nodes, MAC layer, packet size,… when using different routing protocols. In large scale ad hoc networks because the number of nodes is large and the distance between source and destination may be far, so, routes with large number of hops can be established which itself can result more errors. By evaluating and comparing these four protocols here, we can find three point: 1- if e.g. DSR,… is efficient enough for network application in a special condition 2- which of them has better performance in that case.Our goal in this section is to compare these routing protocols in networks having different number of nodes and topology with respect to the packet size (1024 bytes and uniform packet size), pause time (0,50s) and (802.11b and 802.11g) PHY DSSS (Direct Sequence Spread Spectrum). Behaviour of these protocols has not yet been evaluated for large scale networks with respect to these metrics and different MAC layers

1) *End-to-End Delay*

As we can see from figures(33 to 40), DSR has the largest delay between these four protocols. This is due to aggressive use of caching in DSR. Because number of nodes is large and routes has much more hops than routs in ordinary network, so, the cache size increases which itself can increase delay to choose stale roots. Because OLSR always has routes available due to its proactive nature so it has the best delay.

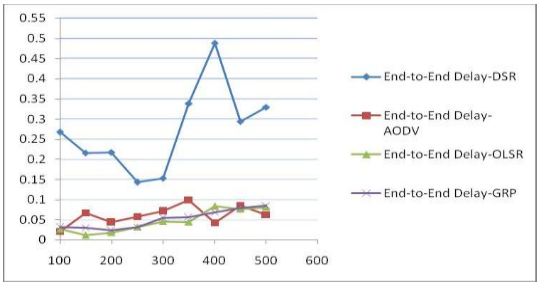


Fig. 33 pt z[1]_1024 -802.11b-p.time[2] 0

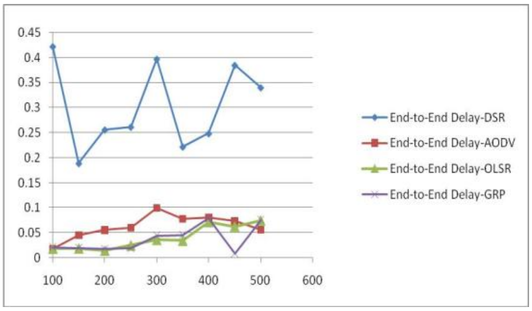


Fig. 34 pt z_uniform-802.11b-p. time 0

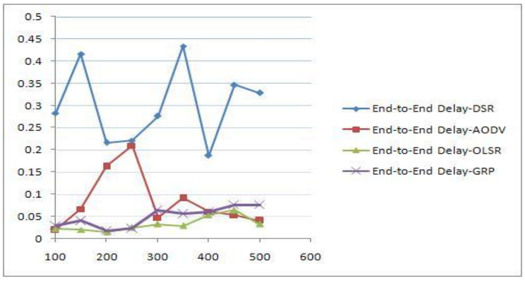


Fig. 35 pt z_1024 -802.11g-p.time 0

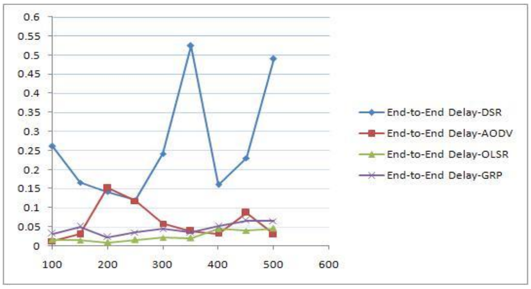


Fig. 36 pt z_uniform-802.11g-p. time 0

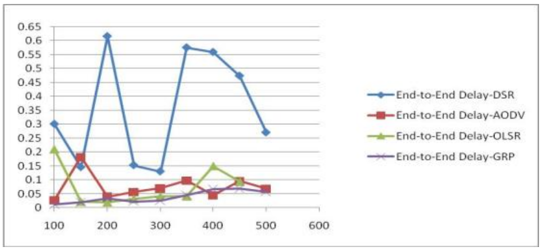


Fig. 37 pt z_1024 -802.11b-p.time50

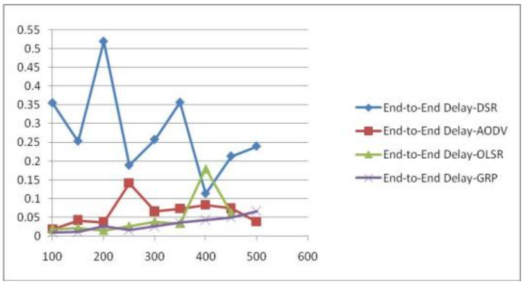


Fig. 38 pt z_uniform-802.11b-p. time 50

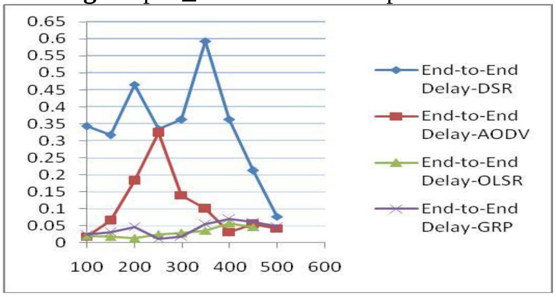


Fig. 39 pt z_1024 -802.11g-p.time 50

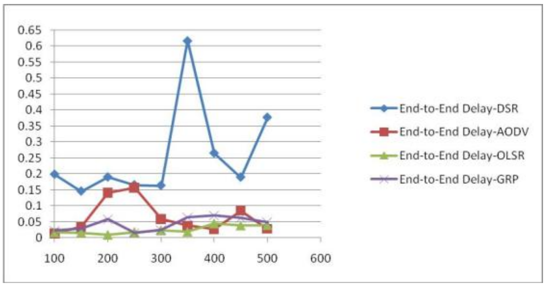


Fig. 40 pt z_uniform-802.11g-p. time50

[1] Packet size

[2] Pause time

Notifying figures(33 to 40) when using 802.11b and 802.11g we can see a lot of changes in delay that is so much non-linear, it means it does not increases or decreases linearly. For GRP, it has lower end-to-end delay than reactive protocols like 50 nodes. End-to-end of AODV is less than DSR, because, in DSR the length of a route is a main criterion and it chooses a route between several routes which found by routing discovery process or stored in a node's cache but AODV selects routes having the least congestion due to respond to the first RREQ, also, ignores the length of a route.It should be said that packet delivery and delay also depends on node density in network. Here, we consider networks which have different node density and different number of nodes that varies from 100-500 (as shown in table 3). we compare networks which has same number of nodes and topology and situation of these networks are the same except the pause time and MAC layer which we name below the figures.

2) *Packet Delivery Fraction*

Reactive protocols (AODV, DSR) has better packet delivery ratio than GRP and OLSR. The packet
delivery of OLSR decreases comparing to when the number of nodes is 50 nodes, because, the number of nodes is large and the routing traffic of OLSR is so high causing the congestion in network increases which itself can result to more errors.

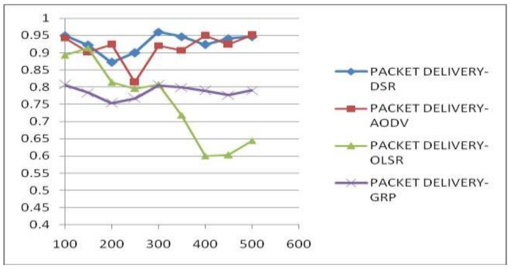


Fig. 41 pt z_1024 -802.11b-p.time 0

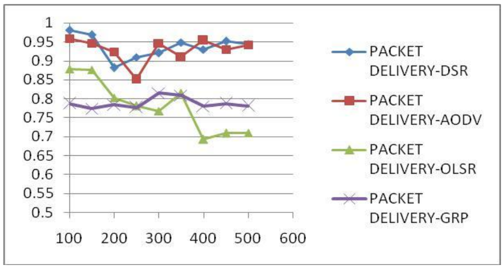


Fig. 42 pt z_uniform-802.11b-p. time 0

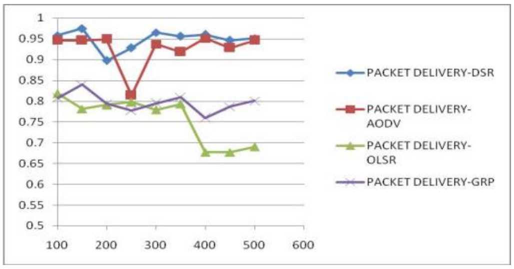


Fig. 43 pt z_1024 -802.11g-p.time 0

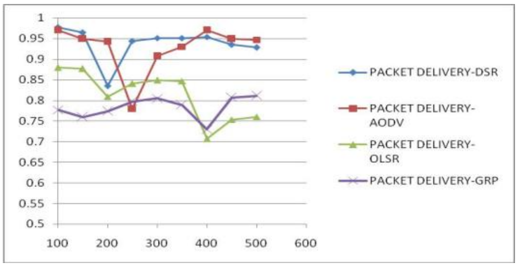


Fig. 44 pt z_uniform-802.11g-p. time 0

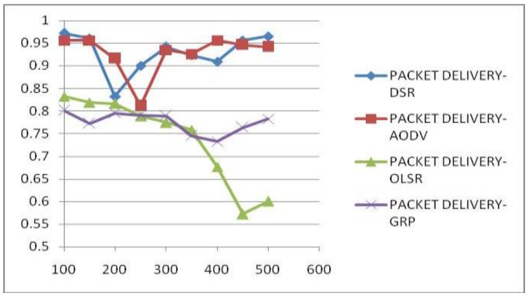


Fig. 45 pt z_1024 -802.11b-p.time50

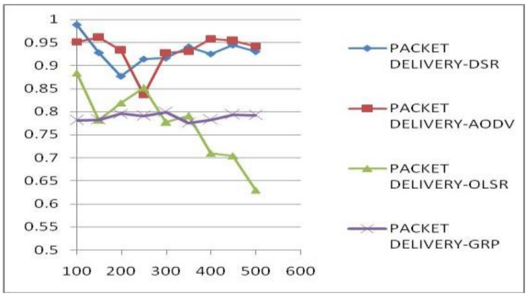


Fig. 46 pt z_uniform-802.11b-p. time 50

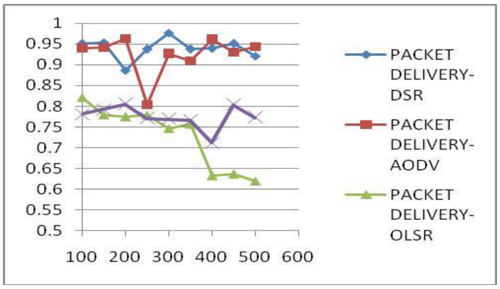


Fig. 47 pt z_1024 -802.11g-p.time 50

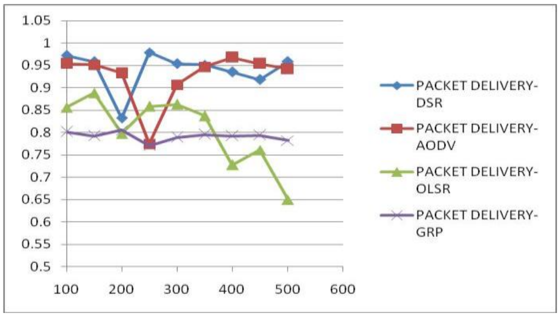


Fig. 48 pt z_uniform-802.11g-p. time50

Also, comparing packet delivery ratio between OLSR and GRP, it can be seen that packet delivery ratio of GRP is more linear and has less changes.

3) *Media Access Delay*

The MAC delay of OLSR is well enough but AODV also has good MAC delay especially when the number of nodes is more than 300.

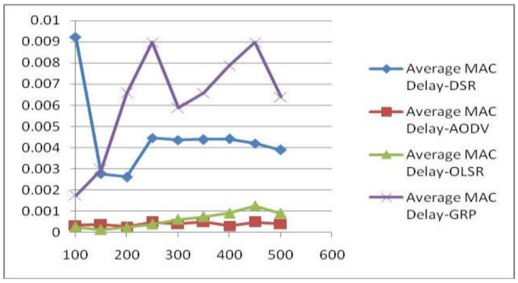


Fig.49 pt z_1024 -802.11b-p.time 0

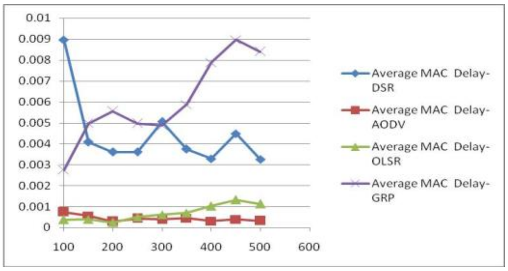


Fig. 50 pt z_uniform-802.11b-p. time 0

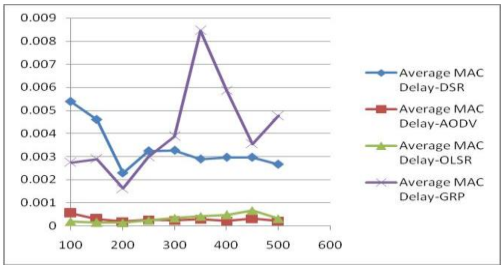


Fig. 51 pt z_1024 -802.11g-p.time 0

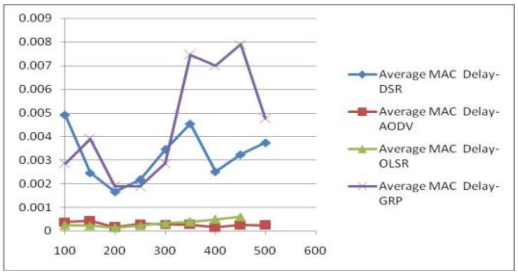


Fig. 52 pt z_uniform-802.11g- p. time 0

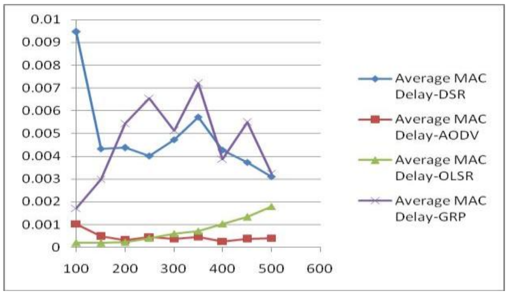


Fig. 53 pt z_1024 -802.11b-p.time50

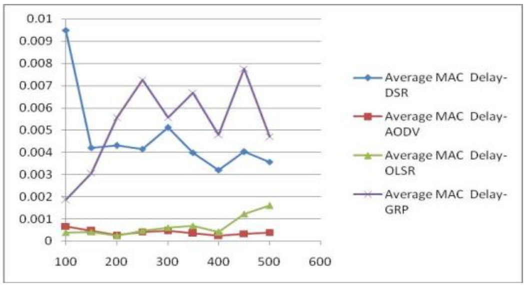


Fig. 54 pt z_uniform-802.11b-p. time 50

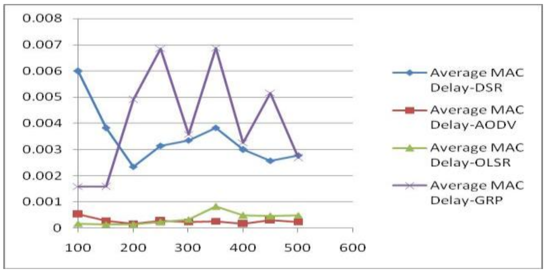


Fig.55 pt z_1024 -802.11g-p.time 50

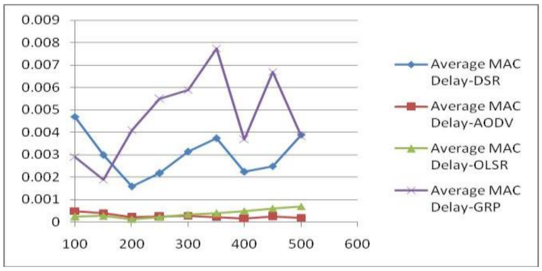


Fig. 56 pt z_uniform-802.11g-p. time50

Comparing MAC delay of one network from figure (57, 59, 61, 63), when we use DSR, the MAC delay of network in 802.11g is less than 802.11b.

4) *Link layer Retransmission*

AODV has the least link layer retransmission and GRP has the most. In large-scale network when we use GRP, the congestion in network increases and its performance become worse. In large-scale network because routes have large number of hops, so, if failure of a link occurs then transmission attempt of DSR increases due to use of source caching.

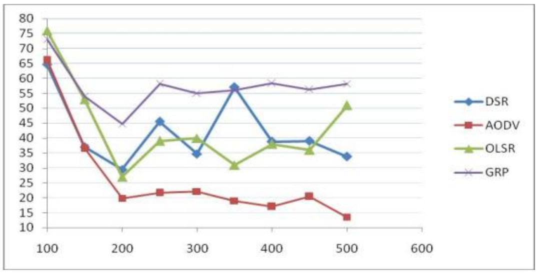


Fig. 57 pt z_1024 -802.11b-p.time 0

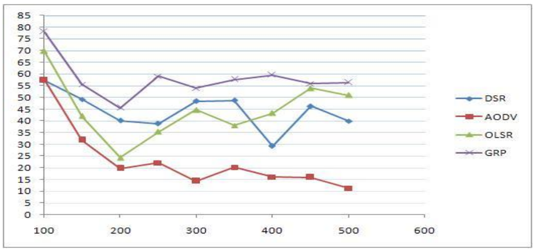


Fig. 58 pt z_uniform-802.11b-p. time 0

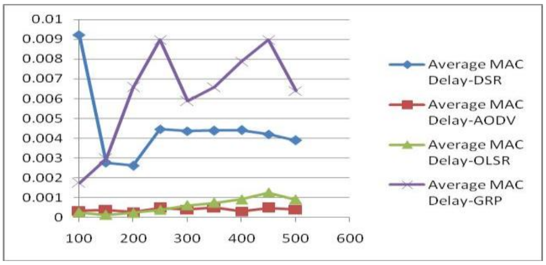


Fig. 59 pt z_1024 -802.11g-p.time 0

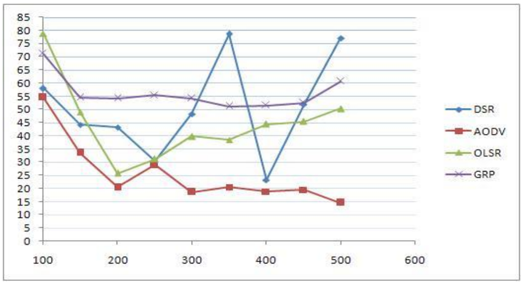


Fig. 60 pt z_uniform-802.11g-p. time 0

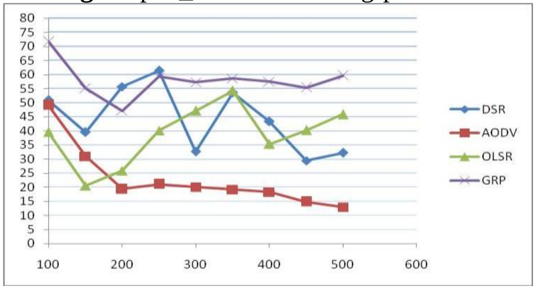


Fig. 61 pt z_1024 -802.11b-p.time50

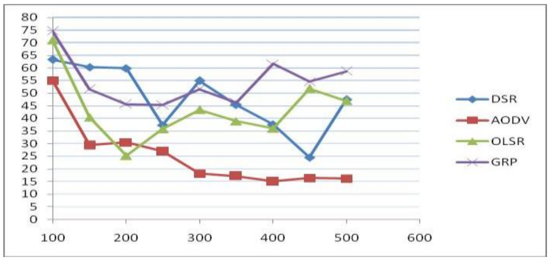


Fig. 62 pt z_uniform-802.11b-p. time 50

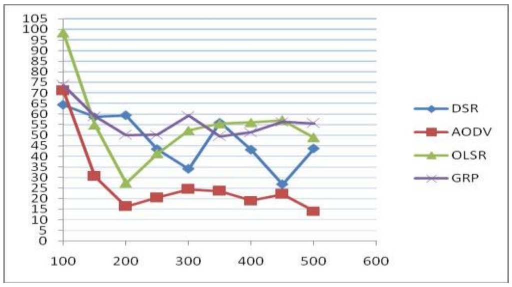


Fig. 63 pt z_1024 -802.11g-p.time 50

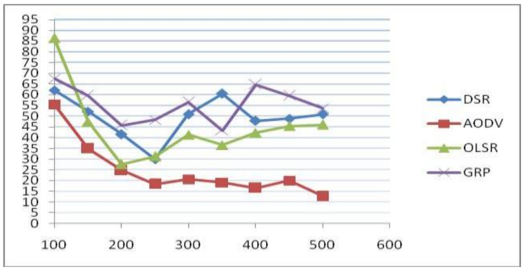


Fig. 64 pt z_uniform-802.11g-p. time 50

VIII. Conclusion

This work is the first attempt towards a comprehensive performance evaluation of four important routing protocols (DSR, AODV, OLSR, GRP) for ordinary and large-scale mobile ad hoc networks. In this paper, using simulation environment (OPNET 14.0) we evaluated the performance of four widely used ad hoc network routing protocols using different packet size patterns (uniform distribution and 1024 bytes) and also, different MAC layer (802.11b, 802.11g) for ordinary and large-scale MANETS. Our work uses four metrics to evaluate the performance of these routing protocols to include additional important performance parameters. For comparative performance analysis, we first simulated each protocol for ad hoc networks with 50 and 100 nodes. In this case OLSR and GRP shows good performance for the End-to-End delay and especially OLSR has the best MAC delay. DSR and AODV outperform OLSR and GRP for packet delivery ratio but reactive protocols show poor performance as compared to OLSR and GRP for the MAC delay. GRP and OLSR performance for End-to-End delay is near together but OLSR. OLSR outperforms GRP for packet delivery ratio, however, GRP has lesser and link layer retransmission.In large-scale network, we evaluate the performance of this protocols for eight different cases (802.11b, 802.11g MAC layer, pause time 0 and 50, different packet size). Our experiment result shows that the MAC layer not only affect the absolute performance of a protocol, but because their impact on different protocols is non-uniform, it can even change the relative ranking among protocols for the same scenario. Detailed characteristics of PHY layer (e.g. Length of signal preamble and header) has a non-negligible effect on the performance of higher layer protocols, and this is true for wireless communication media. From evaluating these four routing protocols, we found that DSR has poor End-to-End delay. The packet delivery ratio of AODV is well enough and also it shows better performance than DSR for End-to-End delay, link layer retransmission and MAC delay. OLSR has the least End-to-end and MAC delay (for most of the time), but its performance for packet delivery ratio decreases more than other protocols with increasing the number of nodes because of more traffic and congestion. GRP has better End-to-End delay than reactive protocols (DSR, AODV).

IX. References

1) IETF Working Group:Mobile Adhoc Networks (manet).http://www.ietf.org/html.charters/manetcharter.html.
2) Marco Conti, Silvia Giordano, ”Multihop Ad Hoc Networking: The Theory”; IEEE Communications Magazine, April 2007.
3) A M,Mohamed O K, “Performance Analysis of MANET Routing Protocols in the Presence of Self-Similar Traffic,” Mohamed O K. Local Computer Networks, Proceedings 31st IEEE Conference, P801-807, 2006.
4) Fu Yongsheng, Wang Xinyu, Li Shanping; “Performance comparison and analysis of routing strategies in Mobile ad hoc networks”; 2008 International Conference on Computer Science and Software Engineering.
5) J.-H.Tarng, B.-W. Chuang and Y.-L. Wen, “A radio-link adaptive routing protocol for mobile ad hoc networks,” The 63rd IEEE Vehicular Technology Conference (VTC ’06-spring), Vol. 2, pp. 678-682, May, 2006.
6) http://en.wikipedia.org/wiki/OFDM.
7) http://en.wikipedia.org/wiki/IEEE_802.11g-2003; http://en.wikipedia.org/wiki/IEEE_802.11b.
8) T.Clausen and P.Jacquet; “Optimized Link State Routing Protocol (OLSR).” RFC 3626, IETF Network Working Group, October 2003.
9) C.Perkins, E.B.Royer and S.Das, “Ad hoc On-Demand Distance Vector(AODV) Routing”, RFC 3561,IETF Network Working Group,July 2003.
10) Charles E.Perkins, Elizabeth M.Royer, Samir R.Das and Mahesh K.Marina, “Performance Comparison of Two On-Demand Routing Protocols for Ad Hoc Networks”; IEEE Personal Communications • February 2001.
11) David B.Johnson David A.Maltz Josh Brooch, “DSR: the Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Networks”, http://www.monarch.cs.cmu.edu/.
12) Arun Kumar B. R.; Lokanatha C. Reddy; Prakash S. Hiremath; “Performance Comparison of Wireless Mobile Ad-Hoc Network Routing Protocols”; IJCSNS International Journal of Computer Science and Network Security, VOL.8 No.6, June 2008.
13) H.Ehsan and Z.A.Uzmi (2004), “Performance Comparison of Ad Hoc Wireless Network Routing Protocols”, IEEE INMIC 2004.
14) R.Misra, C.R. Manda, “Performance Comparison of AODV/DSR On-Demand Routing Protocols for Ad Hoc Networks in Constrained Situation”, IEEE ICPWC 2005.
15) Zhan Huawei; Zhou Yun, “Comparison and Analysis AODV and OLSR Routing Protocols in Ad Hoc Network”.
16) S.Jaap, M.Bechler and L.Wolf, “Evaluation of Routing Protocols for Vehicular Ad Hoc Networks in City Traffic Scenarios,”; International Conference on ITS Telecommunications. France,2005.
17) J.Novatnack, L.Greenwald, H.Arora, “Evaluating Ad hoc Routing Protocols With Respect to Quality of Service,” Wireless and Mobile Computing, Networking and Communications WiMob’2005.
18) C.Mbarushimana and A. Shahrabi; “Comparative Study of Reactive and Proactive Routing Protocols Performance in Mobile Ad Hoc Networks” ; 21st International Conference on Advanced Information Networking and Applications Workshops(AINAW'07) 2007 IEEE.
19) Geetha Jayakumar, Gopinath Ganapathy; “Performance Comparison of Mobile Ad-hoc Network Routing Protocol “;IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.11, November 2007.

20) Alexander Klein ISWPC 2008; “Performance Comparison and Evaluation of AODV, OLSR, and SBR in Mobile Ad-Hoc Networks”.

21) S.Gowrishankar,T.G.Basavaraju, M.Singh, Subir Kumar Sarkar;“Scenario based Performance Analysis of AODV and OLSR in Mobile Ad hoc Networks”; Proceedings of the 24th South East Asia Regional Computer Conference, November 18-19, 2007, Bangkok, Thailand.

22) Z.Fan ; “High throughput reactive routing in multi-rate ad hoc networks” ; ELECTRONICS LETTERS 9th December 2004 Vol. 40 No. 25.

23) Ioannis Broustis, Gentian Jakllari, Thomas Repantis, and Mart Molle, “A Comprehensive Comparison of Routing Protocols for Large-Scale Wireless MANETs”, http://www.cs.ucr.edu/~mart/preprints/iwwan06.pdf.

24) www.OPNET.com.

Analysis of M/M/1 Queueing Model With Applications To Waiting Time of Customers In Banks.

GJCST Classification
D.4.8, E.1

Famule, Festus Daisi

***Abstract*-In this paper, an attempt is made to review the analysis of Stochastic Birth-Death Markov processes which turns out to be a highly suitable modeling tool for many queueing systems in general and M/M/1 queueing model in particular. The model, M/M/1, as a single-channel queueing system with Poisson arrivals and exponential service and with queueing discipline of first come first serve basis, is applied to arrivals and waiting times of customers in Intercontinental Bank PLC, Ile-Ife Branch, Osun State, Nigeria. The queue size of customers including traffic intensity and average number of customers in the system and queue; the service/waiting times of customers including the average time spent in the system and queue by a customer, are all obtained. The traffic intensity obtained is 0.8378 which indicate the probability of a customer queueing or waiting for service on arrival.**

Keywords- stochastic birth-death Markov process, Chapman-Kolmogorov equations, Poisson arrivals, exponential service rates, traffic intensity.

I. NTRODUCTION

1) *Background Information to the Study*

The success of economic policies aiming at economic growth and development of any nation in the world today can only be adequately achieved under a sound banking system. A sound banking system motivates and invites investors from within and outside the country. On the other hand, more efficient and effective investments are basis for a good foreign exchange that in turn generates more external reserves to boost a nation's economy. Any nation with insufficient foreign reserves is bound to borrow from richer and more developed nations to cater for her citizens.Banks can affect liquidity in an economy and they can also directly influence the rate of growth of investments into certain sectors of the economy like agric sector, industrial sector, housing sector, education sector, to mention but a few. Banks invest customers' funds, keep deposits and also give out loans to customers.The recent sweeping measures by the Central Bank of Nigeria with the full support of the Federal Government of Nigeria to sanitize banking sector in the country and the injection of about 620 billion Naira into the industry which have complement the CBN's earlier recapitalization policy are all good examples of the importance of the banking sector to the nation's economy. These measures are presently and gradually reducing the earlier problems of liquidation and distress in this sector.

2) *Statement of the Problems*

In spite of all the efforts of the concerned stake holders in Nigeria most especially the Federal Government and the Central Bank of Nigeria (CBN) to improve the services and effective performance of banks in the country, there are still some avoidable problems that are militating against the success of this sector. One of the most frequent of them is the problem of waiting lines (queues) found in virtually all the banks in the country. Queue is a very volatile situation which always cause unnecessary delay and reduce the service effectiveness of establishments. Long queueing in banks has many negative effects to customers and even the affected banks as well. These negative effects include wasting of man-hours, chaos, unnecessary congestion in banking halls which may lead to suffocation and contraction of communicable diseases. Above all, cases have been witnessed where customers, while waiting too long in banks to lodge in or cash some money got bombarded by armed robbers who killed, mimed, injured both customers and banking officials and get away with huge amount of money from both the customers and the affected banks at large. Also an ill-health or aged customer, while waiting too long on a queue without being attended to in good time, may develop complications, faint or even slum; this may lead to death if proper and urgent medical attention is not provided.

Application of the theory of queue as the development of mathematical models to study and analyze waiting lines with the hope of reducing this social phenomenon in our banking systems would go a long way in improving their services. Through better understanding of queuing situation and the application of appropriate models to deal with it, the bank management will be able to make and take good decisions that will be in the best interest of the overall customers and other stake holders. The end result would be high-volume of customers which will in turn bring more cash, profits, investments, shareholders and the like which will directly or indirectly improve our economy. It is against this background that this research is proposed with the research objectives in the next section on the case study– Intercontinental Bank PLC, Ile-Ife Branch (since the queue problems that are being witnessed in this bank is very much

About- Department of Mathematics and Statistics,Osun State College of Technology, Esa-Oke, Osun State Nigeria E-mail: daisifamule@yahoo.com

similar to what is encountered in all banks across the country).

3) *Research Objectives*

Applying M/M/1 Queueing Model to waiting lines (queues) in Intercontinental Bank PLC, Ile-Ife Branch, we want to:

i. To determine the arrivals and service rates per unit time of customers.

ii. To determine the queue size of customers including traffic intensity,expected number of customers in the system and in the queue.

iii. To estimate the waiting time of customers which include the expected time that a customer will spend in the system and in the queue.

4) *Source Of Data For The Study*

Primary data on arrivals, waiting time, service patterns and departures of customers in Intercontinental Bank PLC, Ile-Ife, Osun State was collected for upwards of a month (i.e. 21 working days) and used as input data into the model (M/M/1) to obtain: the queue size of customers including traffic intensity and average number of customers in the system and in the queue, the waiting time of customers including the average time that a customer will spend in the system and in the queue.

5) *Research Methodology*

The research method will involve derivation of various M/M/m queueing systems generally through Birth-Death Markov Processes and M/M/1 Queuing Model in particular which always exhibits Markov behaviour conformity, as such makes such queuing system analytically(mathematically) tractable.Furthermore, an analysis and interpretation of results obtained, applying M/M/1queueing model, will be done using primary data collected which include arrivals, waiting time, service patterns and departures of customers in Intercontinental Bank PLC, Ile-Ife Branch, Osun State, Nigeria.

II. Model Specification

1) *Birth –Death Markov Processes*

Stochastic birth-death Markov processes turns out to be a highly suitable modeling tool for many queueing processes. Examples of these are any M/M/n queueing models which are our preoccupation in this study with much emphasis on M/M/1 model. Full examination of these models will be considered later in this section.Let $\{N(t), t \geq 0\}$, with parameter set T, be an integer-valued continuous-time discrete stochastic process. The discrete state space of the process comprises (or can assume) non-negative integer values $0, 1, 2, \ldots, \infty$. Here $N(t)$ is interpreted as the random number of members in some population as a function of time. In other words $N(t)$ is viewed as the size of population at time t.By assumption, the classical Markov property is imposed as a restriction on the process $N(t)$, i.e. given the value of $N(s)$, the values of $N(s+t) for\, t > 0$ are not influenced by the values of $N(u)$ for $u < s$. In words, the way in which the entire past history affects the future of the process is completely summarized in the current state of the process, In other words; only the present state gives any information of the future behavior of the process, knowledge of the history of the process does not add new information (The future development of a continuous-time Markov process depends only on its present state and not on its evolution in the past). Expressed analytically the Markov property may be defined as thus: A stochastic process $\{N(t), t \geq 0\}$ with set T and discrete state space

is called a continuous-time Markov process (chain) if for any $m \geq 1$

$$P[N(t_{m+1}) = n_{m+1}/N(t_m) = n_m, \ldots N(t_{1)} = \; n_1] = P[N(t_{m+1}) = n_{m+1}/N(t_m) = n_m] \qquad (1)$$

And it should be valid for all $t_0 < t_1 < t_2 < \cdots < t_m < t_{m+1}\, and\, any\, m$

2) *Transition Probabilities*

In equation (1), set $t_m = s$, $n_m = i$, $t_{m+1} = s + t\, and\, n_{m+1} = j$. then the right-hand side of the equation expresses the probability that the process makes a transition from state i at time s to state j in time t relative to s. Such a probability, denoted by $p_{i,j}(s,t)$, is referred to as a state transition probability for the Markov process. In this research work, we are only concerned with transition probabilities being independent of absolute times$s$, i.e. for all $s > 0$ we have:

$$p_{i,j}(s,t) = p_{i,j}(t) = P\left[N(t) = \frac{j}{N(0)} = i\right] = P\left[N(s+t) = jNs = i\right] \qquad (2)$$

This is called time-homogeneous or stationary transition probabilities. In other words, the intensity of leaving a state is constant in time. It is always natural to make the following definition (for the sake of clarity) which states that the transition probabilities only depends on which state the process is in and not on the time.

Definition: Let $\{N(t),\; t \geq 0\}$ be a discrete Markov process. If the conditional probabilities $P[N(s+t) = j/N(s) = i], for\; s,t \geq 0,$ do not depend on s, the process is said to be time homogeneous. Then we define the transition probability $p_{i,j}(t) = P[N(t) = j/N(0) = i]$ and the transition matrix **P(t)**, whose element with index $(i,j)\; is\; p_{i,j}(t)$

It is generally assumed that the transition probabilities $p_{i,j}(t)$ are well behaved in the sense that they are all continuous and the derivative exists.

Note that $p_{ii}(0) = 1\; and\; p_{ij} = 0\; for\; i \neq j\; so\; that$ $\mathrm{P(t) = I}$

For a Markov process with time-homogeneous transition probabilities the so called Chapman-Kolmogorov equation implies

$$p_{ij}(t+s) = \sum_{k=0}^{\infty} P_{ik}(t) P_{kj}(s) \qquad (3)$$

This equation states that in order to move from state i to j in time $(t+s)$, the queue size process $N(t)$ moves to some intermediate state $\boldsymbol{k}$ in time $\boldsymbol{t}$ and then from $\boldsymbol{k}$ to $\boldsymbol{j}$ in the remaining time $\boldsymbol{s}$. It also says how to compute the long - interval transition probability from a sum of short – interval transition probability components.

An infinitesimal transition probability, denoted by $\boldsymbol{p_{ij}}(\boldsymbol{\Delta t})$**,** specifies the *immediate* probabilistic behavior of a Markov process in that $(\Delta \boldsymbol{t}) \rightarrow \mathbf{0}$. By help of equation (63) it turns out that any transition probability $\boldsymbol{p_{ij}(t)}$ can in principle be determined if the infinitesimal transition probabilities are known. Hence, the overall probabilistic behaviour of a Markov process is ultimately given by the infinitesimal transition probabilities. Together they define the transition kernel of the process.

III. Generalized Markov Birth-Death Process

Suppose now that the population size changes by births and deaths. A birth-death Markov process is characterized by the fact that the discrete state variable changes by at most one, if it changes at all, during an infinitely small time interval. The generalized Markov birth-death process thus satisfies the following criteria:

1. The probability distributions governing the numbers of births and deaths in a specific time interval depends on the length of the interval but not on its starting point
2. The probability of exactly one birth in a small time interval,$\Delta \boldsymbol{t}$ given that the population size at time $\boldsymbol{t}$ is $\boldsymbol{n}$ is $\boldsymbol{\lambda_n}\Delta \boldsymbol{t} + \mathbf{0}(\Delta \boldsymbol{t})$, where $\boldsymbol{\lambda_n}$ is a constant,
3. The probability of one death in small time interval $\Delta \boldsymbol{t}$ given that the population size at time $\boldsymbol{t}$ is $\boldsymbol{n}$ is $\boldsymbol{\mu_n}\Delta \boldsymbol{t} + \mathbf{0}(\Delta \boldsymbol{t})$ where $\boldsymbol{\mu_n}$ is a constant.
4. The probability of more than one birth and the probability of more than one death in a small time interval $\Delta \boldsymbol{t}$ are both $\mathbf{0}(\Delta \boldsymbol{t})$.

Reflecting these facts, the following postulations specify the transition kernel of a general birth-death Markov process:

$$\boldsymbol{P}[\boldsymbol{N}(\boldsymbol{t}+\Delta \boldsymbol{t}) = \boldsymbol{n}+\mathbf{1}/\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}] = \boldsymbol{\lambda_n}\Delta \boldsymbol{t} + \mathbf{0}(\Delta \boldsymbol{t}), \qquad \boldsymbol{n} \geq \mathbf{0}$$

$$\boldsymbol{P}[\boldsymbol{N}(\boldsymbol{t}+\Delta \boldsymbol{t}) = \boldsymbol{n}-\mathbf{1}/\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}] = \boldsymbol{\mu_n}\Delta \boldsymbol{t} + \mathbf{0}(\Delta \boldsymbol{t}), \qquad \boldsymbol{n} \geq \mathbf{1}$$

$$\boldsymbol{P}[\boldsymbol{N}(\boldsymbol{t}+\Delta \boldsymbol{t}) = \boldsymbol{n}/\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}] = \mathbf{1} - (\boldsymbol{\lambda_n}+\boldsymbol{\mu_n})\Delta \boldsymbol{t} + \mathbf{0}(\Delta \boldsymbol{t}), \qquad \boldsymbol{n} \geq \mathbf{1}$$

$$\boldsymbol{P}[\boldsymbol{N}(\boldsymbol{t}+\Delta \boldsymbol{t}) = \boldsymbol{k}/\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}] = \mathbf{0}(\Delta \boldsymbol{t}), \qquad |\boldsymbol{k}-\ \boldsymbol{n}| \geq \mathbf{2} \tag{4}$$

Here $\mathbf{0}(\Delta \boldsymbol{t})$ is a quantity such that $\mathbf{lim}_{\Delta \boldsymbol{t} \rightarrow \infty} \mathbf{0}(\Delta_{\boldsymbol{t}}) = \mathbf{0}$. The first equation handles the case when the state variable increases by one i.e. $\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}+\mathbf{1}$. This is referred to as single birth. Here $\boldsymbol{\lambda_n}$ is proportionality constant such that the product $\boldsymbol{\lambda_n}\mathbf{\Delta t}$ should reflect the probability for a single birth to happen during the infinitesimal time interval $(\boldsymbol{t}, \boldsymbol{t}+\mathbf{\Delta t})$. It is customary to interpret $\boldsymbol{\lambda_n}$ as the instantaneous birth rate. Likewise, the second equation is for the case when the state variable is reduced by one i.e. $\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}-\mathbf{1}$. This is referred to as single death. The product $\boldsymbol{\mu_n}\Delta \boldsymbol{t}$ signifies the probability that a single death takes place. $\boldsymbol{\mu_n}$ denotes the instantaneous death rate. The third equation handles the case when the state variable does not change i.e. $\mathbf{1} - (\boldsymbol{\lambda_n}+\boldsymbol{\mu_n})\mathbf{\Delta t}$ reflects the probability that neither a single birth nor single death occur, i.e. $\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}$, during the infinitely small time interval. Multiple births, multiple deaths and simultaneous births and deaths are taken care of by the $\mathbf{0}(\mathbf{\Delta t})$ terms in the equations. This should be interpreted such that the probability for these events to happen is negligible as $\mathbf{\Delta t} \rightarrow \mathbf{0}$, we say that multiple events are prohibited.

We should note that the transition probabilities from (4) are in general state dependent. This is so since the instantaneous births rate $\boldsymbol{\lambda_n}$ and also the death rate $\boldsymbol{\mu_n}$ may depend on the departing state $\boldsymbol{n}$. A small comment also applies to the second and third equations. Since no death can occur if the state variable is already zero i.e. if $\boldsymbol{n} = \mathbf{0}$, we always define $\boldsymbol{\mu_0} = \mathbf{0}$.

By combining equation (3) with the infinitesimal transition probabilities from (4) and using the notation $\boldsymbol{P_n}(\boldsymbol{t}) = \boldsymbol{P}[\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{n}]$, we can write:

$$\boldsymbol{P_n}(\boldsymbol{t}+\mathbf{\Delta t}) = \sum_{\boldsymbol{k}=\mathbf{0}}^{\infty} \boldsymbol{P}[\boldsymbol{N}\,(\boldsymbol{t}+\Delta \boldsymbol{t}) = \boldsymbol{n}/\boldsymbol{N}(\boldsymbol{t}) = \boldsymbol{k}]\boldsymbol{P}[\boldsymbol{N}(\boldsymbol{t})\boldsymbol{k}] \qquad (\boldsymbol{Bayes'formular})$$

$$= \boldsymbol{P_n}(\boldsymbol{t})[\mathbf{1} - (\boldsymbol{\lambda_n}+\boldsymbol{\mu_n})\Delta \boldsymbol{t} + \mathbf{0}(\Delta \boldsymbol{t})] + \boldsymbol{P_{n-1}}(\boldsymbol{t})[\boldsymbol{\lambda_{n-1}}(\Delta \boldsymbol{t}) + \mathbf{0}(\Delta \boldsymbol{t}) + \boldsymbol{P_{n+1}}(\boldsymbol{t})[\boldsymbol{\mu_{n+1}}(\Delta \boldsymbol{t}) + \mathbf{0}(\Delta \boldsymbol{t})] + \mathbf{0}(\Delta \boldsymbol{t}) \tag{5}$$

By rearranging terms and dividing by $\Delta \boldsymbol{t}$**,** we have

$$\frac{\boldsymbol{P_n}\,(\boldsymbol{t}+\Delta \boldsymbol{t}) - \boldsymbol{P_n}(\boldsymbol{t})}{\Delta \boldsymbol{t}} = -(\boldsymbol{\lambda_n}+\boldsymbol{\mu_n})\boldsymbol{P_n}(\boldsymbol{t}) + \boldsymbol{\mu_{n+1}}\boldsymbol{P_{n+1}}(\boldsymbol{t}) + \boldsymbol{\lambda_{n-1}}\boldsymbol{P_{n-1}}(\boldsymbol{t}) \tag{6}$$

Taking limit as $\Delta \boldsymbol{t} \rightarrow \mathbf{0}$, we obtain the differential equation:

$$\boldsymbol{P'_n}(\boldsymbol{t}) = \frac{\boldsymbol{dP_n}(\boldsymbol{t})}{\boldsymbol{dt}} = -(\boldsymbol{\lambda_n}+\boldsymbol{\mu_n})\boldsymbol{P_n}(\boldsymbol{t}) + \boldsymbol{\mu_{n+1}}\boldsymbol{P_{n+1}}(\boldsymbol{t}) + \boldsymbol{\lambda_{n-1}}\boldsymbol{P_{n-1}}(\boldsymbol{t}), \quad \boldsymbol{n} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots \tag{7}$$

While (7) holds for $\boldsymbol{n} = \mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots$, we also require an equation for $\boldsymbol{n} = \mathbf{0}$. By following a logical argument as above, we can write

$$\frac{\boldsymbol{dP_n}(\boldsymbol{t})}{\boldsymbol{dt}} = -\boldsymbol{\lambda_0}\boldsymbol{P_0}(\boldsymbol{t}) + \boldsymbol{\mu_1}\boldsymbol{P_1}(\boldsymbol{t}) \tag{8}$$

Equation (7) is the general model equation for a birth-death Markov process and it essentially captures the probabilistic dynamics of the process. The equation is a differential equation in the continuous time variable $\boldsymbol{t}$ and a difference equation (also called a recurrence equation).

As already shown, note that the model equation (7) is valid for $t > 0$ *and* $n = 0,1,2, \ldots, \infty$. For $\boldsymbol{t} = \mathbf{0}$ we have a boundary condition and it is customary to define $\boldsymbol{P_n}(\mathbf{0}) = \boldsymbol{\delta_{ij}} = \mathbf{1}\ \boldsymbol{for}\ \boldsymbol{i}$ $\boldsymbol{and} = \mathbf{0}\ \boldsymbol{otherwise}$ (9)

Hence, in zero time the process will certainly not move.

3) *Steady State Solution*

We now examine the process when it is in equilibrium (steady state). Under proper conditions, such equilibrium will be reached after the system has been operating for some time. Equilibrium in turn, implies that the state probabilities $\boldsymbol{P_n}(\boldsymbol{t})$ eventually become independent of $\boldsymbol{t}$ and approach a set of constant values (if it exists) which is denoted by $\boldsymbol{P_n}$, $\boldsymbol{n = 0, 1, 2}$ ... as $\boldsymbol{t \to \infty}$ where $\boldsymbol{P_n = \lim_{t\to\infty} P_n(t)}$. This can be interpreted as steady state probability that there are $\boldsymbol{n}$ users in the system. Also under these circumstances in the steady state, $\boldsymbol{\lim_{t\to\infty} P'_n(t) = 0}$.

Given the above, (7) and (8) are then transformed to:

$$\boldsymbol{0 = -(\lambda_n + \mu_n)P_n + \mu_{n+1}P_{n+1}\lambda_{n-1}P_{n-1},}$$
$$\boldsymbol{n = 1, 2, 3, \ldots}$$
$$\boldsymbol{or\ \ (\lambda_n + \mu_n)P_n = \mu_{n+1}P_{n+1} + \lambda_{n-1}P_{n-1}} \qquad \textbf{(10)}$$

And $\boldsymbol{0 = -\lambda_0 P_0 + \mu_1 P_1}$

$$\boldsymbol{or\ \ \lambda_0 P_0 = \mu_1 P_1} \qquad \textbf{(11)}$$

Equations (10) and (11) are called the equilibrium – equations or balance equations of birth and death Markov process. They have a natural and useful interpretation. They say that, at equilibrium,

IV. The Probability Flow Out of a State = the

1) *probability of flow in that state*

In other words, for any time $\boldsymbol{t}$ when the system is in equilibrium, the probability of observing a transition out of state $\boldsymbol{n}$ in the next $\boldsymbol{\Delta t}$ must be equal to the probability of observing a transition into state $\boldsymbol{n}$ This key observation will allow us in most cases to generate the correct equilibrium equations without going through the burden of writing down equations like (5).

For the process, another set of balance equations, even easier than (10) and (11) can be obtained directly from (11)

i.e.
$$\boldsymbol{\lambda_0 P_0 = \mu_1 P_1}$$
$$\boldsymbol{\lambda_1 P_1 = \mu_2 P_2}$$
$$\boldsymbol{\lambda_2 P_2 = \mu_3 P_3}$$
$$\cdot \qquad \cdot$$

and in general, $\boldsymbol{\lambda_n P_n = \mu_{n+1} P_{n+1}}$

$$\boldsymbol{for\ n = 0, 1, 2, \ldots} \qquad \textbf{(12)}$$

We can now proceed to solve the balance equation expressing all steady-state probabilities $\boldsymbol{P_n, n = 0, 1, 2, \ldots}$ in terms of one of them and then taking advantage of the fact that

$$\boldsymbol{\textstyle\sum_{n=0}^{\infty} P_n = 1} \qquad \textbf{(13)}$$

this is called the normalization equation.

It is common practice to express $\boldsymbol{P_1, P_2, P_3, \ldots}$ in terms of $\boldsymbol{P_0}$, the steady – state probability of an empty system. Working with (10) and (11) or equivalently (and preferably) with (12) we have (solving recursively):

$$\boldsymbol{P_1 = \frac{\lambda_0}{\mu_0} P_0} \qquad \textbf{(14)}$$

$$\boldsymbol{P_2 = \frac{\lambda_1}{\mu_2} P_1 = \frac{\lambda_1 \lambda_0}{\mu_2 \mu_1} P_0} \qquad \textbf{(15)}$$

and, in general

$$\boldsymbol{P_n = \frac{\lambda_{n-1}\lambda_{n-2}\cdots\cdots\lambda_1\lambda_0}{\mu_n \cdot \mu_{n-1}\ldots\ldots\mu_2\mu_1} P_0} \qquad \textbf{(16)}$$

or, defining the coefficient of $\boldsymbol{P_0}$**,** in (16) as the quantity $\boldsymbol{K_n}$, we have

$$\boldsymbol{P_n = K_n P_0 \qquad for\ n = 1, 2, 3, \ldots}$$

Going back to (6.13)

$$\boldsymbol{\sum_{n=0}^{\infty} P_n = (1 + \sum_{n=1}^{\infty} K_n) P_0} \qquad \textbf{(17)}$$

Or $\boldsymbol{P_0 = \frac{1}{(1 + \sum_{n=1}^{\infty} K_n)}}$ $\qquad$ **(18)**

It follows that the system can reach steady-state (or equilibrium distribution exists) if and only if $\boldsymbol{\sum_{n=1}^{\infty} K_n}) < \infty$. For, otherwise, $\boldsymbol{P_0 = P_1 = P_2 = P_3 = \ldots = 0}$ (i.e. the number of users in the system never “stabilizes”)

2) *Analysis of the M/M/1 Queueing Model*

In this queueing system, the customers arrive according to Poisson process. Inter-arrival times are exponentially distributed with average arrival rate $\boldsymbol{\lambda}$. The service times (i.e. the time it takes to serve every customer) are exponentially distributed with average service rate $\boldsymbol{\mu}$. The service times are mutually independent and further independent of the inter-arrival times. The constant average arrival rate $\boldsymbol{\lambda}$ and the average service rate $\boldsymbol{\mu}$ are in units of customers per unit time. The expected inter-arrival time and the expected service time are $\boldsymbol{1/_\lambda}$ and $\boldsymbol{1/_\mu}$ respectively.

There is only one server; no limit on the system capacity, i.e. the buffer is assumed to be infinite and the queueing discipline is first-come-first served (FCFS).

Since exponentially distributed inter-arrival times with mean $\boldsymbol{1/_\lambda}$ are equivalent, over a time interval, say τ, to a Poisson – distributed arrival pattern with mean $\lambda\tau$, M/M/1 queueing model is often refer to as single-server, infinite-capacity queueing systems with customers arriving according to Poisson process and exponential service times. When a customer enters an empty system, his service starts at once; if the system is nonempty the incoming customer joins the queue. When a service completion occurs, a

customer from the queue, if any, enters the service facility at once to get served.M/M/1 queueing model is a Poisson birth - death process. A birth occurs when a customer arrives and a death occurs when a customer departs. Both processes are modeled as memoryless Markov process. The M designation in M/M/1 actually refers to this memoryless/Markov feature of the arrival and service processes (i.e. they are exponentially distributed).The memoryless property can be further described as thus: with exponential inter-arrival times and exponential service times, the distribution of the time until the next arrival and/or service completion is not affected by the time that elapsed since the last arrival and the last completion.

System State: Due to the memoryless property of the exponential distribution, the entire state of the system, as far as the concern of probabilistic analysis, can be summarized by the number of customers in the system, $\boldsymbol{n}$ (i.e. the customers waiting in the queue and the one being served). – the past/history (how we get there) does not matter. When a customer arrives or departs, the system moves to an adjacent state

(either $\boldsymbol{n+1\ or\ n-1}$.) Let $\boldsymbol{N(t)}$ be the number of customers in the system at time $\boldsymbol{t}$.

Theorem:

The process $\{N(t), t \geq 0\}$ is a birth and death process with birth rate $\lambda_n = \lambda\, for\, all\, n \geq 0$ and with death rate $\mu_n = \mu\, for\, all\, n \geq 1$.

Proof:

Because of the exponential distribution of the inter-arrival times and of the service times, it should be clear that $\{N(t), t \geq 0\}$ is a Markov process. On the other hand, since the probability of having two events (departures, arrivals) in the interval of time $(t, t + \Delta t)$ is $0(\Delta t)$, which is characterized by the fact that the discrete state variable changes by at most one, if it changes at all, during an infinitely small time interval (criteria for the generalized Markov birth-death process), we have the following transitional probabilities:

$$P[N(t+\Delta t) = n+1/N(t) = n] = \lambda\Delta t + 0(\Delta t), \quad n \geq 0$$
$$P[N(t+\Delta t) = n-1/N(t) = n] = \mu\Delta t + 0(\Delta t), \quad n \geq 1$$
$$P[N(t+\Delta t) = n/N(t) = n] = 1 - (\lambda+\mu)\Delta t + 0(\Delta t), \quad n \geq 1$$
$$P[N(t+\Delta t) = n/N(t) = n] = 1 - \lambda\Delta t + 0(\Delta t), \quad n = 0$$
$$P[N(t+\Delta t) = k/N(t) = n] = 0(\Delta t), \quad |k - n| \geq 2. \qquad (19)$$

Here $0(\Delta t)$ is a quantity such that $\lim_{\Delta t \to \infty} 0(\Delta_t) = 0$.

This shows that $\{N(t), t \geq 0\}$ is a birth and death process.

From (6), (7), and (8) with the fact that $\lambda_n = \lambda$ and $\mu_n = \mu\, for\, all\, n$ in the birth and death Markov process, the differential equations for the transitional probability that the system is in state n at time t (for M/M/1 system) is given as:

$$P'_n(t) = -(\lambda+\mu)P_n(t) + \lambda P_{n-1}(t) + \mu P_{n+1}(t), \quad n = 1, 2, 3, \ldots \qquad (20)$$
$$P'_n(t) = -\lambda P_0(t) + \mu P_1(t), \quad n = 0 \qquad (21)$$

As $t \to \infty$, $P'_n(t) \to 0\ and\ P_n(t) \to P_n$

Thus we have:

$$0 = -(\lambda+\mu)P_n + \lambda P_{n-1} + \mu P_{n+1}, \quad n = 1, 2, 3, \ldots$$

Or $(\lambda+\mu)P_n = \lambda P_{n-1} + \mu P_{n+1}$ (22)

And $0 = -\lambda P_0 + \mu P_1, \quad n = 0$

Or $\lambda P_0 = \mu P_1$ (23)

Which are the equilibrium or balance equations for an M/M/1 system

Solving recursively (using (7.3)), we have

$$\lambda P_0 = \mu P_1$$
$$\lambda P_1 = \mu P_2$$
$$\lambda P_2 \qquad (24)$$
$$. \qquad .$$

We define $\rho = \lambda/\mu$ (25)

The quantity ρ is referred to as the traffic intensity (or the utilization factor) since it gives the mean quantity of work brought to the system per unit of time.

Now, from (24),

$$P_1 = \lambda/\mu\, P_0 = \rho P_0$$
$$P_2 = \lambda/\mu\, P_1 = \rho(\rho P_0) = \rho^2 P_0$$
$$P_3 = \lambda/\mu\, P_2 = \rho(\rho^2 P_0) = \rho^3 P_0$$

In general $P_n = \rho^n P_0$ (26)

Since the sum of the probabilities must equal unity i.e. $\sum_{n=0}^{\infty} P_n = 1\ and\ 0 < \rho < 1$

i.e. $\sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} \rho^n P_0 = 1$

we have $P_0 \sum_{n=0}^{\infty} \rho^n = 1$, $\left(\frac{1}{1-\rho}\right) P_0 = 1$

$$\therefore P_0 = 1 - \rho \quad or \quad 1 - \lambda/\mu \qquad (27)$$

That is

$$P_n = \rho^n(1-\rho), \quad n > 0, \rho < 1 \qquad (28)$$

This is called equilibrium distribution (or queue-length of density function of an M/M/1 queue)

The stability condition, $\rho < 1$ simply says that the system is stable if the work that is brought to the system per unit time is strictly smaller than the processing rate (which is 1 here since there is only one server).

It is noteworthy that the queue will be empty infinitely many times when the system is table i.e. when $n = 0$; from (28)

$$P_0 = 1 - \rho > 0 \qquad (29)$$

We observe that the result in (28) which is the density function of the queue-length in a steady-state is a geometric distribution. We then compute (in particular) the mean number of customers as:

$$E(N) = \frac{\rho}{1-\rho} \qquad (30)$$

We can observe that $E(N) \to \infty\ as\ \rho \to 1$ so that in practice, if the system is not stable, then the queue will explode. We also find from (28) that:

$$V(N) = \frac{\rho}{(1-\rho)^2} \qquad (31)$$

Again, we find from (28) that the probability that the queue exceeds, say, k customers, in steady-state is:

$$P(N \geq k) = \rho^k \qquad (32)$$

We have adopted a single-channel queueing system (M/M/1) with Poisson arrivals and exponential service rate and arrivals are handled on a first come first serve basis, in this queueing system the average arrival rate is less than the average service rate (i.e. $\lambda < \mu$). In this case, there would be an unending queue.

The following formulas are developed for this system.

The average (or expected) number of customers in the queue at any time t is:

$$E(N_q) = L_q = \frac{\rho^2}{1-\rho} \qquad (33)$$

The average (or expected) number of customers waiting to be served at time t is:

$$\frac{1}{1-\rho} \qquad (34)$$

The average (or expected) number of customers in the system is:

$$E(N) = L = \frac{\rho}{1-\rho} \qquad (35)$$

The average (or expected) time a customer spends or wait in queue (before service is rendered), also called average waiting time, is:

$$E(W_q) = T_q = \frac{\rho}{\mu(1-\rho)} \qquad (36)$$

The average (or expected) time a customer spends or wait in the system (on the queue and receiving service, also called the average waiting time in the system, is:

$$\boldsymbol{E(W) = T = \frac{1}{\mu(1-\rho)}} \qquad (37)$$

V. Results

Data on arrivals, waiting times, service patterns and departures of customers in Intercontinental Bank PLC, Ile-Ife, Osun State, Nigeria were observed and collected for upwards of 21 working days between the working hours of 8.00 a.m. and 4.00 p.m. on daily basis. The waiting times and service times were obtained respectively by subtracting arrival times from the time service began; and subtracting the time service began from when it ended. On the final analysis, it was found that a total arrival rate of 1,302 customers per a total of 11,304 minutes (waiting times). Also a total service rate of 1,414 customers per a total service times of 10,284 minutes.

The following results are arrived at:

1. The arrival rate $\lambda = \frac{1,302}{11,304} = 0.1152$
2. The service rate $\mu = \frac{1,414}{10,284} = 0.1375$
3. The traffic intensity $\rho = \frac{\lambda}{\mu} = \frac{0.1152}{0.1375}$ $= 0.8378$
4. The average number of customers in the queue (i.e. number of customers waiting for service) $= \frac{\boldsymbol{\rho^2}}{\boldsymbol{1-\rho}} = \boldsymbol{4.327, approx. 4}$ Thus there would be an average of 4 customers waiting for service.
5. The expected number of customers in the system (i.e. in the queue and the one being served) $= \frac{\boldsymbol{\rho}}{\boldsymbol{1-\rho}} = \boldsymbol{5.1652}$
6. The expected number of customers waiting to be served at any time t (when the queue exists) $= \frac{\boldsymbol{1}}{\boldsymbol{1-\rho}} = \boldsymbol{6.1652}$
7. The expected time a customer spends or waits in the queue before being served $\frac{\boldsymbol{\rho}}{\boldsymbol{\mu(1-\rho)}} = \boldsymbol{37.5653},$ $\boldsymbol{approx. 38\ minutes}$ Alternatively, $\frac{average\ no.of\ customers\ in\ the\ system}{\mu(i.e.service\ rate)} = \frac{5.1652}{0.1375} = 37.5650\ (aprox, 38\ mins.$
8. The expected time a customer spends in the system (on queue and being served) $\frac{\boldsymbol{1}}{\boldsymbol{\mu(1-\rho)}} = 44.8380, approx. 45\ mins.$
9. Average time of service $=$ $average\ time\ in\ the\ system - average\ time\ in\ the\ queue$ $= 45mins. - 38mins = 7mins.$
10. The probability of queueing on arrival $=$ $traffic\ intensity$ $or\ utilization\ factor$ $= \rho = \frac{\lambda}{\mu} = 0.8378$ i.e. probability of at least one customer in the system.
11. The probability of no customer in the system or probability of not queueing on arrival is $1 - \frac{\lambda}{\mu}\ or\ (1-\rho)$ $= 1 - 0.8378 = 0.1622$
12. The probability that there are n customers in the system (or the queue length is $P_n = \rho^n(1-\rho)$ Hence, the probability of having a queue which is the probability of having two or more customers in the system is $1 - P_0 - P_1\ (where\ P_0 = \rho^0(1-\rho)\ \&\ P_1 = \rho^1(1-\rho)$ Now, $P_0 = 0.1622\ \&\ P_1 = 0.1359\ \therefore$ $The\ prob. of\ having\ a\ queue\ is$ $1 - 0.1622 - 0.1359 = 0.7019$

VI. Discussion of Results

The probability that a customer who arrives in the bank has to queue which is $1 - P_0$ which corresponds to traffic intensity (or the utilization factor) $\rho = 0.8378$ clearly indicates that there is always a very high possibility that customers would have to wait for every transaction in the bank since the bank officials concerned would always be busy attending to a customer that has earlier arrived. This is corroborated by $result\ 6$ which indicates that up to six customers will always be waiting to be attended to at any time t.It then follows that there will always be queue at any time t since the average time in the queue system (both on queue and receiving service) ($result\ 8$) is greater than the average time in the queue ($result\ 7$) before service is rendered.Also from $result\ 8$ which shows the average time that an individual customer would spend in the bank i.e. $45\ minutes$ before completing his/her transactions is on the higher side than expected in an ideal situation.

With all these results, we cannot say that the service in the bank is all that efficient. This is the situation in virtually most banks in the country. Therefore effort should always be made to reduce queue at least in our banks if it cannot be totally eliminated.

VII. Conclusion and Recommendation

Application of queueing theory is indeed a very useful and an indispensable statistical tool for solving problems of queueing in our banking sector. The queueing problems witnessed in Intercontinental Bank, PLC, Ile-Ife branch are very much peculiar to other banks across the country. This social phenomenon in our banking sector has caused a lot of negative effects to customers and the affected banks. This include excessive wasting of man hours, chaos, unnecessary congestion in banking halls which may lead to suffocation and contraction of communicable diseases (or other health complications), indirect reduction of customers which may in turn bring about less cash, profits, investments and shareholders to the affected banks.As a result of the above and other similar problems, it is recommended that more banking officials (cashiers, accountants, administrators, computer operators, analysts, etc.) should be employed so as to increase the channels of services.More ATM machines should be mounted in every nooks and crannies of cities, towns as well as villages as this will reduce the number of customers that will withdraw cash and check balances in the banks.Effective and efficient internet facilities which are reliable and fast should be installed for quick and easy access to individual statements of account during any bank transaction as this will reduce the waiting time of customers in banks. This should always be complemented by good sources of electricity e.g. electricity stand-by-generator which will at least bail one out of the epileptic power supply as always witnessed in the country at present.

In nutshell, more bank branches should be established along with the existing ones across the country, and adequately equipped with effective human/materials resources; current and effective banking management and operations should always be adopted for efficient and prompt banking services.

VIII. REFERENCES

1) Bharucha-Reid, A.T. 1988. *Elements of the Theory of Markov Processes and Their Applications.* Dover Publications, INC Mineola, New York, NY.
2) Churchman, G.W., R.C. Ackoff and F.C. Arnoff. 1957. *Introduction to Operations Research.* John Willey and Sons: New York, NY.
3) Frode B. Nielsen. 1998. *Queueing Systems: modelling Analysis and Simulation.* Research Report 259, ISBN 82-7368-185-8, ISS 0806-3036. Department of Informatics, University of Oslo.
4) Ivo Adan and Jacques Resing. 2002. *Queueing Theory* – Department of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, The Netherlands.
5) Kothari, C.R. 1978. *Quantitative Techniques.* Vikas Publishing House PVT Ltd., New Delhi-110002.
6) Philippe Nain. 1998. *Lecture Notes on Basic Elements of Queueing Theory – Application to the Modelling of Computer System.* University of Massachusetts, Abreast, MA.
7) Richard Bronson and Govindasami Naadimuthu. 1997. *Operations Research.* Schaum's Outline Series, The McGraw-Hill Companies, New York.
8) Taha, H.A. 1987. *Operations Research: An Introduction, Fourth Edition.* Macmillan Publishing: New York, NY.

IEEE 802.16e Security Vulnerability : Analysis & Solution

GJCST Classification
C.2.0, D.4.6, K.6.5

A.K.M. Nazmus Sakib[1], Dr. Muhammad Ibrahim Khan[2], Mir Md. Saki Kowsar[3]

Abstract*- Data security has become a major issue in most network protocols. For wireless system, security support is even more important to protect the users as well as the network. Due to this importance, different protocol were designed & deployed with network standards in order to add the security. The security sub layer of IEEE 802.16 employs an authenticated client/server key management protocol in which the B.S, the serve, control the distribution of keying materials to the client M.S. This paper analyzes the physical layer threat & MAC layer threat of WiMAX .First give an overview of security architecture of mobile WiMAX network, then investigate different security vulnerability & gives possible solution to overcome them. These vulnerabilities are the possibilities to forge key messages in Multi- and Broadcast operation, some unauthenticated messages which are susceptible to forgery and the unencrypted management communication which reveal important management information. We modify DH key exchange protocol to fit it into mobile WiMAX network as well as eliminate existing weakness in original DH key exchange protocol. Also RSA & Elliptic curve Diffie Hellman key agreement algorithm are discuss which can be used to generate symmetric key between M.S & B.S. Several one way function are presented by using cryptography, which can be used to solve shared key vulnerability in Multi-&Broadcast service. We find the initial network procedure is not effectively secured that makes Man-in-the-middle attacks & Denial of service attack possible.

Keywords-IEEE 802.16e security, multi-and broadcast service, shared key vulnerability, Key agreement, DoS

I. INTRODUCTION

WiMAX- opens the door to thousands of applications that make use of the solid wireless backbone to connect people together. With the high data rate, applications will include voice calls, video transfer, file transfer and many other services.

All those types of applications will require a secure medium to operate and exchange information safely. This is what the IEEE decided to add to the WiMAX standard in its both versions mobile and fixed broadband wireless access. First of all security scheme of Mobile WiMAX are introduced. Afterwards different security threats, Vulnerabilities and possible solutions to solve them are presented

1) *Wimax Security Scheme*

a) *Protocol layer:*

The IEEE 802.16 standard consists of a protocol Stack with well-defined interfaces. The scope of Protocol contains PHY layer and MAC layer [1]. MAC layer consist three sub-layers which is shown in Figure 1. The Service Specific the security sub-layer, Convergence Sublayer (MAC CS) and the MAC Common Part Sub-layer (MAC CPS). The service specific Convergence Sublayer (CS) maps higher level data services to MAC layer service flows and connections. There are two type of CS: packet CS which supports Ethernet, point-to-point protocol (PPP), both IPv4 and IPv6 internet protocols, and virtual local area network (VLAN) and ATM CS which is designed

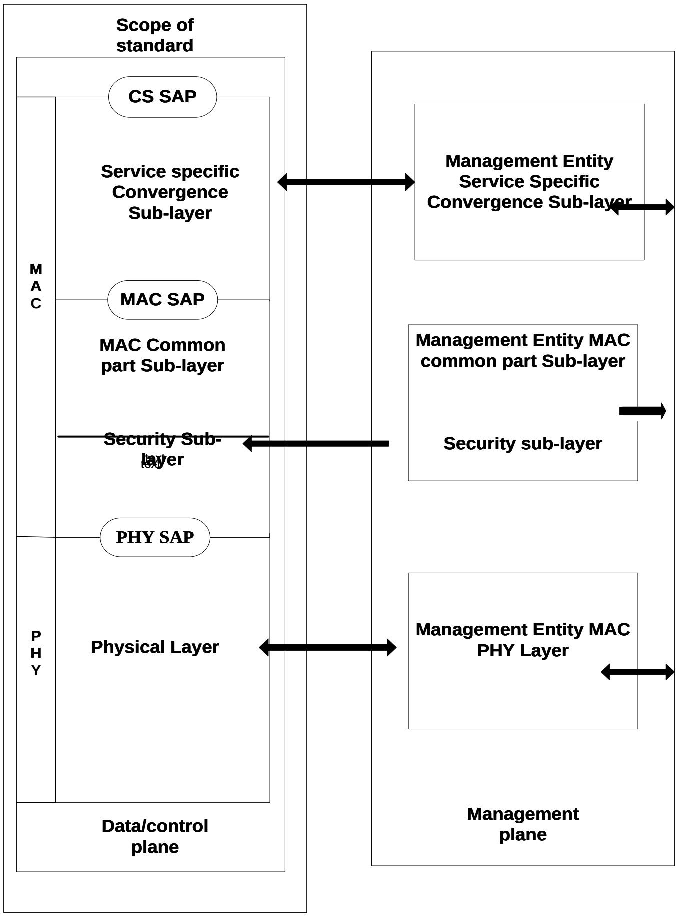

Scope of standard
CS SAP
Service specific Convergence Sub-layer
M A C
MAC SAP
MAC Common part Sub-layer
Security Sub-layer
PHY SAP
P H Y
Physical Layer
Data/control plane
Management Entity Service Specific Convergence Sub-layer
Management Entity MAC common part Sub-layer
Security sub-layer
Management Entity MAC PHY Layer
Management plane

Figure 1: Protocol Layering in 802.16

[1] *About-(telephone: +8801730079790 email: sakib425@gmail.com)*
[2] *About-(telephone: +8801678708081 email: muhammad_ikhancuet@yahoo.com)*
[3] *About- (telephone: +8801713109890 email: . sakikowsar@yahoo.com)*
Department of Computer Science and Engineering Chittagong University of Engineering and Technology Chittagong-4349, Bangladesh text text

Management Entity Service Specific Convergence Sub-layer Management Entity MAC common part Sub-layer Security sub-layer Management Entity MAC PHY Layer CS SAP MAC SAP PHY SAP Service specific Convergence Sub-layer MAC Common part Sub-layer Security Sublayer Physical Layer Data/control plane Management plane M A C P H Y Scope of standard

for ATM network and service [1]. MAC Common Part Sublayer is the core of the standard. MAC CPS defines the rules and mechanisms for system access, bandwidth allocation and connection management. Functions like uplink scheduling, grant, connection control and bandwidth request and automatic repeat request is also defined here. Communications between the the MAC CPS and CS are done by MAC Service Access Point [1]. The security sub-layer is responsible for decryption and encryption of data traveling to and from the PHY layer, and it is also used for authentication and secure key exchange. This Sub-layer lies between PHY layer and MAC CPS [1]. PHY layer, targeted for operation in the 10-66 GHz frequency band, is designed with a high degree of flexibility in order to allow service providers to optimize system deployments with respect to cost, radio capabilities, cell planning and services.

2) *Security scheme*

The Mobile WiMAX system based on the IEEE 802.16e amendment and has more improved security features than previous IEEE 802.16d-based WiMAX network system [2]. Almost all the security issues in Mobile WiMAX are considered in security sub-layer and are shown in Figure 2.

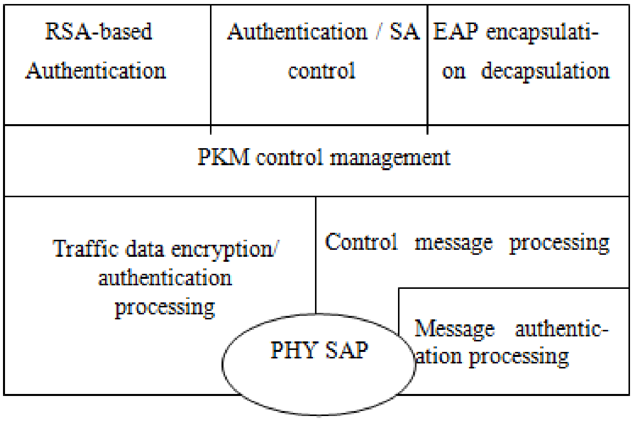

RSA-based Authentication
Authentication / SA control
EAP encapsulation decapsulation
PKM control management
Traffic data encryption/ authentication processing
Control message processing
Message authentication processing
PHY SAP

Figure 2: Security sub-layer

The security sub-layer encompasses three essential Functions: authentication, authorization and encryption [1]. The PKM protocol uses, strong encryption algorithm and RSA public key algorithm, X.509 digital certificates, to carry out key exchanges between BS and M.S [2]. This Privacy protocol has been enhanced to accommodate stronger cryptographic methods such as AES to fit into the IEEE 802.16 MAC [2]. The main objective of the privacy sub layer is to protect service providers against theft of service, rather than guarding network users. Privacy sub layer is above the physical layer, so it only guards data (data link) but does not protect physical layer from intercepted. It is necessary to secure the physical layer

3) *WiMAX Security Process*

WiMAX security process is divided into three steps:

1. Authentication
2. Data Key exchange.
3. Data Encryption.

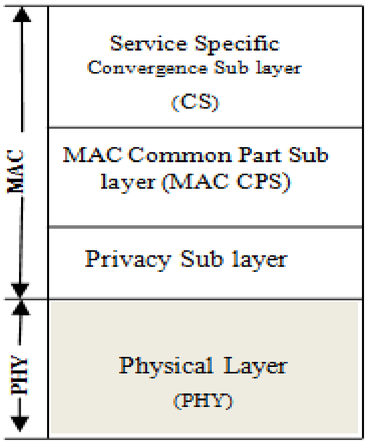

MAC
PHY
Service Specific Convergence Sub layer (CS)
MAC Common Part Sub layer (MAC CPS)
Privacy Sub layer
Physical Layer (PHY)

Figure 3: IEEE 802.16 MAC and Physical Layer

4) *Initial network entry process*

Initial network entry contains four processes: initial Ranging process, M.S Basic Capability negotiation process, PKM authentication process and registration process [2]. It is the most security sensitive processes in IEEE 802.16 [WiMAX] network not only because it is the first phase to establish a connection to the network, but also because many parameters, performance factors and security contexts between B.S and serving M.S are determined during this process. The initial network process is illustrated in Figure 5. After initial network entry, the management communication over the primary and basic management connections remains unencrypted [2]. As most of the management messages are sent on these connections, all management information exchanged between B.S and M.S can be accessed by a adversary [2].

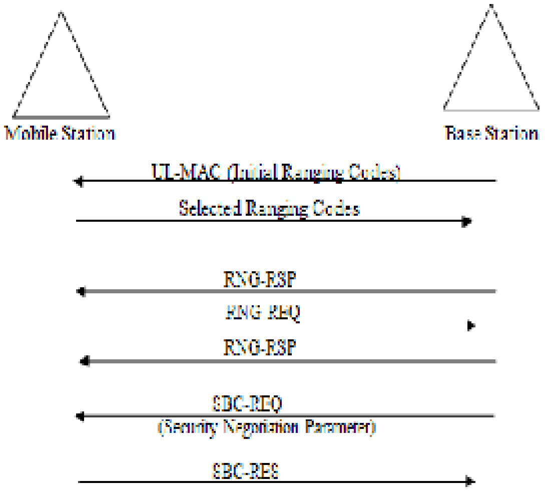


Figure 4: WiMAX initial network entry procedure

The only messages which are encrypted are key transfer messages. But in this case only the key is encrypted, all other information is still sent in the clear. An intruder collecting management information can create detailed profiles about M.S's including capabilities of devices, security settings, associations with base stations and all other information described above. Using the data offered in power reports, registration, ranging and handover messages, an attacker is able to determine the movement and approximate position of the M.S as well. Monitoring the MAC address sent in ranging or registration messages reveals the mapping of connection identifier (CID) and MAC address, making it possible to clearly relate the collected information to user equipment [2].

I. Vulnerabilities in IEEE 802.16e

This section explains vulnerabilities found in Mobile WiMAX by analysis. These vulnerabilities are:

1) *Unauthenticated messages*

There are some unauthenticated Messages include in Mobile WiMAX. Their forgery can interrupt or constrict the communication between M.S and B.S [3]. First it has to be mentioned that a couple of management messages are sent over the broadcast management connection. Broadcasted management messages authentication is difficult since there is no common key to generate message digests. A common key would not completely protect the integrity of the message as M.S sharing the key can forge these messages and generate valid authentication digits [3].

2) *Unencrypted management communications*

Management messages in Mobile WiMAX are still sent in the clear. The consequential risk shall be outlined in this section.When a M.S performs initial network entry it negotiates communication parameters and settings with the B.S [2]. Here a lot of information is exchanged like mobility parameters, security negotiation parameters, configuration settings, power settings, vendor information, MSs Capabilities etc. In the existing system the complete management message exchange in the network entry process is unencrypted and the above mentioned information can be accessed just by listening on the channel [2].

3) *Shared keys in the multi- and broadcast service*

The Multi- and Broadcast service provides the possibility to distribute data to multiple MS with one single message. This saves bandwidth and cost. Broadcasted messages in IEEE 802.16e are encrypted with a shared key [3] [2]. Every member in the group know the key and thus can decrypt the traffic. Also message authentication is also based on the same shared key. This algorithm contains the vulnerability that every group member, besides decrypting and verifying broadcast messages can also encrypt and authenticate messages as if they originated from the legitimate B.S [3]. Another thing which is much more problematic is the distribution of the traffic encryption keys (GTEKs) when the optional Multi- and Broadcast Rekeying Algorithm is used. To transfer a GTEK to all group members, it is broadcasted but encrypted with the group key encryption key (GKEK). Because of broadcasting, the GKEK must also be a shared key and every group member knows it [3]. An adversary group member can use it to generate valid encrypted and authenticated GTEK key update command messages and distribute its own GTEK. Every group member would establish the adversary's key as a valid GTEK. Subsequently, all traffic sent by the legitimate B.S can no longer be decrypted by the MS. From M.Ss point of view, only traffic from the adversary is valid. To establish the adversary's key (M.S), there are several possibilities. If the implementation does not work completely, the key from the latter of two subsequently sent GTEK update command messages may overwrite the former one [3]. Hence the adversary just has to send its GTEK (Group Traffic Encryption Key) update command message after the B.S broadcasted a key update message. If the implementation follows the standard, the keys of both messages are accepted. To be sure the M.S will not establish the legitimate B.Ss key, an adversary could forge some part of the B.Ss GTEK update command message. Such a changed message would not be verified as correct and it is discarded by the M.Ss [3]. After this the adversary can send its own GTEK update command message. Which will be accepted. In a unicast connection this different keying material at the M.S would be detected as the B.S cannot decrypt data sent by the M.S. These results in a TEK invalid message destined to the mobile station which subsequently refreshes its keying material. Since the MBS is unidirectional, the B.S unable to detect that M.S has different GTEKs.

4) *Initial network entry vulnerability*

a) *RNG-RSP vulnerability*

In RNG-RSP vulnerability, the attacker modifies the RNGRSP message and sets the status as failed and re-sends it to M.S. So the M.S has to go for initial ranging again. If the attacker continuously sets the RNG-RSP status as failed, it (M.S) access the network. This leads to the DoS attack. This RNG-RSP vulnerability is solved by Diffie-Hellman (D-H) key agreement [2], which is discussed in the later part of this thesis.

b) *Auth-Request and Invalid vulnerability*

In Auth-Request and Invalid vulnerability, the intruder captures the Auth-Request message and re-sends it to B.S continuously [4]. So the B.S is confused with the continuous request and sets the Auth-Response as failure. Some time the attacker may captures Auth-Response message from B.S and re-sends to M.S [5]. This issue can be solved by either introducing time stamps model. By adding time stamp, B.S and M.S identifies if the authorization message is proper. So the attacker unable to modify the messages.

c) *Rogue BS*

For rogue BS attack, the M.S cannot verify that any authorization protocol messages it receives were generated by an authorized B.S. So any rogue BS can create a response [4]. To solve this issue, the M.S has to authenticate to the B.S.

5) *Denial of Service attack:*

DoS attacks such as unprotected network entry, unprotected management frame, weak key sharing mechanism in multicast and broadcast operations, unencrypted management communication and reset-command message [5]. Some DoS attacks are include the following:

a) *DoS attacks based on Ranging Request/Response (RNG-REG/RNG-RSP) messages*

An intruder can forge a RNG-RSP message to minimize the power level of M.S to make M.S hardly transmit to B.S, thus triggering initial ranging procedure repeatedly [4]. An intruder can also perform a water torture DoS by maximizing the power level of M.S, effectively draining the M.S's battery [5].

b) *DoS attacks based on Mobile Neighbor Advertisement (MOB_NBR_ADV) message*

This message is sent from serving B.S to publicize the characteristics of neighbor base stations to M.Ss searching for possible handovers [4]. This message is not authenticated and it can be forged by an attacker in order to prevent the M.Ss from efficient handovers downgrading the performance or even denying the legitimate service [5].

c) *DoS attacks based on Fast Power Control (FPC) message*

This message is sent from B.S to ask a M.S to adjust its transmission power [4]. This is also one of the management messages which are unprotected. An intruder can intercept and use FPC message to prevent a M.S from correctly adjusting transmission power and communicating with the B.S. He can also use this message to perform a water torture DoS attack to drainthe M.S's battery [5].

d) *DoS attacks based on Authorization-invalid (Auth-invalid) message*

The Auth-invalid is sent from a B.S to a M.S when Authorization key(AK) shared between B.S and M.S expires or B.S is unable to verify the CMAC/HMAC properly [4]. This message is unprotected by HMAC and it has PKM identifier equal to zero [4]. Thus, it can be used as DoS tool to invalidate legitimate M.S.

e) *DoS attacks based on Reset Command (RES-CMD) message*

This message is sent to request a M.S to reinitialize its MAC state machine, it allow a B.S to reset a non-responsive or malfunction M.S [4]. This message is protected by HMAC but is still potential to be used to perform a DoS attacks [5]. To prevent DoS attacks, first need to fix the vulnerabilities in the initial network entry & secure authentication technique.

III. Solution suggested

1) *Time stamp model for secure Authentication*

Time stamping (T.S) is the process of securely keeping track of the creation and modification time of a document. Here Security means that no one — not even the owner of the document should be able to change it once it has been recorded provided that the timestamper's integrity is never compromised. This technique is based on digital signatures and hash functions. First a hash is calculated from the data. A hash is a sort of digital fingerprint of the original data: a string of bits that is different for each set of data. If the original data is changed then this will result a different hash value. Anyone trusting the timestamper can then verify that the document was not created after the date that the timestamper vouches. It can also no longer be repudiated that the requester of the timestamp was in possession of the original data at the time given by the timestamp. Steps are given bellow:

1. M.S sends communication request to B.S.

2. B.S generates the hash (H1) of the data & sends it to the M.S.

3. M.S now adds the T.S with H1 & generates hash H2.Then H2 is encrypted with the private key of M.S. Now send encrypted

H2 & T.S of M.S to B.S.

4. Now B.S has to add its data with the T.S of M.S & generate hash H3.Now decrypt H2 (Which was encrypted by private key of M.S) by the public key of M.S. If H3=H2 then continue further communication, otherwise cease the communication immediately. This model can also apply to eliminate the possibility of Man-In-The-Middle attack.

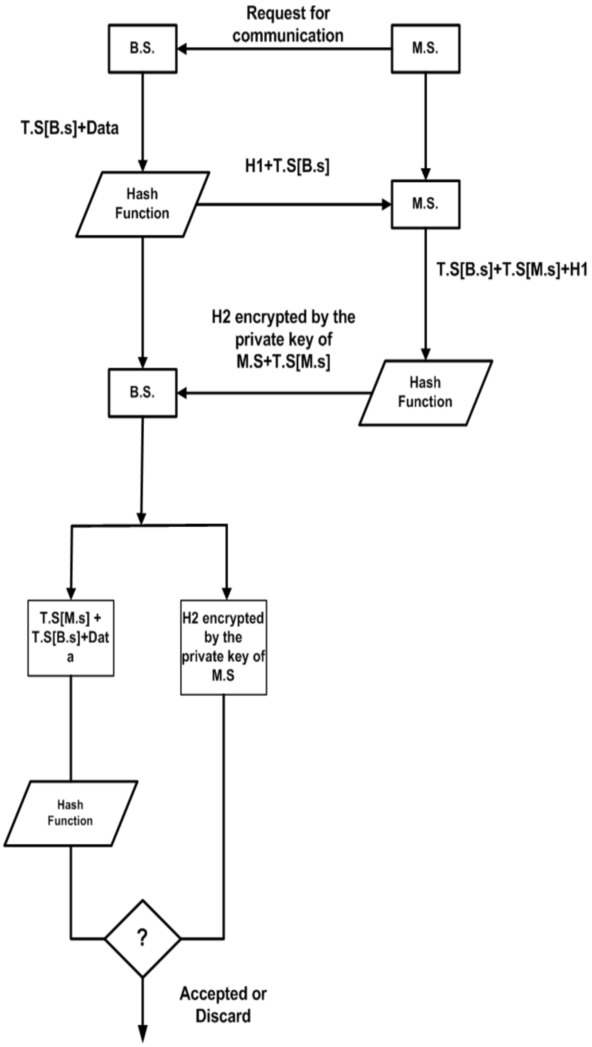

Request for communication
B.S.
M.S.
T.S[B.s]+Data
H1+T.S[B.s]
Hash Function
M.S.
T.S[B.s]+T.S[M.s]+H1
H2 encrypted by the private key of M.S+T.S[M.s]
B.S.
Hash Function
T.S[M.s] + T.S[B.s]+Dat a
H2 encrypted by the private key of M.S
Hash Function
?
Accepted or Discard

Figure 5: Secure Authentication process by using Time Stamp model

2) *DH- Diffie-Hellman Key Agreement*

It is a key agreement algorithm which helps M.S & B.S to agree on a shared secret between them without the need to exchange any secret/private information [13]. The DH standard is specified RFC 2631[6].

a) *Key agreement Algorithm*

To establishing shared secret between M.S & B.S, both must agrees on public constants p and g. Where p is a prime number and g is the generator less than p.
Step: 1 Let x and y be the private keys of M.S and B.S respectively, Private keys are random number less than p.
Step: 2 Let gx mod p and gy mod p be the public keys of devices M.S and B.S respectively
Step: 3 M.S and B.S exchanged their public keys.
Step: 4 The end M.S computes (gy mod p)x mod p which is equal to gyx mod p.
Step: 5 The end B.S computes (gx mod p)y mod p which is equal to gxy mod p.
Step: 6 Since K = gyx mod p=gxy mod p, shared secret = K.

b) *Mathematical Explanation- DH*

From the properties of modular arithmetic,
x mod n * y mod n ≡ x * y (mod n)
We can write;
(x1 mod n)*(x2 mod n)*… *(xk mod n) ≡ x1 * x2 * …* xk (mod n),
if xi=x, where i = 1, 2, 3… k (x mod n)k ≡ xk mod n , (gx mod p)y mod p = gxy mod p and (gy mod p)x mod p = gyx mod p, For all integers gxy=gyx, Therefore shared secret K=gxy
mod p=gyx mod p.
Since it is practically impossible to find the private key x or y from the public key [7] gx mod p or gy mod p, it is impossible to obtain the shared secret K for a attacker .

c) *One-Way function in DH*

For M.S, Let x be the private key and a = gx mod p is the public key, Here, a = gx mod p is one-way function. The public key a is obtained easily in the forward operation, but finding 'x' given a, g and p is the reverse operation and it will take exponentially longer time and is practically impossible. This is called discrete logarithm problem [7] .

3) *ECDH – Elliptic curve Diffie Hellman*

ECDH- a variant of DH, is a key agreement algorithm. To generate a shared secret between M.S and B.S using ECDH [14] [16], both have to agree up on Elliptic Curve domain parameters. An overview of ECDH is given below.

a) *Key Agreement Algorithm*

Establishing a shared secret between M.S and B.S
Step 1: Let dX and dY be the private key of M.S and B.S respectively, Private keys are random number which is less than n, where n is a domain parameter.
Step 2: Let QX = dX*G and QY = dY*G be the public key of M.S and B.S respectively, G is a domain parameter
Step 3: M.S and B.S exchanged their public keys
Step 4: The end M.S computes K = (aK, bK) = dX*QY
Step 5: The end B.S computes L = (aL, bL) = dY*QX
Step 6: Since K=L, shared secret is aK

b) *Mathematical Explanation (ECDH)*

To prove the agreed shared secret K and L at M.S and B.S
K = dX*QY = dX*(dY*G) = (dY*dX)*G = dY*(dX*G) = dY*QX = L, Hence K = L, therefore aK = aL Since it is practically not possible to find the private key dX or dY from the public key QX or QY, it is impossible to obtain the shared secret for a third party [15] [16].

4) *RSA*

It is a public key algorithm which is used for Encryption, Signature and Key Agreement. It (RSA) typically uses keys of size 1024 to 2048 [7]. The RSA standard is specified as

RFC 3447, RSA cryptography Specifications Version 2.1 [17][3] . Overviews of RSA algorithms are given below.

a) *Parameter generation*

Step 1: Consider two prime numbers a and b.
Step 2: Find n=a*b, Where n is the modulus which is made public. The length of n is considered as the RSA key length.
Step 3: Choose a random number 'e' as a public key in the range $0<e<(a-1)(b-1)$ such that $\gcd(e,(a-1)(b-1))=1$.
Step 4: Find private key d such that $ed\equiv1(\text{mod } (a-1)(b-1))$.
Encryption
Consider that B.S needs to send a message to M.S securely.
Step 5:. Let e be M.S's public key. Since e is public, B.S has access to e.
Step 6: To encrypt the message M, represent the message as an integer in the range $0<M<n$.
Step 7: Cipher text C = Me mod n, where n is the modulus.
Decryption
Step 8: Let C be the cipher text received from B.S.
Step 9: Calculate Message M = Cd mod n, where d is M.S's private key and n is the modulus.

b) *Key Agreement (RSA)*

Public key cryptography involves mathematical operation on large numbers and these algorithms are considerably slow compared to the symmetric key algorithm [7] [17]. They are so slow that it is unable to encrypt large amount of data. Public key encryption algorithm such as RSA can be used to encrypt small data such as keys which used in private key algorithm [7]. RSA is thus used as key agreement algorithm.
Key agreement algorithm:
Establishing shared secret between B.S and M.S
Step 10: Generate a random number, key to B.S.
Step 11: Encrypt by RSA encryption algorithm using M.S's public key and pass the cipher text to M.S.
Step 12: M.S decrypt the cipher text using M.S's private key to obtain the key.

c) *One-Way function in RSA*

Consider key generation equation Step 4, $ed\equiv1(\text{mod}(a-1)(b-1))$ and n=a*b ,Where e is the public key d is the private key. a and b are kept private but n is made public. Since e is public, anybody who has access to a and b could easily generate the private key d using the above equation in Step 4. The security of RSA depends on the difficulty to factorize n to obtain the prime numbers a and b [17] [7]. n is easily obtained by multiplying a and b but the reverse operation of factorizing n to obtain prime numbers a and b is practically impossible if a and b are large numbers.

D.4 Encryption process by using Public Key Cryptography:

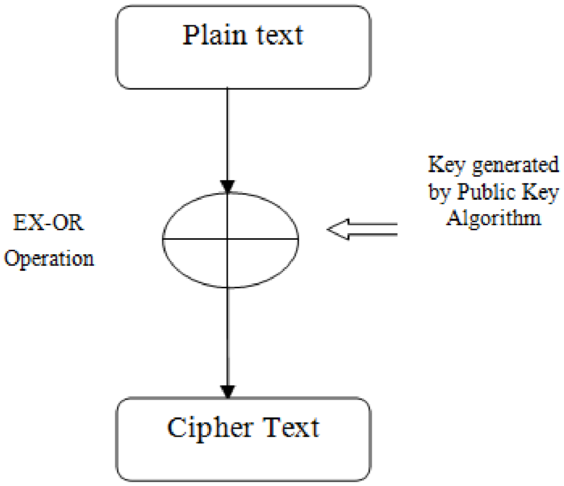


Figure: 6. Encryption Process by using Key, Generated by Public Key Algorithm.

This encryption will be symmetric key encryption process & and it is suggested to use 'Vernam Cipher' encryption process rather than DES or AES to encrypt initial management communication. Where key will be used as a random number for encryption .Because of the use of symmetric key encryption as well as Vernam Cipher which required only to performed bitwise Exclusive-OR operation, it will not introduce any traffic overhead in the network. Encryption process is described in figure 6

5) *Shared keys in Multi- and Broadcast Service*

Secure encryption of the data transferred via the MBS is very Difficult. A shared key cannot be used because every group member can forge messages when having the current symmetric keys [8]. But what can be avoided is the distribution of forged key update command messages allowing an attacker to take control over the data content on a M.Bs connection [8]. By using modular arithmetic we can solve this. We can generate GTEKs (Group Traffic Encryption Key) as part of a one way DH chaining function. But in one way DH function we can say that there is no inverse function or it is almost impossible to derive the inverse. B.S first generates a random number as initial key GTEK0 and other GTEKs are generated by applying a one way DH function to the previous GTEKs respectively. This is iterated n times.

GTEK0 = random ()
GTEK1 = f (GTEK0)
GTEK2 = f (GTEK1)
- -
GTEKn = f (GTEKn-1)

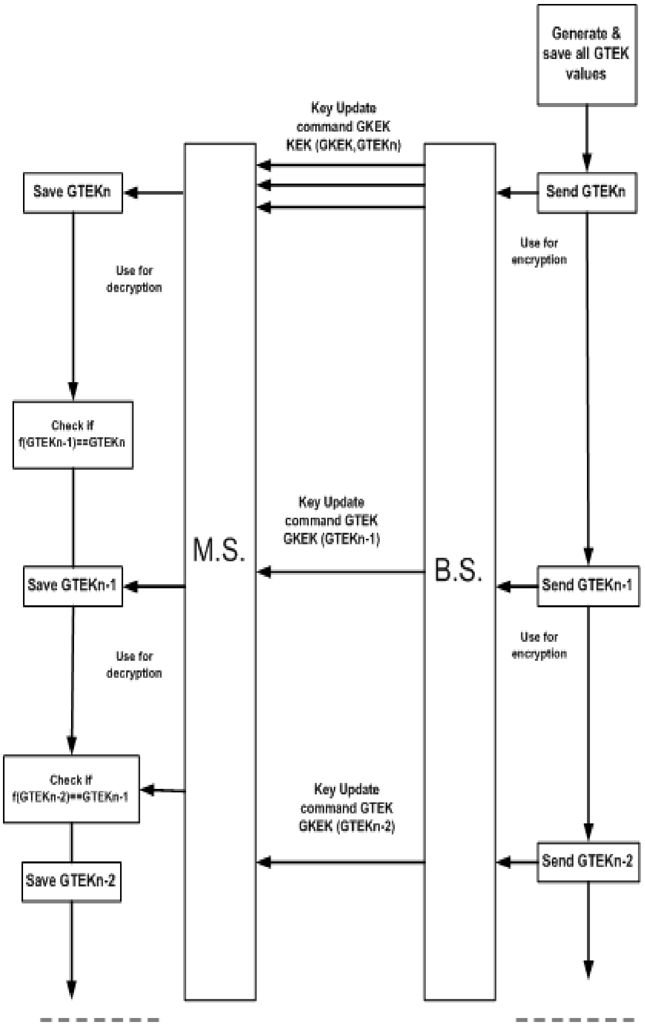

Generate & save all GTEK values
Key Update command GKEK KEK (GKEK,GTEKn)
Save GTEKn
Send GTEKn
Use for decryption
Use for encryption
Check if f(GTEKn-1)==GTEKn
Key Update command GTEK GKEK (GTEKn-1)
M.S.
B.S.
Save GTEKn-1
Send GTEKn-1
Use for decryption
Use for encryption
Check if f(GTEKn-2)==GTEKn-1
Key Update command GTEK GKEK (GTEKn-2)
Save GTEKn-2
Send GTEKn-2

Figure 7: Avoiding Key forgery by generating GTEK as a part of a one way chaining function

This DH chain function is used to verify each GTEK by applying the same one way function to the previous one. To achieve this chained authentication the last GTEK has to be distributed to each MS in a secure way as it is the only key in the chain which cannot be authenticated by another one [8]. To distribute GTEKn in the GKEK update command message is a unicast message and encrypted by a M.S related key. M.S can verify the integrity of new GTEK by applying the one way DH function to it [8]. If this authentication is positive, the current GTEK can be overwritten and the received one is established. If the authentication fails, the MS discards the message and requests a new GTEK [8]. To apply this algorithm, the key GKEK update command message must be capable of transporting GKEK and GTEK keys together [8]. Design of the key update command message already includes both keys so only a little modification is necessary here. Additionally, the GTEK state machine at B.S must generate the GTEK DH chain and store all the keys. The GTEK state machine at M.S must add the functionality to authenticate GTEK keys by calculating the DH function and comparing it to the previous key [8].

IV. Outcomes

Approaches	Existing System	After applying the proposed solution
Unauthenticated messages	Exist	Eliminated
Management massages	Unencrypted	Encrypted
DoS attack	Possibility High	Possibility Low
Shared key	Vulnerable	Secure
Eavesdropping	High	Low
Masquerading M.S: Identity theft B.S: Rogue station	Possible	Not Possible
Initial network entry	Vulnerable	Secure

V. Conclusion

Security mechanism is an expensive process; it requires extensive level of research, performance evolution & implementation outcomes. The IEEE 802.16e open the door for wireless mobility, vulnerability as well, because there are be no constraints for an attacker. In such a situation, more issues like, B.S to M.S key management, roaming user authentication & voice migration arise. IEEE 802.16 has the capability to attain success in wireless communication arena. But this technology is still under development and need more academic research and time to achieve a maturity level.

VI. Reference

1) Lang Wei-min, Wu Run-sheng, Wang jian-qiu : A Simple Key Management Scheme based on WiMAX: PLA Institue of Communication Command, Institute of Physics and Communication & Electronics.
2) Mir Md. Saki Kowsar, Muhammad Sakibur Rahman: Wi- MAX Security Analysis and Enhancement,Department ofComputer Science and Engineering Chittagong University of Engineering and Technology Chittagong-4349, Bangladesh.
3) Shinsaku Kiyomoto, Jun Kurihara, Toshiaki Tanaka, Andreas Deininger :Security Vulnerabilities and Solutions in Mobile WiMAX, KDDI R&D Laboratories, 2-1-15, Ohara, Fujimino-shi, Saitama 356-8502, Japan.
4) Trung Nguyen, Prof. Raj Jain : A survey of Wimax security threats.
5) Perumalraja Rengaraju, Chung-Horng Lung, Yi Qu, Anand Srinivasan," Analysis on Mobile WiMAX Security", IEEE TIC-STH 2009.
6) RSA Laboratories; http://www.rsa.com/rsalabs/node.asp?id=2193

7) RFC 2631, Diffie-Hellman Key Agreement Method, June 1999, Available at http://tools.ietf.org/html/rfc2631
8) Taeshik Shon, Wook Choi: An Analysis of Mobile Wi-MAX Security: Vulnerabilities and Solutions, First International Conference, NBiS 2007, LNCS, Vol. 4650, pp. 88-97,2007
9) Tao Han, Ning Zhang, Kaiming Liu, Bihua Tang : Analysis of Mobile WiMAX security: vulnerabilities and Solutions,Yuan'an Liu Key Lab. Of Universal Wireless Communications, Ministry of Education (Beijing University of Posts and Telecommunications)
10) Evren Eren : WiMAX Security Architecture – Analysis and Assessment, University of Applied Sciences Dortmund.
11) Charles P. Pfleeger, "Security in Computing" VOL No. 2
12) Datta A., He C., Mitchell J.C., Roy A., Sundararajan M.: 802.16e Notes, Electrical Engineering and Computer Science Departments, Stanford University, CA, USA, 2005, available: http://www.iab.org/liaisons/ieee/EAP/802.16e Notes.pdf Yuksel E.: Analysis of the PKMv2 Protocol in IEEE 802.16e-2005 Using Static Analysis Informatics and Mathematical Modeling, Technical University Denmark, DTU, 2007.[Online]. Available: http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=5159
13) RFC 2631, Diffie-Hellman Key Agreement Method, June 1999, Available at http://tools.ietf.org/html/rfc2631
14) Anoop MS, Elliptic Curve Cryptography - An Implementation Guide, January 2007, Available at http://hosteddocs.ittoolbox.com/AN1.5.07.pdf
15) Certicom, Standards for Efficient Cryptography, SEC 1:Elliptic Curve Cryptography, Version 1.0, September 2000, Available at http://www.secg.org/download/aid-385/sec1_final.pdf
16) Certicom, Standards for Efficient Cryptography, SEC 2: Recommended Elliptic Curve Domain Parameters, Version 1.0, September 2000, Available athttp://www.secg.org/download/aid-386/sec2_final.pdf
17) RSA Laboratories, PKCS#1 v2.1: RSA CryptographyStandard, June 2002, ttp://www.rsa.com/rsalabs/node.asp?id=2125

Analyzing Complexity for NP-Complete Problem Through DNA Computing Algorithm

Shalini Rajawat [1], Dr Vijay Singh Rathore [2],Naveen Hemrajani [1],Ekta Menghani[3]

GJCST Classification
F.1.3, F.2.1, F.2.2

***Abstract*- Adleman and Lipton adopted a brute-force search strategy to solve NP-complete problems by DNA computing i.e., a DNA data pool containing the full solution space must first be constructed in the initial test tube (*t*0), and then correct answers are extracted and/or false ones are eliminated from the data pool step by step. Thus, the number of distinct DNA strands contained in the initial test tube (*t*0) grows exponentially with the size of the problem. The number of DNA strands required for large problems eventually swamps the DNA data storage, which makes molecular computation impractical from the outset. Lipton's brute-force search DNA algorithm is limited to about 60 to 70 variables and thus it is believed that DNA computers that use a brute-force search algorithm can not exceed the performance of electronic computers. Since then, studies on DNA computing have focused on reducing the size of the data pool. A few new algorithms, such as the breadth-first search algorithm , Genetic algorithm , random walking algorithm , have been proposed and tested. With the breadth-first search algorithm, the capacity of a DNA computer can be theoretically increased to about 120 variables, but even so, DNA computers are still not capable of competing with electronic computers. Previously, we solved the SAT problem using a DNA computing algorithm based on ligase chain reaction. In the present study, we solve the SAT problem with the same DNA computing strategy using alternative biotechnical operations. Here we report some new results on the universality and space complexity of this DNA computing algorithm.**

Keywords-DNA computing, NP-Complete, space complexity, time complexity

I. DNA COMPUTING ALGORITHM

Without becoming too specific, we can assume that none of the clauses of *F* has both the positive form and negative form of the same variable and that *F* does not have two or more clauses consisting of the same three literals. The program for solving a 3-SAT problem with *n* variables and *m* clauses is shown in *Program. 1*. In the computing process, *tj* contains all of the sequences that satisfy clauses *C*1 to *Cj*. Strings that do not satisfy *C*1 to *Cj* can not be produced because the corresponding variable DNA is absent in *tjk*, or can not be amplified by PCR because they are broken by a restriction enzyme in *tjk*. After *m* steps of such operations guided by the SAT formula, all correct strings that satisfy all of the clauses will be generated.The computation time is *O* (9*m*+3*n*) because *Split*, *U-ligate*, *Cut*, *Amplify*, *Merge* and *Detect* commands are executed at most *m*, 3*n*, 3*m*, 3*m*, *m* and *m* times in the program, respectively.Therefore, the NP-complete problem can be solved in an amount of time that is proportional to the size of the problem.

II. IMPLEMENTATION OF THE ALGORITHM

Biotechnological implementation of the DNA algorithm is shown in Fig. (**1**). The commands are described in detail below:

(1) *PCR amplification of x*0 *v-xi v* was performed in a total volume of 50μL, using 100 nmol/L of each primer *P*0 and *Pi*, 10ng of ligation product *x*0 *v-xi v*, 200 mmol/L of each of the 4 dNTP, and 2.5U *Taq* DNA polymerase in 1X PCR Buffer supplemented with MgCl2
at a final concentration of 1.5 mM (all from Promega). Amplification was carried out on a Biometre T1 thermal controller as follows:
predenaturing at 94°C for 1 min, followed by 20 cycles of denaturing at 94°C for 20s, annealing at 62°C for 20s, and extension at 72°C for 20s, and a final extension at 72°C for 1 min.

(2) *U-ligation of variables xj v to x*0 *v-xi v* was performed in a volume of 20μL, containing 100ng PCR product *x*0 *v-xi v*, 1X PCR buffer and 2U *USER* enzyme (*NEB*). This mixture was incubated at 37□for 30 min to cut the uracil base. Next, 1μmol *xj v*, 1X *Taq* DNA ligase buffer and 80U *Taq* DNA ligase (*NEB*) were added,and the mixture was heated to 95 c for 5 min, gradually cooled to 55 c, and incubated at 55 c for 30 min to ligate *xj v* and *x*0 *v-xi v*.

(3) *Restriction cutting of x*0 *v-xi v*... was performed in a volume of 20μL containing 100ng PCR product *x*0 *v-xi v*..., 1X restriction buffer and 20U restriction enzyme (*NEB*) selected according to *Table* 1, and this mixture was incubated at the temperature recommended by the manufacturer for 60 min to cut strings containing *xi v*.

```
/* Program 1: Solve 3-SAT on a DNA computer */
Function DNA3SAT (F, xi, m, n)
Begin
Empty (t0) /* Begin with an empty test tube t0 */
/* Step 1 to m: for each clause of F, Cj =v3 k = (Ljk)*/
For j = 1 to m
Split (tj-1, tj1, tj2, tj3) /* Split test tube tj-1 equally to tj1, tj2, tj3 */
```

[1] About[1]-Department of Computer Science & Engineering, Suresh GyanVihar University, Jaipur, INDIA (rajawatshalini@yahoo.co.in; naven_h@yahoo.com
About[2]-Department of Computer Science & Engineering, Karni College, Jaipur INDIA. (vijaydiamaond@gmail.com)
About[3]-Department of Biotechnology, Mahatma Gandhi Institute of Applied Sciences, Jaipur, INDIA (ekta.menghani@rediffmail.com)

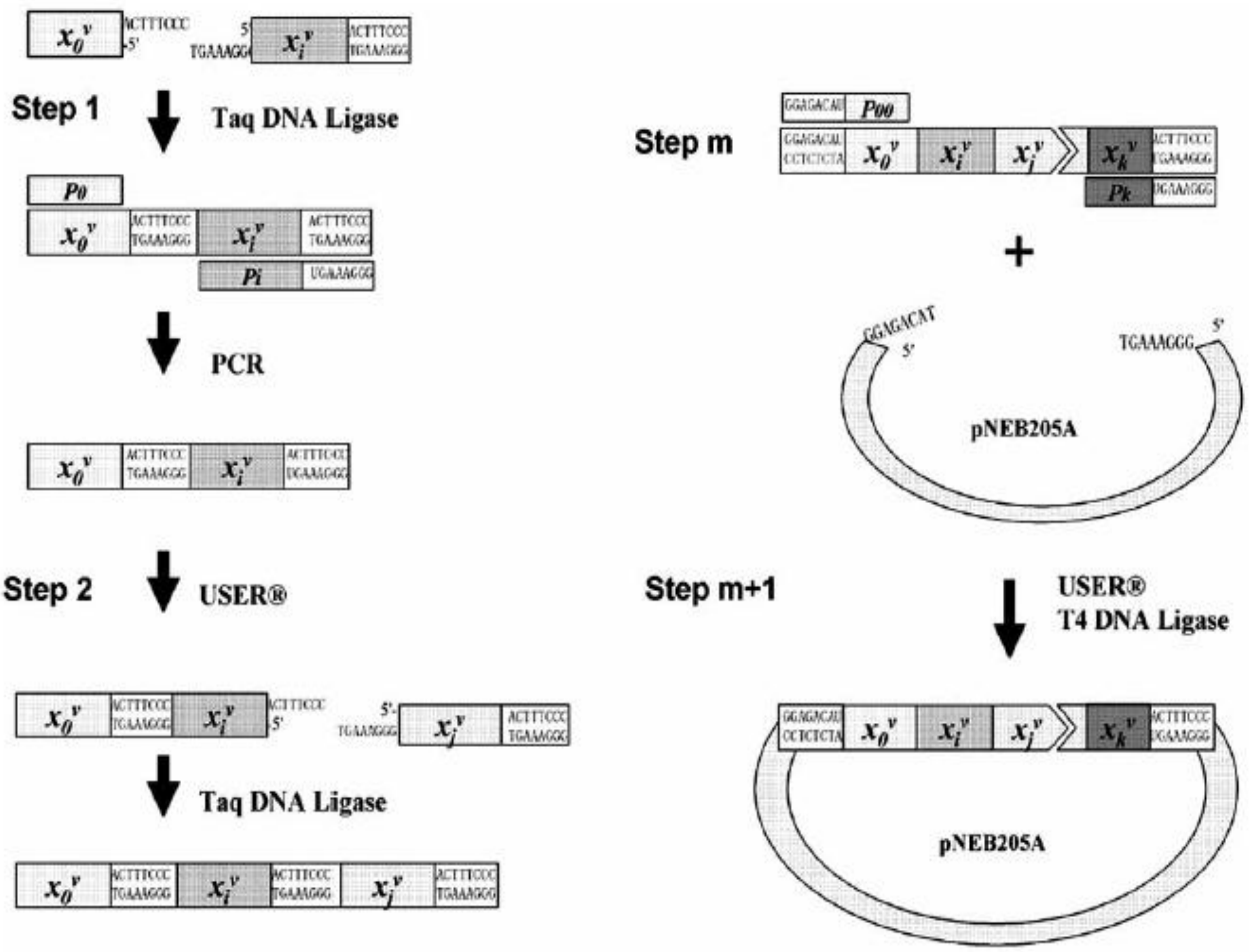


Fig. (1). Biotechnological operations in each step of the DNA computing process:

Step 1: *Ligating* of *x*0 *v* and *xi v* and PCR amplification of the product *x*0 *v*-*xi v*;
Step 2: *U-ligating xj v: USER-cutting* the ligation product *x*0 *v*-*xi v* to regenerate the sticky end and *ligating* of *xj v* to get *x*0 *v*-*xi v*-*xj v*;
Step *m*: At the end of computation, PCR is used to amplify the final answer DNA *x*0
v-*xi v*-*xj v*-…-*xk v* with primer *P*00 and *Pk*;
Step *m*+1: Cloning of PCR product of the final answer DNA into the pNEB205A vector.

```
For k = 1 to 3 /* for each literal of Cj */
i = index (Ljk)
If First_occurrence (F, xi) then /* If Ljk (xi or ~xi) is the
first occurrence
of xi in F */
For l = 1 to 3
U-ligate (tjl, xi
0); U-ligate (tjl, xi
1)
Next l


End If
Restriction_cutting (tjk, NOT Ljk) /* Cut NOT Ljk with
restriction
enzyme */
PCR_ amplification (tj1, tj2, tj3, P0, Pi)
/* PCR Amplify DNA in tj1, tj2 and tj3 with primer P0 and
Pi */
Next k
Merge (tj, tj1, tj2, tj3) /* Merge test tube tj1, tj2, tj3 to tj */
Next j
/*Step m+1: detect the result by sequencing*/
Return Detect (tm)
End Function
```

III. Results of the Lab Experiment

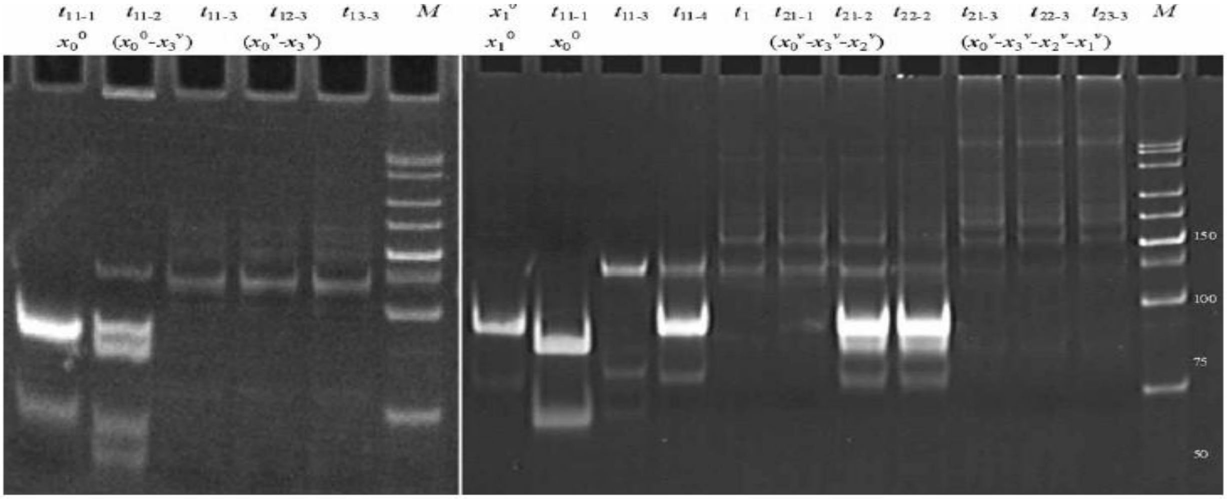


Fig. (2). Results of the DNA computing process (Steps **0** to **2**).

M is a 25-bp DNA ladder (MBI) DNA bands show *x*0 0 in test tube *t*11-1, *ligating* product in test tube *t*11-2 (*x*0 0- *x*3 *v*) and PCR products in test tubes *t*11-3, *t*12-3, *t*13-3, *t*11-3, *t*11-4, *t*1, *t*21-1, *t*21-2, *t*22-2, *t*21-3, *t*22-3 and *t*23-3, respectively.

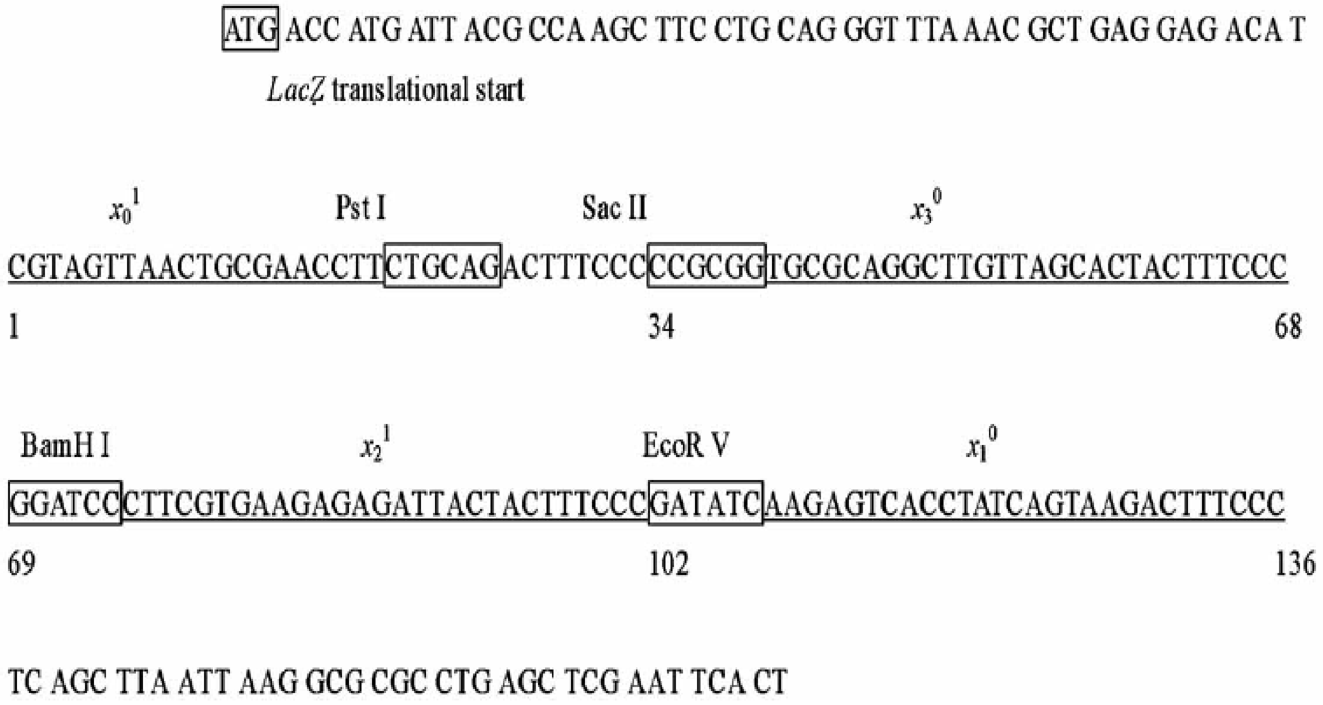

ATG ACC ATG ATT ACG CCA AGC TTC CTG CAG GGT TTA AAC GCT GAG GAG ACA T
LacZ translational start
x_0^1
Pst I
Sac II
x_3^0
CGTAGTTAACTGCGAACCTTCTGCAGACTTTCCCCCGCGGTGCGCAGGCTTGTTAGCACTACTTTCCC
1
34
68
BamH I
x_2^1
EcoR V
x_1^0
GGATCCCTTCGTGAAGAGAGATTACTACTTTCCCGATATCAAGAGTCACCTATCAGTAAGACTTTCCC
69
102
136
TC AGC TTA ATT AAG GCG CGC CTG AGC TCG AAT TCA CT

Fig. (3). Result of DNA cloning and sequencing of the final answer DNA. The variable names and values are marked, the binding positions of primers *P*0 to *P*3 are underlined, the restriction sites are boxed and marked, *Pst I-Sac II-BamH I-EcoR V*, and the unique answer is *x*0 1-*x*3

IV. Evaluation of the Space Complexity of the DNA Algorithm

For a given *m*-clause random SAT formula, *F*, the first *j* clauses is a SAT formula with *j* clauses, say *Fj*. In the computing process, when *j* grows from 1 to *m*, the number of different DNA strands in *tj*, say *Nj*, equals to the number of true assignment of *Fj*. The space complexity of this algorithm is the maximum number of DNA strands produced in test tubes *tj*, or the maximum number of partial assignments of *Fj*, say *max*{*Nj*, *j* = 1,...,*m*}, which is always smaller than the full solution space (2*n*).

We have examined the performance of our algorithm by computer simulation. We implemented on a HP Proliant workstation running a program that simulates our SAT solving algorithm on a family of random generated 3-CNF SAT formulas. In order to generate sample formulas, we wrote a program that give a range for the number of variables, *n*1 to *n*2, and a range for the clause/variable ratios, *r*1 to *r*2, constructs formulas of *n* variables and *m* clauses, where *n*€ [*n*1, *n*2], *m/n*€ [*r*1**,** *r*2]. When picking up a clause, three literals are repeatedly selected independently with equal probability, while keeping the clause free from complementary literals and identical literals. In both conventional and molecular computing studies on SAT problem, what we interested in are *hard* SAT problems. The clause/variable ratio of the hardest instances of 3CNF-SAT is around 4.3 [1, 16]. There are several algorithms that very efficiently solve the SAT problem if the clause/variable ratio is even slightly off from the critical point near 4.3. In order to test our SAT solving algorithm on both *easy* and *hard* problems, we generated at random fifty thousand instances of 3CNF-SAT problems with number of variables *n*€(5, 50) (more than one thousand for each *n*) and clause/variable ratio *m/n*€ (1, 50), and then investigated how the number of partial assignments changes while the algorithm runs. Because *cutting* operation helps decreasing of the partial assignments, and the more frequent a variable occur, the more *cutting* operation it brings to the computing. So we adopted a *cutting-first* strategy, the clauses are scored and sorted so that the *cutting* operations would happen as soon as a new variable is ligated to the solution.

Once a 3-SAT formula is generated, say *F*=*C*1^*C*2^…*Cm*, the clauses, *Cj*, are scored by, *Wj* = *π*3 *k*=1 log (*qi*), *j* = 1,…, *m*, Where *qi* is the occurrence number in *F* of the variable (*xi*) for theliteral *Ljk*. Then the clauses are sorted descendingly according to *Wj*, *F* was then transformed into an equivalent form, *F'* = *C*1'^*C*2'^…*Cm'*; where *Cj*€ [*C*1**,** *C*2, …, *Cm*], *W*1' >*W*2'>...>*Wm*'.Then we solve *F'* using the sequential version of our algorithm running on electronic computer and computes the maximum number of partial assignments (*γ*) that are required, outputs the exponent ratio *p*= (log2*γ*)/*n*. The average and maximum ratio *γ* for the maximum number of necessary partial assignments in solving random 3- SAT problems is shown in Fig. (**4**). The number of assignments in the initial pool generated by the brute force algorithm is 2*n*, so theratio *γ* for Lipton's brute force algorithm is a constant, □*Lipton*=1.0. The observed ratio *γ* for our algorithm decreases almost linearly with the increasing of *n* and *m/n* ratio. When *n* = 50, the overall average and maximum for the maximum number of DNA strands required is 20**.**4198*n* and 20**.**48*n*, respectively. If this relation 20**.**48*n* holds true or decreases further in 3-SAT instances with more variables,our algorithm will make solution of large and hard 3-SAT problemwith much smaller amount of DNA than the conventional bruteforce method. The observed average and maximum exponent ratio for this algorithm decreases logarithmically with the increasing of *n*. The regression equation of the maximum ratio to the number of variables *n* is,= 1.2902 -0.1788 *ln* (*n*) When *n* is set to be 100 and 200, the predicted maximum number of DNA strands required is 1.13E+14 and 4.39E+20, *i.e.*, the amount of DNA strands required are respectively within several nanomole and micromole. These requirements are surely possible with current biotechniques. If this relation holds true, this algorithm will make the solution of large 3-SAT problem possible with much smaller amount of DNA than the conventional brute-force method. Thus, based on the analysis in *section 3* and *section 6*, we proposed conjecture: For the class of SAT problems generated by our program for random generated 3-CNF SAT formulas can be solved on a DNA computer with time complexity O (9m) and space complexity2 [1.2902-0.1788 ln (n)]n.

V. Discussion

Even though the sample SAT problem solved here is very small, the proposed DNA computing algorithm has several advantages. Firstly, it eliminates the need to construct a full-solution 0-*x*2 1-*x*1 0. DNA library. The first test tube (*t*0) is empty instead of containing the full-solution data pool, and the other test tubes *tj* (*j*=1 to *m*) contain only strings that satisfy clauses *C*1 to *Cj*, which greatly reduces the number of DNA strands needed in the DNA computation, and makes it possible to extend this approach to solve large SAT problems, and possibly to other large NP-problems[2]. The maximum number of variables it can deal with depends mainly on how many cycles of *U-ligation* and *amplification* can be performed to extend the DNA strands without any serious error. In the present study, we performed 3 steps of extension and obtained a 4-word DNA solution. Although the process can theoretically proceed for as many steps as desired, the actual number of steps should be determined by further experiment in practice; Secondly, our DNA computation algorithm is error-tolerant. In this algorithm, we adopted *U-ligating*, *PCR amplifying* and *restriction cutting* as basic operations. As far as we know, ligase and restriction enzyme are both the most precise DNA operation enzymes available[3]. A one-base-pair mismatch in the restriction sites or in the sticky ends is enough to prevent them from cutting or ligating DNA molecules. The intrinsic highly accurate DNA sequence-recognition ability of DNA ligase and restriction enzyme makes them the most suitable tools for use in DNA computing. The same operations have been used successfully to solve the max clique problem [7], and those authors pointed out that the major errors in this computation arise from two sources. The first is the production of single stranded DNA (ssDNA) during PCR. This ssDNA cannot be cut by restriction enzymes. The second source of errors is incomplete cutting of double-strand DNA (dsDNA) by restriction enzymes, which also leads to incorrect answers. We used the selected restriction enzymes in 10-fold over digestion and found that they work well enough for our purpose in one cycle of digestion-PCR. Thanks to the combination of restriction digestion and PCR, this procedure gives an exponential amplifier with a larger exponent for uncut strands than for cut strands. Repeating the digestion-PCR process should therefore reduce the amount of noise arising from incomplete digestion [4-5]. In addition, the *U-ligating* operation we used to extend DNA strands not only helps to resist errors, but also increases the practical capacity of the DNA computer, since it is not only fast, easy and effective, but also prevents unwanted DNA strands from being generated by avoiding mistaken-ligation. Thirdly, the variables in the solution DNA are linked in the order of their position in the SAT formula instead of their indices. Compared with previous algorithms in which the variables are usually connected in the order of their indices [4-8], this feature of our algorithm makes it much easier to handle and possible to implement DNA computing without generating the full solution pool. In our algorithm, it is not necessary to sort the variables and literals, while we can reorder the clauses and the literals in any way to make the searching space smaller. As noted by Adleman [4], the information storage capacity of DNA is huge. In principle, 1 *mmol* of DNA can encode 2 gigabytes of data. The major advantage of DNA computing lies in its high parallelism. Our algorithms take advantage of the high information density and parallel computing capacity of DNA molecules, resembling *in vitro* evolution without generating an initial data pool that contains every possible answer; the number of DNA strands required increase exponentially with the size of the problem, but the observed average and maximum exponent ratio for this algorithm decreases logarithmically with the increasing of the number of variables (n). So our algorithm is more space efficient and can be scaled-up to solve large SAT problems. Unfortunately, the laboratory operations used in this algorithm are still very slow: it takes an average of about 30 min for each operation and 30 h in total to solve a small 3-SAT problem. Although the operations may be further optimized, it is still not yet possible to exceed the performance of electronic computers.

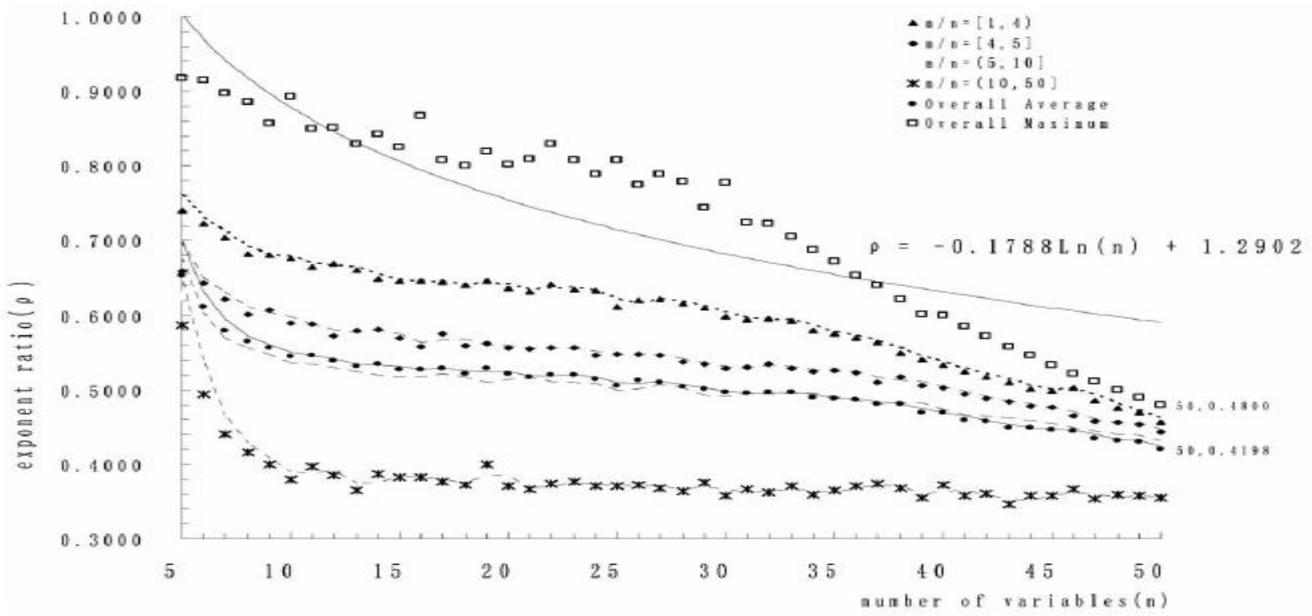


Fig. (4). Average and maximum exponent ratio for different *n* and *m/n* ratio. Data was calculated from fifty thousand random 3CNF–SAT instances with number of variables *n*=(5, 50) and clauses/variable ratio *m/n*=(1, 50).

VI. References

1) Selman B.; Mitchell D.; Levesque H. Generating hard satisfiability problem. *Artif. Intell.*, **1996**, *81*, 17.
2) Wang, X.; Bao, Z.; Hu, J.; Wang, S.; Zhan, A. Solving the SAT problem using a DNA computing algorithm based on ligase chain reaction. *Biosystems*, **2008**, *91*(1), 117.
3) Ouyang, Q.; Kaplan, P.D.; Liu S.; Libchaber, A. DNA solution of the maximal clique problem. *Science*, **1997**, *278*, 446.
4) Adleman, L. Molecular computation of solutions to combinatorial problems. *Science*, **1994**, *266*, 1021.
5) Lipton, R. Using DNA to solve NP-complete problems. *Science*, **1995**, *268*, 542.
6) Yoshida, H.; Suyama, A. DIMACS: Series in Discrete Mathematics and Theoretical Computer Science. Solution to 3-SAT by Breadth-First Search, American Mathematical Society, Providence, **2000**, *RI54*, 9.

7) Li, Y.; Fang, C.; Ouyang, Q. Genetic algorithm in DNA computing: A solution to the maximal clique problem. *Chinese Sci. Bull.*, **2004**, *49*(9), 967.

8) Liu, W.; Gao, L.; Zhang, Q.; Xu, G.; Zhu, X.; Liu, X.; Xu, J. A random walk DNA algorithm for the 3-SAT problem. *Curr. Nanosci.*, **2005**, *1*, 85.

An Implementation of IDATA - Intercept Detection Algorithm for Packet Transmission in Trust Architecture

GJCST Classification
C.2.1, E.3

Dr. S. N. Panda[1], Gaurav Kumar[2]

Abstract - **World is growing with the emerging technologies. The computer networks and packet transmission systems are also growing in parallel, hence to manage and provide security to packet, a secured system is required. Networks seize or simply intercept is one of the challenges in the fast growing world of Cyber Crime. The network establishments are facing various types of threats on routine basis. To efficiently transmit information across a network, there is need of an improved and reliable architecture. An intrusion or intercept refers to an active sequence of events that deliberately try to cause harm, such as rendering system unusable, accessing unauthorized information, or manipulating such information. Security professionals may want to have Intercept Detection Systems record information about both successful and unsuccessful attempts so that security professionals will have a more comprehensive understanding of the events on their networks. The intercept detection systems should be developed with utmost care to avoid any natural or intentional attempts. Moreover, the packet encryption algorithm should be developed in such a way so that cracker is not able to change even a single bit in the confidential data. This paper illustrates the implementation of IDATA - An Intercept Detection Algorithm in the Trust Architecture including the development and execution of a secured and efficient encryption algorithm for efficient and secured data packet transmission.**

Keywords - Cyber Security, Trust Architecture, Intercept Detection, Intrusion Detection, Trust Architecture, E-Transactions, Interception Analysis and Forensics, Packet Encryption, Packet Decryption, Packet Transmission

I. Introduction

Now days, the commercial as well as Defense Applications are facing frequent threats from different source and obviously such highly sensitive applications of public and national interest needs highly secured and consistent architecture so that packets can be transmitted in the network without any peril. Trust is considered as the footing of the relationship which is established by a business organization with their customers, vendors, and employees. All Trust Architectures and Intercept detection technology are not effective. These neither provided security to packet formation nor giving any security during transmission. All Trust Architecture developed till now doesn't provide absolute security and significant features. The VAN sometimes paralyzed and giving a great scope to the intruders/interceptors and other cyber criminals either to damage or alter or misuse the packets during transmission. Most of the fund transfer systems, EDI systems, business applications are using emerging technologies and exposed to vulnerability increases tremendously. Moreover, the cryptographic algorithms used during packet formation and transmission are sometimes responsible for vulnerabilities. Trust is the establishment of confidence that something will or will not occur in a predictable or promised manner. The enabling of confidence is supported by identification, authentication, accountability, authorization, and availability. To develop the trust between multiple parties, a set of principles or rules is to be offered so that the security of the entire model can be improved. According to the ITU-T X.509, Section 3.3.54, trust is defined as: "Generally an entity can be said to 'trust' a second entity when the first entity makes the assumption that the second entity will behave exactly as the first entity expects."

II. Intercept Detection Systems (IDS)

An intrusion-detection system (IDS) refers to the tools, methods, and resources to help identify, assess, and report unauthorized or unapproved network activity. The intrusion detection part of the name is a bit of a misnomer, as an IDS does not actually detect intrusions - it detects activity in traffic that may or may not be an intrusion. Intrusion detection is typically one part of an overall protection system that is installed around a system or device - it is not a stand-alone protection measure.

About[1]- Professor & Principal, Regional Institute of Management and Technology [RIMT],Mandi Gobindgarh, Punjab E-mail: panda.india@gmail.com

About[2]- Sr. Lecturer, Computer Applications Chitkara Institute of Engineering and Technology, Rajpura, PunjabE-mail: kumargaurav.in@gmail.com

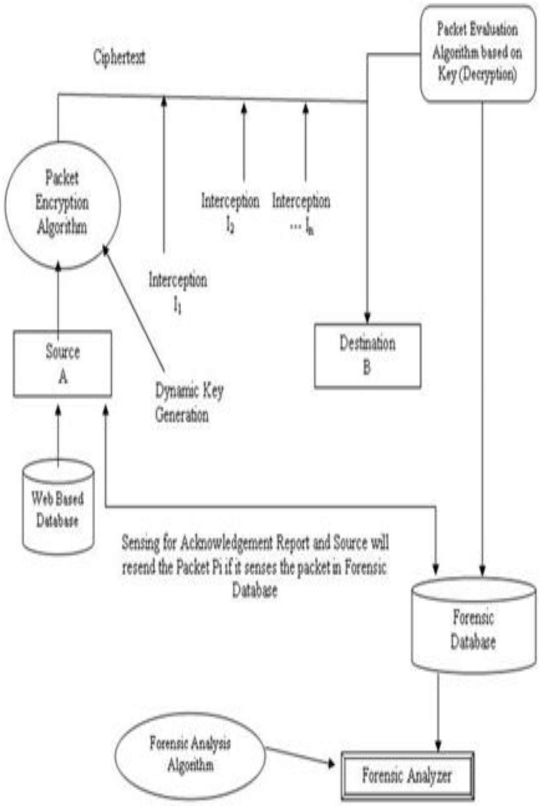

Packet Evaluation Algorithm based on Key (Decryption)
Ciphertext
Packet Encryption Algorithm
Interception I_2
Interception ... I_n
Interception I_1
Source A
Destination B
Dynamic Key Generation
Web Based Database
Sensing for Acknowledgement Report and Source will resend the Packet Pi if it senses the packet in Forensic Database
Forensic Database
Forensic Analysis Algorithm
Forensic Analyzer

Figure 1 : The Proposed Trust Architecture for Intercept Management

1) *Algorithmic Approach*

Step 1: Initialize & Activate Packet P_i at Source S_i for transmission to Destination D_i

Step 2: Packet Encryption Module PE_k based on Dynamic Key k Generation, once the Packet moves from Source S_i

$C_i := PE_k (P_i)$

Step 3: Transmission of Encrypted Packet C_i using specified Path/Route R_i

$C_i \rightarrow D_i [R_i]$

Step 4: Packet Authentication on Decryption

The most common types of threats fall into categories such as:

- Actual or attempted unauthorized probing of any system or data
- Actual or attempted unauthorized access
- Introduction of viruses or malicious code
- Unauthorized modification, deletion, or disclosure of data
- Denial of service attacks

III. Proposed Algorithms

The proposed architecture consists of various phases which will include algorithms for encryption and decryption of data packet alongwith the technique to analyze the overall interception patterns.

```
IF ( $C_i$ = $PD_k$ ($C_i$)    // Packet Decryption Module $PD_k$ to decrypt the packet at destination
        BEGIN
DEST [i] := $PD_k$ ($C_i$)
Successful Delivery of Packet
ACK sent to Source $S_i$    // Acknowledgement ACK is delivered to Source in case of Success
END
ELSE
BEGIN
A record will be inserted in the Forensic Database. The Interception Table will consist of the Structure (Id, Interception Type, Timestamp of Interception).    // Acknowledgement ACK is sent to Forensic Database in case of Failure Attempt
Source $S_i$ senses the Forensic Database.
        Select All Records from Forensic Database
        IF (true)Then
        print "Failure Delivery, Retransmit the packet"
GOTO Step 1
Update Forensic Analyzer Database for taking remedial actions.
END
```

Step 5: Forensic Analyzer

Retrieve Records for analysis of interceptions. In the proposed architecture, an extended flavor of link level encryption will be used to encrypt the entire data packet.The packet encryption algorithm at the originating site encrypts the entire packet including the packet header and

1) *Encryption Algorithm*

Step 1: Activate and Initialize the Packet P_i

Step 2: Generate a Random Key K_R by analyzing number of 1s in Packet.

(a) Develop a routine to count bits in the Data Packet

(b) Set N := Count(P_i) // Count Number of 1's in the Data Packet.

2) *Encryption Algorithm*

Step 1: Activate and Initialize the Packet P_i

Step 2: Generate a Random Key K_R by analyzing number of 1s in Packet.

(a) Develop a routine to count bits in the Data Packet

(b) Set N := Count(P_i) // Count Number of 1's in the Data Packet.

IV. Decryption and Intercept Detection Algorithm

A decryption algorithm at the destination site will check the entire encrypted packet. The received packet will be of specific format and structure in which key is given. By

Analyze the type T_i of Intercept

Perform remedial stroke for avoiding the stored interception type

The proposed architecture consists of various phases which will include algorithms for encryption and decryption of data packet alongwith the technique to analyze the overall interception patterns.

V. PACKET ENCRYPTION ALGORITHM

At the initial stage, the data packet will be transmitted from source to destination over transmission media using efficient cryptographic algorithm to encrypt the entire packet. Cryptography is the process used to make a meaningful message appear meaningless. An algorithm is a set of rules or procedures used to scramble, or encrypt the plaintext to produce Ciphertext. The algorithm applies a key to text.Encryption is the procedure that guarantees secrecy of the data exchanged. Any encryption algorithm depends on some key, and keys are normally generated during authentication phase, so the two phases are strictly connected. provides it a new header. This readable new header also includes a dynamic key-id. The key-id controls the behavior of encryption and decryption mechanism. It specifies the information as the encryption algorithm, the encryption block size, the error checking code and lifetime of the key.

(c) Set K_R :=N // Store N in Random Number K_R

Step 3: Apply XOR (Exclusive-OR) Operation

(a) Set $E_{K\,:}= P_i \oplus_R$

(b) The Encrypted Packet E_K is generated using XOR Operation.

(c) Set $PE_K := E_K$ // Utilize E_K as Encrypted Packet

Step 4: Packet equipped for Transmission

analyzing the structure of encrypted packet, the location of key will be accessed and the packet can be decrypted. In case of interceptions at the transmission line, the details of such attempts will be stored in the web based databases so that interception points and sources can be identified. In case, there is an interception and packet is not matched after decrypting the Ciphertext C_p, a record will be inserted in the forensic database. The pattern/behavior of intercepts will be analyzed using a forensic analyzer. In case of successful decryption and transmission of packet, an acknowledgement will be transmitted to the web based

1) *Algorithm*

```
Step 1: Receive the Encrypted Packet $PE_K$
Step 2: Check the Front $PF_i$ and Rear End $PR_i$ of Packet
        if ($PF_i = PR_i$)
        Accept PFi
        Set $K_R := PF_i$
        else
        goto Step 5
Step 3: Generate the Binary Equivalent of $K_R$
        $PB_i = Binary(K_R)$
Step 4: Perform XOR Operation
        if ($PB_i = PE_K$)
Decryption Successful Accept the Packet Else goto step 5
Step 5: Insert the Record of Corrupt Packet in Forensic Database
```

VI. IMPLEMENTATION AND SIMULATION

The cryptographic algorithm based on XOR logic gate is implemented and simulated and tested using the following

MS Windows Platform

Windows XP,
Turbo C IDE

Linux Platform

Fedora 11
gcc

SOURCE CODE IN C

```
int a[40], b[40], c[40], decrypted[40], choice;
char d[40];
FILE *key, *source, *f1, *f2;
char buf[5];
void main()
{
FILE *fp;
int random, i;
clrscr();
fp=fopen("packet.txt","r");
source=fopen("sourcebin.txt","w+");
key=fopen("keybin.txt","w+");
randomize();
random=rand()%5;
fseek(fp,5*random,0);
fread(buf,5,1,fp);
printf("\n\t\tPacket Received : %5s\n",buf);
printf("Packet\tASCII\tBinary Eq.\t1's\tDynamic Key\n");
buf[5]='\0';
```

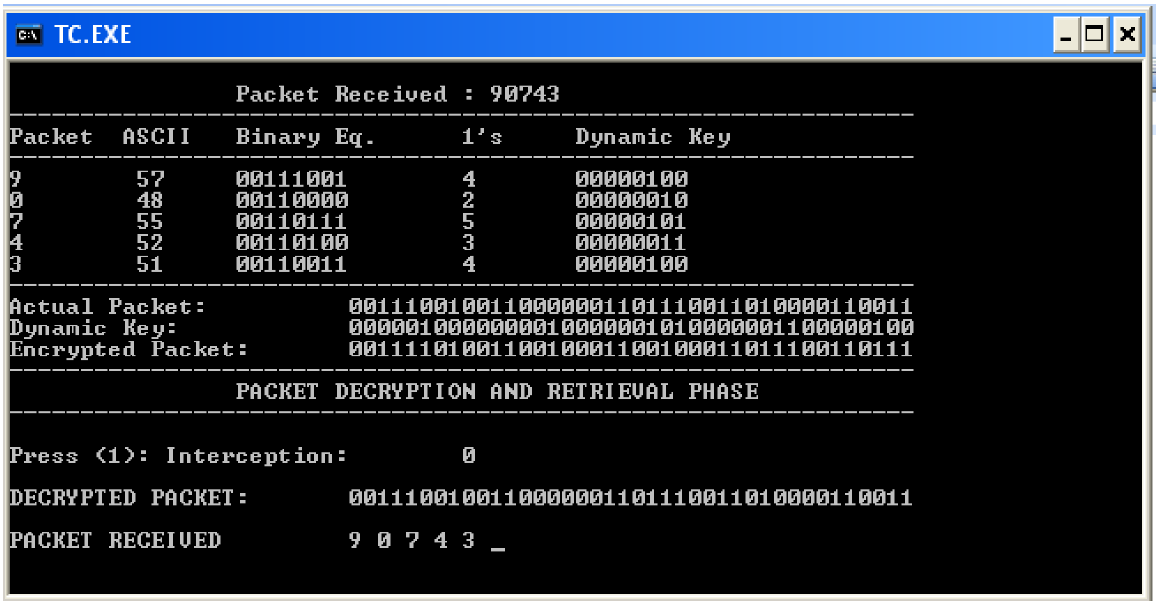
database where the source site can verify the delivery of message.

```
for (i=0;i<5;i++)
{
delay(50);
printf("\n%c\t%3d\t",buf[i], buf[i]);
decbin(buf[i]);
printf("\t");
saveresults_a(buf[i]);
printf("%d",countdecbin(buf[i]));
printf("\t");
decbin(countdecbin(buf[i]));
saveresults_b(countdecbin(buf[i]));
}
fclose(source);
fclose(key);
xor_op();
delay(200);
printf("\t\t");
textattr(128+10);
decrypt_packet();
for (i=0;i<40;i++)
{
if (a[i]!=decrypted[i])
{
```

```
printf("\n\t\t\t");
textattr(128+YELLOW);
cprintf("Interception ! Packet sent to Forensic Database");
}
}
getch();
}
int decbin(int number)
{
int x, y;
x = y = 0;
for(y=7; y>=0; y--)
{
  x = number / (1 << y);
  number = number - x * (1 << y);
  printf("%d",x);
}
return 0;
}
int saveresults_a(int number)
{
int x, y;
x = y = 0;
for(y=7; y>=0; y--)
{
  x = number / (1 << y);
  number = number - x * (1 << y);
  fprintf(source, "%d ", x);
}
return 0;
}
int saveresults_b(int number)
{
int x, y;
x = y = 0;
for(y=7; y>=0; y--)
{
  x = number / (1 << y);
  number = number - x * (1 << y);
  fprintf(key, "%d ", x);
}
return 0;
}
int countdecbin(int number)
{
int x, y, count=0;
x = y = 0;
for(y = 7; y >= 0; y--)
{
  x = number / (1 << y);
  number = number - x * (1 << y);
  if (x==1)
  {
  count++;
  }
}
return count;
```

```
}
int xor_op()
{
int i=0, j=0;
f1=fopen("sourcebin.txt","r");
f2=fopen("keybin.txt","r");
do
{
   fscanf(f1, "%d", &a[i]); i++;
} while (i<40);
do
{
   fscanf(f2, "%d", &b[j]); j++;
} while (j<40);
printf("Actual Packet: \t\t");
for (i=0;i<40;i++)
{
delay(50);
printf("%d",a[i]);
}
printf("\nDynamic Key: \t\t");
for (i=0;i<40;i++)
{
delay(50);
printf("%d",b[i]);
}
fclose(f1);
fclose(f2);
for (i=0;i<40;i++)
{
if (a[i]==b[i])
c[i]=0;
else
c[i]=1;
}
printf("\nEncrypted Packet:\t");
for (i=0;i<40;i++)
printf("%d",c[i]);
return 0;
}
int decrypt_packet()
{
int i=0;
int p;
char bb[40];
char xx='1', yy='0';
int k=0, j=8, x;
for(i=0; i<40; i++)
{
if (c[i]==b[i])
{
d[i]=yy;
decrypted[i]=0;
}
else
{
d[i]=xx;
decrypted[i]=1;
}
}
d[40]='\0';
printf("\n");
textcolor(WHITE);
textbackground(BLUE);
cprintf("Press (1): Interception:");
printf("\t");
scanf("%d",&choice);
if (choice==1)
{
decrypted[10]=2;
d[10]='.';
}
printf("\nDECRYPTED PACKET:\t");
for (i=0;i<40;i++)
{
delay(50);
printf("%c",d[i]);
}
printf("\n\nPACKET RECEIVED\t\t");
i=0;
while(i<=40)
{
for (p=0,i=k;i<j;i++,p++)
{
bb[p]=d[i];
}
textattr(128+10);
cprintf("%c ",bin2dec(bb));
k=k+8;j=j+8;i=i+8;
}
return 0;
}
int bin2dec(char *bin)
{
int b, k, m, n;
int len, sum = 0;
len = strlen(bin) - 1;
for(k = 0; k <= len; k++)
{
n = (bin[k] - '0'); // char to numeric value
for(b = 1, m = len; m > k; m--)
{
// 1 2 4 8 16 32 64 ... place-values, reversed here
b *= 2;
}
// sum it up
sum = sum + n * b;
}
return(sum);
}
```

VII. Results Obtained



```
                Packet Received : 90743
----------------------------------------------------------------
Packet   ASCII    Binary Eq.       1's      Dynamic Key
----------------------------------------------------------------
9         57      00111001         4        00000100
0         48      00110000         2        00000010
7         55      00110111         5        00000101
4         52      00110100         3        00000011
3         51      00110011         4        00000100
----------------------------------------------------------------
Actual Packet:            0011100100110000001101110011010000110011
Dynamic Key:              0000010000000010000001010000001100000100
Encrypted Packet:         0011110100110010001100100011011100110111
----------------------------------------------------------------
                PACKET DECRYPTION AND RETRIEVAL PHASE
----------------------------------------------------------------

Press (1): Interception:         0

DECRYPTED PACKET:         0011100100110000001101110011010000110011

PACKET RECEIVED           9 0 7 4 3 _
```

VIII. Conclusion

The business, defense and government applications are required to be deployed in a highly secured and confidential environment which needs secured architecture as well as an efficient method of data packet encryption. Different types of networks are in front of number of threats from increasing intercepts originating from various sources. To secure these applications from unauthorized and illegal access, there is need to secure the network from multiple interceptions using efficient algorithms. The implementation demonstrated in this paper illustrates an efficient algorithm based on Exclusive-OR operation which is a unique method to encrypt any data packet traveling in the network. Using this method, encryption, decryption and traveling of packet can be performed effectively without any complexity. Moreover, the forensic database module keeps track of every invalid or unacceptable decrypted packet. With the prior information of unauthorized access of records in the database, the behavior of intercepts can be analyzed to avoid such attempts in future.

IX. References

1) Cochavy, Baruch, Method of efficiently sending packets onto a network by eliminating an interrupt, US Patent 5797039 Issued on August 18, 1998
2) Dimitris M. Kyriazanos, Neeli R. Prasad, Charalampos Z. Patrikakis, A Security, Privacy and Trust Architecture for Wireless Sensor Networks, 50th International Symposium ELMAR-2008, 10-12 September 2008, Zadar, Croatia
3) Dr. S. N. Panda, Gaurav Kumar, “IDATA – An Effective Intercept Detection Algorithm for Packet Transmission in Trust Architecture” (POT-2010-0006), selected for publication in IEEE Potentials ISSN: 0278-6648.
4) Dr. S. N. Panda, Gaurav Kumar, “Effective Implementation Of Intruder Detection Trust Architecture Using XOR Logic Gate Cryptographic Technique” published in Journal of Ultra Scientist of Physical Sciences, Bhopal, ISSN 0970-9150, Vol. 22 No. 1, April 2010.
5) Donna Andert, Robin Wakefield, and Joel Weise, Professional Services Security Practice, Sun BluePrints™ OnLine - December 2002, Trust Modeling for Security Architecture Development
6) Security, Encryption, Acceleration, http://www.networkintercept.com
7) Youlu Zheng, Shakil Akhtar, Networks for Computer Scientists and Engineers, Oxford University Press, 2009
8) Carl Endorf, Eugene Schultz and Jim Mellander, Intrusion Detection & Prevention, McGraw-Hill, 2004

Cell Segmentation from Cellular Image

Mohammad Ibrahim Khan , Muhammad Kamal Hossen , Md.Sabbir Ali

GJCSTClassification J.3

***Abstract*-To understand the cell movement and cell behavior into different parts of organs in human or animal body, it is necessary to study the cells in culture medium. Fluorescence microscopy is an emerging tool for acquiring this cellular images. The large number of cellular images produced by fluorescence microscopy is unmanageable for human to analyze them manually. Thus, cellular image segmentation is a primary requirement for higher level analysis of medical diagnosis and research. In this paper, a fully automatic method for segmentation of cells from fluorescent microscopy images is proposed. The method first mark the probable foreground and background seeds planted on the image. Then it applies a popular segmentation method, namely watershed segmentation, based on these seeds. The result is further refined based on the gradient value along the initial segmented lines and then again performing watershed transform on the distance transformed image of the previously founded result. The experimental results shows that this approach for segmenting cell images is both fast and robust.**

I. Introduction

In biology, a common problem is the segmentation of cells for counting and feature extraction purposes. Large scale quantification of the dynamical behavior of cell populations in a variety of experimental systems would provide important capabilities for many areas of cell biology. In addition, such quantitative cellular data would be useful in many theoretical contexts.Segmentation is the partitioning of a scene into different meaningful regions, by identifying regions of an image that have common properties while separating regions that are dissimilar [1]. It is often the first, most vital, and most difficult step in an image analysis task. The result of the segmentation usually determines eventual success of the final analysis. For this reason, many segmentation techniques have been developed by researchers worldwide, and there exist almost as many segmentation methods as there are segmentation problems.Cell image segmentation is a necessary first step of many automated biomedical image processing procedures. There certainly has been much research in the area. Automatic segmentation of cell nuclei from fluorescence microscopic cellular images allows the study of individual cell nuclei within their natural tissue context. Compared with manual methods based on drawing the outlines of the nuclei with a mouse, automatic methods need far less interaction, and the result is more objective and easily reproduced. Automation also increases the amount of data that can be processed. Once the objects of interest have been delineated, a large number of descriptive features can be extracted from the objects [2].

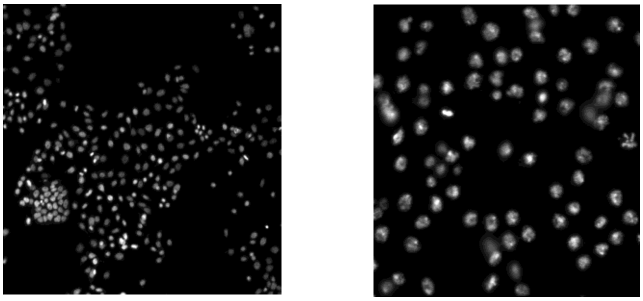

(a) (b)

Fig. 1. Two fluorescent microscopic cellular images (a) human colon cancer cells and (b) drosophila's cells

Cell segmentation is one of the most challenging problems due to both the complex nature of the cells and problems inherent to video microscopy. Segmentation of cells from cellular image is quite difficult for the following reasons:-

1. since the cells are unstained, as the stain would be harmful to the living cells. The contrast is thus quite low.
2. The cell images are also acquired with auto-focus, which sometimes yields a poorly focused image.
3. In tissue culture environment, cells are non-rigid , irregularly shaped bodies. The cells external environment influences their shapes, which in turn affects their locomotory behavior and ultimately how they function.
4. Cells that naturally migrate within organisms can take on a variety of different sizes and shapes, and can migrate at different rates, depending on their current functional state. As the cell changes shape during locomotion, the contrast between the cell boundary and the background varies continually.
5. In addition, equipment related factors which contribute to the quality of the image, such as uneven illumination and electronic or optical noise, also play an important role in the effective segmentation of a digital image.

There are many algorithms used for cell segmentation, and some of them segmented an image based on the object while some can segment automatically. Now-a-days, no one can point out which the optimal solution is due to different constraints. A model-based contour tracing approach used to the problem of automatically segmenting a Scanning Electron Microscope(SEM) image of cells [3]. This method forces the contours to be smooth by using a model-based approach, such as matching ellipses to edge data. It define gradient vector for each pixel in the gradient map image. It traces the contours of the cell by using history-based

About- Department of Computer Science & Engineering, Chittagong University of Engineering & Technology Chittagong-4349, Bangladesh muhammad_ikhancuet@yahoo.com kamal_cuet@yahoo.com sabbircse05@yahoo.com

prediction & data- based prediction. Finally it use culling algorithm for reducing noise contours or noise within cell. However, there are some non-elliptic cells, which can not be detected correctly and efficiently using this method.Another popular method is to use thresholding , based on histogram characteristics of the pixel intensities of the image[4]. In order to obtain a satisfactory segmentation result by thresholding, a uniform background is required. Many background correction techniques exist[5], but they may not always result in an image suitable for further analysis by thresholding. So, one have to use thresholding technique in a different way. The [6] developed a multistage segmentation strategy, using two image features associated with cell regions, namely, intensity level and local variation of intensity. The first step applies a global threshold to the local variation of intensity. This step segments a region of the image, consisting of a cell and the nearby surrounding background, from the distant background. The region of the image that is segmented is referred to as the approximate region. It is then further segmented by applying a global intensity threshold to the approximate region. At this stage, the cell has been segmented from the background. Smoothing and filling schemes are implemented to obtain a cell boundary representation. The problem in this approach is, it assumes that the gray levels of the object and background are normally distributed but in reality, this may not happened because- (1) The transition between object and background may be diffuse, making an optimal threshold level difficult to find. (2) The image background intensity is often uneven due to auto fluorescence from the tissue and fluorescence from out-of-focus objects. A popular region growing method, which has proved to be very useful in many areas of image segmentation and analysis, is the so-called watershed algorithm[7]. If the intensity of the image is interpreted as elevation in a landscape, the watershed algorithm will split the image into regions similar to the drainage regions of this landscape. The watershed borders will be built at the crests in the image. In a gradient magnitude image, water will start to rise from minima representing areas of low gradient, i.e. the interior of the objects and the background, and the watershed borders will be built at the maxima of the gradient magnitude. However, if watershed segmentation is applied directly to the gradient magnitude image, it will almost always result in over-segmentation, owing to the intensity variations within both objects and background. So, one have to apply watershed algorithm in a different way. The [8] combine watershed algorithm with thresholding technique for segmentation efficiency. The segmentation of the image is implemented in three levels. Initial automatic segmentation is taking place at first level, where a fuzzy threshold is performed on the image and then a fuzzy gray weighted distance transform is applied. Then it uses the extended h-maxima transform[9] to find suitable seed points for the watershed algorithm. Segmentation on poorly focused images is in second level where it use a fast geometric active contour model based on the level set algorithm. However, The problems of segmenting clustered objects and choosing a suitable threshold level for objects with unsharp edges will remain. It also use complex geometric computation.Edge-based segmentation techniques, which try to connect local maxima of the gradient image, often run into problems when trying to produce closed curves. That is why region-based methods, such as region growing or watershed, that group similar pixels are often used. Another group of methods that do not have the problem of being required to produce closed curves are methods related to snakes or active shape models. From a rough marking of the border or a seed inside the object of interest a curve expands until it finds a strong edge. The function describing the expansion consists of different energy terms attracting the curve to edges. The approach with expanding curves has been used for cell nuclei segmentation[10]. The problems with this method are defining suitable energy terms and, again, the problem of constructing automatic seeding methods, which are restricted to one unique seed per nucleus.There doesn’t exist any segmentation method that will alone produce a satisfactory result on images of fluorescence-stained nuclei in tissue if (1) the nuclei are clustered, (2) the image background is variable and (3) there are intensity variations within the nuclei. Thus our proposed segmentation method give careful attention for the above facts to get efficient and correct segmentation result. We add to this a new method that automatically extracts cells from microscopic imagery, and does so in four phases. Phase 1 applies some morphological operations on the cellular image to identify and mark foreground and background seeds of objects with an overall accuracy of >97%. Phase 2 of the method uses a well known segmentation algorithm, called Watershed Algorithm, on this seeds of objects to identify cells, quickly but results some over-segmentation. In phase 3, we refine the previous result by using edge strength of the cells. In the final phase, we further refine the result based on the shape of the cells. The method takes only four input parameters and takes less than 1 minute to operate on a microscopic image.

II. Segmentation Strategy

Our new improved cell segmentation method requires few input parameters and gives stable results. Morphological filtering on the intensity image is used for finding object seeds. Morphological filtering of the gradient magnitude image is used for finding background seeds. Seeded watershed segmentation is then applied to the gradient magnitude image, and region borders are created at the crest lines in the gradient image. More than one seed in an object means that the object will be divided into more than one region, i.e. we will start with over-segmentation. After watershed segmentation, we merge neighbouring objects and only keep those borders that correspond to strong edges. This step will also remove objects with poor contrast. If the nuclei are tightly clustered, no edge is present where they touch, and they will therefore not be separated. Objects found by the first steps of the segmentation process are further separated on the basis of shape. Shape-based cluster separation using the distance transform is applied to all objects found by the previous steps, but only those separation lines that go through deep enough valleys in the distance map are kept. After this step, we get our desired segmented cellular image. Fig. 2 shows the block diagram of our segmentation method.

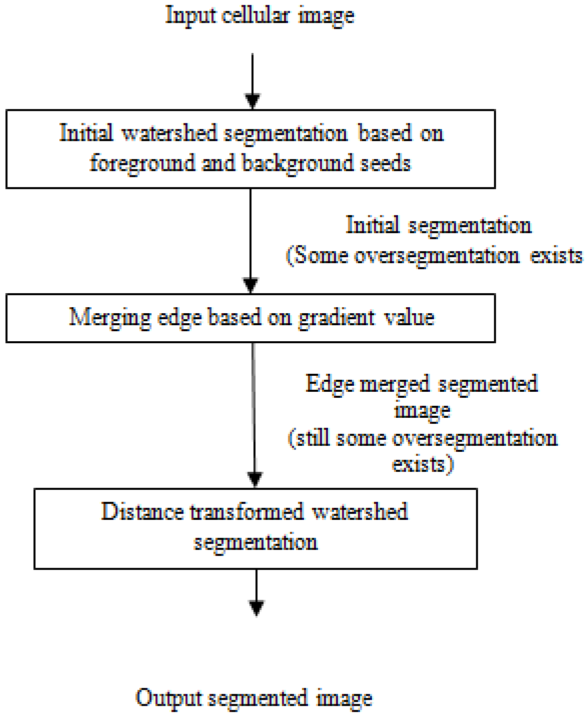

Input cellular image
Initial watershed segmentation based on foreground and background seeds
Initial segmentation
(Some oversegmentation exists
Merging edge based on gradient value
Edge merged segmented image
(still some oversegmentation exists)
Distance transformed watershed segmentation
Output segmented image

Fig. 2. Block diagram of our new improved cell segmentation strategy

Methodology of our new improved segmentation algorithm is given below:-

1) Firstly, we need to remove noise from the image. Thus, the image of size 512*512 was smoothed by a 3*3 Gauss filter. No other pre-processing was necessary. (Noise is a random unpatterned variation of intensity in the image or it is think of as unwanted pixels in the image that do not part of any foreground/background object)

2) Then we need some seeds that marks probable foreground regions(cells) planted on the image. The images, we consider here, contains bright objects on a darker background. Hence, each object of interest contains at least one local intensity maximum. We define foreground seeds in the original image using the extended h-maxima transform. The extended h-maxima transform filters out the relevant maxima using a contrast criterion. All maxima whose heights are smaller than a given threshold level h are suppressed. The extended h-maxima transformation can be implemented using sorted pixels and searching for local maxima with a given contrast compared with the local neighbourhood.

3) Now we need to define our background seeds. For this purpose we have done the followings:-

i) First, we have find the gradient magnitude of the original image.(Gradient magnitude image is those that have a high intensity value in the edge of a object and low intensity value on the other places of the image.)

ii) Just as the objects can be seeded by extended h-maxima in the original image, the background can be seeded by extended h-minima in the gradient magnitude image, i.e. local minima deeper than a certain depth h .This step will also remove the small components.

4) These seeds serve as starting points in the watershed algorithm applied to the gradient magnitude image. Watershed segmentation can be understood by interpreting the intensity image as a landscape. A hole is drilled in every minima of the landscape, and the landscape is submerged in water. Water will then start to fill the minima, creating catchment basins. As the water rises, water from neighbouring catchment basins will meet. At every point where two catchment basins meet, a dam, or watershed, is built. These watersheds are the segmentation of the image. Watershed segmentation can be implemented with sorted pixel lists.

5) Now we need to merge the area with weak borders. If too many seeds are created in the seeding step, some objects will have more than one seed. These objects will be oversegmented after the watershed algorithm, because each seed results in one region. However, if two seeds are in the same object, the magnitude of the gradient at the region Boundaries will usually be low. Thus, by comparing the gradient magnitude image and previously founded segmented image, we remove those segmented lines where gradient value is low. So, in this way we remove oversegmentation/under segmentation.

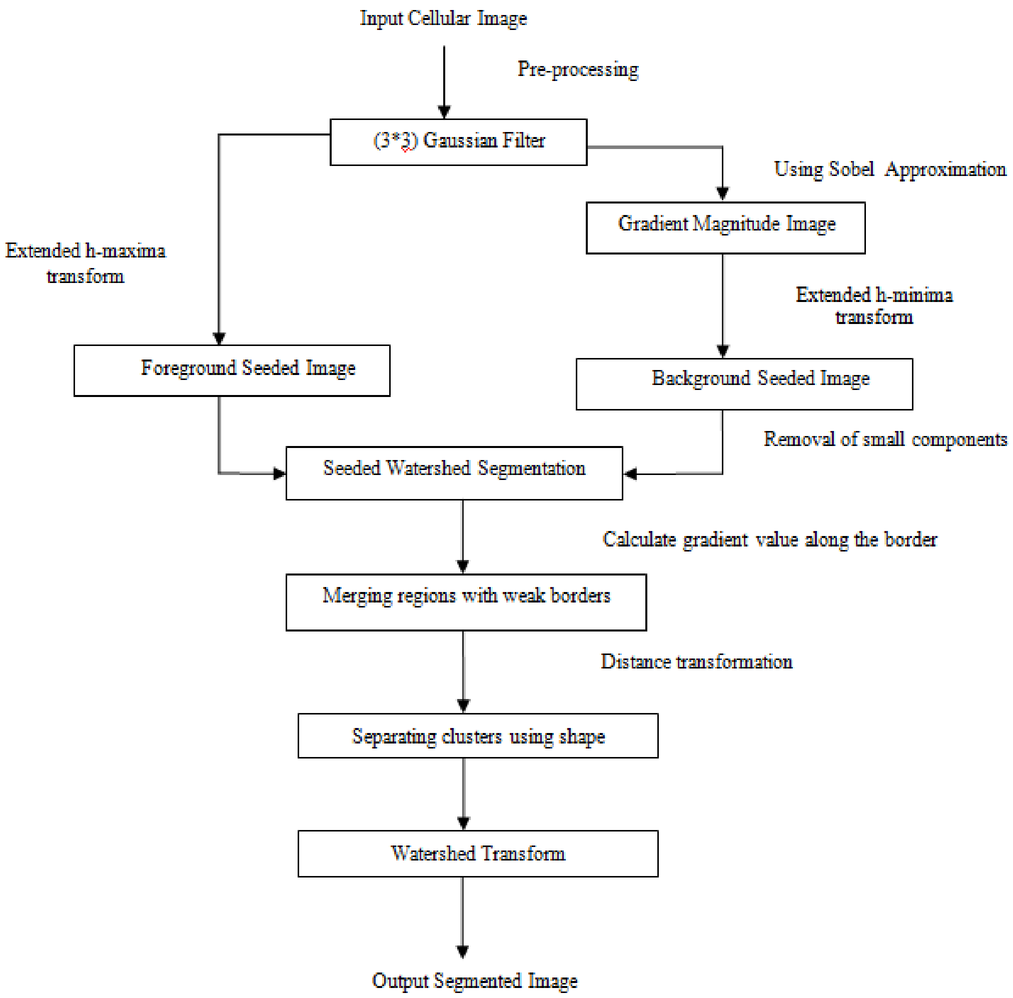

Input Cellular Image
Pre-processing
(3*3) Gaussian Filter
Using Sobel Approximation
Gradient Magnitude Image
Extended h-maxima transform
Extended h-minima transform
Foreground Seeded Image
Background Seeded Image
Removal of small components
Seeded Watershed Segmentation
Calculate gradient value along the border
Merging regions with weak borders
Distance transformation
Separating clusters using shape
Watershed Transform
Output Segmented Image

Fig. 3. Flow diagram of our cell segmentation strategy

6) The clustered cells will be separated using shape. To do so, we use the seeded and watershed result image as binary input to distance transformation. The distance transform of a binary image assigns to each object pixel the distance to the closest background pixe

7)Taking the inverse of the distance image, the distance maxima serve as regional minima for watershed segmentation. Now we again apply watershed segmentation in the resulting image. After this step, the resulting segmented image will be appear. Fig. 3 shows the flow diagram of our new improved cell segmentation method.

III. Implementation & Testing

The implementation was done in MATLAB 7 software. It has rich image processing toolbox and enough functions

for implementation efficiency. Once the four input parameters h_1, h_2, s, t were set, the experiments needed no human interaction. The speed of the segmentation depends on image size and the number of objects in the image.

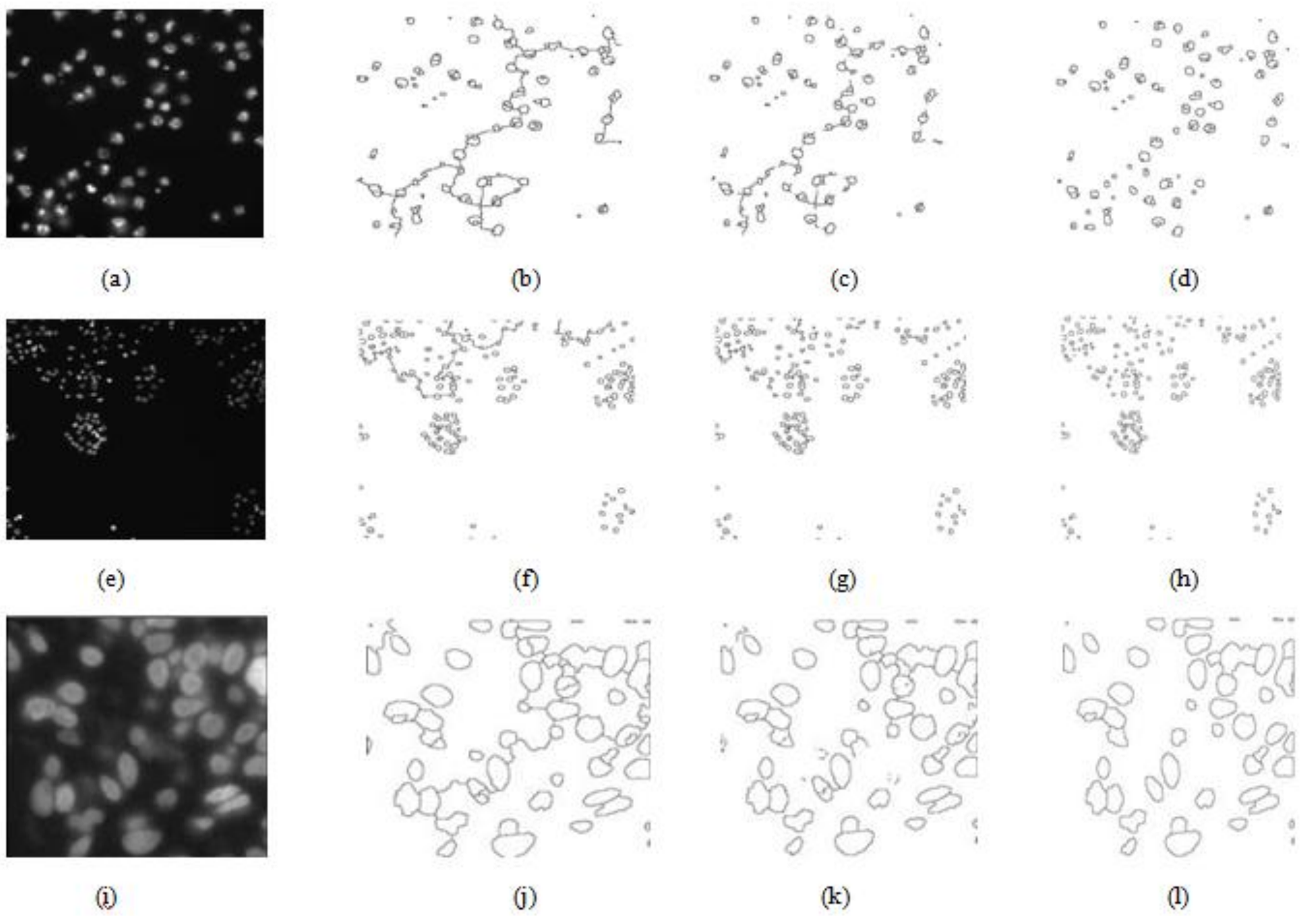


Fig. 4. The three-stage cell segmentation method: (a),(e),(i) are three fluorescent microscopic cellular images of Drosophila, Human colon cancer and slice of tumor cells, respectively. (b),(f),(j) are the results after initial seeded watershed segmentation. (c),(g),(k) are the results after merging edge based on gradient value. (d),(h),(l) are the results after final segmentation stage. Parameters are: (1) For Drosophila cells, h_1=7, h_2=9,s=13, t =11 (2) For Human colon cancer cells, h_1=7, h_2=7,s=11, t =17 (3) For Tumor cells, h_1=7, h_2=9,s=13, t =16 .

IV. Experimental Results

Fig. 4 shows the results after each segmentation stages applied on three different cellular images. In the first stage of the algorithm, initially watershed segmentation is applied on the image based on foreground and background seeds, three parameters are specified – foreground seeds height (h_1), background seeds height (h_2) and structured element size(s). A low $h_1$will result in many seeds, often more than one seed per object. A high h_1 will result in fewer seeds, and some objects may not get a seed at all. Owing to a subsequent merging step based on gradient magnitude, we use a rather low h_1 value to ensure that each object gets at least one seed. The previous step results some oversegmented line. These oversegmented lines are seen in those place where the gradient value of the gradient magnitude image is low. In the next stage , we choose a threshold value(t) for merging edge.An optimum value is chosen so that the resulting oversegmented line is removed.The second stage also results some undesirable segmented line due to the shape of the cells. To remove these undesirable lines, we perform a distance transformed watershed segmentation which requires no parameters.

V. Conclusion

We have presented a three stage segmentation approach for fluorescent microscopic cellular images of live and unstained cells. The basic idea behind successful segmentation is to efficiently restrict the parameter, especially the foreground seeds height(h_1) and merging threshold(t). Very little pre-processing is needed, even if the background variation in the image is large. The input parameters are at present manually set for a test image, and the same parameters are thereafter used for fully automatic segmentation of images created under the same imaging conditions. As only four input parameters are required, this can be done quickly.Methods for automatic parameter approximation are subjects for future work. When a new type of specimen is imaged, adjustment of input parameters will only be necessary if image dynamics or nuclear size changes. The segmentation method can be useful for many different segmentation tasks where a simple foreground/background threshold is not sufficient. Further processing, such as removal of nuclei that are damaged or under-segmented, by a size threshold, or more advanced statistical methods, may improve the result. Automatic segmentation is not only faster than manual segmentation, it is also observer-independent and reproducible.The segmentation can be used by medical clients to validate the effectiveness of strain imaging in locating and distinguishing benign and malignant cells. Once the interested cell region and background region are defined, quantitative measures such as contrast or area can be calculated.

VI. Acknowledgment

The authors wish to thank Dr. Md. Ali Akber Dewan for his valuable suggestions in implementing the method. The cellular images was acquired from the Board of Bio-Medical Benchmark, Harvard University(http://broad.harvard.edu/bbbc).

VII. References

1) M. D. Levine, *Vision in Man and Machines*, New York: McGraw-Hill, 1985.
2) Rodenacker, K. & Bengtsson, E. , *A feature set for cytometry on digitized microscopic images*, Anal. Cellular Pathol., Publication Year-2003, Page(s): 25, 1–36.
3) Joost Vromen, Brendan McCane, "Red blood cell segmentation using guided contour tracing", Presented at SIRC 2006 - The 18th Annual Colloquium of the Spatial Information Research Centre, University of Otago, Dunedin, New Zealand, November 6th-7th 2006.
4) Sahoo, P.K., Soltani, S., Wong, A.K.C. & Chen, Y.C., A survey of thresholding techniques. Comp. Vision, Graphics Image Proc. Publication Year-1988, Page(s):41, 233–260.
5) Lindblad, J. & Bengtsson, E., "A comparison of methods for estimation of intensity nonuniformities in 2D and 3D microscope images of fluorescence stained cells", Proceedings of the 12th Scandinavian Conference on Image Analysis (SCIA) (ed. by I. Austvoll), 2001, pp. 264–271. NOBIM,Norway.
6) Kenong Wu, David Gauthier, and Martin D. Levine, "Live Cell Image Segmentation" in *IEEE Transactions on Biomedical Engineering*, January 1999.
7) Beucher, S., The watershed transformation applied to image segmentation. Scanning Microsc. Publication Year-1992, Page(s):6, 299–314.
8) Chunming Tang and Ewert Bengtsson "Segmentation and Tracking of Neural Stem Cell" in Info. & Commun. Eng. Coll., Harbin Engineering University, Postfach 15 00 01, Harbin, China.
9) Soille, P., Morphological Image Analysis: Principles and Applications. First edn. Springer-Verlag, Berlin Heidelberg New York (1999) 170-171.
10) Garrido, A. & Pérez de la Blanca, N., Applying deformable templates for cell image segmentation. Pattern Recognition, Publication Year-2000, Page(s): 33, 821–832.

Novel Respiratory Diseases Diagnosis by Using Fuzzy Logic

GJCST Classification
I.2.3, J.3

Abbas K. Ali, Xu De Zhi, Shaker K. Ali

***Abstract*-In this paper we design an expert system to diseases diagnosis by using Fuzzy set depending on doctor's opinions. Approach: Using fuzzy set to diseases diagnosis depending on opinion of 20 doctors, Results: it has been to diagnose three types of respiratory diseases (primary kinds of respiratory diseases) (pneumonia (PEN), tuberculosis (TB) and normal influenza (INF)), there are four symptoms X- ray, Respiratory rate (RR), Cough (CO) and Fever (F) which indicate as input of the fuzzy logic and the output will be a range of the risks and type of respiratory diseases.**

Key words- expert system; respiratory diseases; fuzzy set .

I. Introduction

Respiratory diseases are common diseases in human life, the goal of our work is diagnosing a diseases with rapidly, economically and without risks than traditional diagnostic systems. This system allows determining if there is a need for the biopsy and it gives the user a range and types of this diseases. Diagnosis of a disease is a problem in medicine because some patients may have similar symptoms but the doctor may diagnose different diseases, so this work will help doctor when he or she has fuzziness in that thinking process [1,10].Fuzzy logic controller (FLC) was initiated in 1965 by Lotfi Zadeh as a new way of representing vagueness in everyday life [6].The architecture of the proposed fuzzy system consists of three main blocks: the fuzzification step, the fuzzy rule base, the fuzzy inference engine [13].Diagnosis is based on indirect evidence too, the presence of symptoms, and the knowledge of the medical mechanisms that relate presumed causes to observed effects. The problems of diagnosis do not only arise from the incompleteness of this knowledge, but also and most immediately from the theoretical and practical limitations associated with the reversal of the chain of implications that lead from an initial cause to its observable effects [8,9].The natural evolution of various diseases, the obscure nature of medical data and the intrinsic ambiguity of medical problems require a consistent framework that can handle uncertainty by allowing variable and multiple class memberships and facilitating approximate reasoning. This inevitably makes the fuzzy logic (FL) a valuable tool for depicting medical concepts by treating them as fuzzy sets [11,12].The system was developed by aid of the Mat lab 6.5. The rest of paper Experiments is organized as follow: Section 2 give the background information including respiratory diseases and fuzzy set, Section 3 will explain experiments use to diagnose and the result obtain after that. In section 4 we conclude the paper by summary of result and mentioning about future work.

II. Back ground

1) *Respiratory diseases*

Respiratory diseases are an inflammation of the lung that is most often caused by infection with bacteria, viruses, or other organisms. Healthy people can usually fight off respiratory diseases. However, people who are sick,including those who are recovering from the flu (influenza) or an upper respiratory illness, have weakened immune systems that make it easier for bacteria to grow in their lungs, There are three primary types of respiratory diseases are (pneumonia , tuberculosis and normal influenza) we will try to diagnosis in our work [3,7].

2) *Fuzzy set*

Medicine is one field, in which the applicability of fuzzy set theory, within this field it is the uncertainty found in the process of diagnosis of disease that has most frequently been the focus of applications of fuzzy set theory [4].In other word real word knowledge is characterized by in completeness, in accuracy and in consistency.Makes it is possible to define in exact medical entities as fuzzy set ,it is provides an excellent approach for approximating medical text, furthermore fuzzy logic provides reasoning methods for approximate inference this paper surveys the utilization of fuzzy logic on the basis of three medical application

C. algorithm :-

The algorithm is:

```
Call our fuzzy logic function (see fuzzy logic section).
              If output1 then
 Print (“Pneumonia”)
      Else if output2 then
            Print (“Tuberculosis”)
              Else if output 3 then
                                 Print (“Normal Influenza”)
            Else
                                           Print (“unknown disease”)
   End if
          End if
      End if
```

III. Metod

There are 40 cases that were collected from Xiang Ya second Hospital in (Changsha, Hunan, China). Our

About- This paper is supported by the national natural science foundation of china (60970096)

experiments have been very beneficial for medical treatment and perm solving many problems in an easier manner.

1) *Input*

Our work is an expert system for diagnosing the respiratory diseases by using fuzzy set for which inputs will be used symptoms of this disease (X- ray , Respiratory rate (RR), Cough (CO), Fever (F)) **a**nd the output will be pneumonia (PEN), tuberculosis (TB) and normal influenza (INF) As show in figure 1 :-

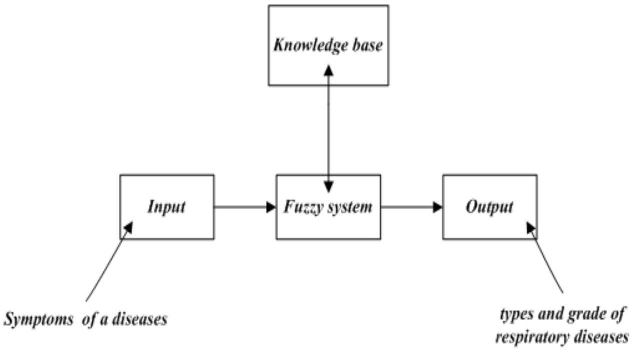


Figure 1 show the schematic diagram of the fuzzy system

The inputs of the diseases (symptoms) must be as numerical values, by built data base of the symptoms depending on the doctor's guess grade of all four symptoms and also the result measurement of temperature of the fever. All these symptoms will be as numerical as shown in table1. The inputs ranges will be change as shown in the bellow:

1) **X- ray** is has three levels (low, medium, high) and the values will be (1-9). As shown in figure 2.
2) **Respiratory rate (RR)** it also depends on doctor's opinion; we also give it three levels (low, medium, high) and value (30 -40). As shown in figure 3
3) **Cough, (C)** the level of cough we proposition it is has three levels is (low, medium, high), and values will be (1-9). As shown in figure 4.
4) **Fever (F)** doesn't need to be guess, we can check it by using a Thermometer, so we proposition three levels too, and the value is (98,102). As shown in figure5.

2) *Output*

The out puts as we said the grade of the diseases as show in figure 6. The number of probabilities are 3×3×3×3= 81which mean we need 81 rules as shown in table 2 [1], in fuzzy logic every rule has result (type of respiratory diseases and level this disease), the inputs (symptoms) of the fuzzy set will be 4 inputs ,the fuzzy will decide the type and grade of risk of the respiratory diseases, it will make decisions depend on the rules that we made, so it will give the percentage of the respiratory diseases; pneumonia (PEN), tuberculosis (TB) and normal influenza (INF) as shown in table 3.If for example, X-ray is low and respiratory rate is high and cough level is high and fever is high then the result is normal influenza, and level of influenza is high, other diseases have low value, so we can deduce the output by these rules as show in the algorithm

Table 1. the range of symptoms

Symptoms	Range
Cough,(C)	1-9
Respiratory rate (RR)	30-40
Fever (F)	98-102
Chest X- ray (CH)	1-9

Table2 Collection rules of fuzzy logic

No of rules	inputs				outputs		
	X-ray	RR	CO	F	PNE	TB.	INF
Rule1	L	L	L	L	L	L	L
Rule2	L	L	L	L	L	L	L
..	..	..	..	..	..	..	..
Rule81	H	H	H	H	H	L	L

The input will pass three stage (fuzzifiction, rule evaluation, and defuzzifiction) .Inputs are given as real crisp values and the output is a fuzzy value. The accuracy of rules should be clarified at this defuzzifuction stage. Firstly, the minimum amount of each rule is recognized and then the maximum amount between them is chosen.

For instance:

X-ray =5.00, RR=3.00, CO=5.77, BT=98.00 the result will be INF (4.99) (medium) .

$\alpha 1$= min (M ,L ,M ,L)
= min (1 ,1 ,0.90 ,1)
= 0.90

$\alpha 2$ = min (M,L,H,L)
= min (1,1,0.10,1)
= 0.10

Using the Mamdani inference (max, min) [2], the system's membership function is:

max ($\alpha 1$, $\alpha 2$)= 0.90

as show in rules this case is medium of influenza the final result is 4.99.

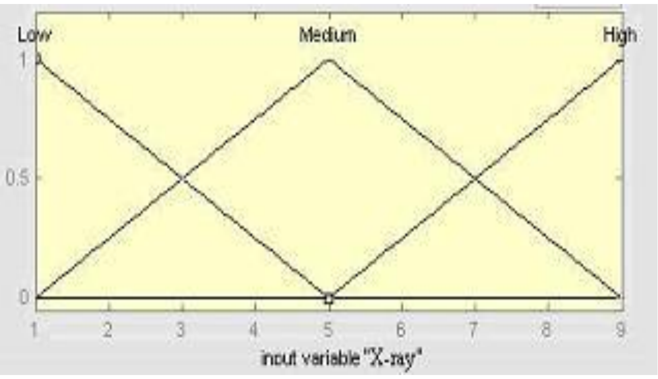


Figure 2 show the membership function X-ray

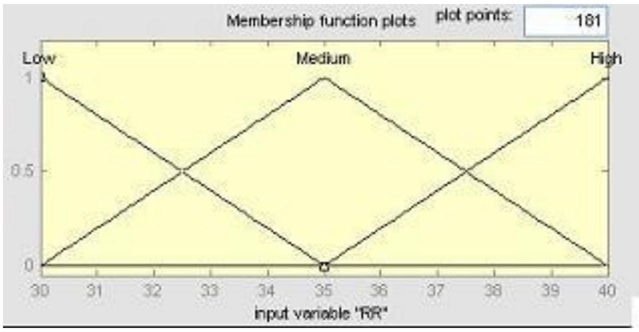


Figure 3 show the membership function respiratory rat

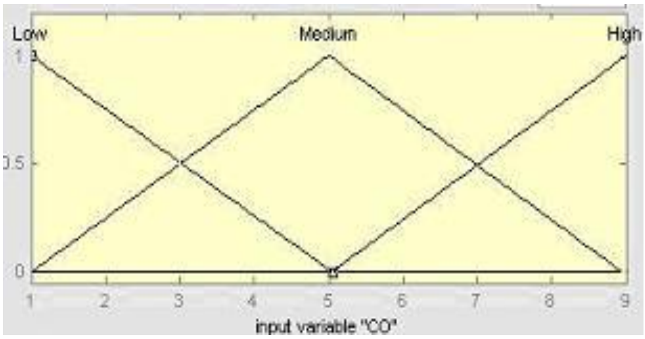


Figure 4 show the membership function cough

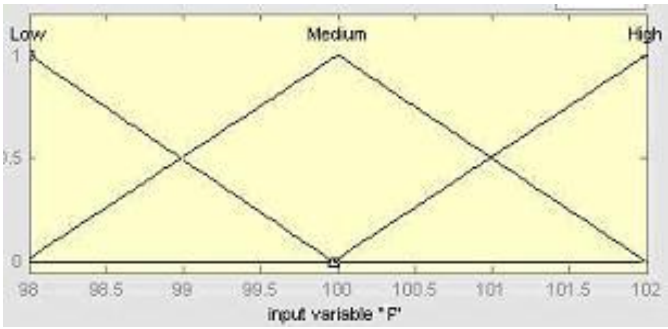


Figure 5 show the membership function fever

Table 3 Results of some of the training and test cases-respiratory diseases

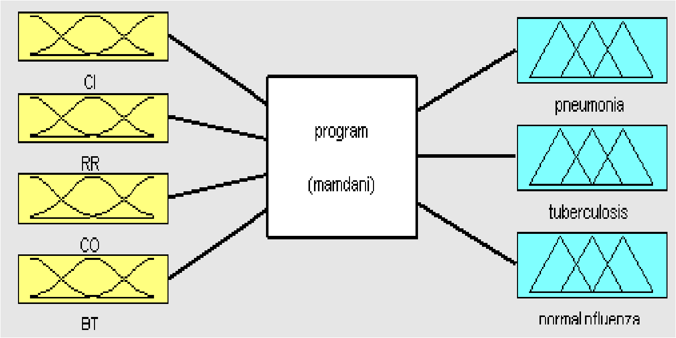


Figure 6 show the outline model of respiratory diseases fuzzy expert system

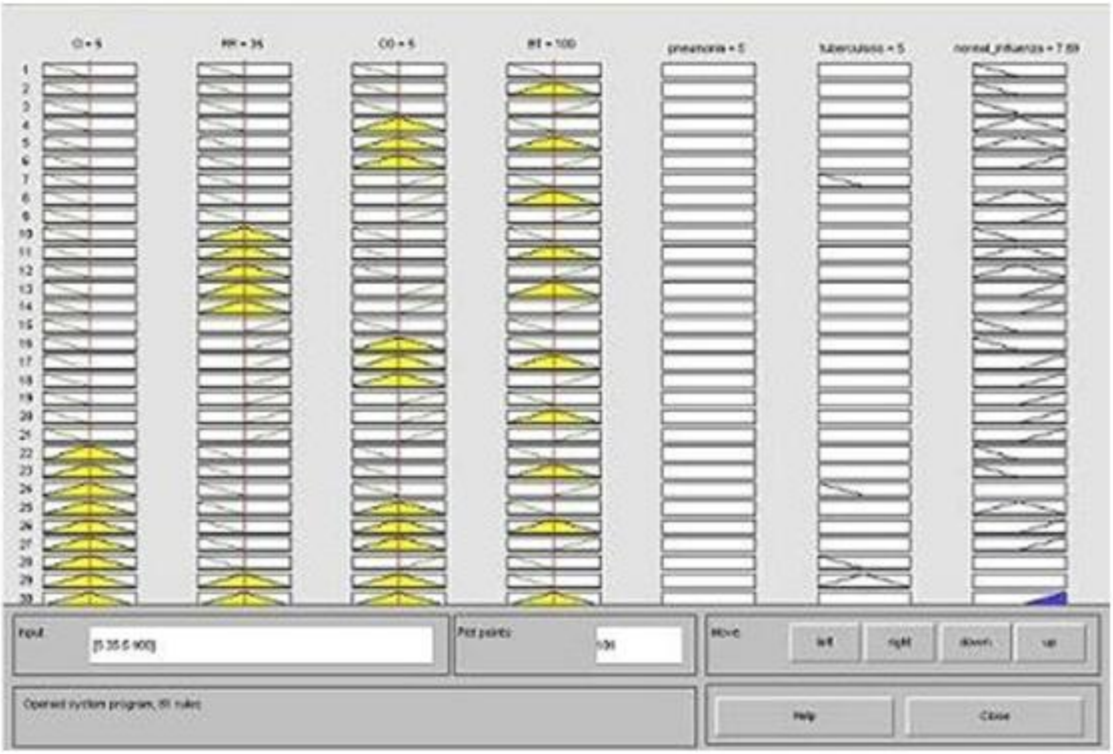
Table 3 Results of some of the training and test cases-respiratory diseases

X-ray	RR	CO	BT	Result
1.00	30.00	1.00	98.00	INF(2.33)L
5.00	30.00	5.77	98.00	INF(4.99)M
1.00	37.00	4.50	100.50	INF(6.94)H
4.50	35.50	1.50	100.00	PIN(2.80)L
4.00	38.55	1.50	99.00	PIN(4.79)M
9.00	31.50	1.50	101.50	PIN(6.16)H
5.55	30.00	1.00	101.00	TB(3.38)L
5.01	35.50	5.50	98.50	TB(5.00)M
8.5	35.50	8.50	98.50	TB(6.71)H

Defuzzification's centre of gravity formula is used for calculating the certain output amount:-

$$D^{\bullet} = \frac{\int D\cdot\mu_{middle}(D)dD}{\int \mu_{middle}(D)dD} \qquad (2)$$

As it is shown in figure 7, the amount 0.81 indicates intensity. The repertory diseases diagnosis for this field is the normal situation while the system reports a small disease risk.



Input: [5 35 5 100]

Plot points: 101

Opened system program, 81 rules

Figure 7 shows the repertory diseases diagnosis

IV. Conclusion

In this paper we will help doctor to diagnosis of diseases and away from fuzziness in the thinking process a doctor, the classification accuracies obtained by expertise of 20 doctors and other studies, these expertise of the disease has been modelled by fuzzy logic.This will help to give a more realistic solution to the problem. The effectiveness of the developed algorithm will be tested,In future work, can apply this method to diagnosis other diseases.

V. Acknowlegements

This work is supported by National Natural Science Foundation of China under grant No.60970096 and 90818004, the Postdoctoral Science Foundation of China under Grant No.20080440988 and the Natural Science Foundation of Hunan Province, China Under Grant No.09JJ4030.

VI. References

1) M Anish R., Dilip K. P., Nilav B., K.P. Sankaranarayanan, N. Sudhahar "Diagnosis of the diseases—using a GA-fuzzy approach "Elsevier, Information Sciences 162 (2004) pp.105~120.
2) mehid.neshat,mehdi ,Yaghobi, Designing a Fuzzy Expert System of Diagnosing the Hepatitis B intensity Rate and Comparing it with Adaptive Neural Network Fuzzy System, International Conference in Modeling Health Advances pp.1~6,2009 .
3) Word health organization department of communication diseases surveillance and response http://www.who.int/emc Christian Schuh,"
4) Christian Schuh "Managing Uncertainty with Fuzzy-Automata and Control in an Intensive Care Environment" ,Springer, Theor. Adv. and Appl. of Fuzzy Logic, ASC 42, pp. 263~271, 2007.
5) Orhan Er & Feyzullah Temurtas & A. Çetin Tanrıkulu," Tuberculosis Disease Diagnosis Using Artificial Neural Networks", Springer Science , Journal Business Media, LLC pp. 299~302 ,2008.
6) CHUEN C. LEE ,"Fuzzy logic control system: Fuzzy logic controller –part I ",IEEE Transaction on systems ,man, and cybernetics , Vol.20, No.2 ,pp.404~418 ,March/April 1990 .
7) Introduction of pneumonia http://www.healthscout.com/ency/1/guides/000064_1.html
8) Mohammad S., Kamran K., Behrooz M., Ali R., Nasser S." A Neuro-Fuzzy Approach to Diagnosis of Neonatal Jaundice" ACM International Conference Proceeding Series;Vol. 275,pp. 1~4, 2006.
9) M.Neshat, M.Yaghobi, M.B.Naghibi, A.Esmaelzadeh, " Fuzzy Expert System Design for Diagnosis of liver disorders", IEEE Computer Society (2008),pp. 252~256.

10) Ismail SARITAS," A Fuzzy Expert System Design for Diagnosis of Prostate Cancer", ACM ,2003, PP.: 345~351.
11) Steimann F. "On the use and usefulness of fuzzy sets in medical AI." Elsafir Artif. Intell. In Med. pp.131~137 2001.
12) Ioannis G. , Ludmil M.," An interpretable fuzzy rule-ased classification methodology for medical diagnosis",Elsevier Artificial Intelligence in Medicine (2009) 47, 25~41.
13) Fabrizio A., Edoardo G., Stefano P., Cesare A.," A Fuzzy Logic Approach to Decision Support in Medicine", SCI 2002,pp.1 ~5,July 2002 – Orlando, USA.

Children Abnormal GAIT Classification Using Extreme Learning Machine

M. Pushpa Rani[1], G.Arumugam[2]

GJCST Classification
I.2.6, K.3.2

***Abstract*-Analyzing human gait has earned considerable interest in recent computer vision researches, as it has immense use in deducing the physical well-being of people. Detection of unusual movement patterns can be performed using Support Vector Machines classification with T-Test pre-normalization. Support Vector Machine classifiers are powerful tools, specifically designed to solve large-scale classification problems. Almost all recent works broadly uses SVM method for gait analysis because of its remarkable learning ability. But when dealing with time complexity there exists a limitation with the SVM. As the computation cost for the SVM is high, the recently developed Extreme Learning Machine (ELM) is being used for the gait classification as a better option in this paper . ELM avoids problems like local minima, improper learning rate and over fitting commonly faced by previous iterative learning methods and completes the training very fast. The multi category classification performance of ELM with T-Test is evaluated with the Virginia gait dataset. The results indicate that ELM produces better classification accuracies with reduced training time and implementation complexity when compared to SVM.**

Keypoints-Extreme learning machine, Gait analysis, SVM classification.

I. Introduction

Human gait is the way locomotion is achieved using human limbs. Different gaits are characterized by differences in limb movement patterns, overall velocity, forces, kinetic and potential energy cycles, and changes in the contact with the surface (ground, floor, etc). Early diagnosis of gait disease and rehabilitation assessment is possible with gait analysis.Most doctors diagnose gait diseases based on their own judgments by comparing many curves created by certain gait analysis system and it is inexact in many cases. In recent years, doctors gain a more objective and exact disease assessment by means of machine learning technology and this has gained much utilization of gait analysis [1].Artificial neural network (ANN) has strong non-linear learning ability, with which it is able to distinguish normal gait and diseased gait [2]. But, the ANN often gets into a local minimum and overstrains training samples which may reduce the accuracy of classifier. Also many ANN techniques are not suitable for gait analysis, as they demand high volume of data for training and it is very much time consuming in case of acquiring huge gait database. Also, this way of implementation will result in increase in the system complexity; greater computational burden and longer training time. Support Vector Machine (SVM) is widely applied in pattern recognition because of its remarkable learning ability. Begg [3] applied SVM for the gait classification of the children and the also for the elderly. Kamruzzaman [4] used SVM to distinguish the gait of children with cerebral palsy. As like the case in ANN, SVM also faces some difficulties, especially the time compatibility. Hence in order to overcome these difficulties, we propose a new training algorithm called the Extreme Learning Machine (ELM) for gait analysis in this paper . The performance of the ELM algorithm is tested using the Virginia gait dataset. For the Virginia data set, the results indicate that ELM can perform direct classification for these multi category microarray problems in a fast and efficient manner. ELM produces higher classification accuracies than those obtained by SVM with a more compact network and shorter training time in case of the Virginia data set .

II. Related Work

The study of human gait has created much interest in many application areas. As a result many researches emerged in recent years and of which, a few studies related to Gait classification are noted herewith.A.Bobick et.al.,[5] projected a view-based approach to the representation and recognition of human movement. The basis of the representation is a temporal template-a static vector-image where the vector value at each point is a function of the motion properties at the corresponding spatial location in an image sequence. Using aerobics exercises as a test domain, the representational power of a simple, two component versions of the templates is explored: The first value is a binary value indicating the presence of motion and the second value is a function of the regency of motion in a sequence. Then the author develops a recognition method matching temporal templates against stored instances of views of known actions. The method automatically performs temporal segmentation, is invariant to linear changes in speed, and runs in real-time on standard platforms. Ross Cutler et.al.,[6] described new techniques to detect and analyze periodic motion as seen from both a static and a moving camera. By tracking objects of interest, we compute an object's self-similarity as it evolves in time. For periodic motion, the self-similarity measure is also periodic and we

About[1]-Ms.M.Pushpa Rani , Associate Professor in Computer Science ,Mother Teresa Women's University, India. (telephone: +919942035825 email: pushpa_john@yahoo.com).
About[2]- Dr G Arumugam , Professor & Head ,Department of Computer Science, Madurai Kamaraj University, India. (telephone: +919443090394 (e-mail: gurusamyarumugam@gmail.com).

apply Time-Frequency analysis to detect and characterize the periodic motion. The periodicity is also analyzed robustly using the 2D lattice structures inherent in similarity matrices.Sheng-Wu Xiong et.al.,[7] proposed in their paper that fuzzy support vector machines based on fuzzy c-means clustering. They apply the fuzzy c-means clustering technique to each class of the training set. During the clustering with a suitable fuzziness parameter q, the more important samples, such as support vectors, become the cluster centers respectively. G.-B. Huang et.al.,[8] explained in their paper about the recently developed Extreme Learning Machine (ELM). ELM is used for direct multi category classification problems in the cancer diagnosis area. ELM avoids problems like local minima; improper learning rate and over fitting commonly faced by iterative learning methods and completes the training very fast. The authors have evaluated the multi-category classification performance of ELM on three benchmark microarray datasets for cancer diagnosis, namely, the GCM dataset, the Lung dataset and the Lymphoma dataset. The results indicate that ELM produces comparable or better classification accuracies with reduced training time and implementation complexity compared to artificial neural networks methods like conventional back-propagation ANN, Linder's SANN, and Support Vector Machine methods like SVM-OVO and Ramaswamy's SVM-OVA. ELM also achieves better accuracies for classification of individual categories.C.-K. Siew et.al.,[9] gives an idea on ELM. In this paper they presented Extreme Learning Machine (ELM) for Single-hidden Layer Feed-forward Neural-networks (SLFNs) which randomly chooses hidden nodes and analytically determines the output weights of SLFNs. The ELM avoids problems like local minima, improper learning rate and over fitting commonly faced by iterative learning methods and completes the training very fast. The author have evaluated the multi-category classification performance of ELM on five different data sets related to bioinformatics namely, the Breast Cancer Wisconsin data set, the Pima Diabetes data set, the Heart-Statlog data set, the Hepatitis data set and the Hypothyroid data set. A detailed analysis of different activation functions with varying number of neurons is also carried out which concludes that Algebraic Sigmoid function outperforms all other activation functions on these data sets. The evaluation results indicate that ELM produces better classification accuracy with reduced training time and implementation complexity compared to earlier implemented models.Ju Han et al., [16] proposed a new spatio-temporal gait representation, called the Gait Energy Image (GEI), for individual recognition by gait. Unlike other gait representations which consider gait as a sequence of templates (poses), GEI represents human motion sequence in a single image while preserving temporal information. To overcome the limitation of training templates, we propose a simple model for simulating distortion in synthetic templates and a statistical gait feature fusion approach for human recognition by gait. Experimental results show that 1) GEI is an effective and efficient gait representation and 2) the proposed recognition approach achieves highly competitive performance with respect to the published major gait recognition approaches. This paper presents a systematic and comprehensive gait recognition approach, which can work just as fine as other complex published techniques in terms of effectiveness of performance while providing all the advantages associated with the computational efficiency for real-world applications.

Davrondzhon Gafurov [17] presented an overview of biometric gait recognition is given. Depending on the way the gait data is captured, biometric gait verification and identification is categorized into three classes (MV-based, WS-based and FS-based). The primary advantage of MV based gait biometric is in being captured from the distance. The main advantage of WS-based and FS-based gait biometric is in providing unobtrusive user authentication and identification. In a multi- modal biometric system, gait helps to increase the accuracy of the system too. However, there are many factors that can negatively influence the accuracy of a gait recognition system. An investigation of these factors is very important towards developing robust systems. With respect to gait security, studies also indicated that gait biometric is robust against minimal effort impersonation attacks. However, impostors who know their closest person in the database or the gender of the users in the database can be a threat to a gait authentication system. Multi-modal biometric systems combine evidences from several biometric modalities to establish more reliable and accurate identification.

III. Methodology

The block diagram of the proposed system is as shown in Figure 5.1. After suitable preprocessing, the Salient Gait Features are extracted and statistical based feature selection is performed using the extracted features. These selected salient gait features are then subjected to normalization and ranking methodologies followed by a classification algorithm. The ranking and normalization methodologies dealt in this chapter are Principle Component Analysis (PCA) and T-Test. The two classification algorithms discussed here are Support Vector

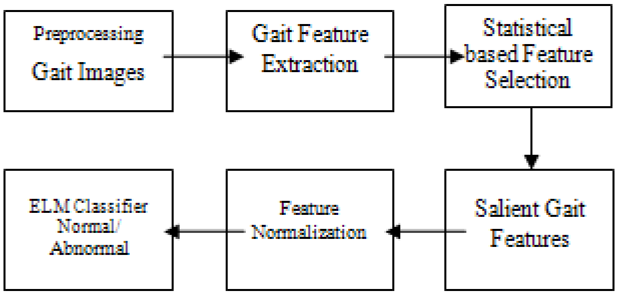

Preprocessing Gait Images
Gait Feature Extraction
Statistical based Feature Selection
Salient Gait Features
Feature Normalization
ELM Classifier Normal/ Abnormal

Figure1.Block Diagram of Gait Analysis using ELM

The two ranking methodology dealt in this paper are Principle Component Analysis (PCA) and T-Test. The two classification algorithm is Support Vector Machine (SVM) and Extreme Learning Machine (ELM).

1) *Short introduction of SVM*

SVM is usually used for classification problems introduced by Vapnik [10]. For binary classification, an SVM is used to find an optimal separating hyper plane (OSH) which generates a maximum margin between the two categories of data. To construct an OSH, SVM maps data into a higher dimensional feature space. SVM performs this nonlinear mapping by using a kernel function. Then, SVM constructs a linear OSH between the two categories of data in the higher feature space. Data vectors which are nearer to the OSH in the higher feature space are called support vectors (SVs) and will contain all the information required for classification. The theory of SVM can be briefed as follows [10].Consider training set $D = \{(x_j, y_i)\}_{i=1}^{L}$ with each input in $x \in R^n$ and an associated output

$y_i \in \{-1, +1\}$. Initially,each input x is mapped into a higher dimension feature space F, by $z=\varphi(x)$ via a nonlinear mapping $\varphi: R^n \rightarrow F$. When the data are linearly non-separable in F, there exists a vector $w \in F$ and a scalar b which defines the separating hyper plane as:

$$y_i(w'.z_i + b) \geq 1 - \xi_i, \forall i \qquad (1)$$

where $\xi(\geq 0)$ are called slack variable. The hyper plane that optimally separates the data in F is the one that

$$minimise \frac{1}{2}.w'.w + C.$$

$$subject\ to\ y_i(w'.z_i + b) \geq 1 - \xi_i, \xi_i \geq 0, \forall i \qquad (2)$$

where C is called as the regularization parameter that determines the tradeoff between the maximum margin and the minimum classification error. By constructing a Lagrangian, the optimal hyper plane according to the equation (2) may be shown as the solution of

$$maximize\ W(\alpha) = \sum_{i=1}^{L} \alpha_i - \frac{1}{2} \sum_{i=1}^{L} \sum_{j=1}^{L} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$subject\ to\ \sum_{i=1}^{L} y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, \forall i \qquad (3)$$

where $\alpha_1, \ldots.., \alpha_L$ are the nonnegative Lagrangian multipliers. The data points x_i that correspond to $\alpha_i > 0$ are SVs. The weight vector w is then given by

$$w = \sum_{ieSVs} \alpha_i y_i z_i \qquad (4)$$

For any test vector $x \in R^n$, the classification output is then given by

$$y = sign(w.z + b) = sign(\sum_{ieSVs} \alpha_i y_i K(x_i, x) + b) \qquad (5)$$

To construct an SVM classifier, a suitable kernel function and its parameters need to be chosen. So far, no analytical or empirical studies have been established to prove the superiority of one kernel over another. Hence, in this study, the following three kernel functions have been applied to build the SVM classifiers:

1) Linear kernel function, $K(x,z) = \langle x,z \rangle$;

2) Polynomial kernel function $K(x, z) = (\langle x, z \rangle + 1)^d$ is the degree of polynomial;

3) Radial basis function, $K(x, z) = \exp\left\{-\frac{||x-z||^2}{2\sigma^2}\right\}, \sigma$ is the width of the function.

2) *Introduction of FCM algorithm*

FCM algorithm proposed by Dunn [11] and extended by Bezdek [12] is one of the most well-known methods in cluster analysis. FCM partitions a set of *s*-dimensional vectors $X=\{X_1, X_2, \ldots, X_n\}$into c clusters, where $X_j = (X_{j1}, X_{j2}, \ldots., X_{js})'$ represents the *j*th sample for *j=1,….,n*. The *i*th cluster is supposed to have a center vector $v_i = (v_{i1}, \ldots, v_{is})$ and FCM aims to determine the cluster centers v_i , where $1 \leq i \leq c$. For the j^{th} sample X_j and the *i*th cluster center v_i , there is a membership degree $u_{ij} (\in [0,1])$ indicating in which degree the sample X_j belongs to the cluster center vector v_i , which results in a fuzzy partition matrix $U=(u_{ij})_{nXc}$. The objective function *J* is defined as follows:

$$J(U, v_1, \ldots, v_c, X) = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^m d_{ij}^2 \qquad (6)$$

$$d_{ij} = (\sum_{k=1}^{5} (v_{ik} - x_{jk})^2)^{\frac{1}{2}} \qquad (7)$$

$$v_i = \sum_{j=1}^{n} u_{ij}^m X_j / \sum_{j=1}^{n} u_{ij}^m \qquad (8)$$

$$u_{ij} = (\sum_{k=1}^{c} ({}^{d_{ij}}\!/_{d_{kj}})^{\frac{2}{(m-1)}})^{-1}, m \neq 1 \qquad (9)$$

subject to$\sum_{i=1}^{c} u_{ij} = 1, \forall j = 1, \ldots., n$, where *m* (usually set to be 2) in (11) and (12) is used to adjust the weight effect of membership values.
The FCM algorithm [13] may be described as follows:

1) Choose an integer c and a threshold value ε . Initialize the fuzzy partition matrix *U* by *c*×*n* random numbers in the interval [0,1] ;
2) Compute $v_i (i = 1, \ldots, c)$ according to equation (8);
3) Compute all d_{ij} and u_{ij} according to (7) and (9) respectively. Thus update the fuzzy partition matrix *U* by the new computed u_{ij} .

4) Compute the objective function *J*. If the difference between two adjacent *J values* is less than the given threshold ε , then stop. Otherwise go to step 2.

3) *SVM algorithm*

SV plays a decisive role in SVM, but non-SVs are usually inoperative. In F-SVM, the treatment is not the same for each sample. Training samples that are affirmatively not SVs are not involved in SVM, and only the samples that have a weak relationship with each cluster are chosen to be trained in SVM. By using this method, the classification accuracy can be increased and the training time may considerably be reduced[7],[14].The selection of training data is executed by the FCM which may cluster samples non-compulsively. The membership degree u_i in FCM indicates with what degree a sample belongs to the cluster center vector v_i. If there exists one $u_{i_0} (i_0 \in \{1, \ldots, c\})$ which is bigger than a threshold λ (80% in this study), it is clearly that the sample is far from OSH and its probability to be a SV is small. So, this sample is not involved in SVM. The block diagram of F-SVM is shown in Figure 2

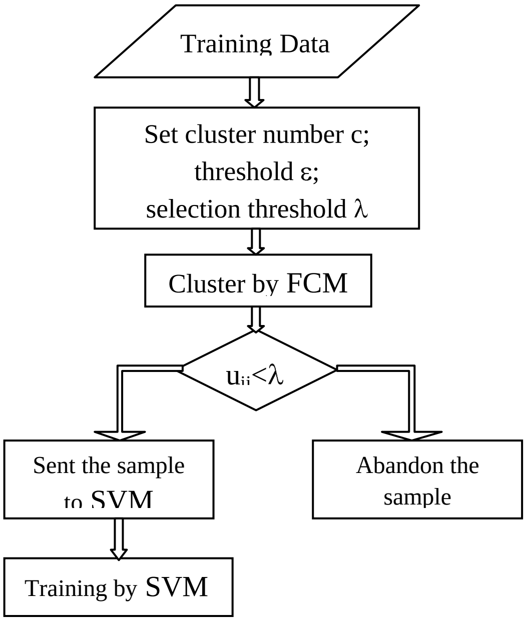


Figure 2 block diagram of the F-SVM.

The process of F-SVM is described as follows:
1) Choose a cluster number *c* (2 in this study) and a threshold value ε for FCM. Choose a selection threshold λ (80% in this paper).
2) Cluster the training samples by applying FCM and calculate the membership degree u_{ij} according to equation (9). If u_{ij}< λ, ∀i∈{1,….,c}, sent the sample to SVM. Otherwise, abandon it.
3) Classify the selected training sample set using SVM. Calculate each classification output according to (7).

IV. Extreme Learning Machine

A new learning algorithm called the Extreme Learning Machine (ELM) is a Single-hidden Layer Feed forward neural Networks (SLFNs) supervised batch learning. The output of an SLFN with ~N hidden nodes (additive or the RBF nodes) can be represented by

$$f_{\tilde{N}}(X) = \sum_{i=1}^{\tilde{N}} \beta_i G(a_i, b_i, X),\quad X \in R^n,\quad a_i \in R^n, \qquad (10)$$

where ai and bi are the learning parameters of hidden nodes and β_i is the weight connecting the ith hidden node to the output node. $G(a_i,b_i,X)$ is the output of the ith hidden node with respect to the input x. For the additive hidden node with the activation function g(x):R→R (e.g., sigmoid or threshold), $G(a_i,b_i,X)$ is given by

$$\mathrm{G(ai, bi, X)} = \mathrm{g}(\mathrm{a_i}.\mathrm{X} + \mathrm{b_i}), \mathrm{b_i} \in \mathrm{R} \qquad (11)$$

Where ai is the weight vector connecting the input layer to the ith hidden node and bi is the bias of the ith hidden node. ai .x denotes the inner product of vectors ai and x in R. For a RBF hidden node with an activation function g(x):R→R(e.g., Gaussian), $G(a_i,b_i,X)$ is given by

$$\mathrm{G(ai, bi, X)} = g\left(\mathrm{b_i}\left||\mathrm{x} - \mathrm{a_i}|\right|\right), \mathrm{b_i} \in \mathrm{R}^+ \qquad (12)$$

where ai and bi are the center and impact factor of the ith RBF node. Here R^+ indicates the set of all positive real values. The RBF network is a special case of SLFN with the RBF nodes in its hidden layer. Each RBF node has its own centroid and impact factor and its output is given by a radially symmetric function of the distance between the input and the center.The learning algorithms use a finite number of input-output samples for training. Here, we consider N arbitrary distinct samples $(x_i,t_i)\in R^n$ x R^m, where x_i is an n x 1 input vector and t_i is an m x 1 target vector. If an SLFN with $\tilde{N}$ hidden nodes can approximate these N samples with zero error, it then implies that there exist β_i, ai, and bi such that

$$f_{\tilde{N}}(X_j) = \sum_{i=1}^{\tilde{N}} \beta_i G(a_i, b_j, X_j) = t_j,\ j = 1, \ldots., N \qquad (13)$$

Equation () can be written compactly as
$H\beta = T$ (14)

Where

$H(a_1, \ldots\ldots, a_{\tilde{N}}, b_1, \ldots\ldots, b_{\tilde{N}}, X_1, \ldots\ldots, X_{\tilde{N}}) =$

$$\begin{bmatrix} G(a_1, b_1, X_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, X_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, X_N) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, X_N) \end{bmatrix}_{N \times \tilde{N}} \quad (15)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (16)$$

H is called the hidden layer output matrix of the network [15]; the i^th column of H is the i^th hidden node's output vector with respect to inputs x_1, x_2,…, x_N and the j^th row of H is the output vector of the hidden layer with respect to input x_j.In real applications, the number of hidden nodes, $\tilde{N}$, will always be less than the number of training samples, N, and, hence, the training error cannot be made exactly zero but can approach nearer to a nonzero training error . The hidden node parameters a_i and b_i (input weights and biases or centers and impact factors) of SLFNs need not be tuned during training and may simply be assigned with random values according to any continuous sampling distribution. Equation (5) then becomes a linear system and the output weights _ are estimated as

$$\tilde{\beta} = H \dagger T \quad (17)$$

Where $H \dagger$ the Moore-Penrose is generalized inverse [15] of the hidden layer output matrix H. The ELM algorithm which consists of only three steps, and can then be summarized as

ELM Algorithm: Given a training set $\aleph = \{(X_i, t_i) | X_i \in R^n, t_i \in R^m, i = 1, \ldots, N\}$ activation function g(x), and hidden node number $\tilde{N}$,

1) Assign random hidden nodes by randomly generating parameters (a_i,b_i) according to any continuous sampling distribution, i=1,…., $\tilde{N}$
2) Calculate the hidden layer output matrix H.
3) Calculate the output weightβ: $\tilde{\beta} = H \dagger T$

The universal approximation capability of ELM has been analyzed in Huang et al. [16] using an incremental method and it has been shown that single SLFNs with randomly generated additive or RBF nodes with a wide range of activation functions can universally approximate any continuous target functions in any compact subset of the euclidean space R^n. $g(x) = \frac{1}{1+e^{-\lambda x}}$ is the sigmoidal function used as activation function in ELM

V. Experimental Results

This study mainly deals with the Performance analysis of the t-test ELM classification method for gait classification as normality and abnormality. Several experiments are carried out to test the validity of t-test ELM. A comparative analysis of the results of our proposed method is also done with t-test SVM. The experimental data used in this study are obtained from the gait database in Virginia University. There are totally 158 gait samples present in the database. These samples belong to 68 children with normal gait and 88 children with abnormal gait. The ages of these children range from 2 years to 13 years. Four features of gait samples are selected for classification and they are stride length, cadence, leg length and age.

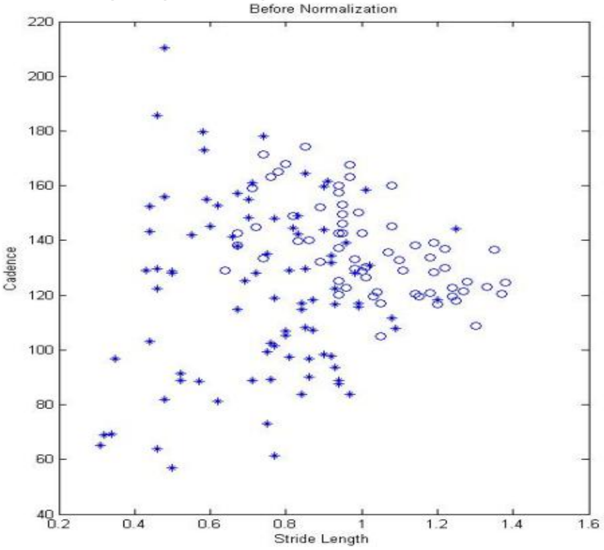


Figure 3: (a) Distribution of raw data before normalization.

In this study, the t-test is applied to normalize the gait samples. Figure 3.(a & b) shows the distribution of data before and after normalization. As shown in Figure3.b, the overlap of two sample sets is effectively reduced after normalization, which helps to improve the classification accuracy.

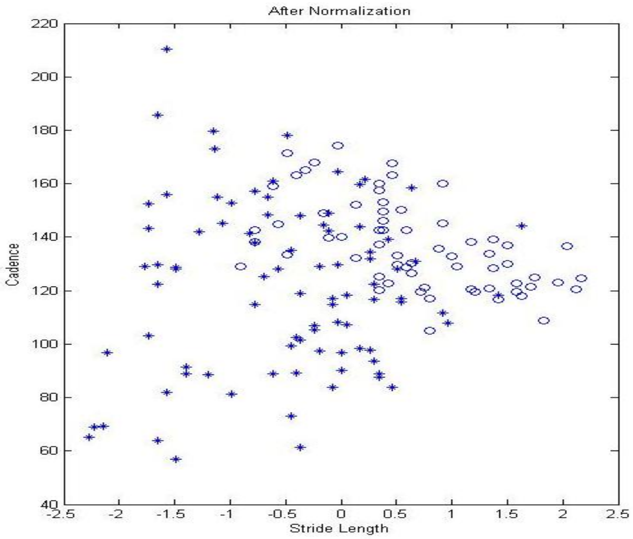


Figure 3: (b) Distribution of raw data after normalization.

The generalization capability of SVM enhances along with the increase of C (Regularization parameter). This is because the regularization parameter C may adjust the ratio of confidence interval and empirical risk. The generalization capability of SVM is weak when C is small, because a small C indicates that the punishment for empirical risk is small and the empirical risk is large. When C exceeds a certain value, the complexity of a classifier reaches the allowed maximum in the feature space. In this case, the SVM has almost no change of the empirical risk and generalization capability. The accuracy of using standard SVM, SVM with PCA and SVM with T-Test are measured and tabulated in Table1

When dealing with Classification Accuracy of ELM, it achieves better results. The accuracy achieved by ELM is 95 percent or more. This is consistent with our hypothesis that ELM performs better in the multi category classification applications where the number of classes is large.For all data sets, ELM takes much less total training time than does the SVM algorithm. As mentioned before, the SVM-algorithm has to build binary classifiers to distinguish between every two class combinations. For the data sets used with the number of categories classified decreasing, the difference between ELM and SVM is also decreasing. ELM takes significantly lower training time. For all data sets, it can be seen that the number of hidden nodes for ELM is always smaller than the number of support vectors for SVM, indicating a more compact network realized by ELM. The accuracy of ELM with PCA and T-Test are tabulated in Table1.

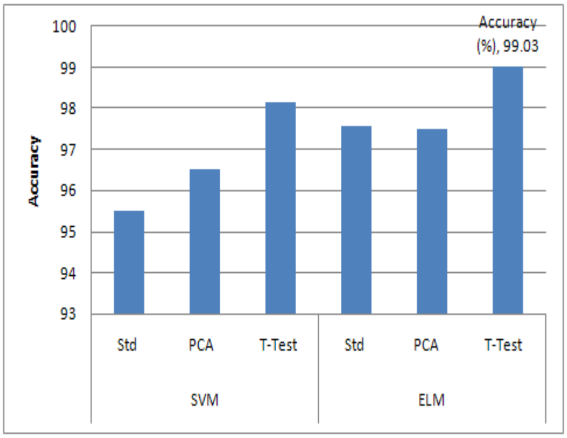


Figure 4. Comparison of SVM and ELM accuracy.

In reference with that ELM with PCA gives result with accuracy of 97.98% and when T-Test is used it gives the maximum accuracy of 99.21%. The increase in percentage of accuracy when using ELM is plotted in graph above.The performance of ELM can be estimated with the use of total training time. When the number of hidden nodes for ELM and support vectors of SVM is compared it can be seen that the number of hidden nodes for ELM is always smaller than in support vectors indicating a more compact network realized by ELM. For each network, there are five modules each consisting of 10 hidden nodes. This means that, for each experiment, up to 4,600 hidden nodes are needed for the training process of SANN, while, for ELM, the network needs less than 50 hidden nodes. Thus reducing the number of hidden nodes reduces the training time. The total training time for ELM and SVM is shown in Table 2. This shows that SVM took around 410 seconds to train the dataset completely. For the same dataset ELM took only 125 seconds.

Table1.SVM and ELM accuracies

Algorithm		Accuracy (%)
SVM	Std	95.51
	PCA	96.51
	T-Test	98.15
ELM	Std	97.56
	PCA	97.49
	T-Test	99.03

Table2. Total training time used by SVM and ELM

Algorithm	Time(in sec)
SVM	410
ELM	125

This result clearly shows that ELM takes a much smaller training time than SVM.

VI. Conclusion

In this paper, a fast and efficient classification method called the ELM algorithm is used to classify the abnormal gait. ELM can perform the multi category classification directly, without any modification. Study results are consistent with our proposition that, when the number of categories for the classification task is large, the ELM algorithm achieves higher classification accuracy than the other algorithms with less training time and a smaller network structure. It can also be seen that ELM achieves better and more balanced classification for individual categories as well as very less training time comparative to SVM. In future some advanced neural network techniques can be used to train the ELM classifier which may enhance the classification accuracy further with reduced training time.

VII. References

1) Simon S. R., Quantification of human motion: gait analysis-benefits and limitations to its application to clinical problems. Journal of Biomechanics, 2004(37): 1869-1880.

2) Barton J G, Lees A, An application of neural networks for distinguishing gait patterns on the basis of hip-knee joint angle diagrams. Gait and Posture, 1997(5): 28-33.
3) R. K. Begg, M. Palaniswami, and B. Owen, Support Vector Machines for automated gait classification. IEEE Trans. Biomed. Eng., vol. 52, no. 5, pp. 828-823, May 2005.
4) Joarder Kamruzzaman, Support Vector Machines and Other Pattern Recognition Approaches to the Diagnosis of Cerebral Palsy Gait. IEEE Trans. Biomed. Eng., vol. 53, no. 12, pp. 2479-2489, Dec 2006.
5) Bobick and W. Davis. The recognition of human movement using temporal templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(3),March 2001.
6) Ross Cutler and Larry Davis. Robust real-time periodic motion detection, analysis, and applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(8):781 – 796, 2000.
7) Sheng-Wu Xiong, Hong-Bing Liu and Xiao-Xiao Niu, Fuzzy support vector machines based on FCM clustering. Proceedings of 2005 International Conference on Machine Learning and Cybernetics, Aug. 2005, vol. 5, pp. 2608- 2613.
8) G.-B. Huang, Q.-Y. Zhou, and CK. Siew, “Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks,” Proc. Int’l Joint Conf. Neural Networks (IJCNN ’04), July 2004.
9) G.-B. Huang and C.-K. Siew, “Extreme Learning Machine with Randomly Assigned RBF Kernels,” Int’l J. Information Technology, vol. 11, no. 1, 2005.
10) Vladimir N. Vapnik, The Nature of Statistical Learning Theory. New York: Springer-Verlag, 1995, 187 pp.
11) Dunn and J.C., Some recent investigations of a new fuzzy partition algorithm and its application to pattern classification problems. J. Cybernetics, vol. 4, pp. 1–15, 1974.
12) Bezdek and J.C., Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York, 1981.
13) Wang, Y. Wang, and L. Wang, Improving Fuzzy c-Means Clustering based on Feature-Weight Learning. Pattern Recognition Letters, 25 (10), pp. 1123–1132, 2004.
14) Dae-Jong Lee, Jong-Pil Lee, Pyeong-Shik Ji, Jae-Woon Park and Jae-Yoon Lim, Fault Diagnosis of Power Transformer Using SVM and FCM. 2008 IEEE International Symposium on Electrical Insulation, June 2008, pp. 112-115.
15) G.-B. Huang, “Learning Capability and Storage Capacity of Two- Hidden-Layer Feedforward Networks,” IEEE Trans. Neural Networks, vol. 14, no. 2, pp. 274-281, 2003.
16) Ju Han, Bir Bhanu, "Individual Recognition Using Gait Energy Image", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No. 2, February 2006.
17) Davrondzhon Gafurov, "A Survey of Biometric Gait Recognition: Approaches, Security and Challenges", NIK-2007 conference.

Health Information System Implementation: The Role of Business Process Management on Successful Implementation

Abdullah S. Al-Mudimigh

GJCST Classification
J.1, J.3

***Abstract*-Integrated HIS provides significant improvements in efficiency across an environment of health care. However, there are also risk associated with implementation of integrated HIS but the actual implementation of integrated health information systems (HIS) is a challenging issue. Implementing integrated HIS are increasingly seen as the way to achieve quality and continuity in treatment, reducing clinical errors, and supporting health care professional in providing care. With the hurried growth in popularity of integrated HIS package, an increasing number of health care sectors are making the decision to implement an integrated HIS. The critical success factors for integrated HIS implementation include top management support, a clear business vision and issues specific to HIS such as implementation strategy and software configuration. However, some of the more important factors are the issues related to re-engineering business processes and the integration of various core processes to the HIS. This paper investigates the role and impact of business process management in successful integrated HIS implementation. The paper starts by defining business process management and the integrated HIS critical success factors. The impact of business process management on successful integrated HIS implementation is then assessed through looking at the experiences of several health care sectors. The paper concludes by documenting best practices for capitalizing on business process management for successful integrated HIS implementation.**

*Key Words:*Critical Success Factors (CSF), HIS, Business Process Management (BPM).

I. Introduction

The recent push for healthcare reform has caused health departments to focus on ways to streamline their processes in order to deliver high quality care while at the same time reducing costs. Integrated Health Information Systems (HIS) software packages (synonyms are hospital information systems, health information management, clinical information systems, healthcare information systems, health information technology, and health application systems) seek to integrate the complete range of a health departments processes and functions in order to present a holistic view of the health care from a single information and IT architecture. I can say that the health departments increasingly recognize the value of sharing information among all stakeholders. Few health care sectors will dispute the value that HIS software can bring to their business. However, most health organizations are not putting in place the procedures tomanage the changes and customizations they need to make to HIS for establishing better services to their patients and staff. Most health sectors are too busy building and running the technical aspects of their HIS package to recognize the need, and long-term value of change and business process management. These values extend well beyond application development and in fact it provides the backbone for successful installations and operation of an HIS. Health care are becoming engrossed in building and running the technical aspects of their HIS to recognize the need and long-term value of change and business process management. Integrated HIS is the umbrella for integrating sets of health departments applications that allow them to manage almost all aspects of operations. The value of this holistic view extends well beyond application development and in fact provides the backbone forsuccessful installations and operation of an integrated HIS software package. Integrated HIS packages have made a tremendous contribution to the world of health care [12]. Indeed, the value that HIS packages can bring to health sectors is clear to many health organizations and few will dispute its potential. However, there are also hazards associated with implementation of integrated HIS. Their failure is high and may cause negative effects on staff and patients [1,2,11]. These software packages are huge and complex systems and warrant careful plan and execution to ensure successful implementation. In other words, integrated HIS implementation is much more than implementation of hardware or software systems; they affect how a health care conducts itself. The success of an integrated HIS implementation has often been attributed to two facts: the HIS package is configured and running and the whole project is (more or less) on time and within budget [6]. However, this is a narrow view of HIS implementation focusing on the hard aspects and reducing it to mere software or IT project. Many integrated software implementation failures have been due to the lack of focus on 'the soft issues', i.e. the business process and change management [5,14]. The role and impact of business process management (BPM) in successful integrated HIS implementation is crucial, and is the subject of our investigation here.

About- PhD Department of Information System College of Computer and Information Sciences King Saud University
E-mail – mudimigh@ksu.edu.sa

II. INTEGRATED HIS CRITICAL SUCCESS FACTORS

Due to the complex and integrated nature of HIS software package, the large investments involved (time and money), and the relatively high implementation failure rates [1,2,3]. It is imperative for health care sectors to study the experiences of others, and learn from their practices and success factors.A literature review was conducted to understand the critical success factors in successful integrated software implementation [17]. The review covered numerous published books and articles. The review concluded the identified CSFs fall under one of four main categories, namely: executive leadership, changing of the existing processes, deploying change management, and the IT infrastructure. These CSF categories are presented in Figure 1.

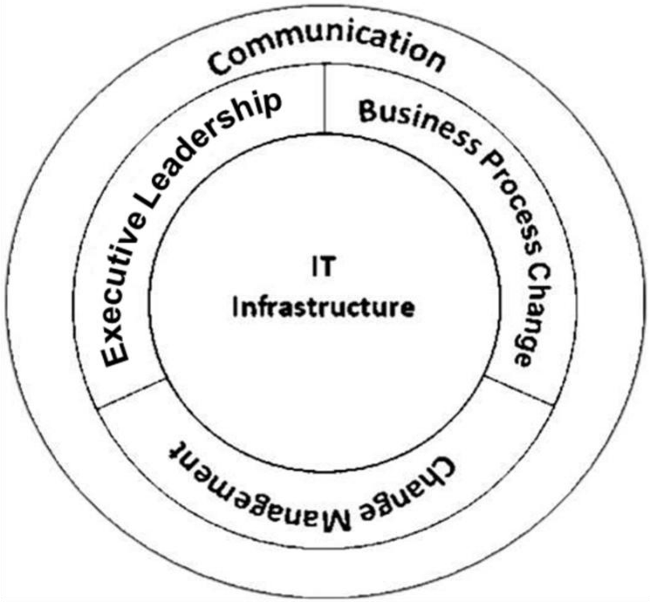


Figure 1: CSF Categories for successful integrated HIS

The following is a brief overview of each of these categories:

1). Executive Leadership: Executive must be a part of integrated HIS implementations. The leadership is a high level official who works for institutionalization of the project, is a talent and good communicator, and has political awareness and influential contacts. The IT literatures have clearly demonstrated that for IT protects to succeed executive support is critical. However, executive in many health care sectors still view the installation of an integrated HIS as primarily a technological challenge and assign its responsibility to the IT departments. This is seen as a risky act due to HIS's profound health care implications.

2). Business Process Change (BPC): Implementing an Integrated HIS package involves changing the existing business processes to the best business process standard. Integrated HIS was built on best practices that are followed in the health care sectors, and to successfully install HIS package, all the processes in a health departments must conform to the HIS package.

During the HIS preparation phase, health care sectors face a question as to whether to implement the HIS software "AS IS" or modify the product to the specific needs of their requirements. Indeed, it has been recommended by practitioners and expert people that a hospital has to change its processes to conform to the integrated HIS package. In fact, this need to change the hospital's business processes is seen as one of HIS's major benefits.

3). Change management: One of the main obstacles facing integrated HIS implementation is resistance to change. There will be resistance from users (for instance, from nurses, phlebotomists other paramedical staff, etc) who may feel that feeding information into the computerized system is additional work and not their primary responsibility or core competence. "Health care constantly evolving. Wave after wave of new technologies, insurance model, information systems, regulatory changes, and institutional arrangements buffet the system and the people in it. But people and institutions, for the most part, do not like change. It is painful, difficult and uncertain" [2]. To successfully implement integrated HIS package, the way health departments do business will need to change and the ways people do their jobs will need to change too. Its success is largely dependent on the commitment of health management, IT staff, and program staff to implementing integrated HIS that will change the way they do their jobs [9]. Thus, change management is essential for preparing a company for the introduction of a HIS , and its successful implementation. However, change management has to be structured within an overall Business Process Management methodology to achieve its goals.

4). IT Infrastructure: Implementing integrated HIS presents an immediate information architectural challenges that has organizational implications. Adequate hardware, communication, and networking infrastructure is required for HIS application. Integrated HIS can't be without sophisticated information technology infrastructure. Three primary attributions of success were identified from the descriptive statistics: willingness to change to new computer applications, effort, and persistence. In addition to the infrastructure, clearly, the software configuration has a critical influence on the implementation process and outcome [4].Clearly, three out of four of these main categories fall under the umbrella of Business Process Management (BPM). If anything, this strongly re-iterates the fact that integrated HIS is not merely software implementation or an IT project. Thus, to ensure successful HIS implementation and running, health care sectors must pay sufficient attention to BPM.

III. BUSINESS PROCESS MANAGEMENT (BPM)

A business process is set of interrelated activities which have definable inputs and, when executed, result in an output that adds value form a customer perspective. Business processes are quite simply the way work is done in any organization. They are cross-functional and go across the organizational functions, e.g. order fulfillment which spans all organizational functions from customer order to final delivery. BPM is a structured approach to understand, analyze, support, and continuously improve fundamental process such as manufacturing, marketing, communications and others major elements of a company's operations.

BPM is a wide and encompassing system that starts with top management understanding and involvement, focuses on process improvement across the supply chain, instills a structured approach to change management, and emphasizes people management and development.

IV. BPM FOR SUCCESSFUL HIS IMPLEMENTATION

The importance and impact of BPM on integrated HIS success will be demonstrated in this section through assessing the experiences of six hospitals.

1) *Case Study Hospitals*

The case studies analyzed in this paper are shown in Table 1.

2) *Bpm Elements*

As noted earlier, BPM has several main pillars. The following are highlights to demonstrate their importance in successful integrated HIS implementation

a. *Executive Leadership*

The experience of all six hospitals highlight the importance of having executive leadership directly involved in planning and implementing an HIS. KFSHRC's executive was instrumental in overseeing its integrated HIS project and the entire board reviewed and approved the plans. At KFMC and Dallah hospital, the decision to implement an integrated HIS was also made at the board level and the senior management team input was very important when selecting a suitable vendor.Executive support and commitment does not end with initiation and facilitation, but must extend to the full implementation of an integrated HIS. KFSHRC, KFMC, SGHG, and Dallah hospital noted that HIS implementation is about people, not technology. The organization went through a major transformation, and the management of this change was carefully planned (from a strategic viewpoint) and meticulously implemented [10]. All the case studies analyzed have shown that the key to a smooth rollout is the effective change management from top. Intervention from management has been necessary to crucial for the adequate resourcing of the project, to taking fast and effective decisions, resolve conflicts and bring everybody to the same thinking, to promote company-wide acceptance of the project, and to build co-operation among the diverse groups in the organization, and in many times across national borders. Executive needs to constantly monitor the progress of the project and provide direction to the implementation teams

.

Table1. Case Studies of successful integrated HIS

Hospital	Major integrated HIS Results
King Faisal Specialist Hospital & Research Centre (KFSHRC)	Improving the quality of patient care. Makes data retrieval faster Makes management decision faster. Provides better service to Patients. More expansion and increase of activities in the same resource.
King Fahad Medical City (KFMC)	Comprehensive Performance Reports. Powerful Search Facility for patient records. Use efforts and time more effectively and productively. Lower inventory holding. Better decision.
Dallah Hospital	Reduce operation cost. Access to timely and complete information Cut the costs of operational systems, improved the reliability of customer service, and assured timely delivery and follow-up. Data integration and standardization.
Al-Noor Specialist Hospital	Makes data retrieval faster. Provides better service to Patients. Improved inventory record accuracy Enhanced data visibility. Reduction in operation costs.
Dr. Abdul Rahman Al-Mishari Hospital	Comprehensive Performance Reports. increased revenue and the decreased costs Information can be accessed in real-time, meeting one of the prime objectives of the project.
Saudi German Hospitals Group (SGHG)	Improving the quality of patient care. Well controlled inventory system. Improved financial control. Reduce administrative costs.

b. Process Management and Improvement

The two main areas in process management and improvement that directly affected integrated HIS success were business process change, performance measurement, and putting in place the appropriate process management structure.Business Process Change – Proper business processes, re-engineering and accurate definition of workflows incorporating global best practices will improve the effectiveness and efficiency of the hospital and in turn provide better patient care. The most common reason that hospitals walk away integrated HIS projects is that they discover that the software does not support one of their important business processes. At that point there are two things they can do:

a). They can change the business process to accommodate the HIS software, which will mean deep changes in long-established ways of doing business (that often provide better services to patients and staff) and shake up important peoples' roles and responsibilities. Or they can modify the HIS software to fit the process, which will slow down the project, introduce risky bugs into the system and make upgrading the software to the integrated HIS vendor's next release excruciatingly difficult, because the customizations will need to be torn apart and rewritten to fit with the new version. Without exception, all six hospitals agreed that BPC is one of the main critical success factors for integrated HIS success. Rather than attempting to modify the software, KFSHRC, KFMC, Dr. Abdul Rahman Al-Mishari hospital, Al-Noor specialist hospital, Dallah hospital, and SGHG redesigned their business processes to be consistent with the software. This has proved to be critical to the project's success. The others undertook a mix of BPC and HIS software re-adjustment. Within this context, KFSHRC and KFMC have strongly emphasized on the criticality of structured project management approaches for integrated HIS success.

b). Performance Measurement – It has been said that you can not manage what you do not measure, and this is especially true in the case of integrated HIS implementation. Health sectors must be able to establish a clear and well defined performance measurement system to allow them to assess the development, and/or problems, that are occurring. KFSHRC, KFMC, and Dallah hospital noted that having a well established measurement system was crucial in their HIS project management initiative to allow for measuring and publicizing success stories for motivation, assessing progress, assigning and redirecting resources, and instilling an overall system of continuous improvement for the integrated HIS life cycle.

c). Process Management Structure - KFMC put someone "in charge" and centralized the management structure of the project in order to avoid duplication of effort. KFMC considered their project a success because of a centralized management structure. This has been implemented by KFSHRC, and Dallah hospital all saw this as an important factor in managing the HIS implementation efficiently. However, even those with no 'HIS Process Leader' still maintained this focus by appointing a 'champion'. The project leader for the HIS project was clearly a "champion" for the project, and that role was critical to marketing the project throughout the organization.

c. Change Management

The main hurdle faced by all the organizations was resistance to change. There will be resistance from users (for instance, from nurses, phlebotomists other paramedical staff, etc) who may feel that feeding information into the computerized system is additional work and not their primary responsibility or core competence. Indeed, staff were reluctant to learn new techniques or the IT department was reluctant to change due to attachment to its product; this was one of the main hurdles faced during the integrated HIS implementations [15,16]. For users, the implementation of HIS means that their computer-related job tasks is completed in totally different computer environment. The complexity of these systems results in enormous learning curves and behavioral changes for user, implementers, and organizations. A variety of reactions by individuals, ranging from resisting to enthusiastically embracing HISs, are demonstrated, and unexpected difficulties often arise during all phases of implementation. Consequently, HIS users need to make sense of, and understand, their reactions to this technology, and their changing computer environment and computer-related job tasks. The attribution of HIS performance are important because they can either positively or negatively influence user's learning, confidence levels, effort, persistence, and use of these systems. Unfortunately, our understanding of individuals' reactions to HISs, and why they elect to use or avoid them, is limited [9].Four elements which can help reduce the resistance are tremendous executive support; training and education, placement of best people on implementation; and heavy involvement of people from the field. The main approaches to achieve the sought-after people involvement and commitment are an open environment, characterized by open communication and trust. Dr. Abdul Rahman Al-Mishari hospital, AL-Noor specialist hospital, KFSHRC, Dallah hospital, and KFMC agreed that effective communications should tell everyone in advance what is happening including the scope, objectives, and activities of the project, and admit that there will be change. Dallah hospital and KFMC saw an open and honest information policy helping the user to become acquainted with the new situation, to build up confidence in the project and its members, and finally to accept the project.Open communication and ethical behavior generate trust. KFMC highlighted the relationships of trust among the project members as a main success factor for HIS package. KFSHRC noted that trust can be built up with intensive communication, coaching, delegation of responsibility, personal care and attention, among other things.

d. People Management And Development

People management is clearly a subset of change management. However, some specific issues have been shown to directly affect the success of integrated HIS implementation, and were mainly in the area of people development. The implementation of integrated HIS package requires a whole new set of skills and expertise that organizations must pay extra attention to where these skills will come from. Two main streams have emerged and all six organizations have used a mix of both:Training and re-skilling - Rigorous and continuous training and showing tangible benefits are the answers to overcome the initial resistance. Training is critical in an integrated HIS project. The most effective HIS possible will not improve health departments if their employees do not know how to use it.

Installing an integrated HIS package without adequate staff preparation could yield drastic consequences. In this respect, KFSHRC and KFMC noted that the costs of training and support are often under-estimated, and these costs may be many times greater than originally anticipated. At Al-Noor specialist hospital, Dr. Abdul Rahman Al-Mishari hospital, Dallah hospital, and SGHG one of the critical workforce requirements for the project was the ability to obtain and train analysts with both "business" and technology knowledge. However, retaining these professionals was a significant problem because of their market value. SGHG, KFMC and KFSHRC invested heavily in training and re-skilling their developers in integrated HIS package software design and methodology. Dr. Abdul Rahman Al-Mishari hospital considered their project a success because of investments in training and support required to overcome technical and procedural challenges in design and implementation.

(b) Using external consultants - With new technology,it is often critical to acquire external expertise, including vendor support, to facilitate successful implementation. Hundreds of companies provide integrated information systems services. Those services may include all or some combination of these offerings:

- Road Map
- Change management
- HIS package selection
- Business process planning or changing
- HIS implementation
- Training
- HIS maintenance and support

Quite simply, when they didn't have needed expertise internally, KFMC brought in the consultants they needed. They stressed that good consultants improve throughput time and quality. The success of a project depends strongly on the capabilities of the consultants because the consultant is the only one with in-depth knowledge of the integrated software

V. Best Practices for Capitalizing on BPM for Successful HIS

This study relieves that the best practices for capitalizing on business process management for successful HIS implementation, are the following:The success of a major project like an integrated HIS implementation hinges on the sustained commitment form executive. An overall commitment that is visible, well defined and felt is a sure way to ensure a successful outcome.

1..Implementing an integrated HIS is not a matter of changing software systems; rather it is a matter of repositioning the health care sector and transforming the business practices.

2..Training - whole departments must be retrained, jobs redefined, and procedures discarded or rebuilt from scratch.

3..Performance Measurement - Because the successful implementation of an integrated HIS is contingent upon an accurate assessment of the associated organizational changes, there is a need to investigate the organizational consequences of HIS software.

4..Selecting the right employees to participate in the implementation process and motivating them is critical for the implementation's success

VI. Conclusion

Healthcare departments involve complex processes that span diverse groups and organizations. The implementation of Health Information System (HIS) to manage and automate the processes has increasingly played an important role in improving the efficiency of healthcare enterprises. However, most of the Health Information System (HIS) implementations are big failures considering the time taken or the desired results achieved. However, the benefits of implementing a fully integrated Health Information System (HIS) – better patient care, increased efficiency, lower costs, etc. – can be enormous. But the price tag can also be large, and the time-to-payback long. Overall, it can be concluded that integrated HIS is far form being an IT project, and is more of an integrated clinical development approach that changes the way health departments do business, and the way work is done. Consequently, to implement HIS successfully, health departments must treat it like a change management project and focus on an integrated approach of Business Process Management. This paper has investigated the role and impact of business process management in successful integrated HIS implementation.

VII. References

1) Anderson, J. “Increasing the Acceptance of Clinical Information Systems”, MD Computing, Jan-Feb, 1999, pp. 62-65.
2) Beynon-Davies, P., Lloyd- Williams, M. “When health information systems fail”, Topic in
3) Health Information Management, Vol. 20, No 1, 1999, pp 66- 79.Cain, M. and Mittman, R. “Diffusion ofInnovation in Health Care”, Oakland, CA: California Health Care Foundation, May 2002.
4) Holland, C. and Light, B. “A Critical Success Factors Model For ERP Implementation”, IEEE Software, May- June, 1999.
5) Kelly, S.; Holland, C. and Light, B. “Enterprise Resource planning: A BusinessApproach ToSystems Development”, In: Proceedings of AMCIS, Milwaukee,WI, USA, 1999.
6) Rosemann, M. and Wises, J.“Measuring the performance of ERP software: a Balanced Scorecard approach”, Wellington, New Zealand, 1-3 December 1999.
7) Sumner, M. “Critical Success Factors in Enterprise Wide Information Management Systems projects”, In: Proceedings of AMCIS, Milwaukee, WI, USA, 13-15 August 1999.
8) Venrura, P. “Poor Project Planning Leads to ERP Failure”, Arabian Computer News, Vol. 18, July, 2003.

9) Wild, E., Hastigs, T., Gubernick, R., Ross, D., Fehrenbach, S. “Key Elements for Successful Integrated Health Information Systems: Lessons From the States”, *Journal of Health Management Practice*, November, 2004, pp. 36-47.

10) Brenda L. Killingsworth, ElaineSeeman. "An Integrative Health Information Systems, Approach for Facilitating Strategic Planning In Hospitals", in Proceeding of the Southern Association of Information Systems Conference, 2005.
11) Kim, K.K, & Michelman, J.E. "AnExamination of Factors for the Strategic Useof Information Systems in the Healthcare Industry", MIS Quarterly June, 1990.
12) Johnson K.F., "Integrated Systems Brings Hospital Data Together",Health Progress, October, 1987.
13) Hejna, W. J., and Hosking, J.E., "Five Critical Strategies for Achieving Operational Efficiency".*Journal of Healthcare Management*, September 2004.
14) K. Dhinesh Kumar, H. Roth, L. Karunamoorthy, "Critical Success Factor for the Implementation of Integrated Automation Solutions with PC Based Control, in: Proceeding of the 10th Mediterranean Conference on Control and Automation MED 2002, Lisbon, Portugal, July 9-12, 2002.
15) Uttam Kumar Tripathi, Knut Hinkelmann and Daneila Feldkamp, "Life Cycle for Change Management in Business Process using Semantic Technologies", *Journal of Computers*, Vol. 3, No. 1, January, 2008.
16) Hartini Ahmad, Arthur Francis and Mahamed Zairi, "Business Process Re-Engineering: Critical Success Factors in Higher Education", Business Process
Management Journal Vol. 13, No. 3, 2007.
17) Cecilia Kennedy, "Critical Success Factors for Implementing a Clinical Information System", Health Care Industry, September, 2000.
18) Public Health Informatics Institute "Evaluation Toolkit for Integrated Health Information Systems", February, 2007.

Global Journals Inc. (US) Guidelines Handbook 2011

www.GlobalJournals.org

FELLOW OF INTERNATIONAL CONGRESS OF COMPUTER SCIENCE AND TECHNOLOGY (FICCT)

- FICCT' title will be awarded to the person after approval of Editor-in-Chief and Editorial Board. The title 'FICCT" can be added to name in the following manner e.g. **Dr. Andrew Knoll, Ph.D., FICCT, Er. Pettor Jone, M.E., FICCT**
- FICCT can submit two papers every year for publication without any charges. The paper will be sent to two peer reviewers. The paper will be published after the acceptance of peer reviewers and Editorial Board.
- Free unlimited Web-space will be allotted to 'FICCT 'along with subDomain to contribute and partake in our activities.
- A professional email address will be allotted free with unlimited email space.
- FICCT will be authorized to receive e-Journals - GJCST for the Lifetime.
- FICCT will be exempted from the registration fees of Seminar/Symposium/Conference/Workshop conducted internationally of GJCST (FREE of Charge).
- FICCT will be an Honorable Guest of any gathering hold.

ASSOCIATE OF INTERNATIONAL CONGRESS OF COMPUTER SCIENCE AND TECHNOLOGY (AICCT)

- AICCT title will be awarded to the person/institution after approval of Editor-in-Chef and Editorial Board. The title 'AICCTcan be added to name in the following manner:

  eg. **Dr. Thomas Herry, Ph.D., AICCT**
- AICCT can submit one paper every year for publication without any charges. The paper will be sent to two peer reviewers. The paper will be published after the acceptance of peer reviewers and Editorial Board.
- Free 2GB Web-space will be allotted to 'FICCT' along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted with free 1GB email space.
- AICCT will be authorized to receive e-Journal GJCST for lifetime.
- A professional email address will be allotted with free 1GB email space.
- AICHSS will be authorized to receive e-Journal GJHSS for lifetime.

ANNUAL MEMBER

- Annual Member will be authorized to receive e-Journal GJCST for one year (subscription for one year).
- The member will be allotted free 1 GB Web-space along with subDomain to contribute and participate in our activities.
- A professional email address will be allotted free 500 MB email space.

PAPER PUBLICATION

- The members can publish paper once. The paper will be sent to two-peer reviewer. The paper will be published after the acceptance of peer reviewers and Editorial Board.

Process of submission of Research Paper

The Area or field of specialization may or may not be of any category as mentioned in 'Scope of Journal' menu of the GlobalJournals.org website. There are 37 Research Journal categorized with Six parental Journals GJCST, GJMR, GJRE, GJMBR, GJSFR, GJHSS. For Authors should prefer the mentioned categories. There are three widely used systems UDC, DDC and LCC. The details are available as 'Knowledge Abstract' at Home page. The major advantage of this coding is that, the research work will be exposed to and shared with all over the world as we are being abstracted and indexed worldwide.

The paper should be in proper format. The format can be downloaded from first page of 'Author Guideline' Menu. The Author is expected to follow the general rules as mentioned in this menu. The paper should be written in MS-Word Format (*.DOC,*.DOCX).

The Author can submit the paper either online or offline. The authors should prefer online submission.

Online Submission: There are three ways to submit your paper:

(A) (I) Register yourself using top right corner of Home page then Login from same place twice. If you are already registered, then login using your username and password.

(II) Choose corresponding Journal from "Research Journals" Menu.

(III) Click 'Submit Manuscript'. Fill required information and Upload the paper.

(B) If you are using Internet Explorer (Although Mozilla Firefox is preferred), then Direct Submission through Homepage is also available.

(C) If these two are not convenient, and then email the paper directly to dean@globaljournals.org as an attachment.

Offline Submission: Author can send the typed form of paper by Post. However, online submission should be preferred.

Preferred Author Guidelines

MANUSCRIPT STYLE INSTRUCTION **(Must be strictly followed)**

Page Size: 8.27" X 11'"

- Left Margin: 0.65
- Right Margin: 0.65
- Top Margin: 0.75
- Bottom Margin: 0.75
- Font type of all text should be Times New Roman.
- Paper Title should be of Font Size 24 with one Column section.
- Author Name in Font Size of 11 with one column as of Title.
- Abstract Font size of 9 Bold, "Abstract" word in Italic Bold.
- Main Text: Font size 10 with justified two columns section
- Two Column with Equal Column with of 3.38 and Gaping of .2
- First Character must be two lines Drop capped.
- Paragraph before Spacing of 1 pt and After of 0 pt.
- Line Spacing of 1 pt
- Large Images must be in One Column
- Numbering of First Main Headings (Heading 1) must be in Roman Letters, Capital Letter, and Font Size of 10.
- Numbering of Second Main Headings (Heading 2) must be in Alphabets, Italic, and Font Size of 10.

You can use your own standard format also.

Author Guidelines:

1. General,

2. Ethical Guidelines,

3. Submission of Manuscripts,

4. Manuscript's Category,

5. Structure and Format of Manuscript,

6. After Acceptance.

1. GENERAL

Before submitting your research paper, one is advised to go through the details as mentioned in following heads. It will be beneficial, while peer reviewer justify your paper for publication.

Scope

The Global Journals Inc. (US) welcome the submission of original paper, review paper, survey article relevant to the all the streams of Philosophy and knowledge. The Global Journals Inc. (US) is parental platform for Global Journal of Computer Science and Technology, Researches in Engineering, Medical Research, Science Frontier Research, Human Social Science, Management, and Business organization. The choice of specific field can be done otherwise as following in Abstracting and Indexing Page on this Website. As the all Global

Journals Inc. (US) are being abstracted and indexed (in process) by most of the reputed organizations. Topics of only narrow interest will not be accepted unless they have wider potential or consequences.

2. ETHICAL GUIDELINES

Authors should follow the ethical guidelines as mentioned below for publication of research paper and research activities.

Papers are accepted on strict understanding that the material in whole or in part has not been, nor is being, considered for publication elsewhere. If the paper once accepted by Global Journals Inc. (US) and Editorial Board, will become the copyright of the Global Journals Inc. (US).

Authorship: The authors and coauthors should have active contribution to conception design, analysis and interpretation of findings. They should critically review the contents and drafting of the paper. All should approve the final version of the paper before submission

The Global Journals Inc. (US) follows the definition of authorship set up by the Global Academy of Research and Development. According to the Global Academy of R&D authorship, criteria must be based on:

1) Substantial contributions to conception and acquisition of data, analysis and interpretation of the findings.

2) Drafting the paper and revising it critically regarding important academic content.

3) Final approval of the version of the paper to be published.

All authors should have been credited according to their appropriate contribution in research activity and preparing paper. Contributors who do not match the criteria as authors may be mentioned under Acknowledgement.

Acknowledgements: Contributors to the research other than authors credited should be mentioned under acknowledgement. The specifications of the source of funding for the research if appropriate can be included. Suppliers of resources may be mentioned along with address.

Appeal of Decision: The Editorial Board's decision on publication of the paper is final and cannot be appealed elsewhere.

Permissions: It is the author's responsibility to have prior permission if all or parts of earlier published illustrations are used in this paper.

Please mention proper reference and appropriate acknowledgements wherever expected.

If all or parts of previously published illustrations are used, permission must be taken from the copyright holder concerned. It is the author's responsibility to take these in writing.

Approval for reproduction/modification of any information (including figures and tables) published elsewhere must be obtained by the authors/copyright holders before submission of the manuscript. Contributors (Authors) are responsible for any copyright fee involved.

3. SUBMISSION OF MANUSCRIPTS

Manuscripts should be uploaded via this online submission page. The online submission is most efficient method for submission of papers, as it enables rapid distribution of manuscripts and consequently speeds up the review procedure. It also enables authors to know the status of their own manuscripts by emailing us. Complete instructions for submitting a paper is available below.

Manuscript submission is a systematic procedure and little preparation is required beyond having all parts of your manuscript in a given format and a computer with an Internet connection and a Web browser. Full help and instructions are provided on-screen. As an author, you will be prompted for login and manuscript details as Field of Paper and then to upload your manuscript file(s) according to the instructions.

To avoid postal delays, all transaction is preferred by e-mail. A finished manuscript submission is confirmed by e-mail immediately and your paper enters the editorial process with no postal delays. When a conclusion is made about the publication of your paper by our Editorial Board, revisions can be submitted online with the same procedure, with an occasion to view and respond to all comments.

Complete support for both authors and co-author is provided.

4. MANUSCRIPT'S CATEGORY

Based on potential and nature, the manuscript can be categorized under the following heads: Original research paper: Such papers are reports of high-level significant original research work.

Review papers: These are concise, significant but helpful and decisive topics for young researchers.

Research articles: These are handled with small investigation and applications

Research letters: The letters are small and concise comments on previously published matters.

5. STRUCTURE AND FORMAT OF MANUSCRIPT

The recommended size of original research paper is less than seven thousand words, review papers fewer than seven thousands words also. Preparation of research paper or how to write research paper, are major hurdle, while writing manuscript. The research articles and research letters should be fewer than three thousand words, the structure original research paper; sometime review paper should be as follows:

Papers: These are reports of significant research (typically less than 7000 words equivalent, including tables, figures, references), and comprise:

(a)Title should be relevant and commensurate with the theme of the paper.

(b) A brief Summary, "Abstract" (less than 150 words) containing the major results and conclusions.

(c) Up to ten keywords, that precisely identifies the paper's subject, purpose, and focus.

(d) An Introduction, giving necessary background excluding subheadings; objectives must be clearly declared.

(e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition; sources of information must be given and numerical methods must be specified by reference, unless non-standard.

(f) Results should be presented concisely, by well-designed tables and/or figures; the same data may not be used in both; suitable statistical data should be given. All data must be obtained with attention to numerical detail in the planning stage. As reproduced design has been recognized to be important to experiments for a considerable time, the Editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned un-refereed;

(g) Discussion should cover the implications and consequences, not just recapitulating the results; conclusions should be summarizing.

(h) Brief Acknowledgements.

(i) References in the proper form.

Authors should very cautiously consider the preparation of papers to ensure that they communicate efficiently. Papers are much more likely to be accepted, if they are cautiously designed and laid out, contain few or no errors, are summarizing, and be conventional to the approach and instructions. They will in addition, be published with much less delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and to make suggestions to improve briefness.

It is vital, that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

Format

Language: The language of publication is UK English. Authors, for whom English is a second language, must have their manuscript efficiently edited by an English-speaking person before submission to make sure that, the English is of high excellence. It is preferable, that manuscripts should be professionally edited.

Standard Usage, Abbreviations, and Units: Spelling and hyphenation should be conventional to The Concise Oxford English Dictionary. Statistics and measurements should at all times be given in figures, e.g. 16 min, except for when the number begins a sentence. When the number does not refer to a unit of measurement it should be spelt in full unless, it is 160 or greater.

Abbreviations supposed to be used carefully. The abbreviated name or expression is supposed to be cited in full at first usage, followed by the conventional abbreviation in parentheses.

Metric SI units are supposed to generally be used excluding where they conflict with current practice or are confusing. For illustration, 1.4 l rather than 1.4 × 10-3 m3, or 4 mm somewhat than 4 × 10-3 m. Chemical formula and solutions must identify the form used, e.g. anhydrous or hydrated, and the concentration must be in clearly defined units. Common species names should be followed by underlines at the first mention. For following use the generic name should be constricted to a single letter, if it is clear.

Structure

All manuscripts submitted to Global Journals Inc. (US), ought to include:
Title: The title page must carry an instructive title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) wherever the work was carried out. The full postal address in addition with the e-mail address of related author must be given. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining and indexing.
Abstract, used in Original Papers and Reviews:
Optimizing Abstract for Search Engines
Many researchers searching for information online will use search engines such as Google, Yahoo or similar. By optimizing your paper for search engines, you will amplify the chance of someone finding it. This in turn will make it more likely to be viewed and/or cited in a further work. Global Journals Inc. (US) have compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Key Words
A major linchpin in research work for the writing research paper is the keyword search, which one will employ to find both library and Internet resources.
One must be persistent and creative in using keywords. An effective keyword search requires a strategy and planning a list of possible keywords and phrases to try.

Search engines for most searches, use Boolean searching, which is somewhat different from Internet searches. The Boolean search uses "operators," words (and, or, not, and near) that enable you to expand or narrow your affords. Tips for research paper while preparing research paper are very helpful guideline of research paper.

Choice of key words is first tool of tips to write research paper. Research paper writing is an art.A few tips for deciding as strategically as possible about keyword search:

- One should start brainstorming lists of possible keywords before even begin searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in research paper?" Then consider synonyms for the important words.
- It may take the discovery of only one relevant paper to let steer in the right keyword direction because in most databases, the keywords under which a research paper is abstracted are listed with the paper.
- One should avoid outdated words.

Keywords are the key that opens a door to research work sources. Keyword searching is an art in which researcher's skills are bound to improve with experience and time.

Numerical Methods: Numerical methods used should be clear and, where appropriate, supported by references.

Acknowledgements: Please make these as concise as possible.

References

References follow the Harvard scheme of referencing. References in the text should cite the authors' names followed by the time of their publication, unless there are three or more authors when simply the first author's name is quoted followed by et al. unpublished work has to only be cited where necessary, and only in the text. Copies of references in press in other journals have to be supplied with submitted typescripts. It is necessary that all citations and references be carefully checked before submission, as mistakes or omissions will cause delays.

References to information on the World Wide Web can be given, but only if the information is available without charge to readers on an official site. Wikipedia and Similar websites are not allowed where anyone can change the information. Authors will be asked to make available electronic copies of the cited information for inclusion on the Global Journals Inc. (US) homepage at the judgment of the Editorial Board.

The Editorial Board and Global Journals Inc. (US) recommend that, citation of online-published papers and other material should be done via a DOI (digital object identifier). If an author cites anything, which does not have a DOI, they run the risk of the cited material not being noticeable.

The Editorial Board and Global Journals Inc. (US) recommend the use of a tool such as Reference Manager for reference management and formatting.

Tables, Figures and Figure Legends

Tables: Tables should be few in number, cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g. Table 4, a self-explanatory caption and be on a separate sheet. Vertical lines should not be used.

Figures: Figures are supposed to be submitted as separate files. Always take in a citation in the text for each figure using Arabic numbers, e.g. Fig. 4. Artwork must be submitted online in electronic form by e-mailing them.

Preparation of Electronic Figures for Publication

Even though low quality images are sufficient for review purposes, print publication requires high quality images to prevent the final product being blurred or fuzzy. Submit (or e-mail) EPS (line art) or TIFF (halftone/photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Do not use pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings) in relation to the imitation size. Please give the data for figures in black and white or submit a Color Work Agreement Form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution (at final image size) ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs) : >350 dpi; figures containing both halftone and line images: >650 dpi.

Color Charges: It is the rule of the Global Journals Inc. (US) for authors to pay the full cost for the reproduction of their color artwork. Hence, please note that, if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a color work agreement form before your paper can be published.

Figure Legends: Self-explanatory legends of all figures should be incorporated separately under the heading 'Legends to Figures'. In the full-text online edition of the journal, figure legends may possibly be truncated in abbreviated links to the full screen version. Therefore, the first 100 characters of any legend should notify the reader, about the key aspects of the figure.

6. AFTER ACCEPTANCE

Upon approval of a paper for publication, the manuscript will be forwarded to the dean, who is responsible for the publication of the Global Journals Inc. (US).

6.1 Proof Corrections

The corresponding author will receive an e-mail alert containing a link to a website or will be attached. A working e-mail address must therefore be provided for the related author.

Acrobat Reader will be required in order to read this file. This software can be downloaded

(Free of charge) from the following website:

www.adobe.com/products/acrobat/readstep2.html. This will facilitate the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent with the proof.

Proofs must be returned to the dean at dean@globaljournals.org within three days of receipt.

As changes to proofs are costly, we inquire that you only correct typesetting errors. All illustrations are retained by the publisher. Please note that the authors are responsible for all statements made in their work, including changes made by the copy editor.

6.2 Early View of Global Journals Inc. (US) (Publication Prior to Print)

The Global Journals Inc. (US) are enclosed by our publishing's Early View service. Early View articles are complete full-text articles sent in advance of their publication. Early View articles are absolute and final. They have been completely reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after sending them. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the conventional way.

6.3 Author Services

Online production tracking is available for your article through Author Services. Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The authors will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript.

6.4 Author Material Archive Policy

Please note that if not specifically requested, publisher will dispose off hardcopy & electronic information submitted, after the two months of publication. If you require the return of any information submitted, please inform the Editorial Board or dean as soon as possible.

6.5 Offprint and Extra Copies

A PDF offprint of the online-published article will be provided free of charge to the related author, and may be distributed according to the Publisher's terms and conditions. Additional paper offprint may be ordered by emailing us at: editor@globaljournals.org .

Informal Tips for writing a Computer Science Research Paper to increase readability and citation

Before start writing a good quality Computer Science Research Paper, let us first understand what is Computer Science Research Paper? So, Computer Science Research Paper is the paper which is written by professionals or scientists who are associated to Computer Science and Information Technology, or doing research study in these areas. If you are novel to this field then you can consult about this field from your supervisor or guide.

Techniques for writing a good quality Computer Science Research Paper:

1. Choosing the topic- In most cases, the topic is searched by the interest of author but it can be also suggested by the guides. You can have several topics and then you can judge that in which topic or subject you are finding yourself most comfortable. This can be done by asking several questions to yourself, like Will I be able to carry our search in this area? Will I find all necessary recourses to accomplish the search? Will I be able to find all information in this field area? If the answer of these types of questions will be "Yes" then you can choose that topic. In most of the cases, you may have to conduct the surveys and have to visit several places because this field is related to Computer Science and Information Technology. Also, you may have to do a lot of work to find all rise and falls regarding the various data of that subject. Sometimes, detailed information plays a vital role, instead of short information.

2. Evaluators are human: First thing to remember that evaluators are also human being. They are not only meant for rejecting a paper. They are here to evaluate your paper. So, present your Best.

3. Think Like Evaluators: If you are in a confusion or getting demotivated that your paper will be accepted by evaluators or not, then think and try to evaluate your paper like an Evaluator. Try to understand that what an evaluator wants in your research paper and automatically you will have your answer.

4. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

5. Ask your Guides: If you are having any difficulty in your research, then do not hesitate to share your difficulty to your guide (if you have any). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work then ask the supervisor to help you with the alternative. He might also provide you the list of essential readings.

6. Use of computer is recommended: As you are doing research in the field of Computer Science, then this point is quite obvious.

7. Use right software: Always use good quality software packages. If you are not capable to judge good software then you can lose quality of your paper unknowingly. There are various software programs available to help you, which you can get through Internet.

8. Use the Internet for help: An excellent start for your paper can be by using the Google. It is an excellent search engine, where you can have your doubts resolved. You may also read some answers for the frequent question how to write my research paper or find model research paper. From the internet library you can download books. If you have all required books make important reading selecting and analyzing the specified information. Then put together research paper sketch out.

9. Use and get big pictures: Always use encyclopedias, Wikipedia to get pictures so that you can go into the depth.

10. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right! It is a good habit, which helps to not to lose your continuity. You should always use bookmarks while searching on Internet also, which will make your search easier.

11. Revise what you wrote: When you write anything, always read it, summarize it and then finalize it.

12. Make all efforts: Make all efforts to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in introduction, that what is the need of a particular research paper. Polish your work by good skill of writing and always give an evaluator, what he wants.

13. Have backups: When you are going to do any important thing like making research paper, you should always have backup copies of it either in your computer or in paper. This will help you to not to lose any of your important.

14. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several and unnecessary diagrams will degrade the quality of your paper by creating "hotchpotch." So always, try to make and include those diagrams, which are made by your own to improve readability and understandability of your paper.

15. Use of direct quotes: When you do research relevant to literature, history or current affairs then use of quotes become essential but if study is relevant to science then use of quotes is not preferable.

16. Use proper verb tense: Use proper verb tenses in your paper. Use past tense, to present those events that happened. Use present tense to indicate events that are going on. Use future tense to indicate future happening events. Use of improper and wrong tenses will confuse the evaluator. Avoid the sentences that are incomplete.

17. Never use online paper: If you are getting any paper on Internet, then never use it as your research paper because it might be possible that evaluator has already seen it or maybe it is outdated version.

18. **Pick a good study spot:** To do your research studies always try to pick a spot, which is quiet. Every spot is not for studies. Spot that suits you choose it and proceed further.

19. Know what you know: Always try to know, what you know by making objectives. Else, you will be confused and cannot achieve your target.

20. Use good quality grammar: Always use a good quality grammar and use words that will throw positive impact on evaluator. Use of good quality grammar does not mean to use tough words, that for each word the evaluator has to go through dictionary. Do not start sentence with a conjunction. Do not fragment sentences. Eliminate one-word sentences. Ignore passive voice. Do not ever use a big word when a diminutive one would suffice. Verbs have to be in agreement with their subjects. Prepositions are not expressions to finish sentences with. It is incorrect to ever divide an infinitive. Avoid clichés like the disease. Also, always shun irritating alliteration. Use language that is simple and straight forward. put together a neat summary.

21. Arrangement of information: Each section of the main body should start with an opening sentence and there should be a changeover at the end of the section. Give only valid and powerful arguments to your topic. You may also maintain your arguments with records.

22. Never start in last minute: Always start at right time and give enough time to research work. Leaving everything to the last minute will degrade your paper and spoil your work.

23. Multitasking in research is not good: Doing several things at the same time proves bad habit in case of research activity. Research is an area, where everything has a particular time slot. Divide your research work in parts and do particular part in particular time slot.

24. Never copy others' work: Never copy others' work and give it your name because if evaluator has seen it anywhere you will be in trouble.

25. Take proper rest and food: No matter how many hours you spend for your research activity, if you are not taking care of your health then all your efforts will be in vain. For a quality research, study is must, and this can be done by taking proper rest and food.

26. Go for seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

27. Refresh your mind after intervals: Try to give rest to your mind by listening to soft music or by sleeping in intervals. This will also improve your memory.

28. Make colleagues: Always try to make colleagues. No matter how sharper or intelligent you are, if you make colleagues you can have several ideas, which will be helpful for your research.

29. **Think technically:** Always think technically. If anything happens, then search its reasons, its benefits, and demerits.

30. Think and then print: When you will go to print your paper, notice that tables are not be split, headings are not detached from their descriptions, and page sequence is maintained.

31. Adding unnecessary information: Do not add unnecessary information, like, I have used MS Excel to draw graph. Do not add irrelevant and inappropriate material. These all will create superfluous. Foreign terminology and phrases are not apropos. One should NEVER take a broad view. Analogy in script is like feathers on a snake. Not at all use a large word when a very small one would be sufficient. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Amplification is a billion times of inferior quality than sarcasm.

32. Never oversimplify everything: To add material in your research paper, never go for oversimplification. This will definitely irritate the evaluator. Be more or less specific. Also too, by no means, ever use rhythmic redundancies. Contractions aren't essential and shouldn't be there used. Comparisons are as terrible as clichés. Give up ampersands and abbreviations, and so on. Remove commas, that are, not necessary. Parenthetical words however should be together with this in commas. Understatement is all the time the complete best way to put onward earth-shaking thoughts. Give a detailed literary review.

33. Report concluded results: Use concluded results. From raw data, filter the results and then conclude your studies based on measurements and observations taken. Significant figures and appropriate number of decimal places should be used. Parenthetical remarks are prohibitive. Proofread carefully at final stage. In the end give outline to your arguments. Spot out perspectives of further study of this subject. Justify your conclusion by at the bottom of them with sufficient justifications and examples.

34. After conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print to the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects in your research.

INFORMAL GUIDELINES OF RESEARCH PAPER WRITING

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form, which is presented in the guidelines using the template.
- Please note the criterion for grading the final paper by peer-reviewers.

Final Points:

A purpose of organizing a research paper is to let people to interpret your effort selectively. The journal requires the following sections, submitted in the order listed, each section to start on a new page.

The introduction will be compiled from reference matter and will reflect the design processes or outline of basis that direct you to make study. As you will carry out the process of study, the method and process section will be constructed as like that. The result segment will show related statistics in nearly sequential order and will direct the reviewers next to the similar intellectual paths throughout the data that you took to carry out your study. The discussion section will provide understanding of the data and projections as to the implication of the results. The use of good quality references all through the paper will give the effort trustworthiness by representing an alertness of prior workings.

Writing a research paper is not an easy job no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record keeping are the only means to make straightforward the progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear

· Adhere to recommended page limits

Mistakes to evade

- Insertion a title at the foot of a page with the subsequent text on the next page

- Separating a table/chart or figure - impound each figure/table to a single page
- Submitting a manuscript with pages out of sequence

In every sections of your document

· Use standard writing style including articles ("a", "the," etc.)

· Keep on paying attention on the research topic of the paper

· Use paragraphs to split each significant point (excluding for the abstract)

· Align the primary line of each section

· Present your points in sound order

· Use present tense to report well accepted

· Use past tense to describe specific results

· Shun familiar wording, don't address the reviewer directly, and don't use slang, slang language, or superlatives

· Shun use of extra pictures - include only those figures essential to presenting results

Title Page:

Choose a revealing title. It should be short. It should not have non-standard acronyms or abbreviations. It should not exceed two printed lines. It should include the name(s) and address (es) of all authors.

Abstract:

The summary should be two hundred words or less. It should briefly and clearly explain the key findings reported in the manuscript--must have precise statistics. It should not have abnormal acronyms or abbreviations. It should be logical in itself. Shun citing references at this point.

An abstract is a brief distinct paragraph summary of finished work or work in development. In a minute or less a reviewer can be taught the foundation behind the study, common approach to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Yet, use comprehensive sentences and do not let go readability for briefness. You can maintain it succinct by phrasing sentences so that they provide more than lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study, with the subsequent elements in any summary. Try to maintain the initial two items to no more than one ruling each.

- Reason of the study - theory, overall issue, purpose
- Fundamental goal
- To the point depiction of the research
- Consequences, including definite statistics - if the consequences are quantitative in nature, account quantitative data; results of any numerical analysis should be reported
- Significant conclusions or questions that track from the research(es)

Approach:

- Single section, and succinct
- As a outline of job done, it is always written in past tense
- A conceptual should situate on its own, and not submit to any other part of the paper such as a form or table
- Center on shortening results - bound background information to a verdict or two, if completely necessary
- What you account in an conceptual must be regular with what you reported in the manuscript
- Exact spelling, clearness of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else

Introduction:

The **Introduction** should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable to comprehend and calculate the purpose of your study without having to submit to other works. The basis for the study should be offered. Give most important references but shun difficult to make a comprehensive appraisal of the topic. In the introduction, describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will have no attention in your result. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here. Following approach can create a valuable beginning:

- Explain the value (significance) of the study
- Shield the model - why did you employ this particular system or method? What is its compensation? You strength remark on its appropriateness from a abstract point of vision as well as point out sensible reasons for using it.
- Present a justification. Status your particular theory (es) or aim(s), and describe the logic that led you to choose them.
- Very for a short time explain the tentative propose and how it skilled the declared objectives.

Approach:

- Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done.
- Sort out your thoughts; manufacture one key point with every section. If you make the four points listed above, you will need a least of four paragraphs.
- Present surroundings information only as desirable in order hold up a situation. The reviewer does not desire to read the whole thing you know about a topic.
- Shape the theory/purpose specifically - do not take a broad view.
- As always, give awareness to spelling, simplicity and correctness of sentences and phrases.

Procedures (Methods and Materials):

This part is supposed to be the easiest to carve if you have good skills. A sound written Procedures segment allows a capable scientist to replacement your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt for the least amount of information that would permit another capable scientist to spare your outcome but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section. When a technique is used that has been well described in another object, mention the specific item describing a way but draw the basic principle while stating the situation. The purpose is to text all particular resources and broad procedures, so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step by step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

- Explain materials individually only if the study is so complex that it saves liberty this way.
- Embrace particular materials, and any tools or provisions that are not frequently found in laboratories.
- Do not take in frequently found.
- If use of a definite type of tools.
- Materials may be reported in a part section or else they may be recognized along with your measures.

Methods:

- Report the method (not particulars of each process that engaged the same methodology)
- Describe the method entirely
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures
- Simplify - details how procedures were completed not how they were exclusively performed on a particular day.
- If well known procedures were used, account the procedure by name, possibly with reference, and that's all.

Approach:

- It is embarrassed or not possible to use vigorous voice when documenting methods with no using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result when script up the methods most authors use third person passive voice.
- Use standard style in this and in every other part of the paper - avoid familiar lists, and use full sentences.

What to keep away from

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings - save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part a entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Carry on to be to the point, by means of statistics and tables, if suitable, to present consequences most efficiently.You must obviously differentiate material that would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matter should not be submitted at all except requested by the instructor.

Content

- Sum up your conclusion in text and demonstrate them, if suitable, with figures and tables.
- In manuscript, explain each of your consequences, point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation an exacting study.
- Explain results of control experiments and comprise remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or in manuscript form.

What to stay away from

- Do not discuss or infer your outcome, report surroundings information, or try to explain anything.
- Not at all, take in raw data or intermediate calculations in a research manuscript.

- Do not present the similar data more than once.
- Manuscript should complement any figures or tables, not duplicate the identical information.
- Never confuse figures with tables - there is a difference.

Approach

- As forever, use past tense when you submit to your results, and put the whole thing in a reasonable order.
- Put figures and tables, appropriately numbered, in order at the end of the report
- If you desire, you may place your figures and tables properly within the text of your results part.

Figures and tables

- If you put figures and tables at the end of the details, make certain that they are visibly distinguished from any attach appendix materials, such as raw facts
- Despite of position, each figure must be numbered one after the other and complete with subtitle
- In spite of position, each table must be titled, numbered one after the other and complete with heading
- All figure and table must be adequately complete that it could situate on its own, divide from text

Discussion:

The Discussion is expected the trickiest segment to write and describe. A lot of papers submitted for journal are discarded based on problems with the Discussion. There is no head of state for how long a argument should be. Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implication of the study. The purpose here is to offer an understanding of your results and hold up for all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of result should be visibly described. Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved with prospect, and let it drop at that.

- Make a decision if each premise is supported, discarded, or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."
- Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work
- You may propose future guidelines, such as how the experiment might be personalized to accomplish a new idea.
- Give details all of your remarks as much as possible, focus on mechanisms.
- Make a decision if the tentative design sufficiently addressed the theory, and whether or not it was correctly restricted.
- Try to present substitute explanations if sensible alternatives be present.
- One research will not counter an overall question, so maintain the large picture in mind, where do you go next? The best studies unlock new avenues of study. What questions remain?
- Recommendations for detailed papers will offer supplementary suggestions.

Approach:

- When you refer to information, differentiate data generated by your own studies from available information
- Submit to work done by specific persons (including you) in past tense.
- Submit to generally acknowledged facts and main beliefs in present tense.

Administration Rules Listed Before Submitting Your Research Paper to Global Journals Inc. (US)

Please carefully note down following rules and regulation before submitting your Research Paper to Global Journals Inc. (US):

Segment Draft and Final Research Paper: You have to strictly follow the template of research paper. If it is not done your paper may get rejected.

- The **major constraint** is that you must independently make all content, tables, graphs, and facts that are offered in the paper. You must write each part of the paper wholly on your own. The Peer-reviewers need to identify your own perceptive of the concepts in your own terms. NEVER extract straight from any foundation, and never rephrase someone else's analysis.

- Do not give permission to anyone else to "PROOFREAD" your manuscript.

- Methods to avoid Plagiarism is applied by us on every paper, if found guilty, you will be blacklisted by all of our collaborated research groups, your institution will be informed for this and strict legal actions will be taken immediately.)
- To guard yourself and others from possible illegal use please do not permit anyone right to use to your paper and files.

CRITERION FOR GRADING A RESEARCH PAPER (COMPILATION) BY GLOBAL JOURNALS INC. (US)

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals Inc. (US).

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

Index

save our planet

Global Journal of Computer Science and Technology

Visit us on the Web at www.GlobalJournals.org | www.ComputerResearch.org
or email us at helpdesk@globaljournals.org

9 770116 58698 2 61427>

ISSN 9754350