January - June, 2010 Vol. 2 No. 1 ISSN 0973-5658

BVICAM's IJIT

BVICAM's International Journal of Information Technology

CONTENTS

Bharati Vidyapeeth's
Institute of Computer Applications and Management
A-4, Paschim Vihar, Rohtak Road, New Delhi-63

Email : bijit@bvicam.ac.in, Website : http://www.bvicam.ac.in

Volume 2, Number 1

January - June, 2010

BVICAM's International Journal of Information Technology (BIJIT) is a bi-annual publication of Bharati Vidyapeeth's Institute of Computer Applications and Management, A-4, Paschim Vihar, Rohtak Road, New Delhi – 110063.

ISSN 0973 – 5658

Address for Correspondence:

Prof. M. N. Hoda
Chief Editor – BIJIT
Director, Bharati Vidyapeeth's
Institute of Computer Applications and Management,
A-4, Paschim Vihar, Rohtak Road, New Delhi – 110063 (INDIA).
Tel./Fax: 91 – 11 – 25275055 E-Mail: bijit@bvicam.ac.in
Visit us at www.bvicam.ac.in

Published and printed by Prof. M. N. Hoda, Chief Editor – BIJIT and Director, Bharati Vidyapeeth's Institute of Computer Applications and Management, A-4, Paschim Vihar, New Delhi – 63 (INDIA). Tel. / Fax: 91 – 11 – 25275055, E-Mail: bijit@bvicam.ac.in, Visit us at www.bvicam.ac.in

BVICAM's
International Journal of Information Technology (BIJIT)

BIJIT is a bi-annual publication of
Bharati Vidyapeeth's
Institute of Computer Applications and Management
A-4, Paschim Vihar, Rohtak Road, New Delhi – 110063 (INDIA)
Tel.: 91 – 11 – 25275055 Fax: 91 – 11 – 25255056 E-Mail: bijit@bvicam.ac.in
Visit us at www.bvicam.ac.in

Editorial

It is a matter of both honor and pleasure for us to put forth the third issue of BIJIT; the BVICAM's International Journal of Information Technology. This issue of the journal presents a compilation of ten papers that span a broad variety of research topics in various emerging areas of Information Technology and Computer Science. Some application oriented papers, having novelty in application, have also been included in this issue, hoping that usage of these would enrich the knowledge base and facilitate the overall economic growth. This issue shows our commitment in realizing our vision "*to achieve a standard comparable to the best in the field and finally become a symbol of quality"*.

As a matter of policy of the Journal, all the manuscripts received and considered for the Journal by the editorial board are double blind reviewed by at-least two referees. Our panel of expert referees posses a sound academic background and have a rich publication record in various prestigious journals representing Universities, Research Laboratories and other institutions of repute, which, we intend to further augment from time to time. Finalizing the constitution of the panel of referees, for double blind review(s) of the considered manuscripts, was a painstaking process, but it helped us to ensure that the best of the considered manuscripts are showcased and that too after undergoing multiple cycles of review, as required.

The ten papers that were finally published were chosen out of more than eighty papers that we received from all over the world for this issue. We understand that the confirmation of final acceptance, to the authors / contributors, is delayed, but we also hope that you concur with us in the fact that quality review is a time taking process and is further delayed if the reviewers are senior researchers in their respective fields and hence, are hard pressed for time.

We wish to express our sincere gratitude to our panel of experts in steering the considered manuscripts through multiple cycles of review and bringing out the best from the contributing authors. We thank our esteemed authors for having shown confidence in BIJIT and considering it a platform to showcase and share their original research work. We would also wish to thank the authors whose papers were not published in this issue of the Journal, probably because of the minor shortcomings. However, we would like to encourage them to actively contribute for the forthcoming issues.

The undertaken Quality Assurance Process involved a series of well defined activities that, we hope, went a long way in ensuring the quality of the publication. Still, there is always a scope for improvement, and so we request the contributors and readers to kindly mail us their criticism, suggestions and feedback at bijit@bvicam.ac.in and help us in further enhancing the quality of forthcoming issues.

Editors

C O N T E N T S

EECHDA: Energy Efficient Clustering Hierarchy and Data Accumulation For Sensor Networks

Dilip Kumar[1], T. C. Aseri[2] and R. B. Patel[3]

Abstract - A wireless sensor network with a large number of tiny sensor nodes can be used as an effective tool for gathering data for various applications under different situations. One of the major issues in wireless sensor network is developing an energy–efficient routing protocol which has a significant impact on the overall lifetime of the sensor network. Clustering sensor nodes is an effective technique in wireless sensor networks which can increase network energy efficiency, scalability and lifetime. In this paper, we have proposed an energy–efficient clustering based protocol for wireless sensor networks. We have considered a set of cluster heads for control and management of the network. On rotation basis, a cluster head receives data from the neighboring nodes and transmits the aggregated data to the base station. Adopting this approach, Energy Efficient Clustering Hierarchy and Data Accumulation (EECHDA) is better than existing protocols in terms of energy consumption and network lifetime. Our simulation results demonstrated that EECHDA is able to prolong the time interval of the death of first node in the network.

Index Terms -wireless sensor networks; clustering; energy efficient; aggregation; lifetime

1. INTRODUCTION

With the development of the information society, sensors are facing ever more new challenges. Detection and monitoring requirements are becoming more complicated and difficult. They trend from single variable to multiple variables; from one point to a plane; from one sensor to a set of sensors; from simple to complex and cooperative. Networking the sensors to empower them with the ability to coordinate on a larger sensing task will revolutionize information gathering and processing in many situations. Networks of sensors can greatly improve environment monitoring for many civil and military applications. Furthermore, many environments may be unsuitable for humans and thus the use of sensors is the only solution; in some places, although accessible, in general it is more effective to place small autonomous sensors than to use humans for collection of data.

By integrating sensing, signal processing, and communications functions, a sensor network provides a natural platform for hierarchical and efficient information processing. It allows information to be processed on different levels of abstraction, ranging from detailed microscopic examination of specific targets to a macroscopic view of the aggregate behavior of targets. With focus on applications requiring tight coupling with the physical world, as opposed to the personal communication focus of conventional wireless networks, wireless sensor networks pose significantly different design, implementation, and deployment challenges.

As a microelectronic device, the main task of a sensor node is to detect phenomena, carry out data processing timely and locally, and transmit or receive data. A typical sensor node is generally composed of four components [1], [2], [3], [4], [5], [6], [7], [8]: a power supply unit; a sensing unit; a computing/processing unit; and a communicating unit. The sensing node is powered by a limited battery, which is impossible to replace or recharge in most application scenarios. Except for the power unit, the entire network layer in WSNs is responsible for data delivery from source to destination via well-selected routes [9], [10]. Due to the unique characteristics of Wireless Sensor Networks (WSNs), many of the network layer protocols designed for conventional networks may not fit with the requirements of WSNs. The following principles must be considered in WSN network layer protocols:

1. Energy efficiency is always a dominant consideration.
2. Routing is often data centric.
3. Data aggregation/fusion is desirable, but only useful if it does not affect the collaborative efforts among sensor nodes.
4. An ideal sensor network has attribute-based addressing and location awareness.
5. Protocols are most likely application specific.

Depending on how the hierarchical structure is formed, hierarchical protocols can be grouped as reserved tree based, chain based, or clustering based. Among these, the clustering-based approach has received increased attention because of its effectiveness, lower complexity, and flexibility.

In WSNs, a cluster head (CH) is generally a sensor node, which has severe resource limitations, and cluster heads are selected dynamically; therefore, clusters are dynamic within the network, but sensor nodes are often in stationary position. This would reduce the overall energy consumed for data communication over the whole WSN.

Clustering-based schemes also have the advantages of load balancing, and scalability when the network size grows. Challenges faced by such clustering-based approaches include

[1]*Design Engineer, Centre for Development of Advanced Computing (CDAC), A Scientific Society of the Ministry of Communication & Information Technology, Government of India, A-34, Phase-8, Industrial Area, Mohali -160071 (India)*
[2]*Sr. Lecturer, Department of Computer Science & Engineering, Punjab Engineering College (PEC), Deemed University, Sector-12, Chandigarh-160012 (India)*
[3]*Prof. & Head, Department of Computer Science & Engineering, Maharishi Markandeshwar University (MMU), Mullana, Ambala-133203 (India)*
E-Mail: [1]dilipkant@rediffmail.com, [2]trilokchand@pec.ac.in and [3]patel_r_b@yahoo.com

how to select the cluster heads and how to organize the clusters. The clustering strategy could be single-hop cluster or multi hop cluster, based on the distance between the cluster heads and their members. According to the hierarchy of clusters, the clustering strategies can also be grouped into single-level or multilevel clustering.

The principle of data aggregation or data fusion is to minimize traffic load (in terms of number and/or length of packets) by eliminating redundancy. It applies a novel data-centric approach to replace the traditional address-centric approach in data forwarding [11]. Specifically, when an intermediate node receives data from multiple source nodes, instead of forwarding all of them directly, it checks the contents of incoming data and then combines them by eliminating redundant information under the constraints of acceptable accuracy.

In this paper, the main scenario of interest is a cluster based WSN with static homogeneous nodes and energy constrained sensor nodes. All nodes in the network act as sensor nodes collecting information from the environment, apart from they can act as a cluster-head, forwarding the aggregated information to the base station (BS). Each cluster is formed by a set of sensor nodes, one of them assume the role of CH. The cluster head node stores the information it receives and performs the aggregation tasks sending periodical messages to the BS. The proposed routing scheme in this paper is suitable for continuous monitoring of numerous widespread sensors, which are at a large distance from the BS.

The paper is organized as follows: Following the introduction, section 2 summarizes some related work in this area. Section 3 presents our cluster-based hierarchy approach. In Section 4, we perform quantitative analysis for the proposed protocol. Section 5 evaluates the performance of the proposed protocol. Finally, section 6 concludes the paper and provides possible future directions.

2. RELATED WORK

The cluster-based routing protocols are investigated in several research studies. For example, the work in [5] shows that a 2-tier architecture is more energy efficient when hierarchical clusters are deployed at specific locations. In [3], the authors described a multi-level hierarchical clustering algorithm, where the parameters for minimum energy consumption are obtained using stochastic geometry.

Cluster-based approaches are suitable for habitat and environment monitoring, which requires a continuous stream of sensor data. Directed diffusion and its variations are used for event-based monitoring. In [4], authors have described a directed diffusion protocol where query (task) is disseminated into the network using hop-by-hop communication. When the query is traversed, the gradients (interests) are established for the result return path. Finally, the result is routed using the path based on gradients and interests. In [6], a variation of directed diffusion, use rumor routing to flood events and route queries; this approach is suitable for a large number of queries and a fewer events.

In [7], authors have analyzed a method to elect cluster heads according to the energy left in each node. The assumption of global knowledge of the energy left in the whole network makes this method difficult to implement. Even a centralized approach of this method would be very complicated and very slow, as the feedback should be reliably delivered to each sensor in every round.

In [12], it proposes a maximum energy cluster head routing protocol which has self configuration and hierarchical tree routing properties. The proposed protocol improved LEACH in several aspects such as it constructs clusters based on radio range and the number of cluster members and the cluster topology in the network is distributed more equally.

In [13], a novel self –organizing energy efficient hybrid protocol based on LEACH is presented, combining cluster based architecture and multiple-hop routing. Multi-hop routing is utilized for inter-cluster communication between Clusterheads and the base station, instead of direct transmission in order to minimize transmission energy.

LEACH [14][15]16][17][18] is one of the most popular hierarchical routing algorithms for clustering of WSNs. In LEACH, a small number of clusters are formed in a self –organized manner. Thus it is a suitable solution for energy efficiency in the sensor network. Although, LEACH is a sound solution in data gathering, but it has certain issues and have several limitations:

LEACH does not address the problem that some nodes are close to each other and thus redundant data may be transferred to the base station.

Cluster heads are not selected in a distributed manner it is possible that too many CHs are located in a specific area that may not produce good clusters.

On an average five CH nodes transmit the fused data from their cluster to the base station.

In [19] [20], the authors worked on the heterogeneous sensor nodes and evaluated the energy efficiency. The performance measures that have been considered are network lifetime, number of cluster heads, stability, throughput and energy of the system.

3. CLUSTER BASED HIERARCHY ARCHITECTURE

As previously described, LEACH has some issues. In the proposed protocol we have tried to solve these problems. The protocols optimize energy cost when gathering data. In addition, it distributes energy fairly.

The proposed routing scheme is based on the fact that the energy consumed to send a message to a distant node is far greater than the energy needed for a short range transmission. The CHs are responsible for transmitting messages to the distant base station. At one time, only one member of the member node is active and the remaining members are in sleep mode. The task of transmission to the base station is uniformly distributed among all the CHs.

We now describe a few terms that are used in defining our protocol. A CH is a sensor node that transmits an aggregated sensor data to the distant base station. Non-cluster heads are

sensor nodes that transmit the collected data to their cluster head. Each cluster has a head-set that consists of several non-cluster heads nodes. A round consists of two stages: a cluster head election phase and a data transfer phase. In a cluster head election phase, the head-sets are chosen for the pre-determined number of clusters. In the data transfer phase, the CHs nodes transmit aggregated data to the base station.

3.1. Cluster Head Election Phase

In the proposed model, the number of clusters, q, are pre-determined for the wireless sensor network. At the beginning, a set of CHs are chosen on random basis. The sensor nodes closer to the base station can directly send their messages to the base station. Thus they become the member of a cluster. These cluster heads send a short range announces a Join Request message to the nearby nodes. The member nodes receive the advertisements and choose their cluster heads based on the distance. Each member sensor node sends an acknowledgment message to its cluster head. The cluster heads act as local control centers to coordinate the data transmissions in their cluster and are also responsible to send aggregated messages to the distant base station.

3.2. Data Transfer Phase

Once clusters and TDMA-based schedules are formed, data transmission begins. The non-cluster head nodes collect the sensor data and transmit the data to the cluster head, in their allotted timer slots. The cluster-head node must keep its radio turned on to receive the data from the nodes in the cluster. After, some pre-determined time interval, the next non cluster head member becomes a cluster head and the current cluster head becomes a non cluster head member. The energy of the cluster head is drained out more as compared to a non cluster head member; because it has to do more work than the other nodes. In other rounds the higher energy nodes become cluster heads.

4. NETWORK MODEL ANALYSIS

In this section, we describe a radio energy model that is used in the analysis of EECHDA protocol. The energy dissipation, number of frames, time for message transfer, and the optimum number of clusters are analytically determined.

4.1. Radio Energy Dissipation Model

We have used the same radio model as described in [17], where for a shorter distance transmission, such as within clusters, the energy consumed by a transmit amplifier is proportional to d2 where d is the distance between the transmitter unit and the receiver unit. However, for a longer distance transmission, such as from a cluster head to the base station, the energy consumed is proportional to d4. Using the given radio model, the energy consumed to transmit an m-bit message for a longer distance, d, is given by:

$$E_{TL} = E(m, d^4) \quad (1)$$

Similarly, the energy consumed to transmit an m-bit message for a shorter distance is given by:

$$E_{TS} = E(m, d^2) \quad (2)$$

Moreover, the energy consumed to receive the m-bit message is given by:

$$E_{RX} = E(m)_{+\text{EDA}} \quad (3)$$

where E(m) presents the energy consumption of radio dissipation. Additionally, the operation of data aggregation approach consumes the energy as EDA . The constants used in the radio model are given in Table 1.

4.2. Energy Consumption in Election Phase

For a sensor network of N nodes, the optimal number of clusters is given as q. All nodes are assumed to be at the same energy level at the beginning. The amount of consumed energy is same for all the clusters. At the start of the election phase, the base station randomly selects a given number of cluster heads. Initially, the cluster heads broadcast messages to all the sensor nodes in their neighborhood. Next, the sensor nodes receive messages from one or more cluster heads and choose their cluster head using the received signal strength. After this, the sensor nodes transmit their decision to their corresponding cluster heads. Finally, the cluster heads receive messages from their sensor nodes and remember their corresponding nodes. For uniformly distributed clusters, each cluster contains N/q nodes. Using Equation 2 and Equation 3, the energy consumed by a cluster head is estimated as follows:

$$E_{CH} = E_{TS} + (\frac{N}{q} - 1).E_{RX} + E_{DA} \quad (4)$$

The first part of Equation 4 represents the energy consumed to transmit the advertisement message; this energy consumption is based on a shorter distance energy dissipation model. The second part of Equation 4 represents the energy consumed to receive (N/q-1) messages from the sensor nodes of the same cluster. Equation 4 can be simplified as follows:

$$E_{CH} = m.Q.\frac{N}{q} + m.E_{DA}.(\frac{N}{q} - 1) + m.\tau.d^2 \quad (5)$$

Using Equation 2 and Equation 3, the energy consumed by non-cluster head sensor nodes is estimated as follows:

$$E_{NCH} = \{q.E(m) + E(m, d^2)\} \quad (6)$$

The first part of Equation 6 shows the energy consumed to receive messages from q cluster heads; it is assumed that a sensor node receive messages from all the cluster heads. The second part of Equation 6 shows the energy consumed to transmit the decision to the corresponding cluster head. Equation 6 can be simplified as follows:

$$E_{NCH} = \{(q+1).m.Q + q.m.E_{DA} + m.\tau.d^2)\} \quad (7)$$

4.3. Energy Consumption in Data transfer Phase

During data transfer phase, the nodes transmit messages to their cluster head and cluster heads transmit aggregated messages to a distant base station. The energy consumed by a cluster head is as follows:

$$E_{CH/f} = \{m.Q + m.\mu.d^4\} + \{\frac{N}{q}.m.(Q + E_{DA})\} \quad (8)$$

The first part of Equation 8 shows the energy consumed to transmit a message to the distant base station. The second part of Equation 8 shows the energy consumed to receive messages from the remaining (N/q) nodes that are non cluster head nodes. Equation 8 can be simplified as follows:

$$E_{CH/f} = m.(.\mu.d^4 + \frac{N}{q}.(Q + E_{DA})) \quad (9)$$

The energy, ENCH/ f, consumed by a non-cluster head node to transmit the sensor data to the cluster head is given below:

$$E_{NCH/f} = E_{TS} \quad (10)$$

For circular clusters with a uniform distribution of sensor nodes and a network diameter of A, the average value of d2 is given as:

$$E[d^2] = \left(\frac{A^2}{2\pi q}\right)$$

and Equation 10 can be simplified as follows:

$$E_{NCH} = m.(Q + \tau\frac{A^2}{2\pi q}) \quad (11)$$

In one round, Df data frames are transmitted. The number of frames transmitted by each cluster is Df /q. The Df /q frames are uniformly divided among N/q nodes of the cluster. Each cluster head frame transmission needs N/ q −1 non-cluster head frames. For simplification of equations, the fractions G1 and G2 are given as below:

$$G_1 = \left(\frac{1}{\frac{N}{q}}\right)\frac{1}{q} \quad (12)$$

$$G_2 = \left(\frac{\frac{N}{q} - 1}{\frac{N}{q} - 1}\right)\frac{1}{q} \quad (13)$$

The energy consumptions in a data transfer stage of each cluster are as follows:

$$E_{DT} = G_1.D_f.E_{CH/f} \quad (14)$$

$$E_{NDT} = G_2.D_f.E_{NCH/f} \quad (15)$$

4.4. Energy Computed for one round

There are q clusters and N nodes. Each round consists of a cluster head election phase and a data transfer phase. The energy consumed in one iteration of cluster is as follows:

$$E_{CH/iter/cluster} = E_{CH} + E_{DT} \quad (16)$$

$$E_{NCH/iter/cluster} = E_{NCH} + E_{NDT} \quad (17)$$

Since there are N/q nodes in a cluster, the ECH/iter/cluster is uniformly divided among the cluster members, as given below:

$$E_{CH/N} = E_{CH/iter/cluster} \quad (18)$$

Similarly, there are {(N/q)−1} non-cluster head nodes in a cluster. The ENCH/iter/cluster is uniformly distributed among all the non-cluster head members as follows:

$$E_{NCH/N} = \frac{E_{NCH/iter/cluster}}{(\frac{N}{q} - 1)} \quad (19)$$

The start energy, Es, is energy of a sensor node at the initial start time. An estimation of Es is given below:

$$E_S = E_{CH/N} + \left(\frac{N}{q} - 1\right)E_{NCH/N} \quad (20)$$

Using Equation 18, Equation 19 and Equation 20, Es can be described as below:

$$E_S = \left(E_{CH/iter/cluster} + E_{NCH/iter/cluster}\right) \quad (21)$$

4.5. Optimum number of clusters

In a cluster, the energy consumed to transmit an aggregated reading to the base station is as follows:

$$E_C = E_{CH/f} + \left(\frac{N}{q} - 1\right)E_{NCH/f} \quad (22)$$

The first part of Equation 22 is due to the energy consumption by cluster head. The second part of Equation 22 is due to (N/q-1) non-cluster head nodes. The total energy consumed by q clusters is as follows:

$$E_{T/f} = q.E_C \quad (23)$$

The total energy consumed by q clusters is given below:

$$E_{T/f} = q\left\{m\mu d^4 + \left(\frac{N}{q} - 1\right)mQ + \left(\frac{N}{q} - \right)mE_{DA}\right\} + q\left\{\left(\frac{N}{q} - 1\right)\left(mQ + m\tau\frac{A^2}{2\pi q}\right)\right\} \quad (24)$$

The optimum number of q for minimum consumed energy can be determined as follows:

$$\frac{dE_T}{dk} = 0$$

$$q = \sqrt{\frac{N}{2\pi}}\sqrt{\frac{\tau}{\mu.d^4 - Q_{DA}}}A \quad (30)$$

5. SIMULATION RESULTS

Wireless sensor networks (WSNs) contain 200 number of sensor nodes equipped with sensing, computing and communication abilities. Each node has the ability to sense elements of its environment, performs simple computations, and communicates among its peers or directly to an external base station (BS) as shown in Figure 1. Deployment of a sensor network is in random fashion. In this section, we have evaluated the effectiveness performance of the EECHDA through simulations. We have considered first order radio model and the simulation parameters for our model are mentioned in the Table 1.To validate the performance of EECHDA, we have simulated direct and EECHDA wireless sensor network in a field with dimensions M × M as shown in Figure 2. All the sensor nodes are randomly distributed over the sensor field. This means that the horizontal and vertical coordinates of each sensor are randomly selected between 0 and maximum value of the dimension. The base station is located far away from the network. The size of the message that nodes send to their cluster heads as well as the size of the (aggregate) message that a cluster head sends to the base station is set to 50 bytes.

We have simulated EECHDA and Direct protocol in the same environment. The results of EECHDA and Direct simulations are shown in Figures 3 -6.

Figure 3 shows, the variation in the energy consumed per node with respect to the number of clusters and network diameter. The x-axis and y-axis represent the number of clusters and the energy consumed in one round, respectively. Figure 4, shows the energy consumption with respect to the number of clusters. As expected, the energy consumption is reduced when the number of clusters is increased. However, the rate of reduction in energy consumption is reduced for higher cluster sizes. Figure 5, illustrates the energy consumption with respect to the network diameter. The energy consumption is increased when the network diameter is increased in direct transmission protocol. We have also evaluated the network lifetime by examining the round when the first and last node dies in the network. Figures 6(a)-6(b), shows that the proposed protocol offers a much longer lifetime than direct transmission. Here, direct transmission means that each node transmits its data directly to the base station or sink. This extends the network lifetime by 50% in EECHDA over Direct. However, EECHDA requires less energy consumption in cluster configuration than the direct configuration. Thus, the proposed protocol is energy efficient.

Parameters	Symbol	Value
Network area	M x M	(0,0) to (200,200)
Number of nodes	N	200
Location of BS	Outside	(100,100),(100,300)
Data aggregation energy	EDA	5nJ/bit/report
Energy consumed by the amplifier to transmit at a shorter distance	τ	10pJ/bit/m2
Energy consumed by the amplifier to transmit at a long distance	μ	0.0013pJ/bit/m4
Energy consumed in the electronic circuit to transmit or receive the signal	Q	50nJ/bit
Initial energy of node	E0	0.5J
Packet Size	m	50bytes
Number of cluster heads	q	20

Table1: Simulation Parameters

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have developed EECHDA, clustering based network protocol that minimizes energy usage and the quantitative results indicate that the energy consumption can be systematically decreased by including more clusters in networks. Both theoretical analysis and simulation results show that EECHDA has significant gain in network lifetime over direct transmission under the assumption that nodes are randomly and densely deployed. We have also examined the energy of the battery drain rate is less in case of clustered network than the direct transmission in the same network. Simulation results show that the network lifetime is extended by 50% in EECHDA over direct transmission. One of the future works will include the study of an energy efficient algorithm through data accumulation in a mobile sensor network.

REFERENCES

[1]. G. Asada et al., Wireless Integrated Network Sensors: Low Power Systems on a Chip,24th IEEE Eur. Solid-State Circuits Conf., 9–12, The Hague, the Netherlands, 1998.

[2]. J. Hill and D. Culler, A Wireless Embedded Sensor Architecture for System Level Optimization, University of California Berkeley Technical Report, 2002.

[3]. J.M. Kahn, R.H. Katz, and K.S.J. Pister, Next Century Challenges: Mobile Networking for Smart Dust, ACM MOBICOM, 271–278, Seattle, 1999.

[4]. R. Min et al., Low-Power Wireless Sensor Networks, IEEE VLSID 2001, 205–210, India, 2001.

[5]. G.J. Pottie and L.P. Clare, Wireless Integrated Network Sensors: towards Low Cost and Robust Self Organizing Security Networks, Processiding of SPIE 1998, 3577, 86–95, and 1999.

[6]. G.J. Pottie and W.J. Kaiser, Wireless Integrated Network Sensors, Communication ACM, 43 (5), 51–58, 2000.

[7]. J.M. Rabaey et al., Pico Radio Supports Ad Hoc Ultra-Low Power Wireless Networking, IEEE Computer Mag., 33(7), 42–48, 2000.

[8]. A. Sinha and A. Chandrakasan, Dymamic Power Management in Wireless Sensor Networks, IEEE Design Test Computers, 18(2), 62–74, 2001.

[9]. J. F. Kurose, and K. W. Ross, Computer Networking, a Top-Down Approach Featuring the Internet,1st ed., Addison–Wesley, Longman, MA, 2000.
[10]. A.S. Tanenbaum, Computer Networks, Prentice Hall, Upper Saddle River, NJ, 1996.
[11]. B. Krishnamachari, D. Estrin, and S. Wicker, Impact of Data Aggregation in Wireless Sensor Networks, International Workshop Data Aggregation Wireless Sensor Networks, 575–578, Vienna, Austria, July 2002.
[12]. R.S.Chang, C.J Kuo, An Energy Efficient Routing Mechanism for Wireless Sensor Networks, Advanced Information Networking and Applications,2006,20th International Conference on Publications, Vol.2,April 2006.
[13]. J. Zhao, A.T. Erdogan, A Noval Self Organizing Hybrid Network Protocol for Wireless Sensor Networks,Adaptive Hardware and Systems (AHS 2006),First NASA conference, Publication, June 2006.
[14]. I.F.Akylidiz et al.,Wireless Sensor Networks: a Survey, Computer Networks 38(4),393-422, 2002.
[15]. Akkaya and M.Younis, A survey of Routing protocols in Wireless Sensor Networks,in the Elsevier Ad Hoc Network Journal, Sep 2003.
[16]. W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy efficient Communication Protocol for Wireless Microsensor Networks," in Proceedings of the Hawaii International Conference on System Sciences, January 2000.
[17]. W. Heinzelman., Application-Specific Protocol Architectures for Wireless Networks. In PhD thesis, Massachusetts institute of technology, June 2000.
[18]. T. Anker, D. Bickson, D. Dolev, and B. Hod, Efficient Clustering for Improving Network Performance in Wireless Sensor Networks, in Proceedings of the Springer –Verilag Heridelberg (LNCS 4913),221-236, 2008.
[19]. Dilip Kumar, and R.B. Patel, "HCEE: Hierarchical clustered energy efficient protocol for heterogeneous wireless sensor networks", International Journal of Electronics Engineering, Vol. 1, 1, January 2009, pp.123-126.
[20]. Dilip Kumar, T.C Aseri, and R.B. Patel, "A Two Tier Data Aggregation and Clustering Scheme for Heterogeneous Sensor Networks ", in the Proceedings of IEEE International Adavnced Computing Conference (IACC-2009),6-7 March,Patiala,India,2009, pp.2053-2058.

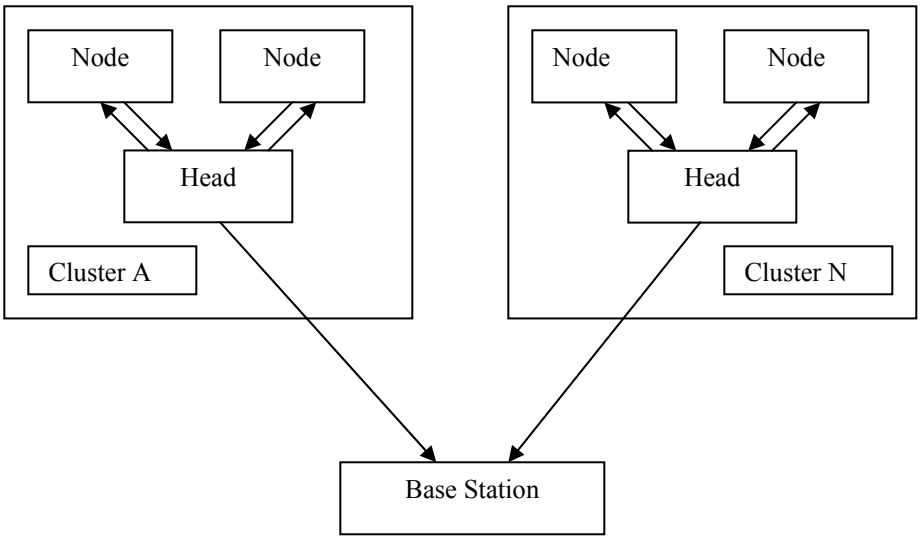


Figure1: Hierarchical clustering architecture

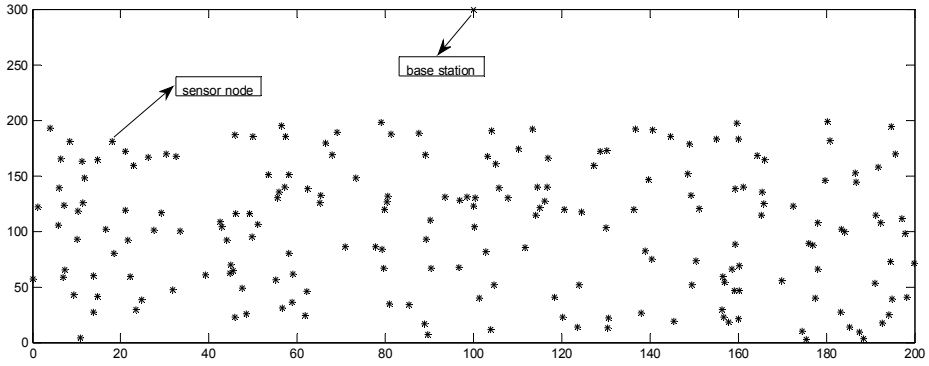


Figure 2: Network model

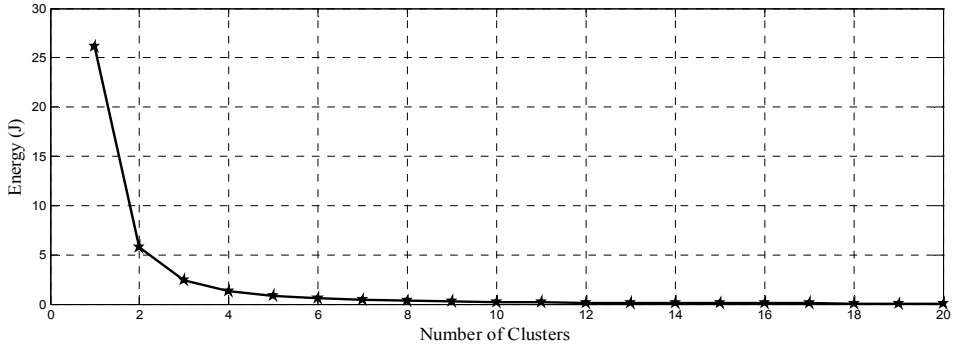


Figure 3: Optimum number of clusters.

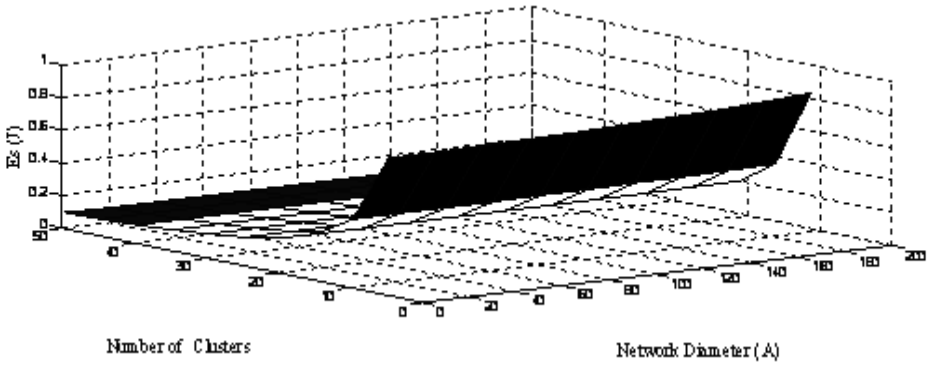


Figure 4: Energy consumed per round with respect to number of clusters.

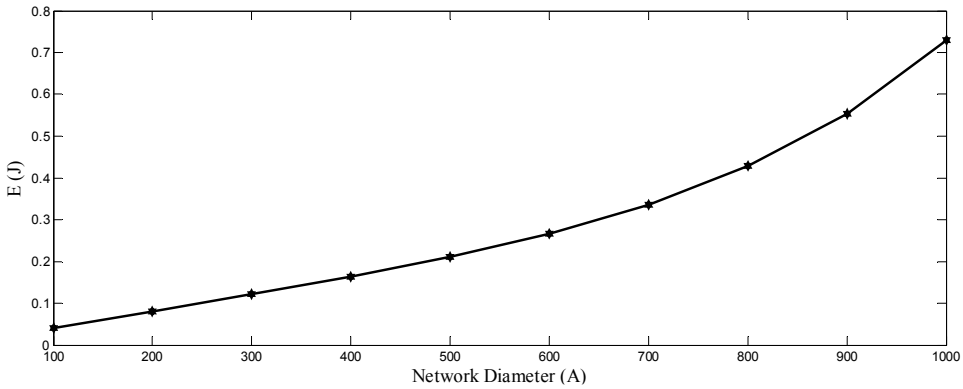


Figure 5: Energy consumed per round in direct transmission over network diameter.

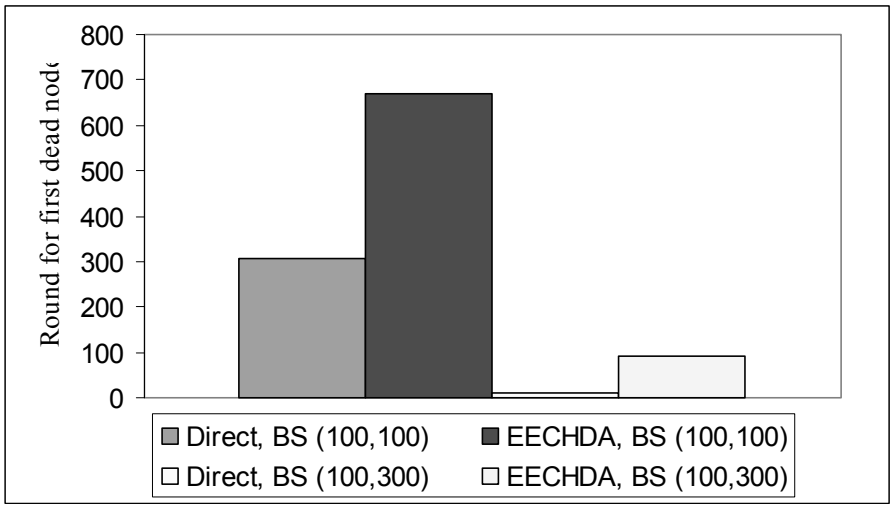


Figure 6 (a): Round for first dead node in EECHDA and Direct.

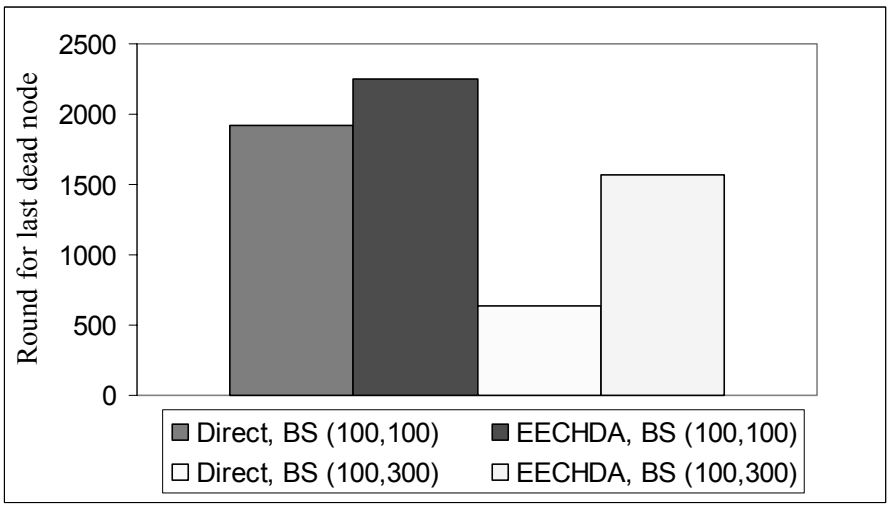


Figure 6(b): Round for last dead node in EECHDA and Direct.

Fuzzy Expert System for Noise Induced Sleep Disturbance and Health Effects

Devendra K. Tayal[1], Amita Jain[2] and **Vinita Gupta[3]**

Abstract - In this paper an effort has been made to develop a fuzzy based model to study the impact of various noise factors on Sleep disturbance and Health. We thoroughly survey the existing literature and identify the deficiencies in the existing models in this field. We then identify various noise factors which can have the significant impact on Sleep and health. The MIMO Expert system developed in this paper gives sleep disturbance, health condition in the morning and health as output variables and noise level, short noise duration, long noise duration, age and Type of noise as the input variables. Appropriate fuzzification and defuzzification strategies have been used and the implementation in MATALAB 7.0.1 has been done. It has been established from work of various researchers that effect of meaningful noise like songs and talks affect sleep and health-conditions badly than meaningless noise like railway noise, roadside noise. Similarly other input variables affect sleep & health condition. These factors have been studied in this paper. The noise level and duration of noise, which are also the prominent factors in deciding effect on hearing output factor have been discussed, for e.g. a noise of low level does not have prominent affect on human being as of high level of noises..

Index Terms - Noise, Expert system, Fuzzy logic

INTRODUCTION

Noise, which is often referred to as unwanted sound, is typically characterized by the intensity, frequency, periodicity (continuous or intermittent) and duration of sound. Sound is the result of pressure changes in the air caused by vibration [2]. Noise effects on people is more than stress. Noise affects millions of people worldwide on a daily basis. Highway noise alone affects more than 18 million people in the United States and 100 million people worldwide [3]. Noise cannot only degrade the living of a person but can also produce some permanent ill-effects like hearing loss[1]. So it is crucial to have a model which can predict the effect of noise on different age groups. There are several factors which can disturb sleep like age, noise duration, noise level, type of noise, physical health, mental health, etc. but the main factors are age, noise duration, noise level and type of noise. It has been found that type of noise disturbs people's sleep in a significant manner. As described in [12], noise-induced sleep disturbance in urban areas degrades the quality of life and therefore must be given a high priority from a public health point of view. World Health Organization (WHO) in this regard has stated "If through living in an area that is too noisy, a person fails to obtain sufficient sleep over long periods of time, the implications for health are obvious". Uninterrupted sleep is known to be a prerequisite for good physiological and mental functioning of healthy persons. During the last four decades, there has been an exponential growth in noise level due to reasons like increase in population, increase in traffic density (both road and air), increase in industrial establishments, and increase in the use of various noise producing devices on several occasions [12]. So making a model on this concept is very necessary and useful. The type of noise is basically a very important factor in deciding the sleep disturbance. As the karaoke songs, people's talk, the noises which get meaning, disturbs people more and in the morning they don't feel good. The effect of this is shown in different ways like headache in morning, tiredness etc. . On the other hand if the noise is meaningless then people get disturbed only when the noise level is high as is shown in the experiments of S. Kuwano and T. Mizunami [8]. So this factor is considered as very important in deciding sleep disturbance and health conditions in the morning. In this paper basically four different types of noises are considered like karaoke songs, people's talk, railway noise[15] and road traffic noise[9]. There are other types of noises also like ventilation noise, air conditioner noise but they are not included in this model as they do not affect people much in all the three output variables. There are also some factors like psychological conditions, physical health etc. which can affect all output variables. This will be focus of our next research.

The effect of noise duration and noise level on the health of human beings is very high, for e.g when the noise duration is very less but the noise level is 75db (A) then hearing loss is also possible, on the other hand if noise duration is long and noise level is room noise then there is no danger to human ears.

2. FUZZY LOGIC

A. Introduction

Boole[5] introduced the beautiful notion of binary sets, which is the foundation of modern digital computer but boolean logic is unable to model the human cognition and thinking process. Because of its rigid boundaries, the two valued logic is not so efficient in mapping real world situations. For handling real world problems Zadeh [6] introduced the concept of 'mathematics of fuzzy or cloudy quantities' followed by his seminal paper 'Fuzzy sets' [7]. Generally, the term fuzzy logic

[1]*H.O.D., Department of Computer Science and Engineering, Indira Gandhi Institute of Technology,GGSIP University, Delhi*
[2]*Lecturer Department of Information Technology, Guru Premsukh Memorial College of Engineering,GGSIP University, Delhi*
[3]*Research Scholar, University School of Information Technology, GGSIP University, Delhi*
E-Mail: [1]dr.tayal@ipu.edu, [2]amita_jain_17@yahoo.com, and [3]vinitagupta27@gmail.com

is used in two different senses [13]. In a narrow sense, fuzzy logic refers to a logical system that generalizes classical two-valued logic for reasoning under uncertainty. In a broad sense, fuzzy logic refers to all of the theories and technologies that employ fuzzy sets, which are classes with unsharp boundaries.

B. Fuzzy Expert System

Expert systems solve problems that are normally solved by human "experts" [17]. The problems that expert systems deal with are highly diverse. As specified in [19], the main paradigm of fuzzy expert-system (fuzzy rule based system) is the fuzzy algorithm, the essential concepts of which are derived from fuzzy logic. It is basically an expert knowledge-based system that contains the fuzzy algorithm in a simple rule base. As depicted in Fig. 1, a fuzzy rule based system is composed of four parts: fuzzifier, knowledge base, inference engine, and defuzzifier [11].

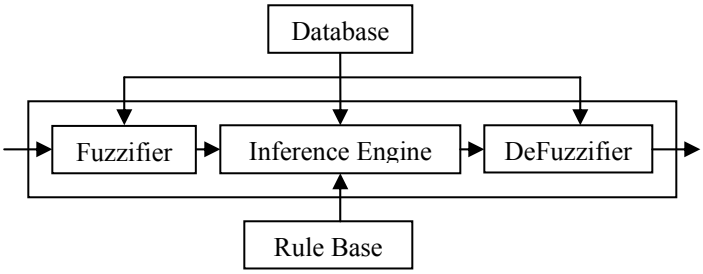


Figure1. The structure of Fuzzy Expert system[11].

In general, a fuzzy rule based system with multi-inputs multiple-output (MIMO) can be represented in the following manner:

IF X1 is A1 AND X2 is A2 AND
.. . AND Xr is Ar
THEN Y1 is C1 AND Y2 is C2 AND
… AND Ys is Cs

Where X1, X2,., Xr are the input variables and Y1, Y2,., Ys are the output variables, Ai (i=1,., m) and Ci (i=1,., s) are fuzzy subsets of the universes of discourse U1, U2,., Ur, and V1, V2,., Vs of X1, X2,., Xr and Y1, Y2,., Ys respectively.

3. PROPOSED FUZZY SYSTEM

In 1975 Mamdani and Assilian proposed MIMO model. Using this model, a new type of Fuzzy Expert System is implemented, the following equations define this system:

V1=F(U1,U2,U3,U5);
V2=F(U1,U4);
V3=F(U1,U3,U5).

Where V1, V2, V3 are output variables sleep disturbance, health effects, health condition in morning res. and U1, U2, U3, U4, U5 are input variables noise level, age, type of noise, long noise duration, short noise duration res. .

For the development of this fuzzy expert system the following steps are followed:

1. System's variables are identified;
2. Ranges of input and output variables are determined.
3. Membership functions for system's variable is selected.
4. Linguistic rules are formed.

The various input and output variables can be considered for this system like stage of sleep, psychological health condition, physical health condition, duration of noise, status of a person and so on but for the sake of simplicity only five input variables viz. long noise duration, short noise duration, noise level, age and type of noise are considered. Similarly only three output variables are considered viz. sleep disturbance, health effects, health condition in morning. The MIMO model of fuzzy system is shown in Fig.2, depicting its various input and output variables. These variables in fuzzy modelling are defined as linguistic variables whose linguistic values are words or sentences in a natural or synthetic language (Zadeh, 1994) [18]. Then in the next step table 1 is formed which shows linguistic variables, their linguistic values and associated fuzzy intervals. For instance corresponding to the linguistic variable Noise-Level, linguistic values are Extremely Extremely Low (EEL) to Very Very High Extremely High (VVHEH), fuzzy interval from 25 dB(A)to145Db(A) is assigned. In third step all linguistic values are expressed in the form of fuzzy sets, which are represented by its membership functions. The Triangular membership function is used as it is simple and computationally efficient. Membership functions for this system are shown in Fig 3 a-h. Finally, through IF-THEN rules, the relationship between input and output variables are formed. A set of rules are illustrated in Table 2. This model is implemented in MATLAB 7.0.1.

4. RESULTS AND DISSCUSSION

In this Mamdani & Sugeno fuzzy model, Sleep disturbance, health effects, health condition in morning are considered to be a function of long noise duration, short noise duration, noise level, age and type of noise. The results are plotted using MATLAB 7.0.1 and are shown in Fig. 4a-I, Fig. 4(a) shows output variable Sleep Disturbance as function of age and noise level and other input variables at their default value. Fig. 4(a) shows output variable Sleep Disturbance as function of age and noise level and other input variables at their default value. Fig. 4(b) shows output variable Sleep Disturbance as function of short noise duration and noise level and other input variables at their default value. Fig. 4(c) shows output variable Sleep Disturbance as function of long noise duration and noise level and other input variables at their default value. Fig. 4(d) shows output variable Sleep Disturbance as function of Type of noise and noise level and other input variables at their default value. Fig. 4(e) shows output variable Health effects as function of age and noise level and other input variables at their default value. Fig. 4(f) shows output variable Health effects as function of short noise duration and noise level and other input variables at their default value. Fig. 4(g) shows output variable Health effects as function of Type of Noise and noise level and other input variables at their default value. Fig. 4(h) shows output variable Health effects as function of Long noise duration and noise level and other input variables at their default value. Fig. 4(i) shows output variable Health condition in morning as

function of Long noise duration and noise level and other input variables at their default value.

5. CONCLUSION

A new type of Fuzzy MIMO Expert system has been successfully implemented using Matlab 7.0.1. This MIMO system predicts the health effect, health condition in the morning and sleep disturbance, taking five different types of input variables viz. type of noise, age, short noise duration, long noise duration, and noise level. This fuzzy expert system can be used for knowing health effects in noisy region. This fuzzy expert system will prove to be a guideline for making new expert systems in future and can be effectively used in medical engineering.

REFERENCE

[1]. Sheela V. Basrur, "Health Effects of Noise", City of Toronto Community and Neighbourhood Services Toronto Public Health Promotion and Environment Protection Office (2000).

[2]. Thompson, S.J., " Noise And Public Health", Health & Environment Digest, vol. 8(4), 1994.

[3]. Cowan, J.P. (1994), "Educating The Public On Environmental And Recreational Noise Exposure", Handbook of Environmental Acoustics, New York, pp. 14-20, 1994.

[4]. Mathworks, "Matlab Fuzzy Logic Toolbox manual"

[5]. Boole G., "The Laws of Thought", New York: Dover Books (Reprinted), 1958.

[6]. Zadeh, L. A., "From Circuit Theory To Systems Theory", Proceedings of the Institute of Radio Engineering, Vol. 50, 1962.

[7]. Zadeh, L. A., "Fuzzy Sets", Information And Control, Vol. 8, 1965.

[8]. Y. sasazawa, P. xin, S. suzuki, T. kawada and M. kuroiwa and Y.Tamura, "Different Effects of Road Traffic Noise And Frogs' Croaking On Night Sleep" , Journal of Sound and Vibration, pn 250(1), 91-99, 2002.

[9]. E. Ohrstrom, A. Skanberg, "Sleep Disturbances From Road Traffic And Ventilation Noise — Laboratory And Field Experiments", Journal of Sound and Vibration, pn 271, 279–296, 2004.

[10]. Karl S. Pearsons, David S. Barber, Barbara G. Tabachnick, and Sanford Fidell, "Predicting Noise-Induced Sleep Disturbance", Acoustical Society of Arnerica, Vol. 97(No.1), January 1995.

[11]. Zaheeruddina, V.K. Jain, "A Fuzzy Expert System For Noise-Induced Sleep Disturbance", Expert Systems with Applications, 30 761–771, 2006, Elsevier Science Publication.

[12]. Zaheeruddin, Vinod K. Jain, and Guru V. Singh , "A Fuzzy Model For Noise-Induced Annoyance", IEEE transactions on systems, man, and cybernetics –Part A: Systems and Humans, Vol. 36(No. 4), July 2006.

[13]. Yager, R. R. and Filev, D. P., "Essentials Of Fuzzy Modelling And Control", Wiley, 1994.

[14]. John Yen and Reza Langari, "Fuzzy Logic Intelligence, Control, And Information", Pearson education, 1st Ed., 1999.

[15]. Anke Marks and Barbara Griefahn, "Railway Noise – Its Effects On Sleep, Mood, Subjective Sleep Quality, And Performance" Somnologie 9: 68–75, 2005.

[16]. E. Ohrstrom, "Longitudinal Surveys On Effects Of Changes In Road Traffic Noise: Effects On Sleep Assessed By General Questionnaires And 3-Day Sleep Logs", Journal of Sound and Vibration 276 (2004) 713–727, Elsevier Science Publication.

[17]. E. Rich and K. Knight, "Artificial Intelligence", TMH, 2nd Ed., 1992.

[18]. Zadeh, L. A., " Soft Computing And Fuzzy Logic", IEEE Software, 11, 1994.

[19]. Zadeh, L. A., "Outline Of A New Approach To The Analysis Of Complex Systems And Decision Processes", IEEE Transactions on Systems, Man and Cybernetics, SMC-3, 28–44, 1973.

TABLE 2 SET OF RULES.

1. If (Noise_level is EL) and (Age is Young) and (Short_Noise_Duration is Short) then (Sleep_Disturbance is ES)(Health_effects is Comfortable) (1)

2. If (Noise_level is EL) and (Age is Young) and (Short_Noise_Duration is Medium) then (Sleep_Disturbance is VVS)(Health_effects is Comfortable) (1)

……

308. If (Noise_level is VVHEH) and (Age is Old) and (Short_Noise_Duration is Long) and (Type_of_noise is Road_traffic) then (Health_effects is beyond_Thershold_of_pain) (1)

309. If (Noise_level is VVHEH) and (Age is Old) and (Short_Noise_Duration is Long) and (Type_of_noise is Karoke_songs) then (Health_effects is beyond_Thershold_of_pain) (1)

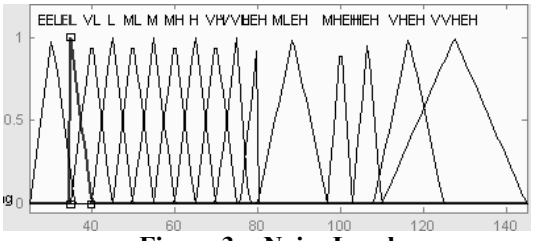


Figure 3a: Noise Level

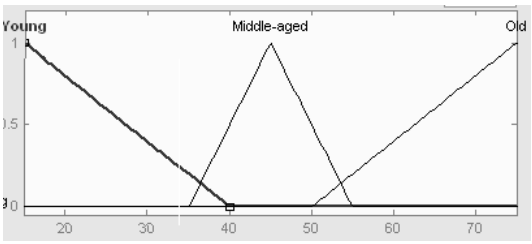


Figure 3b: Age

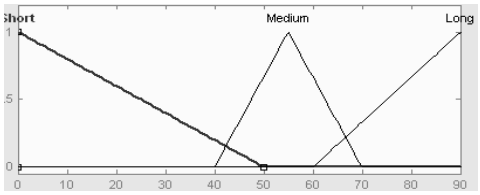


Figure 3c: Short Noise Duration

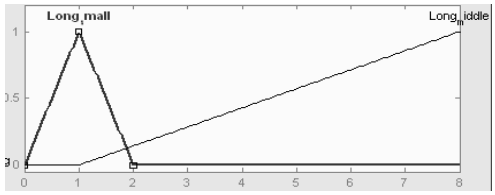


Figure 3d: Long Noise Duration

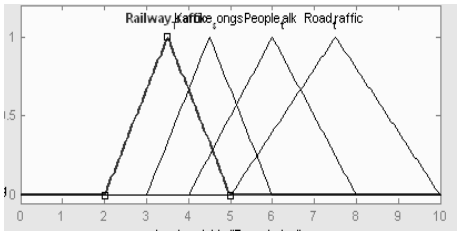


Figure 3e: Type of Noise

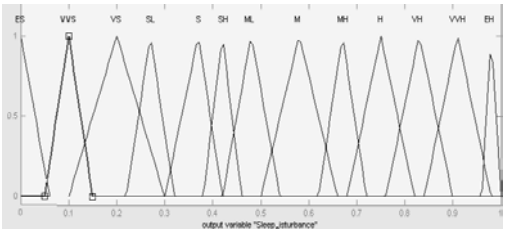


Figure 3f: Sleep Disturbance

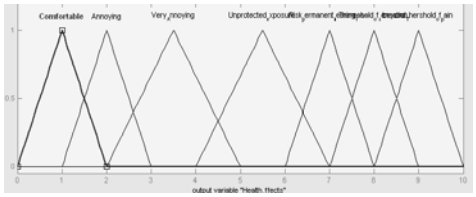


Figure 3g: Health Effects fig

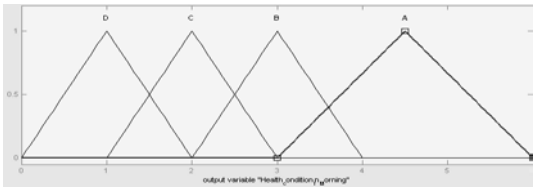


Figure 3h: Health_condition_in_morning

Figure 3: Showing membership functions for all input and output variables

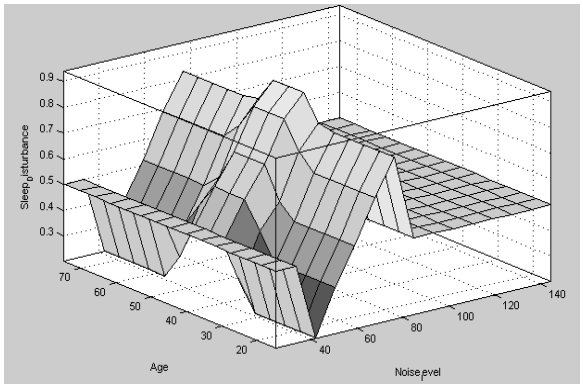


Figure.4a: Output (Sleep Disturbance), inputs noise level and age.

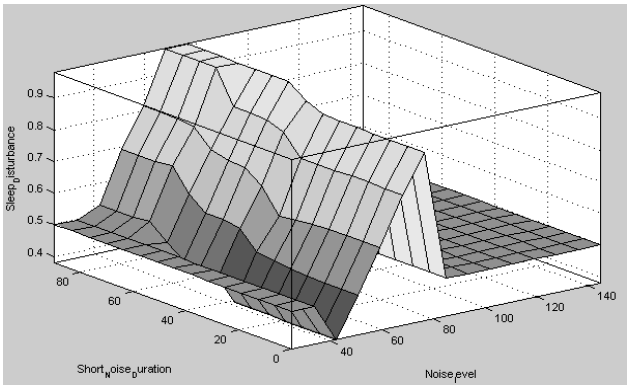


Figure.4b: Output (Sleep Disturbance), inputs noise level and short noise duration.

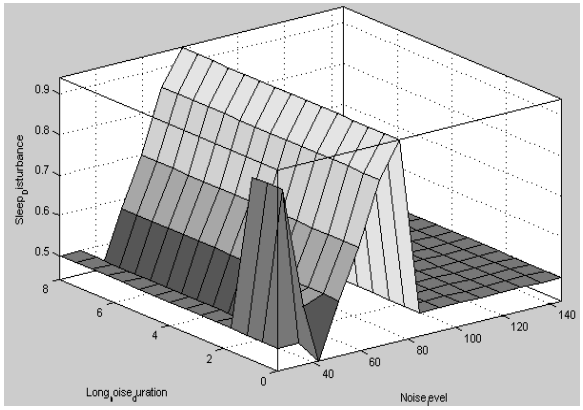


Figure.4c: Output (Sleep Disturbance), inputs noise level and Long noise duration

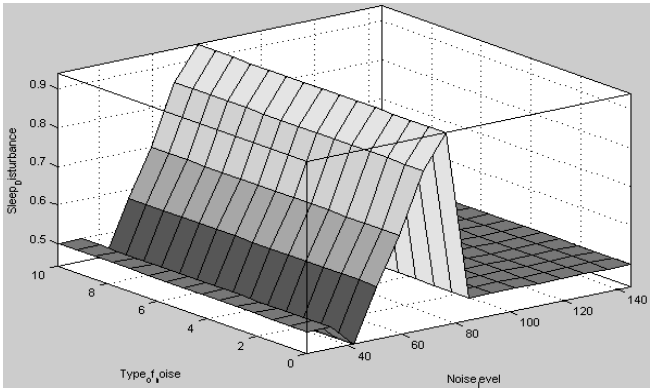


Figure.4d: Output (Sleep Disturbance), inputs noise level and Type of Noise

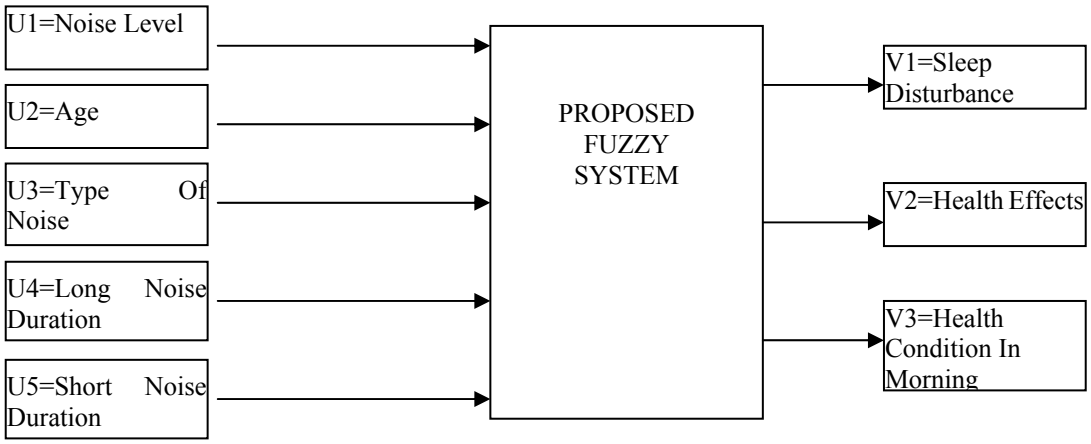


Figure 2: Fuzzy model for this system.

S.NO.	System's variables	Linguistic variables	Linguistic values	Fuzzy intervals
1.	Inputs	Noise level	EEL-Extremely Extremely Low	25 -36 dB(A)
			EL - Extremely Low	35 – 40 dB(A)
			VL - Very Low	35 – 45 dB(A)
			L - Low	40 -50 dB(A)
			ML - Medium Low	45 – 55 dB(A)
			M – Medium	50 – 60 dB(A)
			MH – Medium High	55 – 65 dB(A)
			H – High	60 – 70 dB(A)
			VH – Very High	65 – 75 v
			VVH - Very Very High	70 – 78 dB(A)
			LEH – Low Extremely High	76 – 80 dB(A)
			MLEH – Medium Low Extremely High	80 – 97 dB(A)
			MHEH - Medium High Extremely High	96.8 – 103 dB(A)
			HEH - Medium Extremely High	103 – 110 dB(A)
			VHEH- Very High Extremely High	108 – 125 dB(A)
			VVHEH – Very Very High Extremely High	110 – 145 dB(A)
2.		AGE	YOUNG	15 – 40 years
			MIDDLED-AGED	35 – 55 years

S.NO.	System's variables	Linguistic variables	Linguistic values	Fuzzy intervals
			OLD	50 – 75 years
3.		SHORT NOISE DURATION	SHORT	0 – 50 sec.
			MEDIUM	40 – 70 sec.
			LONG	60 – 90 sec.
4.		LONG NOISE DURATION	LONG-SMALL	0 – 2 Hrs.
			LONG-MIDDLE	1 -8 Hrs.
5.		TYPE OF NOISE	RAILWAY_TRAFFIC	2 – 5
			KAROKE_SONGS	3 – 6
			PEOPLE_TALK	4 – 8
			ROAD_TRAFFIC	5 – 10
6.	OUTPUT	SLEEP_ DISTURBANCE	ES – Extremely Small	0 - 0.06
			VVS – Very Very Small	0.05 - 0.15
			VS - Very Small	0.1 - 0.3
			SL –Small Low	0.22 - 0.32
			S - Small	0.3 - 0.42
			SH - Small	0.38 - 0.46
			ML – Medium Low	0.42 - 0.54
			M - Medium	0.5 - 0.66
			MH - Medium High	0.62 - 0.72
			H - High	0.68 - 0.82
			VH - Very High	0.76 - 0.9
			VVH - Very Very High	0.84 - 0.98
			EH - Extremely High	0.96 – 1
7.		HEALTH_EFFECTS	COMFORTABLE	0 – 2
			ANNOYING	1 – 3
			VERY_ANNOYING	2 – 5
			UNPROTECTED_EXPOSURE	4 – 7
			RISK_PERMANENT_HEARING_LOSS	6 – 8
			THERESHOLD_OF_SENSATION	7 – 9
			BEYOND_THERSHOLD_OF_PAIN	8 – 10
8.		Health_condition_ in_morning	D - could not sleep and have headache	0 – 2
			C - could not sleep well, but do not feel bad	1 – 3
			B - slept well, but feel bad	2 – 4
			A - could not sleep and have headache	3 – 6

Table 1: Inputs and output with their associated fuzzy values.

A Novel Metric for Detection of Jellyfish Reorder Attack on Ad Hoc Network

B. B. Jayasingh[1] and **B. Swathi**[2]

Abstract - Ad Hoc networks are susceptible to many attacks due to its unique characteristics such as open network architecture, stringent resource constraints, shared wireless medium and highly dynamic topology. The attacks can be of different types out of which denial of service is one of the most difficult attacks to detect and defend. Jellyfish is a new denial of service attack that exploits the end to end congestion control mechanism of TCP (Transmission Control Protocol) which has a very devastating effect on the throughput. The architecture for detection of such attack should be both distributed and cooperative to suit the needs of wireless ad-hoc networks that is every node in the wireless ad-hoc network should participate in the intrusion detection. We intend to develop an algorithm that detects the jellyfish attack at a single node and that can be effectively deployed at all other nodes in the ad hoc network. We propose the novel metric that detects the Jellyfish reorder attack based on the Reorder Density which is a basis for developing a metric. The comparison table shows the effectiveness of novel metric, it also helps protocol designers to develop the counter strategies for the attack.

Index Terms - jellyfish attack, Percentage of Late Packets (P_L), Mean Displacement of Packets (M_D), Mean displacement of late packets (M_L), Reorder Entropy(E_R).

1. INTRODUCTION

Dynamic topology, distributed operation, and resource constraints are some of the unique characteristics that exist in the ad hoc networks, which inevitably increase the vulnerability of such network. Many characteristics might be used to classify attacks in the ad hoc networks. MANET not only inherits all the security threats faced in both wired and wireless networks, but it also introduces security attacks unique to itself [1]. As people will be encouraged to use a secured network, it is important to provide MANET with reliable security mechanisms if we want to see this exciting technology become widely used in a next few years. Before the development of any security measure to secure mobile ad hoc networks, it is important to study the variety of attacks that might be related to such networks. With the knowledge of some common attack issues, researchers might have a better understanding of how mobile ad hoc networks could be threatened by the attackers, and thus might lead to the development of more reliable security measures in protecting them.

Routing is one of the most vital mechanisms in the ad hoc networks. Improper and insecure routing mechanisms will not only degrade the performance of the ad hoc networks, but will also render such networks vulnerable to many security attacks. One of the basic elements in the routing mechanism is the routing message, which is used to establish and maintain relationships between nodes in the networks. The importance of the routing message has made it a main target by the attackers to launch attacks against the ad hoc networks [2, 3, 16]. In designing security mechanisms for mobile ad hoc networks, one must consider the attacks variations as well as the characteristics of the attacks that could be launched against the ad hoc networks.

The first JF attack is the packet reordering attack. TCP has a well-known vulnerability to reordered packets due to factors such as route changes or the use of multi-path routing, and a number of TCP modifications have been proposed to improve resistance to misordering including TCP Stack [5] and reorder robust TCP [6].However, no TCP variant is strong enough to resist such malicious and persistent reordering as employed by the JF misordering attack. The mechanism that the jellyfish node uses for attack consists of delivering all received packets, but in scrambled order by placing them in a reordering buffer instead of the canonical FIFO order i.e.JF nodes maliciously re-order packets. Consequently, such persistent re-ordering of packets will result in near zero goodput, despite having all transmitted packets delivered.

We intend to develop a detection algorithm that can detect the jellyfish Reorder attack at a single node. The attack can be effectively detected by deploying the same detection mechanism at all nodes in the ad hoc network. We are assuming that there is no packet loss and duplication. The algorithm detects the persistent reordering employed by the Jellyfish node [17].The algorithm takes sequence number, acknowledgment number and Receive index as inputs. The sequence number and Receive index are used to calculate a value called Reorder density which is the basis for developing a metric that can detect the reordering of the packets that is done by the attacker. The algorithm also checks whether the acknowledgment number, that are generated when the packets are received, are reordered. Thus the new metric and the detection of reordered acknowledgment numbers can detect the Jellyfish Reorder attack effectively.

The rest of the paper focuses on the existing metrics for the packet reordering calculation in the section 2 and the calculations we do by using mathematical formulae's in section 3. We summarized the calculation in the section 4 and the lessons learned able to find a new metric that is proposed in section 5. The comparisons made between the existing metrics

[1]*Dept. of IT, CVR College of Engineering, Vastunagar, Mangalpalli (V), Ibrahimpatan (M),*
RR District – 501510, Hyderbad (AP), India.
[2]*B. Tech (IV-IT), CVR College of Engineering, Vastunagar, Mangalpalli (V), Ibrahimpatan (M),*
RR District – 501510, Hyderbad (AP), India.
E-Mail : [1]*bbjayasingh9@rediffmail.com*
[2]*andbswathi.reddy25@gmail.com*

calculation and the proposed metric is discussed in section 6, following with the conclusion in section 7.

2. EXISTING METRICS FOR REORDERING

Many attackers disobey protocol rules, whereas jellyfish obeys the protocol rules and hence is difficult to detect until after the sting. Jellyfish target closed-loop flows. One example of such a closed loop flow is the TCP flow .Just like any IP service, Jellyfish node can drop packets, Reorder packets, Delay / jitter packets but it is done in a malicious way. Since the Jellyfish Attack maintains compliance with all control plane and data plane protocols, it is difficult to distinguish from congestion and packet losses that occur naturally in a network, and therefore detection and diagnosis is hard, costly ,resource-consuming and time consuming [4].

There are several simple, derived metrics to monitor packet reordering in a network or an end to end connection. In this section, we discus the existing metrics for determining the reordering such as Percentage of Late Packets, Mean Displacement of Packets and the Reorder entropy [8, 9, 10].

We assume that there is no packet loss and duplication for the calculation of these metrics**.** However the calculation of these metrics is based on RD value which is obtained by applying the algorithm in [11, 12, 13, 14,]. Reorder Density (RD) is a discrete density function that is used to detect and capture the nature of reordering in a packet stream.

2.1 Percentage of Late Packets (P_L)

To capture the lateness of packets from their original positions the percentage of late packets is defined as the percentage of packets that exhibit lateness with respect to their expected position, as given by the receive index.

$$P_L = \sum_{i=+1}^{i=D_r} RD[i]$$

P_L =0 corresponds to the case where all the packets are in order. For a sequence with packet reordering, P_L >0.

2.2 Mean Displacement of Packets (M_D)

Packet reordering is associated with two types of events, lateness events and earliness events. In a lateness event, the corresponding displacement is always positive. And a negative displacement is mapped with the earliness event. When calculating the mean displacement of packets, if both late and early packets are included, from equation 1, the mean displacement is zero for all cases.

Equation 1:

$$\sum_{i} (i \times RD[i] = 0$$

Therefore, the mean displacement, when all packets are taken together, is not useful. On the other hand, one can consider the magnitude of displacement of packets, and divide it by the total number of packets to define a mean displacement M_D:

Mean Displacement (M_D)

$$M_D = \left| \sum_{i=-D_r}^{i=+D_r} (|i| \times RD[i]) \right| / \left| \sum_{i=-D_r}^{i=+D_r} RD[i] \right|$$

Mean displacement of late packets (M_L)

RD[*i*] refers to the probability that a packet arrives *i* packets away from its expected position. Thus considering only the late packets, the mean displacement of late packets is given as:

$$M_L = [\sum_{i=1}^{i=+D_r} (i \times RD[i])] / [\sum_{i=1}^{i=D_r} RD[i]]$$

Similarly, the **Mean displacement for earliness** is:

$$M_E = \left| [\sum_{i=-1}^{i=-D_r} (i \times RD[i])] / [\sum_{i=-1}^{i=-D_r} RD[i]] \right|$$

Note here that we divide the total positive (negative) displacement by the total number of late (early) packets. Both M_L and M_E are always none negative values, and $M_L=M_D/2P_L$, $M_E= M_D/2P_E$.

2.3 Reorder Entropy(E_R)

Entropy is a concept that is used to define the randomness or the disorder. As RD is a discrete probability distribution, that of packet displacement (a form of disorder), we define reorder entropy as:

$$E_R = (-1) \times \sum_{i=-D_r}^{i=+D_r} (RD[i] \times \mathrm{Log}_e RD[i])$$

RD [0] = 1

, when no packet ordering is present, the reorder entropy is equal to zero. On the other hand, the packet sequence has the most variance, when packets are displaced uniformly with equal probabilities [15].

3. REORDERING CALCULATION

This section deals with calculation of existing metrics for different number of packets, with and without reorder.

S denotes the Sequence Number of packets. Receive Index (RI) is a value assigned to a packet as it arrives at its destination, according to the order of arrival. Displacement (D) is the difference between RI and the sequence number of the packet. Displacement Frequency FD[k] is the number of arrived packets having a displacement of k. RD is defined as the distribution of the Displacement Frequencies FD[k], normalized with respect to N'. N' is equal to the sum (FD[k])[14].

3.1 Three Packets with Reorder

Consider the sequence of packets (2,3,1). The Tables 1 and 2 show the computational steps when the RD algorithm is applied to the above sequence.

S	2	3	1
RI	1	2	3
D	-1	-1	2
FD[D]	1	2	1

Table 1: showing the calculation of D and FD[D]

The last row (FD[D]) represents the current frequency of occurrence of the displacement D, e.g., column 2 indicates FD[-1] = 2 while column 3 indicates FD[2] = 1. The final sets of values for RD are shown in Table 2.

D	-1	2
FD[D]	2	1
RD[D]	0.66	0.33

Table 2: showing the calculation of RD[D]

P_L= RD[2]=0.33
M_D=(1*0.66+2*0.33)/(0.66+0.33)=1.32
M_L= (2*0.33)/(0.33)=2
M_E= |-1*0.66|/0.66=1
E_R= 0.63

3.2 Three Packets without Reorder

Consider the sequence of packets (1,2,3). The Tables 3 and 4 show the computational steps when the RD algorithm is applied to the above sequence.

S	1	2	3
RI	1	2	3
D	0	0	0
FD[D]	1	2	3

Table 3: showing the calculation of D and FD[D]

The last row (FD [D]) represents the current frequency of occurrence of the displacement D, e.g., column 2 indicates FD[0] = 2 while column 3 indicates FD[0] = 3. The final sets of values for RD are shown in Table 4.

D	0
FD[D]	3
RD[D]	1

Table 4: showing the calculation of RD[D]

P_L= 0, M_D= 0, M_L= 0, M_E= 0, E_R= 0

When the packets sent are in the sequential order i.e. when there is no reordering Percentage of late packets (P_L), Mean Displacement of Packets (M_D), Mean displacement of late packets (M_L), Mean displacement for earliness (M_E) and Reorder Entropy are zero. But when there is any reordering then these values will not be zero and will have some distinct value.

3.3 Five Packets with Reorder

Consider the sequence of packets (5,2,3,1,4). The Table 5 and 6 show the computational steps when the RD algorithm is applied to the above sequence.

S	5	2	3	1	4
RI	1	2	3	4	5
D	-4	0	0	3	1
FD[D]	1	1	2	1	1

Table 5: showing the calculation of D and FD[D]

The last row (FD[D]) represents the current frequency of occurrence of the displacement D, e.g., column 3 indicates FD[0] = 2 while column 4 indicates FD[3] = 1. The final sets of values for RD are shown in Table 6.

D	-4	0	1	3
FD[D]	1	2	1	1
RD[D]	0.2	0.4	0.2	0.2

Table 6: showing calculation of RD[D]

P_L= 0.2+0.2=0.4
M_D=(4*0.2+*0.2+3*0.2)/(0.2+0.4+0.2+0.2)=1.6
M_L= (1*0.2+3*0.2)/(0.2+0.2)=2
M_E= (4*0.2)/0.2=4
E_R= 1.33

3.4 Five Packets without Reorder

Consider the sequence of packets (1,2,3,4,5). The Tables 7 and 8 show the computational steps when the RD algorithm is applied to the above sequence.

S	1	2	3	4	5
RI	1	2	3	4	5
D	0	0	0	0	0
FD[D]	1	2	3	4	5

Table 7: showing the calculation of D and FD[D]

The last row (FD[D]) represents the current frequency of occurrence of the displacement D, e.g., column 3 indicates FD[0] = 3 while column 4 indicates FD[0] = 4. The final sets of values for RD are shown in Table 8.

D	0
FD[D]	5
RD[D]	1

Table 8: showing the calculation of RD[D]

P_L= 0, M_D= 0, M_L= 0, M_E= 0, E_R= 0

When the packets sent are in the sequential order i.e. when there is no reordering Percentage of late packets (P_L), Mean Displacement of Packets (M_D), Mean displacement of late packets (M_L), Mean displacement for earliness (M_E) and Reorder Entropy are zero. But when there is any reordering then these values will not be zero and will have some distinct value.

4. SUMMARY TABLES

In this section we show the table containing the D and RD(D) for different number of packets. D and RD(D) for 3 and 5 number of packets are obtained from the table2 and table6 respectively. Similarly D and RD(D) can be calculated for 8 and 10 packets as shown in Table 9 .

3pkts	D	**-1**	**2**				
	RD(D)	**0.66**	**0.33**				
5pkts	D	**-4**	**1**	**0**	**3**		
	RD(D)	**0.2**	**0.2**	**0.4**	**0.2**		
8pkts	D	**-2**	**-1**	**0**	**1**	**2**	
	RD(D)	**0.125**	**0.125**	**0.5**	**0.125**	**0.125**	
10pkts	D	**-6**	**-4**	**0**	**2**	**5**	**7**
	RD(D)	**0.1**	**0.2**	**0.4**	**0.1**	**0.1**	**0.1**

Table 9: showing the D and RD(D) for 3,5,8,10 packets

The RD values shown in Table 9 are used to calculate the existing metrics i.e. P_L, M_D, M_L, M_E, E_R for different number of packets as shown in Table 10. Mean displacement of packets is calculated for earliness (M_E), lateness (M_L) and combination of both (M_D).

		P_L	M_D	M_L	M_E	E_R
3 packets	With Reorder	0.33	1.32	2	1	0.63
	Without Reorder	0	0	0	0	0
5 packets	With Reorder	0.4	1.6	2	4	1.33
	Without Reorder	0	0	0	0	0
8 packets	With Reorder	0.25	0.75	1.5	1.5	1.386
	Without Reorder	0	0	0	0	0
10 packets	With Reorder	0.3	2.8	4.6	4.6	1.609
	Without Reorder	0	0	0	0	0

Table 10: showing the existing metric values

5. PROPOSED METRIC

Though many Intrusion Detection Systems are available, they are not suitable to detect the attack in ad hoc network because wireless ad-hoc networks don't have any fixed infrastructure and since almost all of current network based IDS sit on the network gateways and routers and analyze the network packets passing through them, these type of network based IDS are rendered ineffective for the wireless ad-hoc networks [7]. Anomaly Detection models of IDS cannot be used for wireless ad-hoc networks, since the separating line between normalcy and anomaly is obscure. A node that transmits erroneous routing information (fabrication) can be either a compromised or is currently out of sync due to volatile physical movement. Hence in wireless ad-hoc networks it is difficult to distinguish between false alarms and real intrusions. So, we have developed a novel metric for detection of the Jellyfish Reorder attack. The metric is calculated by multiplying frequency and reorder density for all the displacements and then taking their summation i.e. $\sum$ FD*RD. Let us analyze this metric for different number of packets in both cases i.e. with reorder and without reorder cases.

5.1 Three Packets with Reorder

With reference to Table 2 there are two Displacement frequencies i.e.2 and 1 and two Reorder density values i.e.0.66 and 0.33.The corresponding displacement frequencies and reorder density are multiplied and then summed which gives a value of 1.65 as shown below.

$\sum$ FD*RD = 2*0.66+1*0.33=1.65

This value is between 1 and 3 i.e. number of packets.

5.2 Five packets with reorder

With reference to Table 6 there are four Displacement frequencies i.e.1, 1, 2 and 1 and four Reorder density values i.e0.2, 0.2, 0.4 and 0.2.The corresponding displacement frequencies and reorder density are multiplied and then summed which gives a value of 1.4 as shown below.

$\sum$ FD*RD = 1*0.2+1*0.2+2*0.4+1*0.2=1.4

This value is between 1 and 5 i.e. number of packets.

5.3 Three Packets without Reorder

With reference to Table 4 there is only one Displacement frequency value i.e.3 and one Reorder density value i.e.1.The corresponding displacement frequency and reorder density is multiplied and then summed which gives a value of 3 as shown below.

$\sum$ FD*RD = 3*1=3

This value is equals to 3 i.e. number of packets

5.4 Five Packets without Reorder

With reference to Table 8 there is only one Displacement frequency value i.e.5 and one Reorder density value i.e. 1 .The corresponding displacement frequency and reorder density is multiplied and then summed which gives a value of 5 as shown below.

$\sum$ FD*RD = 5*1=5

This value is equals to 5 i.e. number of packets.

After analyzing the value of the proposed metric under 2 cases for different number of packets it was found that the value of the metric ($\sum$ FD*RD)

i) is greater than equal to one and less than the number of packets when there is reordering.

1<= $\sum$ FD*RD<Number of packets

ii) is always equals to number of packets when there is no reordering.

$\sum$ FD*RD=Number of packets

6. COMPARISION OF METRICS

The values of the existing metrics are always zero when there is no reorder whereas the proposed metric value is equals to number of packets when there is no reordering. The calculation of the proposed metric is computationally simple because it involves less calculation i.e. it involves simple algebraic operation of addition and multiplication. So the complexity of the algorithm that calculates this proposed metric for determining reordering is comparatively less when compared to the previous metrics.

		P_L	M_D	M_L	M_E	E_R	New Metric $\sum$ **FD*RD**
3 packets	With Reorder	0.33	1.32	2	1	0.63	1.65
	Without Reorder	0	0	0	0	0	3
5 packets	With Reorder	0.4	1.6	2	4	1.33	1.4
	Without Reorder	0	0	0	0	0	5
8 packets	With Reorder	0.25	0.75	1.5	1.5	1.386	2.5
	Without Reorder	0	0	0	0	0	8
10 packets	With Reorder	0.3	2.8	4.6	4.6	1.609	2.4
	Without Reorder	0	0	0	0	0	10

Table 11: showing all the metrics for different number of packets.

7. CONCLUSION

Jellyfish attack is protocol complaint and passive. So it is difficult to detect this attack until after the sting. Though there are existing metrics which can detect the Jellyfish Reorder attack to some extent by incorporating them in the algorithm, but the complexity involved in calculating these metrics increases with the increase in number of packets. The metric that we have proposed to be used in algorithm to detect the Jellyfish Reorder attack is highly effective because the metric is quite simple, efficient and also less time consuming

REFERENCES

[1] T. Karygiannis and L. Owens, "Wireless Network Security, 802.11, Bluetooth and Handheld Devices," NIST Publication, p. 800(48), November 2002.

[2] H. Li, Z. Chen and X. Qin, "Secure Routing in Wired Networks and Wireless Ad Hoc Networks," Univ. of Kentucky, Department of Computer Science, Term-paper, 2003.

[3] K. Sanzgiri, B. Dahill, B. N. Levine, C. Shields, and E. Belding-Royer, "A Secure Routing Protocol for Ad Hoc Networks," in Proc. of 2002 IEEE International Conference on Network Protocols (ICNP), pp. 778-89, Nov. 12-15, 2002.

[4] Jean-Pierre Hubaux , Edward W. Knightly,Impact of Denial of Service Attacks on Ad Hoc Networks Imad Aad, IEEE/ACM Transactions on Networking, Publication Date: Aug 2008 Volume: 16.

[5] K. Fall and S. Floyd, "Simulation-based comparison of Tahoe, Reno and SACK TCP," ACM Computer Communications Review, vol. 5, no. 3, pp. 5–21, July 1996.

[6] M. Zhang, B. Karp, S. Floyd, and L. Peterson, "RR-TCP: A reordering robust TCP with DSACK," in Proceedings of IEEE ICNP, 2003.

[7] Intrusion detection in wireless ad-hoc networks, Proceedings of the 6th annual international conference on Mobile computing and networking, Boston, Massachusetts, United States, Year of Publication: 2000.

[8] B. Ye, A. P. Jayasumana and N. Piratla, "On Monitoring of End-to-End Packet Reordering over the Internet," Proc.Int. Conference on Networking and Services (ICNS'06), Santa Clara, CA, July 2006.

[9] Banka, T., Bare, A. and Jayasumana, A., "Metrics for Degree of Reordering in Packet Sequences," Proc. IEEE 27Local Computer Networks Conf, Nov. 2001, pp. 333-342.

[10] Bare, A. A., "Measurement and Analysis of Packet Reordering," Masters Thesis, Dep. Computer Science,Colorado State University, 2004.

[11] Jayasumana, A., Piratla, N. M., Bare, A. A., Banka, T.,Whitner R., and McCollom, J.,"Reorder Density Function -A Metric for Packet Reordering Measurement," IETF draft.

[12] Piratla, N. M., Jayasumana, A. P., and Bare, A. A., "RD:A Formal, Comprehensive Metric for Packet Reordering,"Proceedings Fourth IFIP-TC6 Networking Conference(Networking 2005), LNCS 3462, Ontario, May 2005, 78-89.

[13] Piratla, N. M. , Jayasumana A. P. ,and Bare A. A., "A Comparative Analysis of Packet Reordering Metrics," Proc. IEEE/ACM 1st Int. Conf. Communication System Softwareand Middleware (COMSWARE 2006), New Delhi, Jan. 2006.

[14] A. Jayasumana, N. Piratla, T.Banka, A. Bare, R. Whitner. "Improved Packet Reordering Metrics",Network Working Group, Colorado State University, June 2008.

[15] Shannon C. E., "A Mathematical Theory of Communication", Bell Sys. Tech. Journal, vol. 27, 1948.

[16] Kristoffer Karlsson IT3, Billy HoIT3,Ad hoc networks: Overview, applications and routing issues, Chalmers University of Technology.

[17] S. Muthukumar, C. Arunkumar, G. Ramesh, Enhancing The Transport Mechanisms In The Presence Of Packet Re-Ordering In Tcp-Pr Algorithm, Dept. of MCA, Adhiparasakthi Engineering College, Melmaruvathur, Tamilnadu

Continued from Page No. 173

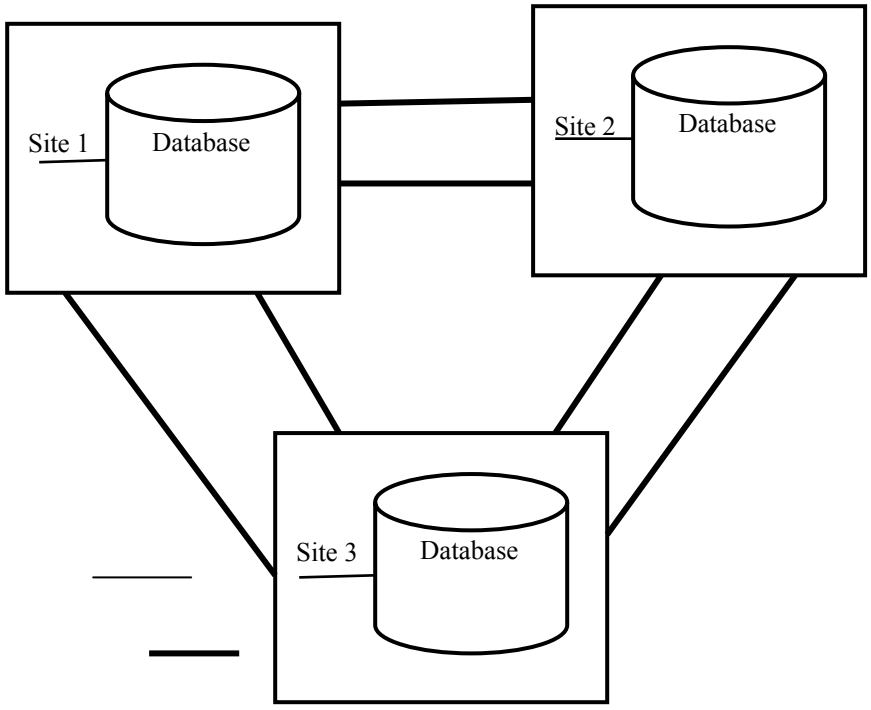


Figure 2: Working Scheme for Replication

Replication Strategies in Mobile Environments

Salman Abdul Moiz[1] and **Lakshmi Rajamani**[2]

Abstract - Transaction management in wireless environment poses challenging issues in preserving data consistency and fault tolerance. As the mobile databases are prone to frequent disconnections, bandwidth limitations, mobility, etc. efficient execution of the transaction may not always be guaranteed. Replicating data at several sites is a powerful mechanism which not only increases performance but also provides fault tolerance for demanding database applications. However the major concern is to keep the replicated copies always consistent. In this paper various replication techniques and their applicability in the mobile environments are presented. Based on the requirements of the services provider and the type of execution model used, one of the replication strategies may be implemented.

Index Terms - Mobile Host (MH), Fixed Host (FH), Transaction, Replication, Synchronous, Asynchronous.

1. INTRODUCTION

Mobility gains more and more importance from a technological as well as social perspective. Since network bandwidth is an expensive resource in mobile environments, the transaction processing should reflect a much concern for bandwidth consumption and constraints than non-mobile environments [2].

The transaction requests is initiated at the Mobile Host (MH) but may be partially or completely executed at the Fixed Host (FH). In the traditional environment once the transaction request is initiated by a mobile host, the respective data item is locked at Fixed host. In case of failures or disconnections the system almost halts as there is no replica to proceed with the transaction processing.

Over the years, data replication is considered as a better solution to increase throughput(more replicas can serve more request), decrease response times (distribute the load and access the local replica) and provide fault tolerance[13]. The major challenge with the implementation of replication strategies is to keep the replicas consistent. In addition the mobility introduces several other challenges for the management & maintenance of replicas in a mobile distributed environment.

Several valuable attempts have been presented for efficient implementation of concurrency control and fault tolerance in mobile environments. However each attempt considers only a subset of the operational requirements.

[1]*Research Scientist, Centre for Development of Advanced Computing, Bangalore.*
[2]*Professor, CSE, University College of Engineering, Osmania University, Hyderabad*
E-Mail: [1]*salmanmca@gmail.com and*
[2]*lakshmiraja@yahoo.com*

For example in [9], the author proposes a scheme, to provide non-blocking protocol with restrained communication. It faces the problem of time lag between the local and global commit. In [7] concept of non-conflicting transactions is introduced such that if a conflicting transaction is detected, it may be aborted. The research in [8], proposes Mobile-2PC, which preserves the traditional 2 Phase Commit protocol while minimizing the impact of unreliable wireless communication. Fragmentation of relations is used to support semantic based concurrency control [4]. In [5] distributed lock management scheme is proposed. Most of these techniques presented produces low throughput due to inefficient implementation of replication strategies. However in [3] authors propose a directory structure to identify replicas rather than migrating the replicas. However the frequent mobility needs the directory structure to be updated. In this paper we propose various replication strategies which may be adopted my application providers based on the fidelity & importance of data.

The remaining part of this paper is organized as follows. Section 2 describes the general architectural view of mobile databases. Section 3 specifies the various replication strategies their advantages and challenges. Section 4 specifies mechanism to uniquely identify a record (replica) in the mobile distributed environment. Section 5 describes the proposed replication strategies of mobile environments and the proposed solutions for the conflict resolution. Section 6 concludes the paper.

2. MOBILE DATABASE ENVIRONMENT

The mobile computing environment generally consists of three entities Fixed Host (FH), Mobile Units (MU) and Base Stations (BS) respectively. Terminals, desktop, servers are the Fixed Host, which is interconnected by means of a fixed network.

Large databases can run on servers that guarantee efficient processing and reliable storage of database. Fixed hosts perform the transaction and data management functions with the help of data base servers (DBS). Mobile units are the portable computers which can retain the network connections through the support of the Base Stations (BS).

Mobile clients may vary from the thin to full clients based on their characteristics.

1. Thin client architecture: In this organization, the resources of the mobile clients are limited. The mobile client takes the request from the user and the execution of the transaction is done at the fixed host.
2. Full client architecture: In this architecture, clients can work in the disconnected mode. Full clients own the responsibility of server functions. They are portable and have enough resources for execution of an application.

In addition to the thin or full client architecture, transactions initiated by the mobile clients may also work on the Flexible client server architecture or the Client agent server architecture. In the flexible client server architecture the roles of client and

servers can be dynamically relocated. In order to increase the throughput of the system, the distinction between clients and servers may be temporarily blurred.

In the client-agent-server architecture, a three tier model introduces an agent similar to a proxy located on the fixed network [1]. In this paper we deal with the crash recovery mechanisms for the thin client and the full client architectures respectively.

3. REPLICATION STRATEGIES

Data Replication is the process that allows building a distributed environment through the management of multiple copies of data, caching one copy on each site. Replication however has the challenges of the replication control. Changes submitted to one replica have to be applied at the other replicas such that the different copies of the database remain consistent despite concurrent updates.

In general, there are two types of replication strategies: Synchronous and Asynchronous replication. In *Synchronous replication* (real time data replication [6]), performs updates on all replicas at the same time. In *Asynchronous replication* (store and forward replication [6]), operations performed on one site is stored locally and later on it is updated to other replicas.

Synchronous replication technology ensures highest level of data integrity but requires permanent availability of participating sites and transmission bandwidth. If one of the participating site holding a replica is not ready the transaction may not proceed. This scheme needs more resources but the replicas will be consistent at any instance of time. Asynchronous replication provides more flexibility than synchronous replication as the database synchronization time interval can be defined which can vary between the applications and from one service provider to another. Moreover a single site could work even if a remote server is not reachable or down.

As disconnections in mobile environments are treated as normal conditions rather than failures, asynchronous replication is better suited. If the mobile host holding a replica is disconnected for a longer time then in synchronous replication the transaction can't proceed unless the mobile host is connected. However techniques described in [11] may help in synchronous replication but it takes more time to keep the consistent copies of the data. Comparatively though asynchronous replication is better suited, efficient concurrency control and fault tolerance mechanisms are to be deployed. A Concurrency control algorithm proposed [10] is better suited for the implementation of concurrency control in mobile environment when the sites are updated based on asynchronous replication policy.

Based on the mobility, there are two approaches to the replication of data in mobile environments. In the first approach the data may be replicated at both the mobile host and the fixed host. Whenever the updates are preformed at mobile host (mobile client) the same has to be consistent with the data on fixed host (mobile server). If Synchronous replication is adopted both the mobile host and fixed host must be in connected mode at least when the updates on replicas are performed. The concurrency control can be effectively implemented using the dynamic timer management mechanism discussed in [10]. Since mobile hosts are prone to disconnections asynchronous replication can be adopted. The replication control is done at the client side. When the mobile host is in disconnected mode and a client replica is updated. These updations has to be performed on the fixed host within the specified time interval (this time interval differs from application to application).

The second approach to the replication deals with mobility. In this mechanism the MSS will maintain the replicas i.e it maintains the list of sites were replicas are present in home location as well it forwards one of the replica to remote location. If any update is performed by the replica in foreign location, the updates are conveyed to MSS which updates the replicas present in that cell.

4. IDENTIFICATION OF REPLICAS

In replicated environment, any database instance my contain the local data or it could be instance of another replica. In this section we followed the approach of Cavelleri et al[6] to distinguish the local site information and the replicated information of other sites.

Any table at any site either mobile host or fixed host can be represented in the following form:

$$T_{a,i} = P_{a,i} \; u \; U_{j=1,...N,\, J \neq i} \; R_{a,j,i}$$

where

a identifies a generic table of the database

i identifies the site (Mobile host or/and fixed host)

$T_{a,i}$ is the entire content of table a on site i

$P_{a,i}$ is the information entered in table a on site i.N is the number of participating sites i.e no of sites where the data is replicated.

$R_{a,j,i}$ is the replicated partitions of table a coming from site j cached into site i

Each site (Mobile host and Fixed host) is uniquely identified by site-id. $P_{a,i}$ and $R_{a,j,i}$ are partitions of table $T_{a,i}$. When an insertion occurs in table a on local site i the information is stored in local partition $P_{a,i}$ and later forwarded to every replicated partition $R_{a,j,i}$ of each remote site j.

Local and replicated sites could be the mobile host and fixed host or it could be one cell to another cell. To identify a record as to whether it is a local or remote partition site record ids are used i.e record id along with the site id will identify a particular record. This mechanism of record identification helps in knowing the origin of a record thereby reducing the overhead of directory structure to be maintained by MSS as discussed in[3].

Consider the following Table stored at site 1, which has local partitions as well as remote partitions. It has three fields A,B,C and the first field specifies the record id (site id + record id). The remote partitions for the purpose of updating of data can be identified using this identifier

Record_Id	A	B	C
(1,1)	10	20	30
(2,1)	23	34	56
(1,3)	33	12	21

Table 1: Schema representation using composite record id

The first two records belong the the site 1 and their record ids are 1, 2 respectively. The third record belongs to site 3 whose record id is 1. If an updation is performed for the third record on data item C. Apart from updating new value of C. The first record and partition 3 will also be updated. This is identified by composite record id.

5. MOBILE DATA OWNERSHIP MODELS

Data ownership models specifies various mechanisms that can be adopted to keep the replicas consistent. These mechanisms may be adopted based on different applications implemented in mobile environments. There are different types of Data ownership models viz., Workload partitioning data ownership model, Master/Slave data ownership model and Update any-where datownership model. The applicability of these models in mobile environments is discussed below:

5.1 Workload Partitioning data ownership Model

This model is implemented by read only access to $R_{a,j,i}$ and read write access to $P_{a,i}$ .The replicated partition of table from site j to will only have read access at site i. In this scheme all replicas can only read the information coming from remote partition. It can preform the write operation if the data item belongs to the local site.

If the mobile host needs to update the data item stored at some fixed host, the write operation has to be performed at fixed host. Even though a replica need to update the data item but since it has read only access the write operation has to be done at only local site for that data item. This mechanism preserves the integrity of the data as the updations to the data item is done at only one site.

If the mobile host moves from one cell-1 to cell-2 and if a write operation has to be performed on a data item then the updates are performed only in cell-1 as the fixed host of that mobile host is in cell-1.

5.2 Master/Slave Data Ownership Model

Master/Slave replication is an asynchronous replication where data is owned by one site (owner) and can be updated by only that site. It follows publisher-subscriber pattern. The other sites who own only read only data are slaves (subscribers).

In this approach whenever a data item has to be updated, the owner is identified using the composite record id scheme (record id + site id) so that the request for the update is given to the respective site and then the site updates the data items. However the site i where updations to the database item is performed has to propagate these updates to all replicas i.e all the slaves need to know the final value of the data item. This is similar to the Thin client architecture of the mobile environment where the mobile host requesting for a service send the request to the fixed host.

When a mobile host moves from once cell to another. The request for write operation for a particular data item may have to sent back to the home cell if the subscriber is in home cell. Concurrency control and failure tolerance is easily implemented as the udpations are done by only one site[10].

5.3 Update Anywhere Data Ownership Model.

In Update Anywhere Data Ownership Model, read write access is available to $R_{a,j,i}$ and $P_{a,i}$. This leads to replication conflicts because any mobile host or fixed host where the replica is present, a write operation on the data item can be performed. The database can be kept consistent by implementing efficient concurrency control techniques for mobile environments [5,10]. However these techniques differ from one replication strategy to another.

In Synchronous replication log based scheme[12] is implemented as the replicas need to be consistent all the time. Whereas in synchronous replication triggers are used to update the replicas after firing of events. The mobility issue can be managed by transfer of log information from one cell to another[10].

6. CONCLUSION

Replicating data at several sites is a powerful mechanism to increase the performance, throughput and can provide fault tolerance. However unlike synchronous replication, asynchronous replication may not keep the database consistent at every moment. If this time lag can be compromised it may require less resources. Further the problem of storage of log on the mobile will not be an issue. Based on the volume of replication and the number of transactions respective ownership models may be chosen. However update anywhere ownership model can make the replication strategies in increasing the throughput but has to be implemented as synchronous replication as any site can update data at any time. When the replication increases, performance will be an issue. Methodologies for the management of huge replicas are needed.

REFERENCES

[1]. Can Turker, Gabriele Zini, "A Survey of Academic & Commercial Approaches to Transaction Support in Mobile Computing Environments", DELOS, Network of Excellence on Digital Libraries, Project No.507618-A, PP. 8-35, 2003.

[2]. Evaggelia Pitoura, Bharat Bhargava, "Revising Transaction Concepts for Mobile Computing", wmcsa, Pp.164-168, 1994 First Workshop on Mobile Computing Systems and Applications, 1994.

[3]. Daniel Barbara Milla, Hector Garcia Molina, "Replicated Data Management in Mobile Environments: Anything New under the Sun?", IFIP Transactions; Vol. A-44 Proceedings of the IFIP WG10.3 Working Conference on

Applications in Parallel and Distributed Computing, PP. 237 – 246, 1994.
[4]. Gary D Walborn, Panos K Chrysanthis, "Management of Mobile Transactions", Proceedings of the 1999 ACM Symposium on Applied computing, PP. 389 – 398, 1999.
[5]. Jin Jing, Omran Bukhres, Ahmed Elmagarmid, "Distributed Lock Management for Mobile Transactions", Proceedings of 15th International conference on Distributed Computing Systems, PP. 118-125, 1995.
[6]. M. Cavalleri, R. Prudentino, U. Pozzoli, G. Veni, "A set of tools for building PostgreSQL distributed database in biomedical environment, Proceedings of the 22nd Annual International conference on "Engineering in Medicine and Biology society", PP. 540-544, 2000.
[7]. Minsoo Lee, Sumi Helal, "High Commit Mobile Transactions", Distributed & Parallel Databases, Vol. 11, Issue: 1, PP.73-92, 2002.
[8]. Nadia Nouali, Anna Doucet, Habiba Drias, "A Two Phase Commit Protocol for Mobile Wireless Environment" Proceedings of 16th Australasian Database Conference.Vol.39 (ADC 2005), PP. 135-143, 2005.
[9]. Patrica Serran-Alvarado, Claudia Roncancio, Michel Adiba, Cyril Labbe, "A Survey of Mobile Transactions" Distributed and ParallelDatabases, 16, 193-230,2004.
[10]. Salman Abdul Moiz, Dr. Lakshmi Rajamani, "Single Lock Manager Approach for acheiving Concurrency Control in Mobile Environments", Proceedings of 13 IEEE International Conference on High Performance Computing, 2007.Springer LNCS 4873, PP.650-660, 2007.
[11]. Salman Abdul Moiz, Dr. Lakshmi Rajamani, "Disconnected Modes of Operations in Mobile Environments", Proceedings of INDIACom-2008, 2nd National Conference on Computing for Nation Development, PP.253-256, 2008.
[12]. Salman Abdul Moiz, Dr. Lakshmi Rajamani, "Log Based Recovery in Mobile Environments", Proceedings of 1st International conference on Advance Computing, ICAC-08 (ACM), ISBN: 978-81-906457-1-3, PP. 530-533.
[13]. Shuqing Wu, Bettina Kemme, "Combining Replica Control with Concurrency Control based on Snapshot Isolation", Proceedings of 21st International Conference on Data Engineering (ICDE'05), PP. 422-433, 2005

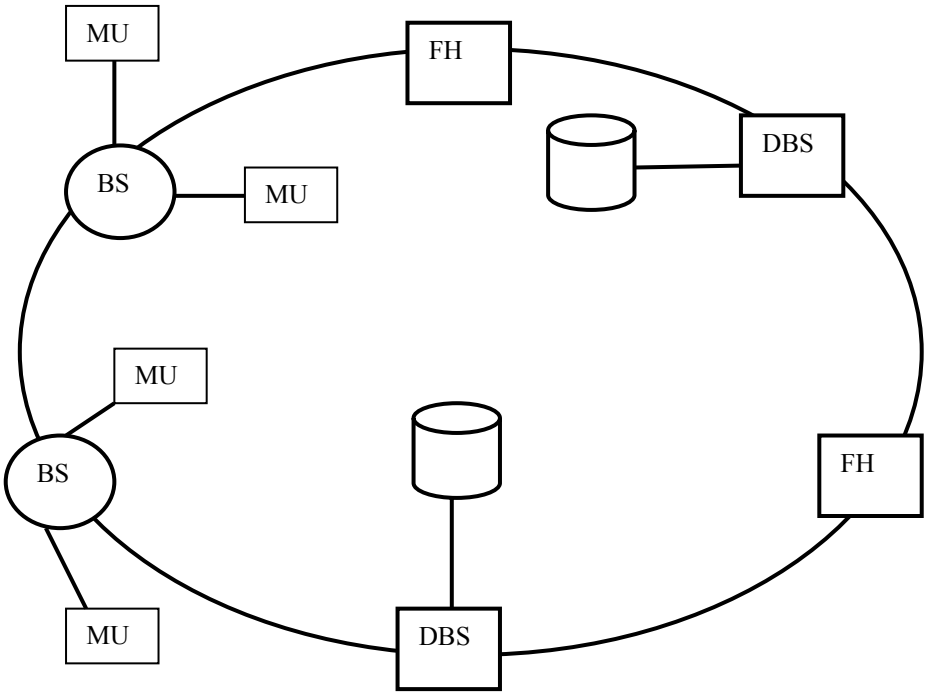


Figure 1: Architectural View of Mobile Environment

Continued on Page No. 169

Management Information System in Indian Universities: A Comparative Study

Sangeeta Gupta[1], H. Bansal[2] and **A. K. Saini[3]**

Abstract - The role of IT in development has been acknowledged worldwide and is expected to bring in major social and economic benefits for the mankind. In the present IT era, like every organization, the management of universities is a key challenge particularly in developing countries like India. Effectiveness of university depends upon its ability to maintain itself internally and adapt to new and dynamic environment. Computerization of universities may substantially improve response and efficiency and ensure savings in terms of time, money and other scarce resources. The present study aims to bring out the status of Management Information System in Indian Universities in terms of the adoption of IT in various functional areas, various issues facing the effective utilization of IT and the level of understanding about MIS among service providers in the University system. The study is based on primary data, which has been collected through a well-designed structured questionnaire from two Central Universities and two State Universities. The data so collected was analyzed to test various hypotheses with the help of SPSS software. Based upon the structural analysis inferences have been drawn which will help University administration to re-engineer their services to make them more effective and efficient.

Index Terms - Information technology (IT), Management Information System (MIS), University Administration, Internet, Intranet, Website, AMC.

1. INTRODUCTION

This paper presents analytical framework for the research study. It throws light on the status of Management Information System in Indian Universities in terms of the adoption of IT in various functional areas, various issues facing the effective utilization of IT and the understanding about MIS among service providing personnel of the University system.

2. UNIVERSITY AS AN ORGANIZATION

University is defined as a body of academic people engaged in the pursuit of academic matters. [1] Universities in present day socio-economic milieu have assumed great importance. A university as an organization consists of teachers, researchers, students, administrators and various sub systems such as examination, finance, personnel, stores, maintenance, planning, etc. [Fig. 1]. In a service organization like a university, the organizational structures is made up of inter-related and inter-dependent parts, and one part or subsystem cannot perform effectively without the other. [2]

[2]*Department of Business Management, GJUS&T, Hisar*
[3]*University School of Management Studies, GGSIPU, New Delhi*

3. MANAGEMENT IN UNIVERSITIES

The University authorities have to perform academic as well as management functions. Some of them are of strategic nature and others are routine. Although university organization has men, machines, materials and money, yet the characteristics of their participants differ from other organizations and therefore, their decision making processes are unique or different in some respects. Their effective management is essential for optimal utilization of resources and for providing maximum benefits and satisfaction to its clientele. [3]

The day-to-day operations of a university System, involve handling of vast quantities of information. Multiplicity of functioning leads to information flow that is highly variable in content, format and importance. The present day set up presents manual collection and retrieval of information that is not only in itself a mammoth exercise but also a time consuming affair. Further, the administration is lesser effective and deficient because of the overburdened resources and facilities. To add to the problem, there is a voluminous paper work to be handled which often means non-availability of related data in time. There are delays in getting the data, which cannot be stored and easily analyzed. Therefore, the universities need to be organized and administered in a truly scientific manner utilizing modern management techniques and tools that are being used in other organizations. It has been observed that the educational institutions generally do not utilize managerial tools for decision making such as computers though they are available physically in the institutions. However, they are commonly found in industry and are being used frequently.

4. ADOPTION OF INFORMATION TECHNOLOGY AS A TOOL

Administrative computing is assuming an ever increasing role as the demands being placed on universities escalate, increasing number of students. As computing costs decline and methods change, universities must develop planning strategies to ensure that the new technologies and procedures are employed effectively while, at the same time, meeting the fundamental goals of the institution. Although computers can play a significant role in university administration, yet sometimes there is reluctance to adopt them because of a variety of reasons and misconceptions about computers. Nevertheless the wise and careful use of computers and Information technology can help maintain quality while keeping the costs in control in today's highly competitive environment. The rapid availability of complete information can result in savings in expenses. Therefore, all functional areas should be inter-linked so that data entered into the system from any of these points may be accessible and used by all concerned. Today the computerized MIS has been accepted as an integral part of a modern university system. [4]

The primary reasons that demand the introduction of computers so as to strengthen the information system are: to increase organizational efficiency through reduction in the overall costs; to provide useful, accurate, complete and timely information to meet the requirements of the various departments requiring such data; to improve managerial effectiveness in planning, allocating and controlling the scarce and expensive resources of the organization; to improve and ensure high quality of service at a reasonable cost; improve the management information system and to reduce clerical workload.[5]

5. RESEARCH OBJECTIVES

The objectives of this paper are:

- To study the status of computerization in the Universities.
- To understand the extent of information availability on the website.
- To identify the desired features in a University MIS and
- To study the linkage between various key factors involved in the information collection and dissemination process in the University.

6. RESEARCH METHODOLOGY

The study is based on primary data, which has been collected through a well-designed structured questionnaire. The sample comprise of all the administrative branches of the university, Directorate of Distance Education, University Library and various teaching departments of Central and State Universities. To begin with, a list of branch officers like Heads of the departments/Chairmen, Director, Librarian, Deputy Registrars, Assistant Registrars, and Section officers/Superintendents etc. was prepared and consequently the sample was selected by using Convenience/Purposive sampling technique. The survey was conducted through face-to-face interview method. The data has been put to analysis by using method of percentages. Apart from this, the use of Chi-square test and correlation analysis has been made for measuring the association between various attributes. The study analyses the result from the survey and reveals what is the status of Management Information System in Indian Universities.

7. DATA ANALYSIS

The following results were revealed after analyzing data

7.1 General Analysis

7.1.1 Status of Computerization

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	154	77.8	78.2	78.2
	No	43	21.7	21.8	100.0
	Total	197	99.5	100.0	
Missing	0	1	.5		
Total		198	100.0		

If Yes, Functioning Has Improved By Using Computers

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	155	78.3	100.0	100.0
Missing	0	43	21.7		
Total		198	100.0		

If No, Can Functioning Improved By Using Computers

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	42	21.2	97.7	97.7
	No	1	.5	2.3	100.0
	Total	43	21.7	100.0	
Missing	0	155	78.3		
Total		198	100.0		

Table 1: Working On Computer in the Section/Department

A reasonably high, 78% respondents indicate that they make use of computers for performing their work while 22% says they do not use computers. However there is a complete unanimity among all the respondents that computer has and can improve the administrative functioning and improve the work performance.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	MS Office	136	68.7	87.7	87.7
	Others	2	1.0	1.3	89.0
	MS Office and Others	17	8.6	11.0	100.0
	Total	155	78.3	100.0	
Missing	0	43	21.7		
Total		198	100.0		

Table 2: Type of Licensed Software

The only licensed software available with the institutions is MS Office, which is loaded in all computers. There is almost no other application software available for serving various specific tasks in the universities.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Stand alone mode	116	58.6	74.8	74.8
	Network	34	17.2	21.9	96.8
	Stand alone mode and Network	5	2.5	3.2	100.0
	Total	155	78.3	100.0	
Missing	0	43	21.7		
Total		198	100.0		

Table 3: Is Computer Connected

About 75% respondents working on stand-alone mode while 22% respondents say that their computers are part of the network.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	27	13.6	20.0	20.0
	No	108	54.5	80.0	100.0
	Total	135	68.2	100.0	

		Frequency	Percent	Valid Percent	Cumulative Percent
Missing	0	63	31.8		
Total		198	100.0		

Table 4: Using Customize Software In The Working

A high, 80% respondents are not using customized software in educational institutions while just 20% are using customized softwares which is very low. The restricted access appears to be mainly on account of high cost of customized software and lack of awareness of available customizes software in education sector.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	129	65.2	85.4	85.4
	No	22	11.1	14.6	100.0
	Total	151	76.3	100.	
Missing	0	47	23.7		
Total		198	100.0		

Table 5: Requirement of Computer Skilled Manpower

About 86% respondents required computer skilled person so that working of their branch will be more effective and they can use resources efficiently while 15% respondents doesn't require computer skilled person which seems to be on account of natural human tendency i.e. fear of loss of power.

7.1.2 Availability of information on Internet

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	70	35.4	36.8	36.8
	No	120	60.6	63.2	100.0
	Total	190	96.0	100.0	
Missing	0	8	4.0		
Total		198	100.0		

Frequency of Website Up-dation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	As And When It Is Required	45	22.7	70.3	70.3
	With In A Week	4	2.0	6.3	76.6
	With In A Month	4	2.0	6.3	82.8
	More Than A Month	11	5.6	17.2	100.0
	Total	64	32.3	100.0	
Missing	0	134	67.7		
Total		198	100.0		

Is Information on Internet Beneficial

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Highly Beneficial	65	32.8	56.5	56.5
	Moderately Beneficial	36	18.2	31.3	87.8
	Not Beneficial	14	7.1	12.2	100.0
	Total	115	58.1	100.0	
Missing	0	83	41.9		
Total		198	100.0		

Table 6: Availability of Information on Internet

Majority of the respondents (63%) say that information is not available on university website however, 37% respondents accept they get some of the information from university website. Among respondents who says information is available on Internet 70% respondents say it is updated whenever new information is there while 30% says it is updated quite late even after more than a month. The respondents who say they are not able to get information from Internet among them, a large number of respondents i.e. 57% agree on the issue that information is beneficial if it is available on Internet, as it would save their time and energy.

7.1.3 Efficiency and Effectiveness expected from a computerized MIS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	166	83.8	89.2	89.2
	To Some Extent	18	9.1	9.7	98.9
	No	2	1.0	1.1	100.0
	Total	186	93.9	100.0	
Missing	0	12	6.1		
Total		198	100.0		

Computerized MIS Will Reduced University Expenditure

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	149	75.3	78.4	78.4
	To Some Extent	34	17.2	17.9	96.3
	No	7	3.5	3.7	100.0
	Total	190	96.0	100.0	
Missing	0	8	4.0		
Total		198	100.0		

Faster Decision-Making through Computerized MIS

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	153	77.3	79.7	79.7
	To Some Extent	35	17.7	18.2	97.9
	No	4	2.0	2.1	100.0
	Total	192	97.0	100.0	
Missing	0	6	3.0		
Total		198	100.0		

Table 7: Computerized MIS Will Streamline University Functioning

There is a high level of agreement among respondents regarding the positive effects of computerized MIS. The table 7 reveals very emphatically that a computerized MIS would streamline University functioning, reduce expenditure, ensure better and faster decision-making as well as help in checking mal-practices.

8. STATISTICAL ANALYSIS

8.1 Cross Tabulation

8.1.1 The null hypothesis that working on computer and University type are disassociated.

			University		
			State University	Central University	Total
Working On Computer In The Branch	Yes	Count	61	93	154
		%	58.7%	100.0%	78.2%
	No	Count	43	0	43
		%	41.3%	.0%	21.8%
Total		Count	104	93	197
		%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	49.188	1	.000		
Continuity Correction	46.795	1	.000		
Likelihood Ratio	65.694	1	000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	48.939	1	.000		
N of Valid Cases	197				

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	-.500	.000
	Cramer's V	.500	.000
N of Valid Cases		197	

Table 8: Working On Computer in the Branch

Applying Pearson Chi-Square test, the value of asymptotic significance (2-sided) with 1 degree of freedom is 0.000 which is less than 0.01(at 99% level of significance), so null hypothesis is rejected means there is significant association between the variables. In the symmetric table Cramer's V is 0.500, which shows these variables have a strong association. It is evident from the table 8 that only 58.7% respondents from State Universities are using computers for their work as compare to 100% respondents of Central Universities.

8.1.2 The null hypothesis that the variables Qualification of the respondents and Knowledge about the term MIS are disassociated.

			Qualification			
			Graduate	Post Graduate	Ph.D	Total
Knowledge About The Term MIS	Yes	Count	25	65	30	120
		% within Qualification	20.8%	54.2%	25.0%	100.0%
		% within Qualification	53.2%	70.7%	76.9%	67.4%
	No	Count	22	27	9	58
		% within Knowledge About The Term MIS	37.9%	46.6%	15.5%	100.0%
		% within Qualification	46.8%	29.3%	23.1%	32.6%
Total		Count	47	92	39	178
		% within Knowledge About The Term MIS	26.4%	51.7%	21.9%	100.0%
		% within Qualification	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.372	2	.041
Likelihood Ratio	6.242	2	.044
Linear-by-Linear Association	5.710	1	.017
N of Valid Cases	178		

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.189	.041
	Cramer's V	.189	.041
N of Valid Cases		178	

Table 9: Knowledge About the Term MIS vs Qualification

Applying Pearson Chi-Square test, the value of asymptotic significance (2-sided) with 2 degree of freedom is 0.041 which is less than 0.05 (at 95% level of significance), so null hypothesis is rejected hence the variables are associated. It means qualifications of the respondents make difference about the knowledge of the term MIS in Central and State University. It is apparent from the table that Cramer's V is 0.189 which shows these variables have a weak association.

However at 99% level of significance the null hypothesis is accepted because the value of significance is 0.041, which is grater than 0.01.

8.1.3 The null hypothesis that there exists no association between Qualification and level of understanding of the term MIS.

	Qualification			
	Grad	Post	Ph,D	Total

			Qualification			
			uate	Graduate		
Level of Understanding	Very Good	Count	1	20	7	28
		% within Level of Understanding	3.6%	71.4%	25.0%	100.0%
		% within Qualification	4.0%	30.8%	23.3%	23.3%
	Good	Count	8	23	10	41
		% within Level of Understanding	19.5%	56.1%	24.4%	100.0%
		% within Qualification	32.0%	35.4%	33.3%	34.2%
	Medium	Count	15	20	8	43
		% within Level of Understanding	34.9%	46.5%	18.6%	100.0%
		% within Qualification	60.0%	30.8%	26.7%	35.8%
	Low	Count	1	2	5	8
		% within Level of Understanding	12.5%	25.0%	62.5%	100.0%
		% within Qualification	4.0%	3.1%	16.7%	6.7%
Total		Count	25	65	30	120
		% within Level of Understanding	20.8%	54.2%	25.0%	100.0%
		% within Qualification	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp Sig. (2-sided)
Pearson Chi-Square	16.880	6	.010
Likelihood Ratio	17.29	6	.008
Linear-by-Linear Association	.905	1	.341
N of Valid Cases	120		

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.375	.010
	Cramer's V	.265	.010
N of Valid Cases		120	

Table 10: Level of Understanding Vs. Qualification

Applying Pearson Chi-Square test, the value of asymptotic significance (2-sided) with 6 degree of freedom is 0.01 which is less than 0.05 (at 95% level of significance), so null hypothesis is rejected means there is significant association between the variables. In the symmetric table Cramer's V is 0.265, which shows these variables have a weak association. Among respondents about 71% postgraduate respondents having very good understanding of the term MIS while 60% of graduate respondents having a medium level of understanding.

8.1.4 The null hypothesis Qualification of the respondents and level of satisfaction from maintenance arrangement has no relationship.

			Qualification			
			Graduate	Post Graduate	Ph.D	Total
Level of Satisfaction From Maintenance Arrangement	Yes	Count	18	41	14	73
		% within Level of Satisfaction From Maintenance Arrangement	24.7%	56.2%	19.2%	100.0%
		% within Qualification	46.2%	55.4%	36.8%	48.3%
	No	Count	21	33	24	78
		% within Level of Satisfaction From Maintenance Arrangement	26.9%	42.3%	30.8%	100.0%
		% within Qualification	53.8%	44.6%	63.2%	51.7%
Total		Count	39	74	38	151
		% within Level of Satisfaction From Maintenance Arrangement	25.8%	49.0%	25.2%	100.0%
		% within Qualification	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	3.566[a]	2	.168
Likelihood Ratio	3.595	2	.166
Linear-by-Linear Association	.639	1	.424
N of Valid Cases	151		

Table 11: Level of Satisfaction from Maintenance arrangement Vs. Qualification

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 18.37.

Applying Pearson Chi-Square test, the value of asymptotic significance (2-sided) with 2 degree of freedom is 0.168 which is higher than 0.05 (at 95% level of significance), so null hypothesis is found to be significantly good i.e. there is no relationship between the variables. Among the respondents

who are satisfied from maintenance arrangement about 76% respondents are of above graduation level.

8.1.5 The null hypothesis the Qualification of the respondents and type of change required in present maintenance arrangement are disassociated.

Refer Table 12 at the end

Applying Pearson Chi-Square test, the value of asymptotic significance (2-sided) with 4 degree of freedom is 0.171 which is larger than 0.05 (

In all the qualification level, a large number of respondents (61.9%, 58.1% and 37% respectively) required training of their own staff to maintain the systems, which shows they do not want to dependent on the outside agencies.

8.1.6 The null hypothesis that respondents have under gone any training for using computer and age are independent.

			Age			Total
			< 35 years	35-50 Years	> 50 Years	
Has Under Gone Any Training For Using Computer	Yes	Count	12	37	42	91
		% within Has Under Gone Any Training for Using Computer	13.2%	40.7%	46.2%	100.0%
		% within Age	54.5%	67.3%	62.7%	63.2%
	No	Count	10	18	25	53
		% within Has Under Gone Any Training For Using Computer	18.9%	34.0%	47.2%	100.0%
		% within Age	45.5%	32.7%	37.3%	36.8%
Total		Count	22	55	67	144
		% within Has Under Gone Any Training For Using Computer	15.3%	38.2%	46.5%	100.0%
		% within Age	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	1.108[a]	2	.575
Likelihood Ratio	1.096	2	.578
Linear-by-Linear Association	.139	1	.709
N of Valid Cases	144		

Table 13: Has Under Any Training for Using Computer Vs. Age

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.10.

Applying Pearson Chi-Square test, the value of asymptotic significance (2-sided) with 2 degree of freedom is 0.575 which is greater than 0.05 (at 95% level of significance), so null hypothesis is accepted i.e. age does not make a difference whether a university employee has under gone any training.

A very low, about 13% of respondents in the age group of Less than 35, have under gone training for using computers in their sections/departments. Such a low percentage indicates that the younger group is computer savvy so they not required any training.

9. CORRELATION

			Requirement of Computer Skilled Manpower	Non Availability of Trained Manpower
Spearman's rho	Requirement of Computer Skilled Manpower	Correlation Coefficient	1.000	.181 *
		Sig. (2-tailed)		.028
		N	151	148
	Non Availability of Trained Manpower	Correlation Coefficient	.181 *	1.000
		Sig. (2-tailed)	.028	
		N	148	192

Table 14: Correlation between Requirement of Computer Skilled Manpower & Non Availability Of Trained Manpower

* Correlation is significant at the 0.05 level (2-tailed).

Applying Spearman's correlation the value of correlation coefficient is 0.181 which is significant at the 95% level of significance, therefore the two variables i.e. requirement of computer skilled manpower and non availability of trained manpower are positively correlated. The problem of non availability of trained manpower indicates that there is a requirement of computer skilled manpower in the sections.

			Having Sufficient Software for Working	Non Availability of Relevant Software
Spearman's rho	Having Sufficient Software For Working	Correlation Coefficient	1.000	.291**
		Sig. (2-tailed)		.000
		N	155	152
	Non Availability of Relevant Software	Correlation Coefficient	.291 **	1.000
		Sig. (2-tailed)	.000	
		N	152	159

Table 15: Correlation between Having Sufficient Software for Working & Non Availability of Relevant Software

** Correlation is significant at the 0.01 level (2-tailed).

Applying Spearman's correlation the value of correlation coefficient is 0.291 which is significant at the 99% level of significance; therefore the two variables i.e. having sufficient

software for working and non availability of relevant software are positively correlated. The problem of non availability of relevant software is exists in the section because they are not having sufficient software for their working.

			Is Requirement of More Computers	Inadequate Computers
Spearman's rho	Is Requirement of More Computers	Correlation Coefficient	1.000	.372**
		Sig. (2-tailed)		.000
		N	152	143
	Inadequate Computers	Correlation Coefficient	.372**	1.000
		Sig. (2-tailed)	000	
		N	143	152

Table16: Correlation between Is Requirement of More Computers & Inadequate Computers

**. Correlation is significant at the 0.01 level (2-tailed).

Applying Spearman's correlation the value of correlation coefficient is 0.372 which is significant at the 99% level of significance, therefore the two variables i.e. requirement of more computers and inadequate number of computer are positively correlated. The problem of inadequacy of computers indicates that there is a requirement of more computers in the sections.

10. FINDINGS & CONCLUSIONS

The above data analysis and the subsequent study carried out gives rise to the following findings:

1. The adoption of computers for various applications has been found to be higher in Central Universities as compared to State Universities. [Table 8].
2. The Universities at present are not using any custom made software for any of the applications. Instead, end users are making use of standard general purpose packages and have developed small applications to meet their day-to-day requirements. [Table 2].
3. The computer networking has not been adopted fully thereby limiting the data sharing and exchange in the university and restricting its usage substantially. [Table 3].
4. The existing practice of system maintenance requires substantial improvements. [Table 12].
5. The website contents require to be updated dynamically on real-time basis to ensure currentness of the information on the website. Further the website should contain meaningful features to reduce the physical visits of the users for various activities. [Table 6].
6. The need and relevance of MIS in University system should be percolated to the lowest level (service providers). This will help in ensuring that all users confirm and adhere to the system requirements as far as data preparation is concerned. [Table 10].
7. There is a strong need to setup a central computing facility in the university to carryout the above tasks and should own the system for its success. In addition, every user section should assign a coordinator to ensure proper coordination of the MIS services with the Central facility. [Table 14].
8. Regular orientation/awareness programmes must be conducted periodically to empower the service providers/users to ensure effective utilization of the system.

11. CONCLUSIONS

The central universities are better placed in terms of adoption of Information technology for various functions. There is an urgent need to focus attention of various factors such as: availability of custom-made application software for optimum hardware investment, proper networking and maintenance, with strong integrated information system approach rather than compartmentalized applications with adequate skilled manpower support.

REFERENCES

[1]. Sharma P.K, Management Information Systems in the Institution of Higher

[2]. Learning, Jaipur, Kuber Associates and Publishers1987.

[3]. Mehendiratta, Pradeep R., "University Administration in India and the USA", Oxford and IBH Publishing Co., New Delhi, 1984.

[4]. Sanyal B.C., "The Use of Computerized Information System to Increase Efficiency in University Management", *IIEP Contributions,* 20.

[5]. Achuthan, Sarla, Binod C Agarwal, et. al., *Computer Technology for Higher Education, vol. I, An Overview*, Concept Publishing Company, New Delhi, 1993.

[6]. Goyal, R.C., "Hand book of hospital personnel management", Prentice Hall of India, New Delhi, 1993.

[7]. Ahmad, S., Management Information System: Tool to Enhance Efficiency of University Management, *University News*, 40(46), November 18-24:12-17, 2002.

[8]. Forkner and McLeod, *Computerized Business System: An Introduction to Data Processing,* New York: John Wiley and Sons, 1973.

[9]. Joshi, M.J. and KulKarni, Management Information System for University Administration, *University News*, 39(44), Oct. 29-Nov.4:11-14, 2001 .

[10]. Kaptan, S.S., Management Information System and University Administration, *University News*, Monday, Aug. 27: 6-13, 1990.

[11]. Nunnally, J., *Psychometric Theory*, New York: McGraw-Hill, 1978.

[12]. Pant, M.M., Internet and Education: With Focus on Higher Education, In *Reading Material of Leadership Development Programme for Principals of Colleges on IT for Development and Management of Colleges,* 2002 .

[13]. Sansanwal, D. N., Information Technology and Higher Education, *University News*, 38(46), November 13:1-6, 2000.
[14]. Cooper, Donald R. *et. al.*, *Business Research Methods*, Eighth edition, TMH, New Delhi, 2003.
[15]. Malhotra, N.K., *Marketing Research*, Third edition, Pearson Education, 2001.
[16]. Kothari, C.R., *Research Methodology: Methods and Techniques*, Second edition, New Age International (P) Limited, New Delhi, 2004.
[17]. Nargundkar, R., *Marketing Research: Text and Cases*, Second edition, TMH, New Delhi, 2004.

			Qualification			Total
			Graduate	Post Graduate	Ph.D	
Type Of Change Required in Present Maintenance Arrangement	AMC	Count	4	12	7	23
		% within Type of Change Required in Present Maintenance Arrangement	17.4%	52.2%	30.4%	100.0%
		% within Qualification	19.0%	27.9%	25.9%	25.3%
	Training of The Own Staff	Count	13	25	10	48
		% within Type of Change Required in Present Maintenance Arrangement	27.1%	52.1%	20.8%	100.0%
		% within Qualification	61.9%	58.1%	37.0%	52.7%
	Visit Basis + Parts	Count	4	6	10	20
		% within Type of Change Required in Present Maintenance Arrangement	20.0%	30.0%	50.0%	100.0%
		% within Qualification	19.0%	14.0%	37.0%	22.0%
Total		Count	21	43	27	91
		% within Type of Change Required in Present Maintenance Arrangement	23.1%	47.3%	29.7%	100.0%
		% within Qualification	100.0%	100.0%	100.0%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	6.407[a]	4	.171
Likelihood Ratio	6.269	4	.180
Linear-by-Linear Association	.451	1	.502
N of Valid Cases	91		

Table 12: Type of Change Required in Present Maintenance Arrangement Vs. Qualification
a. 1 cells (11.1 %) have expected count less than 5. The minimum expected count is 4.62.

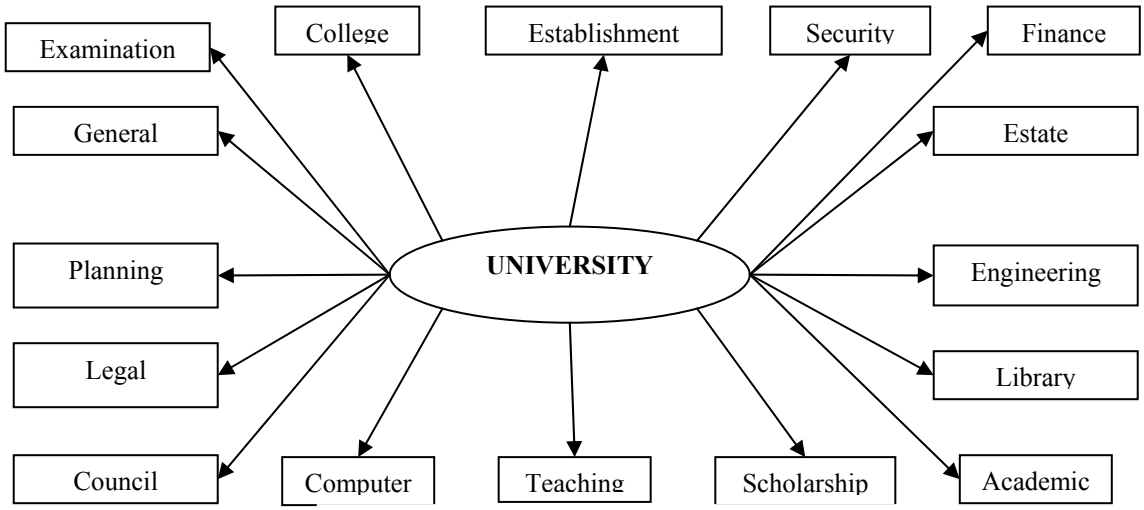


Figure 1: University and its Subsystems

Evolutionary Analytics on Lysosomal Associated Membrane Protein -1 (LAMP-1)

Manish Dwivedi[1], Vijay Tripathi[2], Ashutosh Mani[3] and Dwijendra K. Gupta[4]

Abstract - Lysosomes,the endocytic subcellular compartments play a very important role in the disintegration and recycling of cellular substances. The lysosomal associated membrane proteins LAMP-1 and LAMP-2 are major constituents of the lysosomal membrane. Using different bioinformatics tools, we established the phylogenetic relationship among LAMP-1 proteins from different organisms. The phylogenetic analytics based on ClustalW, MEGA4 and BioEdit softwares showed structural as well as qualitative similarities and dissimilarities of LAMPs and helped us to predict the nature, structure and localization of amino acid of the membrane proteins in the lysosomal membrane. This information can help one explain the molecular basis of different metabolic diseases associated with the lysosomal membrane proteins like Lysosomal storage diseases, I-cell disease etc. and to solve the questions related to biogenesis of lysosomes. This study includes the alignment of the sequences of LAMP-1 by ClustalW and revealed that Ala,, Gly, Leu, Asp and Ser are most frequently occurring amino acids with higher frequency percentage and position from 240 to 420 showed minimal entropy. The entropy rarely touched the scale of two in entropy plot that is a sign of better alignment. It revealed that these proteins were more hydrophobic in N- terminal and C-terminal domain and basically of non-hydrophobic in nature. This work also represented the evolutionary order of LAMP-1 proteins among different mammalians. This study presents the first comparative genomic study and evolutionary analysis of the LAMP-1 proteins across family of organisms with special reference to mammals.

Index Terms -
Endocytic, phylogenetic, LAMP, hydrophobicity, lysosomes.

ABBREVIATIONS

LAMP- Lysosomal associated membrane proteins, NCBI-National Center for Biotechnology Information, BLAST-Basic local alignment search tool, BLOSUM-Blocks of amino acid substitution matrix. MEGA- Molecular Evolutionary Genetics Analysis.

1. INTRODUCTION

Lysosomes represent membrane-bound dense organelles of eukaryotic cells, specialized in breakdown of all four classes of macromolecules. Materials delivered to lysosomes by endocytosis / phagocytosis or autophagocytosis are degraded in these organelles by concerted action of more than 40 hydrolases. The degradation products are then exported to the cytosol through specific transporters and reused in the cellular metabolism [6].

[1,2,3,4]*Center of Bioinformatics, University of Allahabad, Allahabad- 211002, India*
E-Mail: [4]dwijenkumar@rediffmail.com and
[1]mdwivedibio@yahoo.com

The physiological importance of lysosomal metabolite efflux is illustrated by the existence of a group of lysosomal storage diseases with transport defects, such as sialic acid storage disorders and nephropathic cystinosis. These inherited diseases result from defective efflux of sialic acid and cystine from lysosomes, respectively, and they have been linked to mutations in the membrane proteins sialin and cystinosin [20, 21], which are believed to represent sialic acid and cystine transporters. However, most lysosomal transporters, although biochemically characterized, remain unknown at the molecular level.

After the discovery of lysosomes by de Duve and coworkers [3] and conceptualization of inborn lysosomal disease' by Hers, a wide interest in understanding the biology and pathology of lysosomal disorders has led to the discoveries of nearly all lysosomal hydrolases and their encoding genes. However, the knowledge about proteins of the lysosomal membrane, which controls the interchange with other compartments and the cytosol and restricts the aggressive enzymes to the lysosomal interior, remained incomplete [4]. More than 20 distinct transport processes facilitating mainly the export of degradation products across the lysosomal membrane have been characterized functionally [15]. Yet, most of the transport catalyts remain unknown. Similarly, biogenesis of lysosomes and the machinery regulating their interaction with other compartments are still incompletely understood. In a notable report on rat tritosomal membranes, 219 proteins were identified, including 24 novel tentatively lysosomal proteins [14].

Two major lysosomal membrane sialoglycoproteins with apparent Mr ~ 120,000 containing polylactosaminoglycan comprise approximately 0.1-0.2% of total cell proteins. Immunoelectron microscopic examination of HeLa cells localized these two glycoproteins mainly to lysosomes and multivesicular bodies [4]. A number of different cell lines also express these glycoproteins. However, the apparent molecular weights differed between cell lines probably due to differences in the amount of polylactosaminoglycan expressed by each cell line. Fukuda(1988) reported that one of the glycoproteins is very homologous to that of a mouse counterpart, m-lamp- 1 The analogous human form of this glycoprotein is named human lamp-1 (h-lamp-1), while the other glycoprotein, to which the monoclonal antibody was made, is called human lamp-2 (h-lamp-2). Polylactosaminoglycans are heterogenous saccharides often having high molecular weights [18] and represent various antigenic structures such as AB0 blood group antigens, developmental antigens such as mouse F9 antigens and human fetal (i) erythrocyte antigen, and tumor-associated antigens such as sialyl Le". More recently, it has been shown that the lack of polylactosaminoglycan on the human erythrocyte anion transporter causes the glycoprotein to aggregate, resulting in

abnormal membrane structures in a congenital dyserythropoietic anemia-type II [9].

	NCBI Accession code	Lengh (aa)
Mus musculus	gi|13905006|gb|AAH06785.1|	189
Bos Taurus	gi|115497212|ref|NP_001068592.1|	409
Homo sapiens	gi|39645231|gb|AAH07845.2|	248
Rattus norvegicus	gi|6981144|ref|NP_036989.1|	407
Felis catus	gi|156447859|gb|ABU63691.1|	179
Gallus gallus	gi|45384206|ref|NP_990614.1|	414
Xenopus laevis	gi|147902288|ref|NP_001087042.1|	417
Macaca mulatta	gi|109121337|ref|XP_001087801.1|	416
Marmota monax	gi|121044661|gb|ABM46909.1|	322
Cricetulus griseus	gi|1346461|sp|P49129.1	407
Canis lupus familiaris	gi|73989504|ref|XP_534193.2|	413
Sus scrofa	gi|58332862|ref|NP_001011507.1|	413
Pan troglodytes	gi|114650748|ref|XP_001144542.1|	375
Equus caballus	gi|149730523|ref|XP_001495702.1|	400
Monodelphis domestica	gi|126337411|ref|XP_001374132.1	422

Table 1: LAMP-1 sequences with their length and NCBI accession code

2. MATERIALS AND METHOD

In order to search Lysosomal associated membrane proteins (LAMPs)family members we performed *BLAST* [1] by using blastp program in the protein database at *NCBI* [22].*Homo sapiens* LAMP-1 proteins' *gi|39645231|gb|AAH07845.2|* amino acid sequence was selected as query. From the hits 15 sequences, each from different species were selected for further studies. All the sequences were taken in *FASTA* format. The sequences were examined individually and aligned using *CLUSTALW* [11]. *Bioedit version 7.0.9.0* [10] was used for manual editing and analysis of sequences. *Kyte J and Doolittle* [13] method was used to plot hydrophobicity profile. Entropy is then calculated as:

$$H(l) = -\Sigma f(b,l)\ln(f(b,l))$$

where $H(l)$ = the uncertainty, also called ***entropy*** at position l, b represents a residue (out of the allowed choices for the sequence in question), and $f(b,l)$ is the frequency at which residue b is found at position l. The information content of a position l, then, is defined as a decrease in uncertainty or entropy at that position. As an alignment improves in quality, therefore, the entropy at each position (especially conserved regions) should decrease, which gives a measure of uncertainty at each position relative to other positions. Maximum total uncertainty will be defined by the maximum number of different characters found in a column.A window of defined size that was 13 is moved along a sequence, the hydropathy scores are summed along the window, and the average (the sum divided by the window size) is taken for each position in the sequence. The mean hydrophobicity value was plotted for the middle residue of the window. Eisenberg et. al. method [7] was used to plot hydrophobic moment profile with a window size of 13 residues having six residues on either side of the current residue and rotation angle, θ =100 degrees.

$$\mu H = \{[Hn\sin(\delta n)]\text{^}2+[Hn\cos(\delta n)]\text{^}2\}$$

Where μH is the hydrophobic moment, Hn is the hydrophobicity score of the residue H at position n, δ=100 degrees, n is position within the segment, and each hydrophobic moment is summed over a segment of the same defined window length.

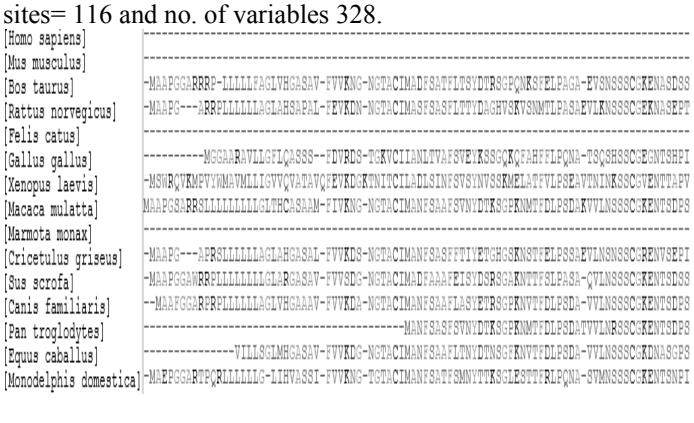
For a conserved region search within the multiple aligned sequences minimum segment length was set to 15 residues, maximum average entropy was set to be 2.0 and the gaps were limited to 2 per segment. Multiple sequence alignment, phylogenetic and molecular evolutionary analyses were conducted using *MEGA version 4* [19]. For pair wise and multiple alignments gap open penalty was -7 and gap extension penalty was -1. *BLOSUM* weight matrix was used for substitution scoring [2]. Hydrophilic gap penalties were used to increase the chances of a gap within a run (5 or more residues) of hydrophilic amino acids; these are likely to be loop or random coil regions where gaps are more common. The multiple alignmets of sequences of LAMP-1 proteins were used to create phylogenetic trees. The evolutionary history was inferred using the *Neighbour-Joining method* [17]. All the characters were given equal weights. The bootstrap consensus tree inferred from 20000 replicates [8] was taken to represent the evolutionary history of the taxa analyzed [8]. Branches corresponding to partitions reproduced in less than 50% bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap tests (20000 replicates) are shown next to the branches [8]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the *phylogenetic tree*. The evolutionary distances were computed using the *poisson correction method* [23] and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). There were a total

of 176 positions in the LAMP-1 proteins' final dataset. Phylogenetic analyses were conducted in *MEGA4* [9].

3. RESULT AND DISCUSSION

Multiple sequence alignment

The Multiple alignment of LAMP-1 proteins (Figure 1) resulted into an alignment having 429 positions. By statistical analysis of multiple aligned sequences it was observed that alanine, glycine, leucine, asparagine, serine, threonine and valine are the most frequently present amino acids with frequency percentage of 8.64, 6.22, 9.85, 7.00, 10.22, 8.45 and 7.71 respectively. While within conserved sites alanine, cystein, phenylalanine, glycine, leucine, asparagine, serine and valine are the most frequently occuring amino acids with frequency percentage of 10.70, 8.10, 8.72, 7.34, 12.00, 6.35, 6.96 and 6.27 respectively. The multiple aligned sequence of LAMP-1 protein was found with No. of conserved sites=99, No. of parsimony informative sites= 207, No. of singleton sites= 116 and no. of variables 328.



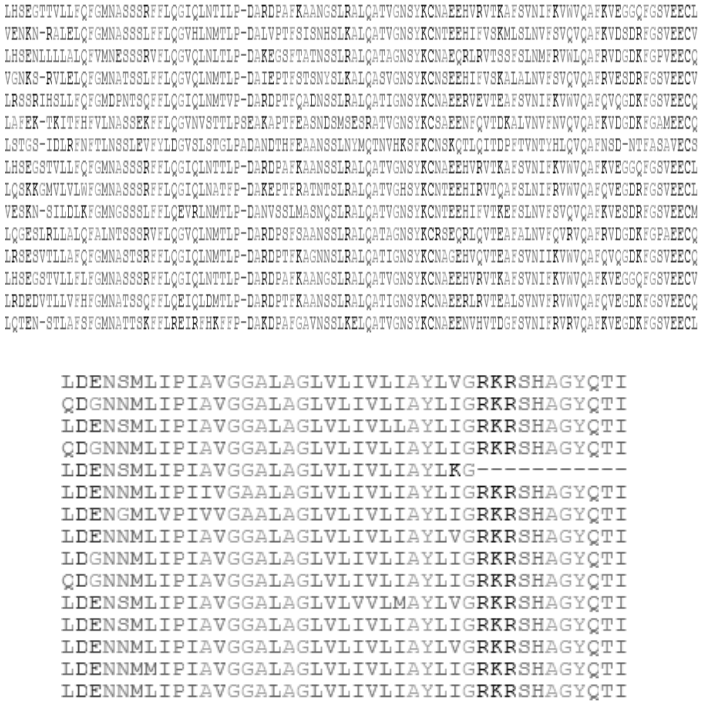


Figure1:.Multipe sequence alignment of LAMP-1 proteins.

Conserved region search

A conserved region search resulted into six regions from position 241 to 257 (segment length = 17), 259 to 276 (segment length = 18), 278 to 295(segment length = 18), 297 to 324 (segment length =28), 326 to 379 (segment length = 54) and 381 to 420 (segment length= 40) with an average entropy of 0.2152, 0.8244, 0.9500, 0.7522, 0.6430 and 0.2481 respectively as shown in figure 2. This conservation has already been upheld by minimal entropy shown by positions 240 to 420.

Region 1: Position 241 to 257
Segment Length: 17
Average entropy (Hx): 0.2152
Consensus:

241 NGTCLLASMGLQLNITY 257

Region 2: Position 259 to 276
Segment Lengths: 18
Average entropy (Hx): 0.8244
Consensus:

259 KKDNTTVTRLLNINPNKT 276

RKDNTTVTRLLNINPNKT
KKDNKTVTRAFNISPNDT
RVDNKTVTREFNVNPNKT
KKDNTTVTRAFNINPSDK
KRDNTTVSGLFSIDPSNT
KKDEKMGLDLLNFIPHNT
TKDGKNGSYVFNIESKGV
RKDNTTVTRLLNINPNKT
RRDNTTVTSVLNINPNET
KKDNMTVTRALNISPNDT
TKDNTTVTRGLNINPNKT
KKDNTTVVKVVSINPNKT
RKDNTTVTRLLNINPNKT
NRDNMTVTRVFNINPNKT
KKDNSSARIIWNINPNKT

Region 3: Position 278 to 295
Segment Length: 18
Average entropy (Hx): 0.9500

Consensus:

278 ASGSCGAQLVTLELHSEN 295

ASGSCGAHLVTLELHSEG
SSGSCGINLVTLKVENKN
FGGNCSATLATLELHSEN
YSGTCGAQLVTLKVGNKS
ATGSCGPQLVTLDLRSSR
ASGMCESTSAFLNLAFEK
VDGNCTNTTAYLSLSTGS
ASGSCGAHLVTLELHSEG
ASGNCSSLLVTLELQSKK
ASGSCSPHVVTLTVESKN
FGGSCSAQLVTLELQGES
AGGSCGAQLVTLELRSES
ASGSCGAHLVTLELHSEG
VNGSCTAQLVTLELRDED
AGGSCSPQVAILELQTEN

Region 4: Position 297 to 324
Segment Length: 28
Average entropy (Hx): 0.7522

Consensus:

297TVLAFQFGMNASSSRFFLQGIQLNMTLP 324

TVLLFQFGMNASSSRFFLQGIQLNTILP
RALELQFGMNASSSLFFLQGVHLNMTLP
LLLALQFVMNESSSRVFLQGVQLNLTLP
RVLELQFGMNATSSLFFLQGVQLNMTLP
HSLLFQFGMDPNTSQFFLQGIQLNMTVP
TKITFHFVLNASSEKFFLQGVNVSTTLP
IDLRFNFTLNSSLEVFYLDGVSLSTGLP
TVLLFQFGMNASSSRFFLQGIQLNTTLP
MVLVLWFGMNASSSRFFLQGIQLNATFP
SILDLKFGMNGSSSLFFLQEVRLNMTLP
RLLALQFALNTSSSRVFLQGVQLNMTLP
TLLAFQFGMNASTSRFFLQGIQLNMTLP
TVLLFLFGMNASSSRFFLQGIQLNTTLP
TLLVFHFGMNATSSQFFLQEIQLDMTLP
STLAFSFGMNATTSKFFLREIRFHKFFP

Region 5: Position 326 to 379
Segment Length: 54
Average entropy (Hx): 0.6430

Consensus:

326 DARDPTFKAANSSLRALQATVGNSYKCNAEEHVRVTKAFSVNIFKVWVQAFKVE 379

DARDPAFKAANGSLRALQATVGNSYKCNAEEHVRVTKAFSVNIFKVWVQAFKVE
DALVPTFSISNHSLKALQATVGNSYKCNTEEHIFVSKMLSLNVFSVQVQAFKVD
DAKEGSFTATNSSLRALQATAGNSYKCNAEQRLRVTSSFSLNMFRVWLQAFRVD
DAIEPTFSTSNYSLKALQASVGNSYKCNSEEHIFVSKALALNVFSVQVQAFRVE
DARDPTFQADNSSLRALQATIGNSYKCNAEERVEVTEAFSVNIFKVWVQAFQVQ
EAKAPTFEASNDSMSESRATVGNSYKCSAEENFQVTDKALVNVFNVQVQAFKVD
DANDTHFEAANSSLNYMQTNVHKSFKCNSKQTLQITDPFTVNTYHLQVQAFNSD
DARDPAFKAANSSLRALQATVGNSYKCNAEEHVRVTKAFSVNIFKVWVQAFKVE
DAKEPTFRATNTSLRALQATVGHSYKCNTEEHIRVTQAFSLNIFRVWVQAFQVE
DANVSSLMASNQSLRALQATVGNSYKCNTEEHIFVTKEFSLNVFSVQVQAFKVE
DARDPSFSAANSSLRALQATAGNSYKCRSEQRLQVTEAFALNVFQVRVQAFRVD
DARDPTFKAGNNSLRALQATIGNSYKCNAGEHVQVTEAFSVNIIKVWVQAFQVQ
DARDPAFKAANGSLRALQATVGNSYKCNAEEHVRVTKAFSVNIFKVWVQAFKVE
DARDPTFKAANSSLRALQATIGNSYRCNAEERLRVTEALSVNVFRVWVQAFQVE
DAKDPAFGAVNSSLKELQATVGNSYKCNAEENVHVTDGFSVNIFRVRVQAFKVE

Region 6: Position 381 to 420
Segment Length: 40
Average entropy (Hx): 0.2481

Consensus:

381DKFGSVEECQLDENNMLIPIAVGGALAGLVLIVLIAYLIG 420

GQFGSVEECLLDENSMLIPIAVGGALAGLVLIVLIAYLVG
DRFGSVEECVQDGNNMLIPIAVGGALAGLVLIVLIAYLIG
DKFGPVEECQLDENSMLIPIAVGGALAGLVLIVLLAYLIG
DRFGSVEECVQDGNNMLIPIAVGGALAGLVLIVLIAYLIG
DKFGSVEECQLDENSMLIPIAVGGALAGLVLIVLIAYLKG
DKFGAMEECQLDENNMLIPIIVGAALAGLVLIVLIAYLIG
NTFASAVECSLDENGMLVPIVVGAALAGLVLIVLIAYLIG
GQFGSVEECLLDENNMLIPIAVGGALAGLVLIVLIAYLVG
DRFGSVEECLLDGNNMLIPIAVGGALAGLVLIVLIAYLIG
DRFGSVEECMQDGNNMLIPIAVGGALAGLVLIVLIAYLIG
DKFGPAEECQLDENSMLIPIAVGGALAGLVLVVLMAYLVG
DKFGSVEECQLDENSMLIPIAVGGALAGLVLIVLIAYLIG
GQFGSVEECVLDENNMLIPIAVGGALAGLVLIVLIAYLVG
DKFGSVEECQLDENNMMIPIAVGGALAGLVLIVLIAYLIG
DKFGSVEECLLDENNMLIPIAVGGALAGLVLIVLIAYLIG

Figure2: Six conserved domains obtained with their length.

Entropy plot

An entropy plot, measure of the lack of the information content and the amount of variability, was generated for all the aligned positions (Fig. 3). The plot shows that entropy rarely touches a scale of two, showing minimal entropy at several positions from position 245 to position 420 where entropy rarely crosses a scale of one, which is a sign of better alignment in the region. Any position before 240 doesn't show much conservedness.

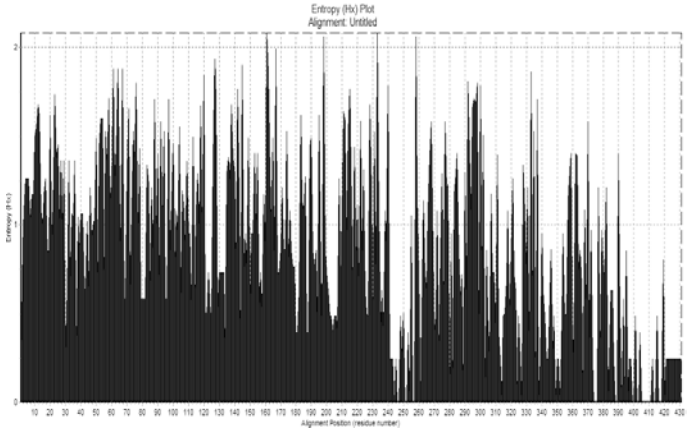


Figure3: Entropy (Hx) Plot

Hydrophobicity profile and hydrophobic moment

A hydrophobicity profile plot shows that mean hydrophobicity of the protein for most of the positions is in all the species is around zero, occassionaly it turns to be positive or negative (Fig.4). N-terminal domain and C-terminal domains are more hydrophobic. Maximum hydrophobicity is observed from positions 1 to 55 and between 390 to 420 positions in *Canis lupus familiaris*. *Macaca mulatta* and *Monodelphis domestica* exhibits high jump in hydrophobicity from position 10 to 40 and 400 to 420 respectively. These proteins are basically of non-hydrophobic in nature as most of the positions are across show a below mean hydrophobicity in the case of most of the organisms studied here.

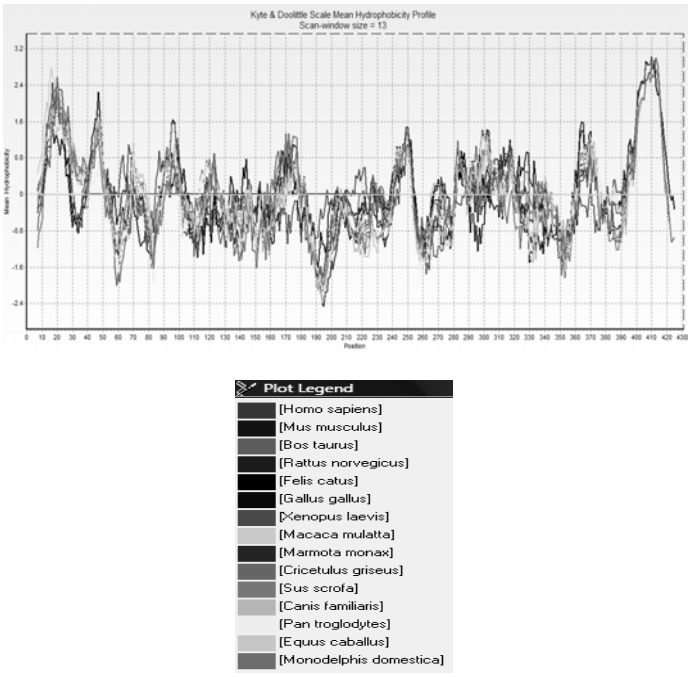


Figure4: Kyte and Dolittle scale mean hydrophobicity profile plot

Phylogeney

The phylogenetic tree constructed by using Neighbour –joining method (Fig. 5-6) shows different organisms on tree nodes branched on the basis of their LAMP-1 proteins. *Xenopus tropicalis* makes a totally diverged branch from the main tree among the selected proteins. Node for Mammalia is supported by lower bootstrap values i.e.99% while the node for primates(*Homo sapiens,Pan troglodytes,Macaca mulatta*) is supported by very high bootstrap value i.e. 100%. This tree gives an idea about the evolutionary order of LAMP-1 proteins. This phylogeny does not seem to be completely consistent with the current view of taxonomy perhaps due to use of a specific protein rather than complete genomes.

4. CONCLUSION

This study presents the first comparative genomic study and evolutionary analysis of the LAMP-1 proteins across family of organisms with special reference to mammals. The study established an overall framework of information for the family of LAMP-1 proteins, which may facilitate and stimulate the study of this gene family across all organisms.

5. FUTURE SCOPE

The evolutionary account of the Lysosomal membrane proteins will help us to reveal the some unexposed aspects of the many metabolic diseases linked to the lysosomal membrane proteins like Lysosomal storage diseases, I-cell disease etc. This work may facilitate to design the drug to over come these diseases.

6. ACKNOWLEDGEMENT

MD is thankful to DST for a Project Junior Research fellowship. The work has been supported by a DBT-BIF Grant to DKG under its BTISNet scheme and DST Project under its NanoScience and Technology (Nanomission).

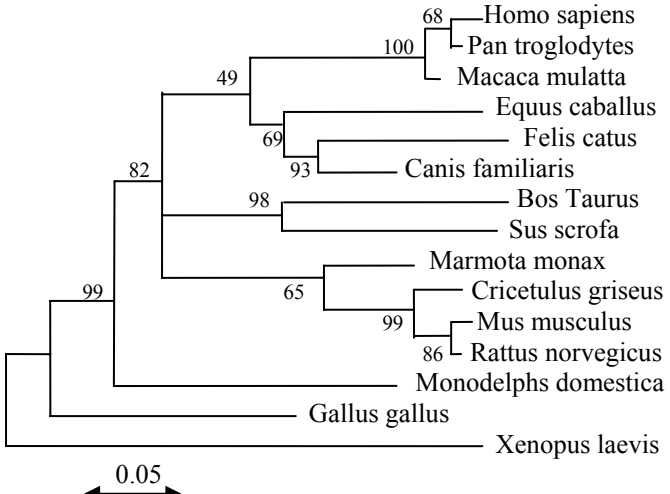


Figure5: Bootstrap consensus phylogenetic tree of LAMP-1 proteins created by Neighbour- joining method showing bootstrap support values on the nodes.

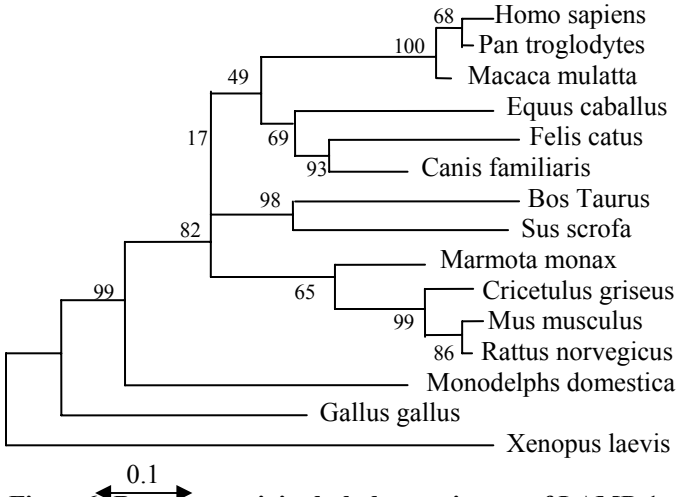


Figure6: Bootstrap original phylogenetic tree of LAMP-1 proteins created by Neighbour- joining method showing bootstrap support values on the nodes.

REFERENCES

[1]. Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman, "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. **25**:3389-3402, 1997.

[2]. Altschul S.F. and Gish G., Local alignment statistics Methods Enzymol. **266**:460-480 1996.

[3]. Bainton DF. The discovery of lysosomes. J Cell Biol ;91:66s–76s,1981.

[4]. Bernd Schroder, Christian Wrocklage, Cuiping Pan, Ralf Jager, Bernd Kosters, Helmut Schafer, Hans-Peter Elsasser,Matthias Mann and Andrej Hasilik, Integra and Associated Lysosomal Membrane Proteins, Journal compilation,1676-1686 , 2007.

[5]. Chi T. Hua, John J. Hopwood, Sven R. Carlsson, Ray J. Harris, and Peter J. Meikle, Evaluation of the lysosome-associated membrane protein LAMP-2 as a marker for Lysosomal storage disorders, Clinical Chemistry 44:10, 2094–2102, 1998.

[6]. Corinne Sagne, Cendra Agulhon, Philippe Ravassard et al Identification and characterization of Lysosomal transporter for small neural amino acids PNAS 98: 7206–721, 2001.

[7]. Eisenberg D. E. Schwarz, M. Komaromy and R.Wall. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. 179(1):125-42, 1984 .

[8]. Felsenstein J ,Confidence limits on phylogenies: An approach using the bootstrap. Evolution **39**:783-791, 1985.

[9]. Fukuda, M. N., Dell, A., and Scartezzini, P., J. Biol. Chem. 262,7195-7206, 1987 .

[10]. Hall, T.A., BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41:95-98, 1999.

[11]. Higgins D., Thompson J., Gibson T. Thompson J. D., Higgins D. G., Gibson T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673-4680, 1994.

[12]. Jonathan N. Glickman and Stuart Kornfeld: Mannose 6-Phosphate-independent Targeting of Lysosomal Enzymes in I-Cell Disease B Lymphoblasts The Journal of Cell Biology, 123, 99-108, 1993.

[13]. Kyte J and Doolittle RF: A simple method for displaying the hydropathic character of a protien. J Mol Biol 157:105, 1982.

[14]. Liu H, Sadygov RG, Yates JR III. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal. Chem;76:4193–420, 2004.

[15]. Mancini GM, Havelaar AC, Verheijen FW. Lysosomal transport disorders. J Inherit Metab Dis;23:278–292, 2000.

[16]. Richard D. Bagshaw, Don J. Mahuran, and John W. Callahan: A Proteomic Analysis of Lysosomal Integral Membrane Proteins Reveals the Diverse Composition of the Organelle, The American Society for Biochemistry and Molecular Biology, Molecular & Cellular Proteomics 4.2, 133-143, 2005.

[17]. Saitou N & Nei M (**1987**) The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution **4**:406-425, 1987.

[18]. Sven R. Carlsson, Jurgen Rothll, Friedrich Pillerz and Minoru Fukuda: Isolation and Characterization of Human Lysosomal Membrane Glycoproteins, h-lamp- 1 and h-lamp-2 ,The Journal of Biological Chemistryl. 263, 18911-18919, 1988.

[19]. Tamura K, Dudley J, Nei M & Kumar S MEGA4: Molecular Evolutionary Genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution **24**:1596-1599, 2007.

[20]. Town, M., Jean, G., Cherqui, S., Attard, M., Forestier, L., Whitmore, S. A., Callen, D. F., Gribouval, O., Broyer, M., Bates, G. P., et al. Nat. Genet. 18, 319–324, 1998.

[21]. Verheijen, F.W.,Verbeek, E.,Aula,N., Beerens, C. E.,Havelaar,A. C., Joosse, M., Peltonen, L., Aula, P., Galjaard, H., van der Spek, P. J. & Mancini, G. M. Nat. Genet. 23, 462–465, 1999.

[22]. www.ncbi.nlm.nih.gov/entrez (National Centre for Biotechnology Information).

[23]. Zuckerkandl E & Pauling L, Evolutionary divergence and convergence in proteins, pp. 97-166 in Evolving Genes and Proteins, edited by V. Bryson and H.J. Vogel. Academic Press, New York, 1965.

An Effective Technique for Data Security in Modern Cryptosystem

Dilbag Singh[1] and **Alit Singh**[2]

Abstract - Present paper provides a conceptual framework on the proposed C-QUBITS Key exchange technique, which is used as a base for the data security through quantum computing in the modern cryptosystem. In the first phase a detailed description of the BB84 Cryptographic protocol is given, which is used as a standard protocol for quantum key distribution in quantum cryptography and the emphasis is also given on the loopholes present in this protocol which makes it less effective than it pretends to be. In the next phase the focus is made on the C-QUBITS technique, which can be used for the exchange of key between the sender and the receiver. Thereafter the key is used for the encryption of the data to be transferred between the two entities. This technique makes use of the concepts of quantum physics like polarization and more importantly C-NOT gate which is mainly used in case of qubits (quantum bits) and it is more effective and secure than the BB84 protocol. In the last phase the focus is made on the information reconciliation and privacy amplification, which is used for error correction carried out between Alice and Bob's keys and for reducing a third party's partial information about the shared secret key between two parties, Alice and Bob respectively. Further the security level in the C-QUBITS technique can be increase by performing the privacy amplification that convert the realized secret key into a smaller length key through some hashing function chosen at random from a known set of hashing functions.

Index Terms - C-qubits algorithm, BB84 protocol, qubits, quantum key distribution, privacy amplification, information reconciliation and hashing function.

1. INTRODUCTION

Modern cryptosystem are specifically designed for use on computers and no longer concern with the written alphabet. The focus is on the use of binary bits. One of the main part of the modern cryptosystem is quantum cryptography. It was born in the early seventies when Stephen Wiesner wrote "Conjugate Coding", which unfortunately took more than ten years to see the light of print[1]. In the mean time, Charles H. Bennett and Gilles Brassard picked up the subject and brought it to fruition in a series of papers that culminated with the demonstration of an experimental prototype that established the technological feasibility of the concept[2]. Quantum cryptographic systems take advantage of Heisenberg's uncertainty principle, according to which measuring a quantum system in general disturbs it and yields incomplete information about its state before the measurement [3]. Eavesdropping on a quantum communication channel therefore causes an unavoidable disturbance, alerting the legitimate users. This yields a cryptographic system for the distribution of a secret random cryptographic key between two parties initially sharing no secret information that is secure against an eavesdropper having at her disposal unlimited computing power. Once this secret key is established, it can be used together with classical cryptographic techniques such as the one-time-pad to allow the parties to communicate meaningful information in absolute secrecy. Advantage of quantum cryptography over traditional key exchange methods is that the exchange of information can be shown to be secure in a very strong sense, without making assumptions about the intractability of certain mathematical problems. Even when assuming hypothetical eavesdroppers with unlimited computing power, the laws of physics guarantee (probabilistically) that the secret key exchange will be secure, given a few other assumptions [4].

[1]*Department of Computer Science & Engineering, Choudhary Devi Lal University, Sirsa,Haryana (India)*
[2]*Department of Computer Science & Engineering, BPS Mahila Vishwavidyalaya, Khanpur Kalan, Sonepat, Haryana (India)*
E-Mail: [1]*dbs_beniwal@rediffmail.com and*
[2]*ghanghas_ajit@rediffmail.com*

2. QUANTUM APPROACH

Main problem of secret-key cryptosystems is secure distribution of keys. It is here that quantum mechanics offers a solution. While the security of public key cryptographic methods can be undermined by advances in technology and mathematical algorithms, the quantum approach will provide unconditional security [12,13]**.** Within the framework of classical physics, it is impossible to reveal possible eavesdropping, because information encoded into any property of a classical object can be acquired without changing the state of the object. All classical signals can be monitored passively. In classical information, one bit of information is encoded in billions of photons, electrons, atoms, or other carriers. You can always deviate part of the signal and perform a measurement on it, whereas in quantum mechanics, any projective measurement will induce disturbances [5].

3. QUANTUM KEY DISTRIBUTION

Key distributed using quantum cryptography would be almost impossible to steal because Quantum key distribution (QKD)[5,6,7] systems continually and randomly generate new private keys that both parties share automatically

A compromised key in a QKD system can only decrypt a small amount of encoded information because the private key may be changed every second or even continuously. To build up a secret key from a stream of single photons, each photon is encoded with a bit value of 0 or 1, typically by a photon in some superposition state, such as polarization. These photons are emitted by a conventional laser as pulses of light so dim that most pulses do not emit a photon. This approach ensures

that few pulses contain more than one photon. Additional losses occur as photons travel through the fiber-optic line. In the end, only a small fraction of the received pulses actually contain a photon [10]. However, this low yield is not problematic for QKD because only photons that reach the receiver are used. The key is generally encoded in either the polarization or the relative phase of the photon.

4. CRYPTOGRAPHIC PROTOCOL BB84

The most common standard protocol for quantum key distribution is called BB84, it was invented by Charles H. Bennet and Gilles Brassard in 1984. It allows two users to establish an identical and purely random sequence of bits at two different locations while allowing revealing of any eavesdropping. BB84 uses two pairs of states, with each pair conjugate to the other pair, and the two states within a pair orthogonal to each other. Pairs of orthogonal states are referred to as a basis. The usual polarization state pairs used are either the rectilinear basis of vertical (0°) and horizontal (90°), the diagonal basis of 45° and 135° or the circular basis of left- and right-handedness. Any two of these bases are conjugate to each other, and so any two can be used in the protocol. Below the rectilinear and diagonal bases are used.

Basis	0	1
|	↑	→
×	↗	↘

The first step in BB84 is quantum transmission. Alice creates a random bit (0 or 1) and then randomly selects one of her two bases (rectilinear or diagonal in this case) to transmit it in. She then prepares a photon polarization state depending both on the bit value and basis, as shown in the table to the left. So for example a 0 is encoded in the rectilinear basis (+) as a vertical polarization state, and a 1 is encoded in the diagonal basis (x) as a 135° state. Alice then transmits a single photon in the state specified to Bob, using the quantum channel. This process is then repeated from the random bit stage, with Alice recording the state, basis and time of each photon sent [8].

In the lab experiment [9], the BB84 protocol encodes single photon polarizations using two bases of the same 2–dimensional Hilbert space:

- rectilinear basis $\{0°: |\rightarrow\rangle, 90°: |\uparrow\rangle\}$
- diagonal basis $\{45°: |\nearrow\rangle, 135°: |\nwarrow\rangle\}$

Only requirement on the involved quantum states is actually that they belong to mutually non-orthogonal bases of their Hilbert space, where each vector of one basis has equal-length projections onto all vectors of the other basis. If a measurement on a system is performed in a basis different from the one the system is prepared in, its outcome is completely random and the system looses all the memory of its previous state.

Any measurement in the diagonal basis on photons prepared in the rectilinear basis will yield random outcomes with equal probabilities and vice versa. On the other hand, measurements performed in the basis identical to the basis of preparation of states will produce deterministic results. The protocol relies on Heisenberg's uncertainty principle, which forbids the measurement of more than one polarization component of one photon.To exchange a secret key in the BB84 protocol[8], Alice and Bob must do as follow:

Alice creates a binary random number and sends it to Bob using randomly the two different bases + (rectilinear) and X (diagonal):

- $|\rightarrow\rangle$ and $|\nearrow\rangle$ both represent 1
- $|\uparrow\rangle$ and $|\nwarrow\rangle$ both represent 0

Therefore, Alice transmits photons randomly in the four polarization states

$$|\rightarrow\rangle, |\uparrow\rangle, |\nearrow\rangle, \text{and } |\nwarrow\rangle.$$

1. Bob simultaneously measures the polarization of the incoming photons using randomly the two different bases. He does not know which of his measurements are deterministic, i.e. measured in the same basis as the one used by Alice.
2. Later, Alice and Bob communicate to each other the list of the bases they used. This communication carries no information about the value of the measurement, but allows Alice and Bob to know which values were measured by Bob correctly.
3. Bob and Alice keep only those bits that were measured deterministically and will disregard those sent and measured in different bases. Statistically, their bases coincide in 50 % of all cases, and Bob's measurements agree with Alice's bits perfectly.
4. Together, they can reconstitute the random bit string created previously by Alice.

4.1 Loophole in BB84 Protocol

Now as we have given a complete description of BB84 protocol. If Eve intercepts the transferred photons two cases are possible.

CASE 1: First one is that the base used by the Alice, Bob and Eve will be same.

CASE 2: Second one is that the base used by the Alice and Bob is same but that used by the Eve is different.

As the base used by all three of them is same in **Case 1** so Eve will be able to correctly guess the value corresponding to the polarized photon. As the base used by Eve and Alice is same so after the interception of the photon, the polarization of the photon won't change so it would be impossible for the Bob to guess that interception took place.

And if suppose 40 photons are send by Alice then on an average in 20 photons (using probability) out of that, the base used by Alice and Bob will be same (Which will form the key) and out of that also in 10 photons base used by Alice, Eve and Bob will be same. So we can conclude that out of 20 photons that will form the key 10 will be known to Eve i.e. **½** of the total key.

NOTE: Here we have not considered the case where the base used by the Alice and Bob will be different as in that case photons won't be considered for being the part of the key (No matter what is the base used by the Eve).

5. PROPOSED C-QUBITS TECHNIQUE

In this the photons will be send in pairs. First photons will be passed through the C-NOT gate (as shown in Fig. -1) and

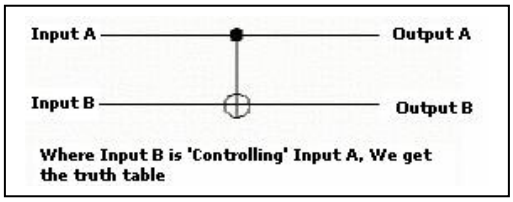


Figure 1 C-NOT gate.

then will be passed through the polarized. Before going any further we would like to explain the working of C-NOT gate.

Input		Output		Type
A (Control)	B (XOR)	A	B	
0	0	0	0	Identity
0	1	0	1	Identity
1	0	1	1	Swap
1	1	1	0	Swap

Table 1: Truth table of the C-NOT gate

The gate will take two inputs and correspondingly give two outputs. Table 1 summaries all input-output possibilities for a C-NOT gate. Output value of B depends on the value of A. If value of A is 0 then Value of B will remain as it is, and if value of A is 1 then value of B will change[12]

The diagram given in Fig. 2 shows how pair of bits is passed through the C-NOT gate and then how polarization takes place[12]. The polarization [9] takes place in the same way as in case of original BB84 algorithm.

5.1 Steps in C-QUBITS Technique

C-QUBITS technique includes the following steps and the actual data to code conversion is given in Table 2[12].

i. Alice creates a binary random number and divides them into pairs and then each pair is passed through **C-NOT gate**. Then it to Bob using randomly the two different bases + (rectilinear) and X (diagonal):

- $|\rightarrow\rangle$ and $|\nearrow\rangle$ both represent 1
- $|\uparrow\rangle$ and $|\nwarrow\rangle$ both represent 0

Therefore, Alice transmits photons randomly in the four polarization states

$$|\rightarrow\rangle, |\uparrow\rangle, |\nearrow\rangle, \text{and } |\nwarrow\rangle.$$

ii. The bases of the pair of photons can be **+X**, **++**, **X+** or **XX**.

iii. Bob simultaneously measures the polarization of the incoming pair of photons using randomly the four possible combinations i.e. **+X**, **++**, **X+** or **XX**. He does not know which of his measurements are deterministic, i.e. measured in the same pair of basis as the one used by Alice.

iv. Later, Alice and Bob communicate to each other the list of the bases they used for each pair of photons.

v. Bob and Alice keep only those pair of bits that were measured deterministically and will disregard those sent and measured in different bases. Statistically, the pair bases coincide in 25 % of all cases, and Bob's measurements agree with Alice's bits perfectly. In those cases only the output B of C-NOT gate is considered for being part of the key.

vi. Together, they can reconstitute the random bit string created previously by Alice.

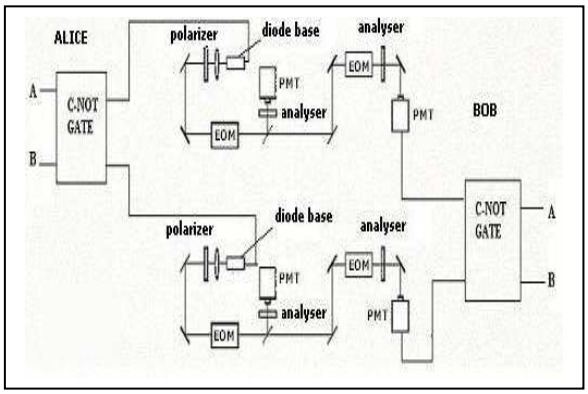


Figure 2: Diagrammatic view of the C-QUBITS technique

Alice	1	0	1	1	1	1	0	1	0	1	0	0
C-NOT gate	1	1	1	0	1	0	0	1	0	1	0	0
Random bases	X	+	X	X	+	X	+	+	+	X	X	+
Alice Polarization	↗	→	↗	↖	→	↖	↑	→	↑	↗	↖	↑
Bob's random Bases	X	+	X	+	+	X	X	+	+	+	X	+
Bob's measurement	↗	→	↗	→	→	↖	↗	→	↑	↑	↖	↑
C-NOT gate	1	0			1	1					0	0
Values Kept		√				√						√

Table 2: Shows the actual data to code conversion.

Note: The table-2 given at the end of the paper shows C-QUBITS in tabular form. The same color cells are used to indicate pairs. Here 3 bits have been deduced which will become part of the final key. This process continues until we get desired number of bits to form the complete key.

5.2 What if Eve intercepts?

Eavesdropper (usually called Eve) intercepts in between to listen to the quantum channel, she can intercept the pair of photons sent by Alice, perform measurements on them and resend them to Bob. However, as Alice alternates her encoding bases at random, Eve does not know the basis to use for her measurement; she must choose her measurement bases at random, as well. Now as there are 4 possible pairs i.e. ++, +X,

XX, X+ Eve will guess the pair correctly one out of every four times. In that case she will be able to send the pair of photons correctly to Bob. But in other 75 % of the cases, though, she measures in the wrong basis and produces errors.
Example, lets assume Alice sends a pair of 1and 0.Now when they are passed through C-NOT gate, value of 1 will remain as it is but value of 0 will change to 1 (refer the C-NOT truth table). Now suppose '**1'** is send in the rectilinear basis i.e. the state → and other 1 in diagonal base i.e. the state (We are only considering the case where Bob will also use are rectilinear base for the first photon and diagonal base for the second base) because for all the other cases the photons wont be considered for being the part of the key. Suppose Eve also measures the first photon in the rectilinear base and the second photon in the diagonal base then she will able to guess the value of that photon perfectly (but this will happen in only 25% of the cases as compared to 50% cases in case of BB84 algorithm). In rest of the 75% cases when Eve will make a mistake in choosing one or both of the random bases, then no matter which polarization Eve detects and re-sends she won't be having any idea of the value of the photons used. **(**Eve won't be able to guess the remaining bits as we are considering only output of B (of the C-NOT gate) for the key, as Eve does not know the value of A on which final value of B depends)[12].

6. COMPARISON OF C-QUBITS TECHNIQUES WITH BB84 PROTOCOL

6.1 BB84 Protocol:

Suppose we have 640 bits (we have taken a large value so that on repeated division we don't get a decimal value). Now if we apply probability then in 320 bits rectilinear base will be applied by Alice and in other 320 bits diagonal base will be applied. Again on an average if Eve intercepts the photons then in ½ of the times the random base used by her, will be same as used by the Alice i.e. again for 320 bits. Now as Bob will also be using the same random bases as by Alice half of the times (on an average) and in 160 of that bits the case will be such that Base used by the Alice, Eve and Bob will be same. So this indicates that out of the 320 bits key that will be generated 160 bits will be known to Eve (although guessing of the remaining 160 bits will be very difficult, which itself explains the power of quantum cryptography).
So the **conclusion** is that:
Total Bits used = 640
No of bits used for the formation of Key =320
No. of bits that Eve could guess =160
i.e. Eve knows half of the key

6.2 C-QUBITS Technique

Suppose we have 640 bits i.e. 320 pair of bits. Here if Eve intercepts the photons then in 1/4 of the times the random base used by her, will be same as used by the Alice because there are 4 possible combinations i.e. ++, +X, XX, X+ i.e. on average in 80 pairs she will guess correctly. Now as Bob will also be using the same random bases for the pair of photons as by Alice 1/4 of the times (on an average) i.e. again 80 pairs and 20 of that pairs the case will be such that the pair of Base used by the Alice, Eve and Bob will be same. So this indicates that only 80 pairs will be considered for the key and of that only the value of B will be considered so out of 640 bits used we will get a key of only 80 bits and out of that only 20 bits will be known to the Eve.
So the **conclusion** is that:
Total Bits used=640 bits or 320 pairs
No of bits used for the formation of Key=80
No of Bits that Eve could guess=20
i.e. Eve will be able to guess only ¼ of the key (Use of C-not gate will make it impossible for Eve to guess the remaining Key)

7. INFORMATION RECONCILIATION

Information reconciliation is a form of error correction carried out between Alice and Bob's keys, in order to ensure both keys are identical. It is conducted over the public channel and as such it is vital to minimize the information sent about each key, as this can be read by Eve. A common protocol used for information reconciliation is the cascade protocol, proposed in 1994. This operates in several rounds, with both keys divided into blocks in each round and the parity of those blocks compared. If a difference in parity is found then a binary search is performed to find and correct the error. If an error is found in a block from a previous round that had correct parity then another error must be contained in that block; this error is found and corrected as before. This process is repeated recursively, which is the source of the cascade name. After all blocks have been compared, Alice and Bob both reorder their keys in the same random way, and a new round begins. At the end of multiple rounds Alice and Bob will have identical keys with high probability, however Eve will have gained additional information about the key from the parity information exchanged [15].

8. PRIVACY AMPLIFICATION

Further to increase the security, Privacy Amplification is performed. **It** is a method for reducing (and effectively eliminating) Eve's partial information about Alice and Bob's key. This partial information could have been gained both by eavesdropping on the quantum channel during key transmission (thus introducing detectable errors), and on the public channel during information reconciliation (where it is assumed Eve gains all possible parity information). Privacy amplification uses Alice and Bob's key to produce a new, shorter key, in such a way that Eve has only negligible information about the new key. This can be done using a hash function,(A hash function is a function from a set of possible inputs, *U*, to a set of outputs, which is usually taken to be $\{1, \ldots, N\}$ for some *N*.) chosen at random from a publicly known set of such functions[11], which takes as its input a binary string of length equal to the key and outputs a binary string of a chosen shorter length. The amount by which this new key is shortened is calculated, based on how much information Eve could have gained about the old key (which is

known due to the errors this would introduce), in order to reduce the probability of Eve having any knowledge of the new key to a very low value[14,17].

9. HOW DOES PRIVACY AMPLIFICATION WORK

In Quantum Key Distribution, to arbitrarily limit the amount of partial information that an eavesdropper can know about a quantum distributed key, the sender and receiver can use privacy amplification. This uses a set of universal hash functions chosen at random to compress both the key size and Eve's knowledge accordingly.

The hash algorithm which defines the family of universal hashes is in the clear. Like if

$$h(x)=(a\_1.x\_1+a\_2.x\_2+a\_3.x\_3+a\_4.x\_4)$$

for an N bit key divided into 4 chunks x_i . The values a_i are randomly generated, and it's this which we don't get how it's transmitted between Alice and Bob and hash function can be publicly communicated. For example, let's say Eve (eavesdropper) knows 1/3 of the key. Alice (sender) and Bob (recipient) can publicly agree to break the key into 3 bit chunks and perform a parity operation on those, to make a key one third the length of the original. Since, at this point, Alice and Bob's keys agree completely, the reduced key will also agree completely without any need for communicating the results of the parity operations. Eve can know that they are performing this hash function, but since she only has 1/3 of the key, she can not perform the hash function on her partial key to get the official reduced key. So the hash function can be made completely public. Of course, Eve could know three consecutive bits, which would allow her to perform the hash function on those to get a bit from the reduced key. So the hash function needs to be chosen intelligently, based on the estimated knowledge of Eve. So, instead, if the hash function took 10 bit blocks for parity check, then Eve's expected knowledge goes down even farther.

To estimate Eve's knowledge, Alice and Bob will look at the error rate in the keys (using information reconciliation). Errors can be caused either by Eve's measurements or by noise. Alice and Bob will attribute all errors to Eve, to be safe. This gives them an idea of how much reduction their hash function must do. When we say the hash function is randomly chosen, the term random just means it is not chosen before hand. If Eve knew that the hash function would use three bits in a row, she could optimize her measurement to be more likely to give three bits in a row. But if she doesn't know how the bits will be grouped for the parity check (they need not be in a row), or if something other than parity will be used, she will have no way to optimize her measurement for the eventual hash function that is used. So 'random', in this case, just means 'decided after the key is sent.

Finally, it might wonder if Eve could do a man-in-the-middle attack, where she intercepts the discussion about the hash function and makes Bob think Alice is using a different has function. She can do this, for sure, but it will not result in Eve learning about the message. It will only keep Alice from communicating her message to Bob [16].

10. PROSPECTS

The current commercial systems are aimed mainly at governments and corporations with high security requirements. Key distribution by courier is typically used in such cases, where traditional key distribution schemes are not believed to offer enough guarantee. This has the advantage of not being intrinsically distance limited, and despite long travel times the transfer rate can be high due to the availability of large capacity portable storage devices. The major difference of quantum cryptography is the ability to detect any interception of the key, whereas with courier the key security cannot be proven or tested. QKD (Quantum Key Distribution) systems also have the advantage of being automatic, with greater reliability and lower operating costs than a secure human courier network.

Factors preventing wide adoption of quantum cryptography outside high security areas include the cost of equipment, and the lack of a demonstrated threat to existing key exchange protocols. However, with optic fibre networks already present in many countries the infrastructure is in place for a more widespread use [17].

11. CONCLUSION

The C-QUBITS technique can be used as a powerful tool for combating the problems of data security and provide more security that previously used BB84 protocol. Further this technique can be make more effective by using the concept of information reconciliation and privacy amplification.

REFERENCES

[1]. W. Die, M.E. Hellman, New Directions in Cryptography, IEEE Transactions on Information Theory 22, 644-654, 1979.

[2]. C. H. Bennett and G. Brassard, "Quantum cryptography: public key distribution and coin tossing," Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing, Bangalore, India, pp. 175-179, 1984.

[3]. C. H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, "Experimental quantum cryptography," J. Cryptol. 5, pp3–28, 1992.

[4]. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," Rev. Mod. Phys.74, pp145–195, 2002.

[5]. G. Brassard and L. Salvail "Secret key reconciliation by public discussion" Advances in Cryptology: Eurocrypt 93 Proc. pp 410-23, 1993.

[6]. C. Gobby, Z. L. Yuan and A. J. Shields, "Quantum key distribution over 122 km telecom fiber," Appl. Phys. Lett. 84, pp 3762–3764 2002.

[7]. D. Gottesman, H. K. Lo, N. L¨utkenhaus and J. Preskill, "Security of quantum key distribution with imperfect devices," Quant. Inf. Comput. 4,pp 325–360 ,2004.

[8]. Frederick Henle, BB84 online demo. <http://monet.mercersburg.edu/henle/bb84/>. An online

demonstration of the original BB84 algorithm from, Bennett et al. 1991.
[9]. Matthias Scholz, Quantum Key Distribution via BB84 an Advanced Lab Experiment, (Oct. 14th 2004).
[10]. K. J. Gordon, V. Fernandez, G. S. Buller, I. Rech, S. D. Cova and P. D. Townsend, "Quantum key distribution system clocked at 2 GHz," Opt. Express 13, pp 3015–3020 2005.
[11].J. F. Dynes, Z. L. Yuan, A. W. Sharpe and A. J. Shields, "A high speed, post-processing free, quantum random number generator," Appl. Phys. Lett. 93, 031109 ,2008.
[12].A.Singh, An efficient quantum cryptography's algorithm for data security, Indian Journal of Engineering & Materials Sciences, Vol. 14, pp. 352-357, October 2007.
[13].Z. L. Yuan, A.W. Sharpe and A. J. Shields, "Unconditionally secure quantum key distribution using decoy pulses," Appl. Phys. Lett. 90, 011118, 2007.
[14].C. H. Bennett, G. Brassard, and J. M. Robert. Privacy amplification by public discussion. SIAM Journal on Computing, 17(2): 210-229, 1988.
[15].G. Brassard and L. Salvail "Secret key reconciliation by public discussion" Advances in Cryptology: Eurocrypt 93 Proc. pp 410-23,1993.
[16].M. N. Wegman and J. L. Carter, "New hash functions and their use in authentication and set equality, Journal of Computer and System Sciences", 22, pp 265-279, 1981.
[17].http://en.wikipedia.org/wiki/Quantum_cryptography/Accessed on 25th October, 2009.

Revival of Tutor Model: A Domain Independent Intelligent Tutoring System (ITS)

Abrar S. Alvi[1] and **M. S. Ali**[2]

Abstract - Ever since the birth of ITS (Intelligent Tutoring System) in late 1970's and early 1980's, it has live through numerous evolutions. Through this paper we opt to explore and revive the Tutor (Pedagogical) Model of ITS. Pedagogy is referred to as the correct use of teaching strategies. The objective is to emphasize on various abstract pedagogies that can be incorporated into an Intelligent Tutoring Systems irrespective of the knowledge domain being taught. Hence the current research and development is to make sculpt that is used to provide tutoring of almost any knowledge domain. Despite of the outgrowth of technology (Computer) based learning, traditional learning/ teaching is still a dominant and a preferred choice not only by the students but also by the teachers. The reasons being reluctance to give up conventional practice, habit of face-to-face teaching and withhold benefit of one to one tutoring. So instead of trying to take over the traditional pattern, we come up with an ITS that bestow each student with a familiar one-to-one tutoring practice, flexibility to learn almost any knowledge domain using same gadget and to learn using contented pedagogies until he masters and governs that knowledge domain.

Index Terms - Knowledge Base, Knowledge Representation, Intelligent Tutoring System, Pedagogy.

1. INTRODUCTION

From the dawn of Internet, technology (Computer) based learning has outgrown to its fullest. [2, 6]. Computer-based learning conjoined with Internet spans a wide spectrum of learning methodologies viz e-learning, distance learning, web based learning, collaborative learning and many more to name. Obvious merits offered by them are self-paced learning of anyone, at anytime and from anywhere. And of course the obvious demerits are feeble urge because of absence of teacher, short in instantaneous communication and frail hold of the learning environment. [3] To overcome the drawbacks of this first generation learning environment a sunrise of second generation learning systems in the form of adaptive educational systems came into reality [3]. Adaptive hypermedia/multimedia system and Intelligent Tutoring System forms the basis of second generation learning environment. To be specific there are systems that are i) Intelligent but not adaptive ii) Adaptive but not really intelligent iii) Adaptive and intelligent [5], see Figure 8. The evident distinctiveness of ITS which facilitated us are one to one teaching with learner being entirely regulated by computer couching/tutoring and triggering students to learn by getting involved in every step of tutoring.

[1]*Asstistant Professor, Department of Information Technology, PRM Institute of Technology & Research, Badnera.*
[2]*Principal, PRM College of Engineering & Management, Badnera.*
E-Mail: [1]abrar_alvi@rediffmail.com, [2]softalis@hotmail.com

Furthermore, Intelligent Tutoring Systems accumulate three basic kinds of knowledge [1]: Domain knowledge (Domain Module), Knowledge about learners (Student Module), and Pedagogical knowledge (Pedagogical Module). [4].

2. BUILDING GENERIC TUTOR MODEL

As the name implies the model will provide tutoring for almost any knowledge domain. We are testing the model for some of the knowledge domain of computer technology domain.

2.1 Domain Model (Knowledge Base Building)

First and foremost step toward initiating the tutor model is working on domain model i.e. knowledge base building or creating learning objects The process of building knowledge has to be done for every domain that you wish to teach. Once the knowledge base (learning objects) is build, it needs to be copied into the model using the administrative interface provided (Figure 2 & Figure 3).

2.2 Knowledge Structuring

The spirit of ITS is Domain module (KB building) and its spirit in turn is Knowledge structuring, done by using a standard way for knowledge representation. Building a knowledge base concentrates on selecting a proper representation for that knowledge as it will be eventually made available to the learner in their learning process. For every knowledge domain we have chosen four level of structuring. (Figure 1). The entire independent knowledge domain will be collected into a common folder (KB). Each knowledge domain is a collection of chapters and each chapter in turn is a collection of topics and every topic is taught using four different pedagogies (Knowledge representation).

3. PEDAGOGICAL MODELING

Pedagogy is the art or science of being a teacher. There exist numerous pedagogies. The one, which we are using, are elaborated in Table 1 and Figure 1.

Pedagogy generally refers to strategies of instruction or a style of instruction. Pedagogy is also sometimes referred to as the correct use of teaching strategies. The main objective is to provide each student with a study experience similar to ideal one-to-one tutoring. It is shown in many expert studies that one of the most effective ways for student to learn is to work with an expert tutor in and individual way [4].

Pedagogical model contains the knowledge of how to teach i.e. teaching or tutoring strategy. It coordinates the whole tutoring process and deals with issues like which topic is to present, when to present a new topic.

Level No.	Pedagogy	Description
1	Theory	Simple Theoretical Description

Level No.	Pedagogy	Description
2	Media Files	Audio and Video tutorials
3	Example	Elaborative Examples
4	Practice	Worksheets

Table 1: Pedagogies used

3.1 Making Learner Meticulous (Examination Process)

Our focus is not mere tutoring the learner but to tutor until the learner actually becomes thorough in the knowledge domain opted by him for learning. To achieve it , we fire the learner with an examination that tests his understanding of the conception, his misconceptions and accordingly decides whether he needs revising the same topic with different pedagogy or can move forward to next topic toward completion of the knowledge domain.

If the learner fails to understand a particular topic by one of the pedagogy, the same topic is taught by more interactive pedagogy, providing a higher probability to be meticulous in that topic and eventually this is done in knowledge domain. The knowledge base already contains number of pedagogy for same topic. The pedagogy should be taken from knowledge base.

3.2 Pedagogy Selection (Result Analysis)

This is the most important phase, as it is the phase that decides whether the student will proceed to next topic or whether he needs to learn the same topic again. If the result of a particular learner is positive (range of percentage will be predefined) then system will continue teaching him new topic with same pedagogy but if result is average or undesired then system will teach the same topic by applying next level of pedagogy and leaner will keep learning by second level. Same process will be applied until completing entire lessons for the selected domain.

The Level-1 pedagogy is set as default for the learner i.e. learner always initiated with theoretical description. If the learner understands the topic, which means clear/pass, the exam appeared for the same topic. The same pedagogy is used to teach next topic but if he disqualifies then level-2 pedagogy is used to teach the same topic.

4. SAMPLE SCREENSHOTS

We have included some of the experimental results showing model working.

4.1. Domain Management

At the click of browse we upload the files with different pedagogies of that topic and at the press of update those files gets copied into the model which eventually are retrieved for tutoring the learner. The button 'Manage Topic Exam Question' lets the administrator build the exam set for that topic which will be used to test the learner after his study of that topic. (KB)

4.2 The User Interface

The model has two users:

i) Admin/Teacher who injects the knowledge domain and
ii) Student/Learner who acquires the tutorials of knowledge domain available in the system.

4.3 Result Analysis

Refer Figure (6, 7).

4.4 Learner's Progress

The learner has completed the first topic using first level of pedagogy and is yet to finish the remaining topics, all the topics are shown proposed with first pedagogy as we have set level 1 as default pedagogy for tutoring.

5. CONCLUSION

This model is sure going to prove pioneering for teachers, who can use it for teaching any knowledge domain. And to make this model newfound for the learner as well, we have experimented teaching the learner using multiple pedagogies and providing tutoring on one to one basis. Our model makes decision and selection of the pedagogies that would best help the learner to understand and govern the domain he has opted for learning.

REFERENCES

[1]. Alla Anohina, Advances in Intelligent Tutoring Systems: Problem-solving Modes and Model of Hints, International Journal of Computers, Comm. & Control Vol. II (2007), No. 1, pp. 48-55.

[2]. W. Fajardo Contreras, E. Gibaja Galindo, E. Marín Caballero and G. Marin Caballero, An Intelligent Tutoring System for a Virtual E-learning Center, Current Developments in Technology-Assisted Education (2006).

[3].] Syed S. Ali and Susan McRoy Reva Freedman, What is an Intelligent Tutoring System? Published in Intelligence 11(3): 15–16 (2000).

[4]. Fernando Salgueiro, Guido Costa, Zulma Cataldi, Fernando Lage, Ramon Garcia-Martinez, Redefinition of basic modules of an Intelligent Tutoring System, The Tutor Module.

[5]. Kinshuk, Does intelligent tutoring have future! , International Conference on Computers in Education (ICCE'02).

[6]. Selwyn Piramuthu, Knowledge-Based Web-Enabled Agents and Intelligent Tutoring Systems.

Continued on Page No. 201

Computational Modeling of Cell Survival Using VHDL

Shruti Jain[1], Pradeep K. Naik[2] and Sunil V. Bhooshan[3]

Abstract - ***The model for cell survival has been implemented using Very High Speed Integrated Circuit Hardware Description Language (VHDL) (Xilinx Tool) taking three input signals: Tumor necrosis factor-α (TNF), Epidermal growth factor (EGF) and Insulin. Cell survival has been regulated by the interaction of five proteins viz P13K, TNFR1, EGFR, IRS and IKK in a network. In the absence of any one, in protein network leads to cell death. For the EGF input signal the proteins like MEK, ERK, AkT, Rac & JNK have been important for regulation of cell survival. Similarly for TNF and Insulin input signal proteins like NFκB, AkT, XIAP, JNK, MAP3K & MK2 and MEK, ERK, AkT, Rac, mTOR & JNK respectively have been important for regulation of cell survival.***

Index Terms - Tumor necrosis factor-α, Epidermal growth factor, Insulin, AkT (PKB).

ABBREVIATIONS

EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; ERK, extracellular-regulated kinase; FADD, Fas-Associated protein with Death Domain; FKHR, Forkhead transcription factor; Grb2, growth factor receptor-bound 2; IGF, insulin-like growth factor; IκB, I Kappa B (nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor); IKK, IκB kinase; IR, insulin receptor; IRS1, insulin receptor substrate 1; JNK1, c-jun NH2 terminal kinase 1; MAP kinases, mitogen-activated protein kinases; MEK, mitogen-activated protein kinase and extracellular-regulated kinase kinase; MK2, mitogen-activated protein kinase-activated protein kinase 2; mTOR, mammalian target of rapamycin; NF-κB, nuclear factor-κB; PI3K, phosphatidylinositol 3-kinase; PKB, Protein Kinase B; p38, P38 mitogen-activated protein kinases; ptEGFR, phospho-to-total EGFR; ptAkT, phospho-to-total AkT; Rac, Ras-related C3 botulinum toxin substrate; SOS, Son of Sevenless; TNF, tumor necrosis factor; TNFR1, tumor necrosis factor receptor 1; TRADD, Tumor necrosis factor receptor associated via death domain; TRAF2, TNF receptor associated factor 2, VHDL, Very High Speed Integrated Circuit Hardware Description Language; XIAP, X-linked Inhibitor of Apoptosis Protein;.

I. INTRODUCTION

Cell signaling pathways interact with one another to form networks. Such networks are complex in their organization and exhibit emergent properties such as bistability and ultra sensitivity [1]. Analysis of signaling networks requires a combination of experimental and theoretical approaches including the development and analysis of models. This work examines signaling networks that control the survival decision treated with combinations of three primary signals [2, 3]; the prodeath cytokine, tumor necrosis factor-α (TNF), and the prosurvival growth factors, epidermal growth factor (EGF) and insulin. TNF induce apoptosis and survival [1, 2, 4], although receptor ligation is rarely enough on its own to initiate apoptosis as is the case with Fas ligand binding. Binding of TNF alpha to TNFR1 [5, 6] results in receptor trimerisation and clustering of intracellular death domains. This allows binding of an intracellular adapter molecule called TNFR-associated death domain (TRADD) via interactions between death domains. TRADD has the ability to recruit a number of different proteins to the activated receptor. Recruitment of TNF-associated factor 2 (TRAF2) can lead to activation of NF-kB and the JNK pathway [6, 7]. EGF is a growth factor that plays an important role in the regulation of cell growth, proliferation, and differentiation. It also increases cancer risk. EGF acts by binding with high affinity to epidermal growth factor receptor (EGFR) on the cell surface and stimulating the intrinsic protein-tyrosine kinase activity of the receptor [8, 9]. Activation of the EGF receptor tyrosine kinase (EGFR) [9, 10, 11] occurs through receptor dimerization, conformational change, and autophosphorylation. Phosphorylated receptors recruit adaptor proteins, and these then activate multiple signaling proteins including extracellular-regulated kinase (ERK) via Ras [12] and the Akt [13] kinase via phosphatidylinositol 3- kinase (PI3K). In general insulin being a peptide hormone plays important role in both metabolic and mitogenic (growth promoting) pathways [14, 15]. Signaling through the insulin pathway is critical for the regulation of intracellular and blood glucose levels and the avoidance of diabetes. Insulin binds to its receptor leading to the auto phosphorylation of the ß-subunits and the tyrosine phosphorylation of insulin receptor substrates (IRS). The role of insulin in cell survival has been proved by Insulin/Glucose Phospho-Antibody Array. It induces cell survival by interacting with Insulin receptors or the cell membrane through several cell signaling pathways such as PI3K/ AkT and MAP kinase [16, 17]. The binding of insulin [18, 19] to the insulin receptor also activates ERK and Akt, but in contrast to EGFR, the insulin receptor is constitutively dimerized, and most insulin induced signaling involves modification of insulin receptor substrate 1 (IRS1) [19, 20], a multidomain adaptor protein.

Regulation of cell survival and cell death are very complicated physiological processes involving a large number of proteins which interact in protein networks. The induction of specific network is executed by input signals like EGF, Insulin and TNF. Therefore, it is very difficult to define and measure the protein- protein interactions in a network leading to cell

[1, 3] Department of Electronics and Communication Engineering
[2]Department of Biotechnology & Bioinformatics, Jaypee University of Information Technology,
Solan-173215, India,
E-Mail: [1]jain.shruti15@gmail.com

survival and cell death experimentally. However, the computational model is useful as a means to assemble and test what we know about proteins networks regulating a physiological process. In this study we have implemented the system model of cell survival considering three input signals (TNF, EGF and Insulin) and 16 proteins using VHDL simulation.

2. MATERIALS AND METHODS

2.1 Experimental

The experimental observation of cell survival from cells treated with ten cytokine combinations of tumor necrosis factor-α (TNF), a pro apoptotic cytokine, in combination with epidermal growth factor (EGF) or insulin, two pro survival growth factors has been worked out by Gaudet et al [3]. They have predicted with the response of cell survival as well as cell death with 94 % accuracy by including eleven different proteins such as MK2, JNK, FKHR, MEK, ERK, IRS, AkT, IKK, pAkT, ptAkT and ptEGFR. All the eleven proteins forms signaling network as represented in Figure1 leads to cell survival. The response of signaling network is regulated by the concentration of cytokines like TNF, EGF and Insulin. Therefore, it is possible to built self consistent compendia cell signaling data based on the above eleven proteins that can be simulated computationally to yield important insites into the control of cell survival.

2.2 Computational Model

The prediction model for cell survival has been implemented using VHDL programming language. We have implemented the signaling system heading by three input signals such as TNF, EGF and Insulin (inputs are same as that of experimental). The block diagram of the signaling system that was modeled is shown in Figure 1.

TRADD recruits TRAF2 and RIP. TRAF2 in turn recruits the multicomponent protein kinase IKK, enabling the serine-threonine kinase RIP to activate it. An inhibitory protein, IκBα, that normally binds to NF-κB and inhibits its translocation, is phosphorylated by IKK and subsequently degraded, releasing NF-κB. NF-κB is a heterodimeric transcription factor that translocates to the nucleus and mediates the transcription of a vast array of proteins involved in cell survival and proliferation, inflammatory response, and anti-apoptotic factors. NF-κB induces the caspase inhibitors IAP1 and IAP2 and pro-survival Bcl-2 family members. Of the major MAPK cascades, TNF induces a strong activation of the stress-related JNK group, evokes moderate response of the p38-MAPK, and minimal activation of the classical ERKs. A general activation scheme involves the activation of receptor tyrosine kinases by growth factors, such as EGF [2, 3], which provides the binding site of the adapter protein Grb2 [19] that in turn localizes Sos to the plasma membrane. Sos activates Ras by exchange of GTP for GDP. The Ras-GTP binds directly to a serine-threonine kinase Raf [17, 18], forming a transient membrane-anchoring signal. Active Raf kinase phosphorylates a dual specificity kinase, MEK, [15, 16] and activates it. MEK can also be phosphorylated by Mos, a protein kinase expressed during meiotic maturation of oocytes and by MEKK1. The mono phosphorylated ERK then rebinds to an active MEK1 for dual phosphorylation and complete activation. The activated MEK phosphorylates ERK1/ERK2. Within the cell, at any time, one may find three low active forms of ERKs: one unphosphorylated enzyme, and two singly phosphorylated forms that contain phosphate either at the tyrosine or threonine residue. Activation of Akt involves growth factor binding to a receptor tyrosine kinase and activation of PI 3-K, which phosphorylates membrane bound PIP2 to generate PIP3. The binding of PIP3 to the PH domain anchors Akt to the plasma membrane and allows its phosphorylation and activation by PDK1. Akt is fully activated following its phosphorylation at two regulatory residues, a threonine residue on the kinase domain and a serine residue on the hydrophobic motif, which are structurally and functionally conserved within the AGC kinase family. Phosphorylation of a threonine residue on the kinase domain, catalyzed by PDK1, is essential for Akt activation. It causes a charge-induced conformational change, allowing substrate binding and increased rate of catalysis. Akt activity is augmented about 10-fold by phosphorylation at the serine residue by PDK2. The activity of Akt is negatively regulated by PTEN and SHIP. The activation of these MAP kinases is mediated by Rac and cdc42, two small G-proteins. The activated cdc42 binds to PAK65 protein kinase and activates it. The activated PAK65 can activate MEKK, which in turn phosphorylates SEK/JNKK and activates it. The active SEK/JNKK phosphorylates JNK/SAPK (at the TPY motif) that in turn binds to the N-terminal region of c-Jun and phosphorylates it. JNKK, an activator of JNK/SAPK, is reported to activate p38, whereas MKK3 activates only p38 and not JNK/SAPK. MEKK1 that stimulates SEK/JNKK1 in the JNK/SAPK cascade has only a trivial effect on p38 activation. In the upstream signaling, Sos stimulates only the ERK pathways without affecting either JNK or p38 cascade. Insulin is the major hormone controlling critical energy functions such as glucose and lipid metabolism. Insulin activates the insulin receptor tyrosine kinase (IR), which phosphorylates and recruits different substrate adaptors such as the IRS family of proteins. Tyrosine phosphorylated IRS then displays binding sites for numerous signaling partners. Among them, PI3K has a major role in insulin function, mainly via the activation of the Akt/PKB. Activated Akt induces glycogen synthesis through inhibition of GSK-3; protein synthesis via mTOR and downstream elements; and cell survival through inhibition of several pro-apoptotic agents, including Bad, Forkhead family transcription factors and GSK-3. Insulin signaling also has growth and mitogenic effects, which are mostly mediated by the Akt cascade as well as by activation of the Ras/MAPK pathway. Cell cycle arrest by the mammalian Target of Rapamycin (m TOR) complex requires the presence of the intact kinase domain of mTOR and, in particular, a conserved serine within this domain, which has been identified as an Akt/PKB-mediated phosphorylation site. Biomarkers indicate that the mTOR pathway is hyperactive in certain types of

cancers, suggesting that mTOR could be an attractive target for cancer therapy. Activated mTOR may provide tumor cells with a growth advantage by promoting protein synthesis, which is the best-described physiological function of mTOR signaling. mTOR regulates Akt activity, a crucial downstream effector in the PI-3K–PTEN pathway, which controls cell proliferation and survival.

3. RESULTS AND DISCUSSIONS

On the basis of block diagram (Figure 1) we have made truth tables of every possible path for cell survival based on individual inputs i.e. TNF, EGF and Insulin. Than we realized the truth tables by Karnaugh Map (K-Map) and got the expression for each input and its individual possible paths. With the help of VHDL tool, we simulated the results of each path and then all the results were combined and got final result of TNF, EGF and Insulin for its cell survival (as shown in Figure 2, Figure 3 and Figure 4). For cell survival the five different proteins i.e. P13K, TNFR1, EGFR, IRS and IKK should present. If any one of them is absent, than cell will die. Figure 2 g, h, i, j and k shows the output signal considering b, c, d, e and f as possible paths taken TNF as input. Path b shows that it consists of three proteins, and for cell survival combination is 101. 101 means first and third protein is present and second protein is absent. Similarly, for path c, d, e and f combinations are 011, 11, 01 and 0 respectively. If any of the path is present than it will lead to cell survival. 'a' shows the five proteins which are compulsory proteins i.e. P13K, TNFR1, EGFR, IRS and IKK. Similarly for Figure 3 f, g, and h are output signals for b, c and d paths. For path b, c and d possible combinations are 01, 11 and 11, respectively. Similarly for Figure 4 f, g, h and i are output signals for b, c, d and e paths. For path b, c, d and e possible combinations are 101, 111, 11 and 101 respectively.

4. CONCLUSION

We have demonstrated that the VHDL programming language can be applied to predict the cell survival with a high level of accuracy using three inputs such as TNF, EGF and Insulin. The computational model has reproduced experimental data with fairly accurate. Understanding the nature of signaling networks that control the cell survival is very significant and theoretical calculations, in particular the simulation process developed using VHDL, seen to be a proper tool for gaining such understanding. The results obtain will give information on how the input signals inducing cell survival should be modulated to achieve desire outputs and thus helps the experimentalists to design proposals regarding possible improvements to cell survival/ cell death.

REFERENCES

[1]. Weixin Zhou 2006 Stat3 in EGF receptor mediated fibroblast and human prostate cancer cell migration, invasion and apoptosis, PhD thesis, University of Pittsburgh.

[2]. Janes Kevin A, Albeck John G, Gaudet Suzanne, Sorger Peter K, Lauffenburger Douglas A, Yaffe Michael B. Dec.9, 2005 A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis; Science 310, 1646-1653.

[3]. Gaudet Suzanne, Janes Kevin A., Albeck John G., Pace Emily A., Lauffenburger Douglas A, and Sorger Peter K. July 18, 2005 A compendium of signals and responses trigerred by prodeath and prosurvival cytokines Manuscript M500158-MCP200.

[4]. Katrien Vermeulen , Dirk R. Van Bockstaele , Zwi N. Berneman 2005 Apoptosis: mechanisms and relevance in cancer; Ann Hematol 84 627-639.

[5]. Nand K Sah, Tarvinder K Taneja and Seyed E Hasnain 2000 Mitochondria can power cells to life and death ; Resonance 74-84.

[6]. Lu C.X., Fan T.J., Hu G.B.and. Cong R.S 2003 Apoptosis-inducing factor and apoptosis Acta Biochim Biophys Sin 35 881−885

[7]. Fan TJ, Han LH, Cong RS, Liang J. 2005 Caspase Family Proteases and Apoptosis Acta Biochimica et Biophysics Sinica 37 (11): 719-727.

[8]. Libermann TA , Razon TA., Bartal AD, Yarden Y., Schlessinger J and Soreq H 1984 Expression of epidermal growth factor receptors in human brain tumors Cancer Res. 44,753-760.

[9]. Arteaga C. 2003 Targeting HER1/EGFR: a molecular approach to cancer therapy Semin Oncol 30, 314.

[10]. Wells A. 1999 Moleculars in focus EGFR receptor Int. J. Biochem. Cell Biol. 31 637-643.

[11]. Hui-Wen Lo, "EGFR signaling pathway in breast cancers: from traditional signal transduction to direct nuclear translocalization" Springer (2005), 95: 211-218.

[12]. Hallberg B, Rayter SI and Downward J 1994 Interaction of Ras and Raf in intact mammalian cells upon extracellular stimulation J Biol Chem. 269(6), 3913-3916.

[13]. Dohoon Kim and Jongkyeong Chung January 2002 Akt: Versatile Mediator of Cell Survival and Beyond Journal of Biochemistry and Molecular Biology 35 No. 1 106-115.

[14]. Morris F. White 1997 The insulin signaling system and the IRS proteins Diabetologia 40, S2–S17

[15]. D T Dudley, L Pang, S J Decker, A J Bridges, and A R Saltiel 1995 A synthetic inhibitor of the mitogen-activated protein kinase cascade Proc Natl Acad Sci U S A. 92(17) 7686–7689.

[16]. Eyers, PA; Craxton, M; Morrice, N; Cohen, P; Goedert, M. 1998 Conversion of SB 203580-insensitive MAP kinase family members to drug-sensitive forms by a single amino-acid substitution Chem Bio.; 5(6) 321–328.

[17]. T Jelinek, A D Catling, C W Reuter, S A Moodie, A Wolfman and M J Weber 1994 RAS and RAF-1 form a signalling complex with MEK-1 but not MEK-2 Mol Cell Biol 14 (12) ,8212-8218.

[18]. Pearson G, Robinson F, Beers Gibson T, Xu BE, Karandikar M, Berman K, Cobb MH 2001 Mitogen-activated protein (MAP) kinase pathways: regulation and physiological functions Endocrine Rev. 22(2) 153-83.

[19]. Baserga R. 1995 The insulin-like growth factor I receptor: a key to tumor growth. Cancer Res, 55, 249-252.

[20]. Blakesley V. A., Scrimgeour A., Esposito D., LeRoith D. 1996 Signaling via the insulin-like growth factor-I receptor: does it differ from insulin receptor signaling? Cytokine Growth Factor Rev, 7: 153-15

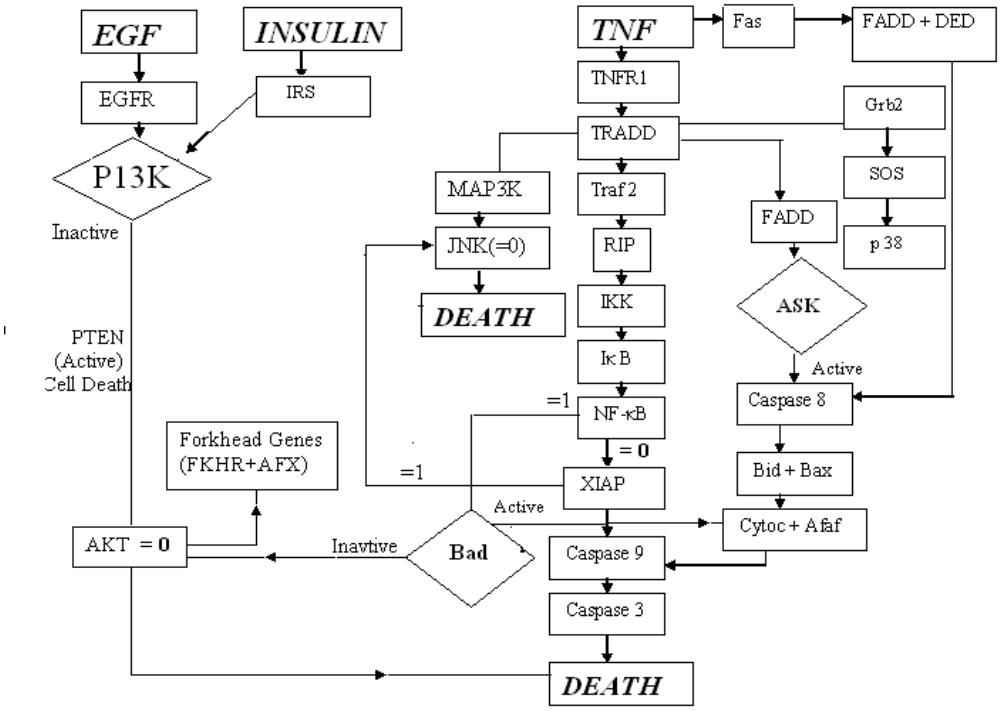


Figure 1: Model for Cell Survival

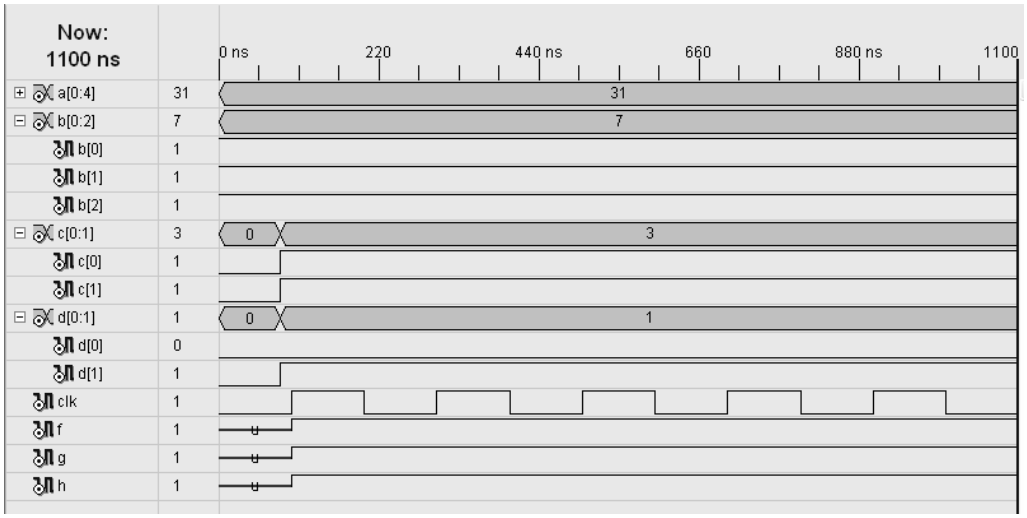


Figure 3: Output signal of cell survival from VHDL simulation considering EGF as input

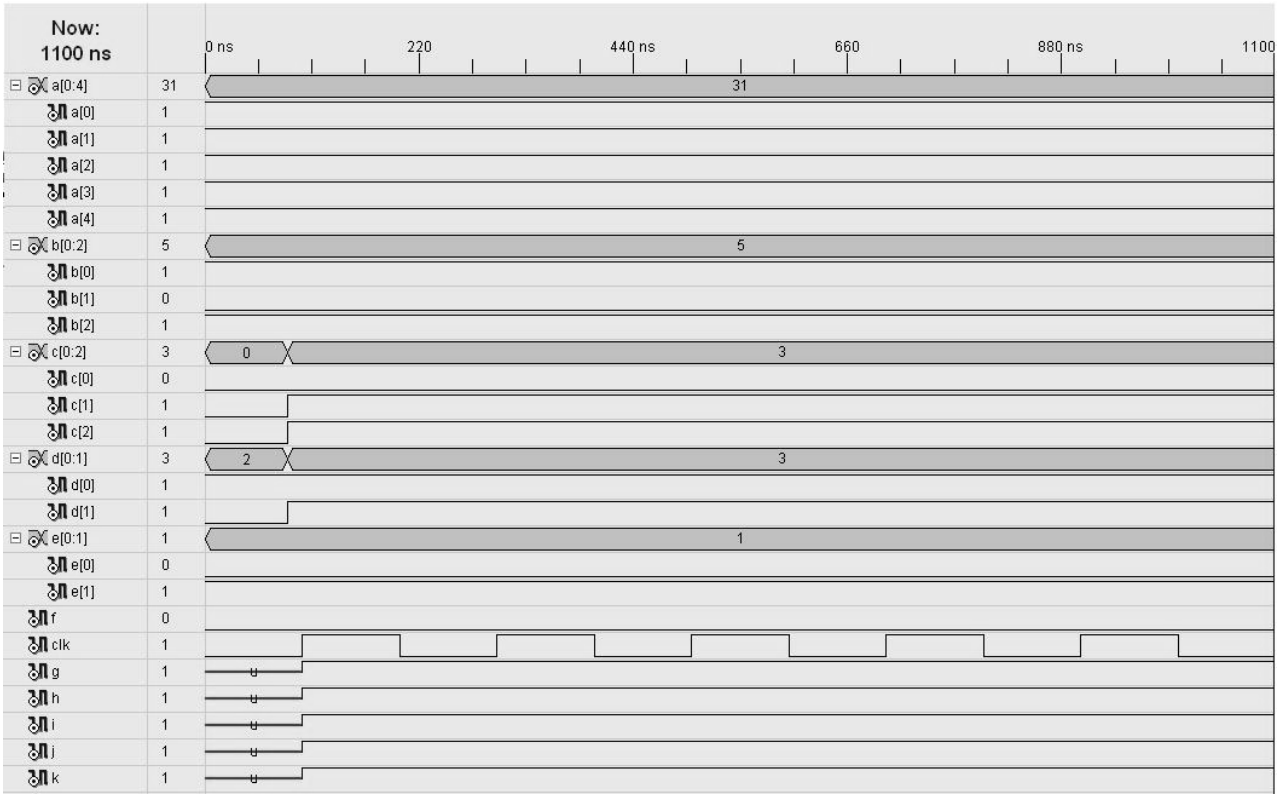


Figure 2: Output signal of cell survival from VHDL simulation considering TNF as input

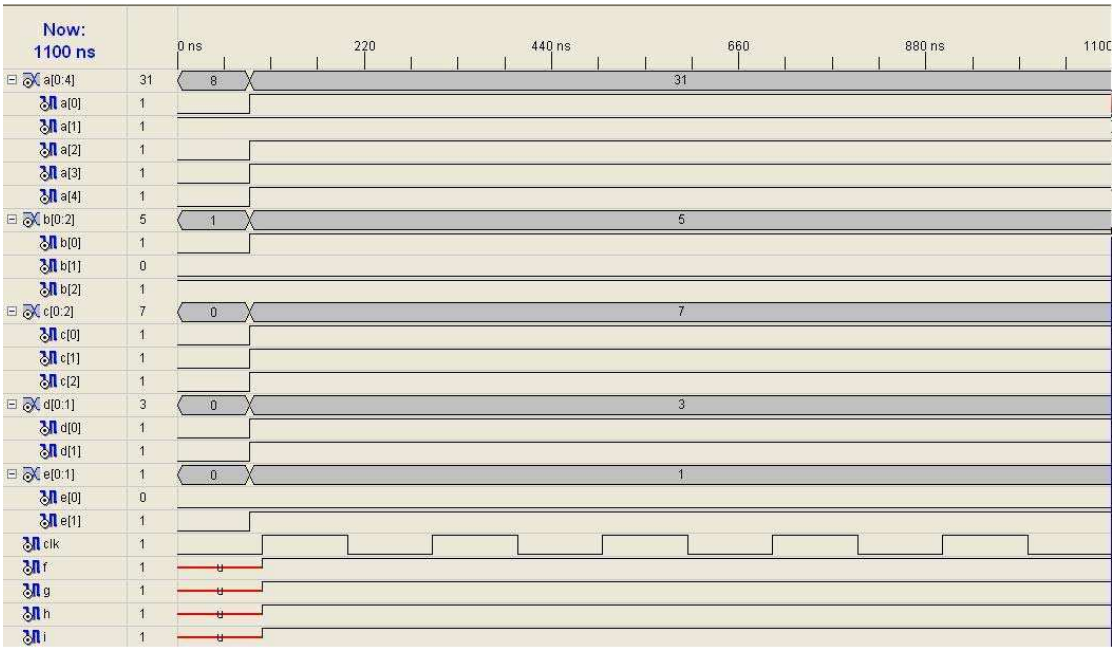


Figure 4: Output signal of cell survival from VHDL simulation considering Insulin as input

Continued from Page No. 195

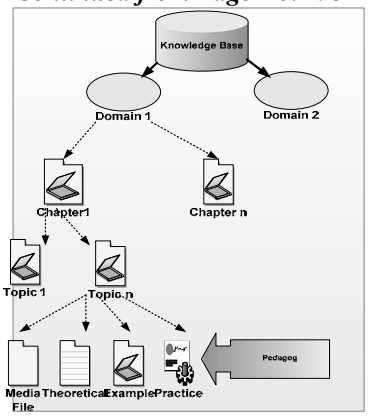


Figure 1: Knowledge Structuring

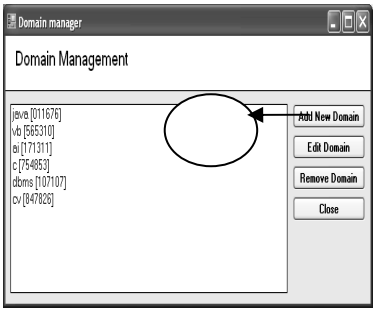


Figure 2: Sample screen shot for adding new domain/editing existing domain.

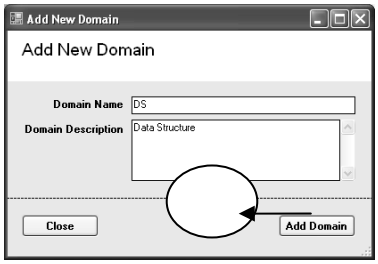


Figure 3: Sample screen shot at the click of A.

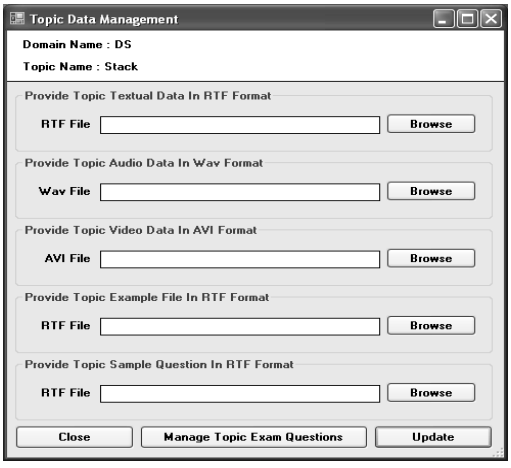


Figure 4: Sample screen shot at click of B.

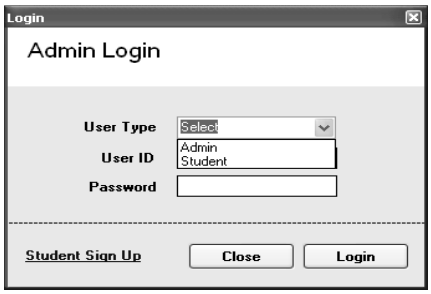


Figure 5: Sample screen shot of lesson planner.

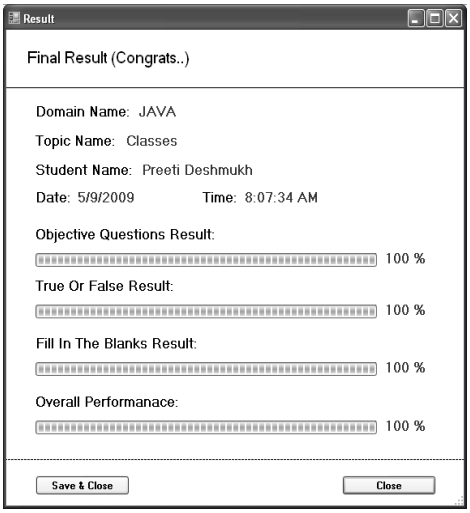


Figure 6: Sample screen shot of Result Generation.

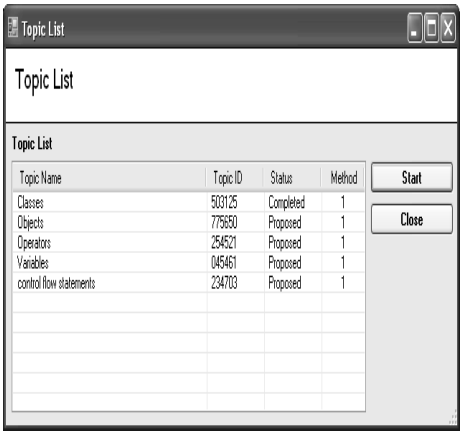


Figure 7: Sample screen shot showing student's study status

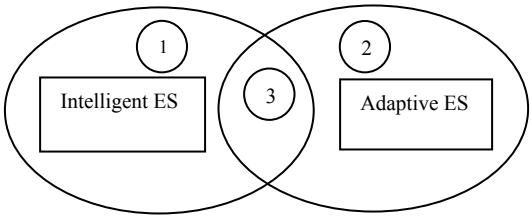


Figure 8: Mapping of Intelligent ES and Adaptive ES.

On Lattice Based Cryptographic Sampling: An Algorithmic Approach

Sunder Lal[1], Santosh Kumar Yadav[2] and Kuldeep Bhardwaj[3]

Abstract - ***In this paper we propose a practical lattice based reduction by sampling to avoid any dependence on Schnorr's Geometric Series Assumption. It is a generalization of Schnorr's RSR algorithm. It is also well defined for bases where this algorithm is not applicable. It demonstrates that the sampling reduction can significantly reduce the length of the base vectors. We also propose a practical sampling reduction algorithm for lattice bases based on work by Schnorr. We report the empirical behavior of these algorithms.***

Index Terms - Sampling algorithm, Best Bound, NTRU Reduction, lattice bases reduction.

1. INTRODUCTION

Lattice bases reduction has been establish as a powerful tool of cryptanalysis. Several cryptosystems have been proposed over the last two decades that are based on the hardness of some lattice problems. The key sizes that need to be selected in order for the system to be secure depend on the efficiency of the best algorithm for computing short vectors in lattices. In 2003 Schnorr presented Random Sampling Reduction (RSR) [8]. It is a new algorithm for computing short vectors in lattices. We assume the Geometrical Series Assumption (GSA), RSR asymptotically outperforms Block Korkine Zolotarev (BKZ) reduction algorithm [7]. However RSR is not a practical algorithm since the choice of parameter in RSR depends on the GSA. The motive of this paper is to present sampling Reduction (SR) as a practical algorithm based on RSR. The experiments to demonstrate that the shortest vector found by SR is significantly shorter than the shortest vector found by BKZ. We also propose two generalization of SR to generate lattice bases with more short vectors On this paper we describe successful attack on low dimensional NTRU lattice bases that require smaller BKZ parameters than previous attacks that used BKZ only.

2. NOTATIONS AND DEFINITIONS

We consider the Euclidean metric on $\mathbb{R}^d$. A lattice L is a discrete subgroup of $\mathbb{R}^d$, its dimension is

$$\dim(L) := \dim(L \otimes_R \mathbb{R}).$$

[1]*Professor and Pro-Vice Chancellor, Dr. B.R. Ambedkar University, Agra (India)*
[2]*Department of Mathematics, Kalindi College, University of Delhi*
[3]*Research Scholar, , Dr. B.R. Ambedkar University, Agra (India)*
E-Mail: [1]sunder_lal2@rediffmail.com,
[2]drskyadav@hotmail.com and [3]kuldeepibs@yahoo.co.in

The first minimum of L is

$$\lambda_1(L) := \min\{\|x\| \mid 0 \neq x \in L\}.$$

For any n-dimensional lattice L, there are ordered bases

$$B = [b_1, ..., b_n] \in \mathbb{R}^{d \times n},\ n \geq 1$$

such that

$$L = L(B) := \{v \mid v = Bx, \text{ for some } x \in \mathbb{R}^n\}.$$

Given an ordered basis B, the set of all bases of L(B) is

$$\{BU \mid U \in \mathbb{R}^{n \times n} \text{ and } \det U = \pm 1\}.$$

We consider integer coefficient lattices only whence $B \in \mathbb{R}^{d \times n}$.

Let $B = \hat{B}R$ be the Gram-Schmidt decomposition of B, i.e. the columns $\hat{b}_j$ of $\hat{B} \in \mathbb{R}^{d \times n}$ are pairwise perpendicular and $R = (\mu_{ij}) \in \mathbb{R}^{n \times n}$ is unit upper triangular. Let

$$\pi_i : \mathbb{R}^d \to lin\{b_1, ..., b_{i-1}\}^{\perp}$$

be the orthogonal projection onto the orthogonal space of the first i–1 base vectors. We denote

$$L_{i,\beta}(B) = L([\pi_i(b_i), ..., \pi_i(b_{\min(i+\beta-1,n)})]).$$

We also consider a generating system of L and parameters (δ, β) with $1/2 < \delta < 1$ and $2 \leq \beta \in \mathbb{R}$, the BKZ algorithm [7] computes a (δ, β)-BKZ reduced basis of L. A (δ, β)-BKZ reduced basis B satisfies

$$|\mu_{ij}| \leq 1/2,\ \forall\ 1 \leq i < j \leq n,$$

(Size condition)

$$\delta \|\hat{b}_i\|^2 \leq \lambda_1(L_{i,\beta}(B)),\ \forall\ 1 \leq i \leq n.$$

(BKZ condition)

In BKZ reduction we obtain the Gram-Schmidt coefficient matrix R as well as $\|\hat{b}_i\|^2$ for i = 1,...,n. L3 reduction is the special case of BKZ reduction with β = 2.

In this paper B = [b1, ..., bn] denotes a (δ, β)-BKZ reduced ordered lattices bases with Gram-Schmidt decomposition

$$B = \hat{B}R,\ \hat{B} = [\hat{b}_1, ..., \hat{b}_n],\ R = (\mu_{i,j}).$$

All lattice points belong to the n-dimensional lattice L = L(B). B is updated in the course of the reduction, L stays always the same.

3. SAMPLING REDUCTION ALGORITHM

Sampling reduction operates on a generating system G of an n-dimensional lattice L. Sampling reduction applies (δ, β)-BKZ reduction to G and obtain the BKZ reduced bases B. The following lemma illustrates to terminate *SR*.

Lemma. *The recursion depth x of SR* $\left(G, \gamma, u_{\max}, \delta, \beta\right)$ *is bound by*

$x \le (n-1)\log_{\gamma}(\delta - 1/4)$. **Proof:** SR operates on (δ, β)-BKZ reduced and thus δ-L^3 reduced bases. Therefore,

$$\|b_1\| \le (\delta - 1/4)^{(1-n)/2}\,\lambda_1(L) \quad [4].$$

BKZ reduction never increases the length of the first vector in the generating system. Each recursion decreases the length of the first base vector by a factor at most $\sqrt{\gamma} < 1$, and b_1 cannot be shorter than $\lambda_1(L)$. Hence,

$$\gamma^x (\delta - 1/4)^{1-n} \ge 1.$$

The input variable $u_{\max} \in \mathbb{N}$ limits the amount of work SR spends on sampling vector.

Algorithm: Sampling Reduction (SR)Input: Generating system G of L, reduction factor γ, search space parameter u_{max}, BKZ parameters (δ, β).

Output: $(B,$ reason) where B is a (δ, β)-BKZ reduced basis of L and reason indicates why the algorithm terminates.

Procedure *SR* $\left(G, \gamma, u_{\max}, \delta, \beta\right)$

$(B, b, R) \leftarrow$ BKZ (G, δ, β)

$/* B = \hat{B}R,\ b = \left(\|\hat{b}_1\|^2, ..., \|\hat{b}_n\|^2\right) */$

if – BESTBOUND $\left(b, u_{\max}, \gamma\right) > u_{\max}$ **then**

return (B, "success probability too small")

else

for $l = 1, ..., 2^{u_{\max}}$ **do**

$v \rightarrow$ SAMPLE (B, R, l)

if $\|v\|^2 \le \gamma \|b_1\|^2$ **then**

return SR ($[v, b_1, ..., b_n]$, γ, u_{max}, δ, β)

end if

end for

return (B, "search space exhausted")

end if

end procedure

4. SAMPLING ALGORITHM

The **Sample** is to generate lattice points that are likely to be short. Because of

$$\|v\|^2 = \sum_{i=1}^{n} v_i^2 \,\|\hat{b}_i\|^2,$$

it is plausible to expect that a lattice point v is short if all Gram-Schmidt coefficients v_i are small. **Sample** enumerates lattice points with all $|v_i| \le 1$.

To be precise, let $2^{u-1} < 1 \le 2^u$. Then v= Sample (B, R, l) = $\sum_{i=1}^{n} v_i \hat{b}_i$ satisfies

$$v_i \in \begin{cases} \left(-\frac{1}{2}, \frac{1}{2}\right), & \text{for } 1 \le i < n-u, \\ (-1, 1], & \text{for } n-u \le i < n, \text{ (SC)} \\ \{1\}, & \text{for } i = n. \end{cases}$$

Let $i \in \{1, ..., n\}$. The choice of v_i in $v = \sum x_j b_j = \sum v_j \hat{b}_j$ does not affect $v_{i+1}, ..., v_n$ since R is unit upper triangular. Therefore, Sample computes (x_i, v_i) by iteration based on $x_n = v_n = 1$. Assume the coefficients $(x_{i+1}, v_{i+1}), ..., (x_n, v_n)$ are already fixed. Then Sample determines the unique $x' \in \mathbb{Z}$ with $\pi_i\left(x' b_1 + \sum_{j=i+1} x_j b_j\right) = v' \hat{b}_i = \sum_{j=i+1}^{n} v_j \hat{b}_j$ and $v' \in (-1/2, 1/2)$. Sampling chooses $(x_i, v_i) = (x', v')$ if l div 2^{n-i-1} is even. Else (x_i, v_i) becomes also unique

$(x' \pm 1, v' \pm 1)$ such that

$v_i \in (-1, -1/2] \cup [1/2, 1)$.

Thus, $\{1, ..., 2^{u_{\max}}\} \rightarrow L(B) : l \mapsto$ **Sample** (B, R, l) is an enumeration of all points in $L(B)$ subject to (SC) with $u = u_{max}$. Inspection of the following algorithm shows the computation of Sample requires $2n$ vector updates and assignments, i.e. $O(n^2)$ arithmetic operations.

Algorithm:

Input: Unit upper triangular matrix

$R = [r_1, ..., r_n] \in \mathbb{R}^{n \times n}$, lattice basis $B = [b_1, ..., b_n] \in \mathbb{R}^{n \times n}$

With Gram-Schmidt decomposition $B = \hat{B}R$, $1 \le l \le 2^{n-1}$.

Output: $u \in L(B)$ subject to (SC).

Procedure: Sample (B, R, l)

$v \leftarrow b_n$, $v = (v_1, ..., v_n)^t \leftarrow r_n$

for $i = n-1, n-2, ..., 1$ **do**

$x \leftarrow [v_i - \frac{1}{2}]$

if l **mod** 2 = 1 **then**

$/* -\frac{1}{2} \le v_i - x \le \frac{1}{2} */$

if $v_i - x \le 0$ **then**

$x \leftarrow x - 1 \qquad /* \frac{1}{2} < v_i - x \leq 1 */$

else

$x \leftarrow x + 1 \qquad /* -1 < v_i - x \leq -\frac{1}{2} */$

end if
end if

$l \leftarrow 1$ **div** 2

$v \leftarrow v - xb_i,\ v \leftarrow v - xr_i$

$/* v_i \leftarrow v_i - x */$

end for
return *v*
end procedure

5. BEST BOUND ALGORITHM

The vectors computed by SAMPLE are likely to be short but they are of course not necessarily shorter than b_1. The algorithm BESTBOUND yields as estimate how many samples are required in the search space

$V_t := \{v_1, ..., v_{2-t}\}$.

If we want to guarantee a success probability

Pr [min $\{\|v\|^2\ v \in V_t\} \leq \gamma \|b_1\|^2] \geq 1/2$.

Let $l \in_R \{1,...,2^u\}$ be by SAMPLE

$(B, R, l) = \sum_{i=1} v_i \hat{b}_i$

are statistically indistinguishable from independent random variables with uniform distribution on the intervals defined by (SC) [1].

BESTBOUND is supposed to return a lower bound for (the $\log_2$ of) the probability that SAMPLE returns a vector shorter than $\sqrt{\gamma}\,\|b_1\|$. The algorithm is based on the following idea: The sampling of a lattice point *v* is a random experiment. We consider some event $(S_{q,k})$ parameterized by $q \in [0, 1]$ and $1 \leq k < n - u_{max} < n$. The probability of $(S_{q,k})$ is strictly increasing in *q*. Let $0 \leq q_\gamma \leq 1$ be maximal s.t. the conditional expected length $E[\|u\|^2 \mid (S_{q,k}] \leq \gamma \|b_1\|^2$. Then the success probability is

$\Pr[\|v\|^2 \leq \gamma \|b_1\|^2] \geq \Pr[\|v\|^2 \leq E[\|v\|^2 | (S_{q_\gamma,k})]$ BestBound $S_{q_\gamma,k})] \Pr[(S_{q_\gamma,k})] = \frac{1}{2} \Pr[(S_{q_\gamma,k})]$.

computes

$\max\left\{ \left\lfloor \log_2 \left(\tfrac{1}{2} \Pr[(S_{q_\gamma,k})] \right) \right\rfloor \mid k = 1,...,n - u_{\max} \right\}$. Consequntly, if SR samples at least $2^{-\text{BestBond}(b, u_{\max}, \gamma)}$ lattice points then the probability to find a sufficiently short vector is at least ½.

The event $(S_{q,k})$. Consider the random experiment v = Sample (*B, R, l*)

$= \sum_{i=1}^{n} v_i \hat{b}_i,\ l \in_R \{1,...,2^{u_{\max}}\}$.

$\sigma \in$ Sym ($\{1,...,n\}$) describes the sorting of the firs $n - u_{\max} - 1$ elements of *b* in non-increasing order, i.e.

$\|\hat{b}_{\sigma(1)}\|^2 \geq \cdots \geq \|\hat{b}_{\sigma(n-u_{\max}-1)}\|^2$ and

$\sigma(i) = i \quad \text{for } i \geq n - u_{\max}$ (1)

Let $q \in [0, 1]$ and $1 \leq k < n - u_{max}$. $(S_{q,k})$ denotes the event

$v_{\sigma(i)}^2 \leq \frac{1}{4} q^{k-i}$ for i = 1,...,*k*–1. $(S_{q,k})$

The randomness assumption on v_i yields

$\Pr[(S_{q,k})] = \prod_{i=1}^{k-1} \Pr\left[|v_{\sigma(i)}| \leq \frac{1}{2} q^{\frac{k-i}{2}} \right] = q^{\frac{k(k-1)}{4}}$. The expected length of *v*. Assume $(S_{q,k})$. For any uniform random variable $x \in (-t, t]$ the expected value of x^2 is $E[x^2] = \frac{1}{3} t^2$.

The v_i are independent random variables uniformly distributed on intervals defined by (SC) and $(S_{q,k})$ whence

$E(b,k,q) := E[\|v\|^2 | (S_{q,k})]$

$= \sum_{i=1}^{n} E[v_i^2 \mid (S_{q,k})] \|\hat{b}_i\|^2$

$= \sum_{i=1}^{k-1} q^{k-i} \frac{\|\hat{b}_{\sigma(i)}\|^2}{12} + \sum_{i=k}^{n-u_{\max}-1} \frac{\|\hat{b}_{\sigma(i)}\|^2}{12}$

$+ \sum_{i=n-u_{\max}}^{n-1} \frac{\|\hat{b}_i\|^2}{3} + \|\hat{b}_n\|^2$.

Algorithm: BestBound

Input: $b = (\|\hat{b}_1\|^2, ..., \|\hat{b}_n\|^2)$, base 2 logarithm $u_{\max}$ of maximum number of samples, reduction factor γ.

Output: $t_{\max} = \max\{t \in \mathbb{R} \mid \Pr[\min\{\|v\|^2 \leq \gamma \|\hat{b}_1\|^2 | v \in V_t\}] \geq 1/2\} \cup \{-\infty\}$.

procedure ExpLength (*l, k, u, q*)

$/* l = (l_1, ..., l_n) */$

return

$\frac{1}{12} \sum_{i=1}^{k-1} q^{k-i} l_i + \frac{1}{12} \sum_{i=k}^{n-u-1} l_i + \frac{1}{3} \sum_{i=n-u}^{n-1} l_i + l_n$

end procedure
procedure LogSuccessProbBound
(*l, k,u, γ*)

if ExpLength (L, k, u, 1) ≤ γ $|| \hat{b}_1 ||^2$ **then**

return -1

else if ExpLength (l, k, u, 0) ≥ γ$|| \hat{b}_1 ||^2$ **then return** $-\infty$

end if

$q_\gamma \leftarrow$ RegulaFalsi(ExpLength(l, k, u, q) = γ $|| \hat{b}_1 ||^2$, q ∈ [0, 1]) **return** $\left\lfloor \frac{k(k-1)}{4} \log_2(q_\gamma) - 1 \right\rfloor$ **end procedure**

procedure BestBound (b, u_{max}, γ)

$$l \leftarrow (|| \hat{b}_{\sigma(1)} ||^2, ..., || \hat{b}_{\sigma(n)} ||^2)$$

/* permutation σ */

return max {LogSuccessProbBound (l, k, u_{max}, γ) | $k = 1, ..., n - u_{max}$} **end procedure**

Numerical Representation

Computation of q_γ. The expected length $E(b, k, q)$ is a polynomial in q with non-negative coefficients Therefore

f: [0, 1] → □ : $q \mapsto E(b,k,q) - \gamma || b_1 ||^2$ is a strictly increasing continuous function that has a root if and only if $f(0) \le 0 \le f(1)$. The unique root q_γ can be efficiently determined with the textbook Regula Falsi algorithm [10] provided such a root exists.

If $f(1) < 0$ then the (unconditional) means value $E[||v||^2]$ is already short enough and we have

Pr $[||v||^2 \le \gamma ||b_1||^2] \ge ½$.

On the other hand, if $f(0) > 0$ then our approach dos not yield a positive lower bound on Pr $[||v||^2 \le \gamma ||b_1||^2]$ for this particular choice of k.

The optimal bound t. BestBound computes the maximum success probability for all k. The computation of $\Pr[(S_{q_\gamma,k})]$ is in our experience fast enough that the cost for computing the probability for all $k = 1,..., n - u_{max} - 1$ is negligible.

6. RESULTS

Followed by Linux system with a 2.4 GHz Pentium 4 processor and 1 GByte RAM. We used a lattice reduction library that is derived from Shoup's NTL [9]. We tested our algorithm with bases in Hermite normal form as proposed by Micciancio [10] for the public keys in his variant of the GGH cryptosystem. They are derived from base vectors uniformly chosen from a cube whence the generated lattices do not have any special structure. The HNF bases were (0.99 β)-BKZ reduced for various values of β. The resulting bases were input to the Sampling Reduction.

A large part of this improvement is gained in the first iterations. With

β = 5, the Sampling Reduction took 1928s, of 2which 71s were spent on the BKZ updates. With β=10, the Sampling reduction ran only for 577s but here 190s were spent in the BKZ updates.

It is noticeable that the very first base vectors are much more improved than the remaining base vectors. Most of the time, the effect of the BKZ updates peters to quickly. In particular, $|| \hat{b}_i ||^2$ does not change significantly beyond base column 20. Since $E[|| v ||^2]$ does not change that much if only few $\hat{b}_i$ become smaller it quickly becomes less likely that a sampled vector is shorter than b_1. This is also reflected in our estimates of the success probability's logarithm, shown in the diagram in Fig. 1(a). The estimates decrease quite rapidly with very recursion.

The value of BestBound actually depends on the choice of u_{max} : If one increments u_{max} then E(b, k, q) grows by $\frac{1}{4} || \hat{b}_{n-u_{max}-1} ||^2$ which means that q_γ and therefore $\Pr[(S_{q_\gamma,k})]$ become smaller.

7. CONCLUSION AND FUTURE TRENDS

In our work we have demonstrated that the Sampling Reduction can significantly reduce the length of the base vectors. We have also proposed two generalizations that further reduce the overall length of the base vectors and that allow the Sampling Reduction to proceed even if jumps in the length of the orthogonalized base vectors disrupt the plain Sampling Reduction. Observing the algorithms and procedure.

We find that our estimates of the success probability are too pessimistic. We plan to test whether it is numerically feasible to calculate the distribution of the length of the sampled vectors directly by convoluting (via FFT) the distributions of the coefficients v_i. The result needs to be verified by further experiments in higher dimensions and approach.The implementa-tion can also be performed in C#.

REFERENCES

[1]. Ajtai, M., Dwork, C.: A public-key cryptosystem with worstcase / average-case equivalence. In: *Proceedings of the 29th Annual Symposium on Theory of Computing (STOC),* ACM Press (1997) 284-293.

[2]. Coppersmith, D., Shamir, A.: Lattice attacks on NTRU. In: *Advances in Cryptology – Eurocrypt' 97.* Volume 1233 of LNCS, Springer (1997) 52-61.

[3]. Lenstra, A.K., Lensta, H.W., Lovasz, L: Factoring polynomials with rational coefficients. *Math. Ann.* **261** (1982) 515534.

[4]. Lenstra, A.K., Verheul, E.R.: Selecting cryptographic key sizes. J. Cryptology 14 (2001) 255-293.

[5]. Micciancio, D.: The shortest vector problem is NP-hard to approximate to within some constant. SIAM Journal on Computing 30 (2001) 2008-2035.

[6]. NTRU Cryptosystems, Inc. : http://www.ntru.com (2004)

[7]. Schnorr, C.P., Euchnere, M.: Lattice basis reduction: Improved practical algorithms and solving subset sum problems. Math. Programming 66 (1994) 181-199.

[8]. Schnorr, C.P.: Lattice reduction by random sampling and birthday methods. In Alt, H., Habib, M., eds: STACS

2003: 20th Annual Symposium on Theoretical Aspects of Computer Science. Volume 2607 of LNCS, Springer (2003) pp. 146-156.

[9]. Shoup, V.: NTL-a library for doing number theory. URL, http://www.shoup.net/ntl/index.html (2004) Release 5.3.2.

[10]. W.H., Tuekolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical Recipes in C. 2nd edn. Cambridge University Press (1992).

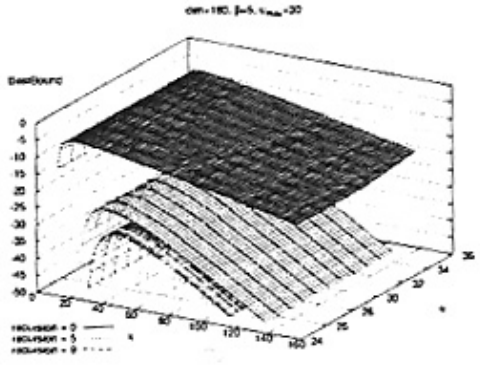


Figure 1(a): Sampling Reduction of HNF Bases

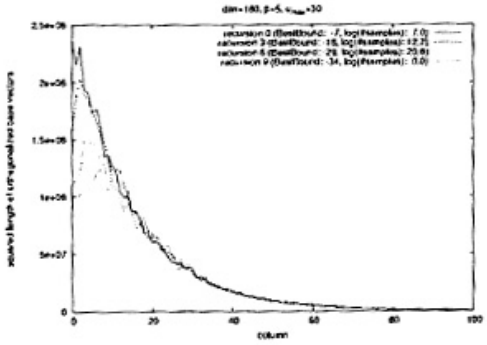


Figure 1(b): Values of Best Bound, $|| \hat{b}_i ||^2$

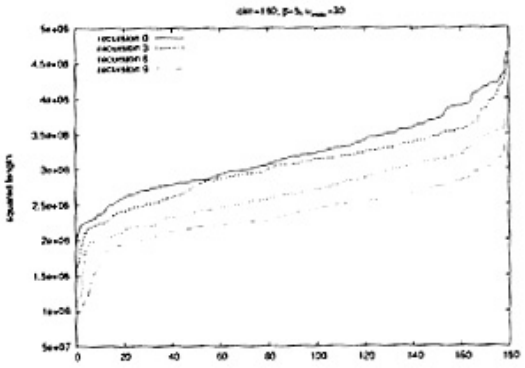


Figure 1(c): Values of BestBound, $|| b_i ||^2$ (sorting in non decreasing order)

BIJIT - BVICAM's International Journal of Information Technology

Paper Structure and Formatting Guidelines for Authors

BIJIT is a peer reviewed refereed bi-annual research journal having **ISSN 0973-5658**, being published since 2009, in both, Hard Copy as well as Soft copy. Two issues; **January – June** and **July – December**, are published every year. The journal intends to disseminate original scientific research and knowledge in the field of, primarily, Computer Science and Information Technology and, generally, all interdisciplinary streams of Engineering Sciences. **Original** and **unpublished** research papers, based on theoretical or experimental works, are published in BIJIT. We publish two types of issues; **Regular Issues** and **Theme Based Special Issues.** Announcement regarding special issues is made from time to time, and once an issue is announced to be a Theme Based Special Issue, Regular Issue for that period will not be published.

Papers for Regular Issues of BIJIT can be submitted, round the year. After the detailed review process, when a paper is finally accepted, the decision regarding the issue in which the paper will be published, will be taken by the Editorial Board; and the author will be intimated accordingly. However, for Theme Based Special Issues, **time bound Special Call for Papers** will be announced and the same will be applicable for that specific issue only.

Submission of a paper implies that the work described has not been published previously (except in the form of an abstract or academic thesis) and is not under consideration for publication elsewhere. The submission should be approved by all the authors of the paper. If a paper is finally accepted, the authorities, where the work had been carried out, shall be responsible for not publishing the work elsewhere in the same form. *Paper, once submitted for consideration in BIJIT, cannot be withdrawn unless the same is finally rejected*.

1. Paper Submission

Authors will be required to submit, MS-Word compatible (.doc, .docx), papers electronically *after logging in at our portal and accessing the submit paper link*, available at http://www.bvicam.ac.in/bijit/SubmitPaper.asp. Once the paper is uploaded successfully, our automated Paper Submission System assigns a Unique Paper ID, acknowledges it on the screen and also sends an acknowledgement email to the author at her / his registered email ID. Consequent upon this, the authors can check the status of their papers at the portal itself, in the Member Area, after login, and can also submit revised paper, based on the review remarks, from member area itself. The authors must quote / refer the paper ID in all future correspondences. Kindly note that we do not accept E-Mailic submission. To understand the detailed step by step procedure for submitting a paper, click at http://www.bvicam.ac.in/BIJIT/guidelines.asp.

2. Paper Structure and Format

While preparing and formatting papers, authors must confirm to the under-mentioned MS-Word (.doc, .docx) format:-

- The total length of the paper, including references and appendices, must not exceed **six (06) Letter Size pages**. It should be typed on one-side with double column, single-line spacing, 10 font size, Times New Roman, in MS Word.
- The Top Margin should be 1", Bottom 1", Left 0.6", and Right 0.6". Page layout should be portrait with 0.5 Header and Footer margins. Select the option for different Headers and Footers for Odd and Even pages and different for First page in Layout (under Page Setup menu option of MS Word). Authors are not supposed to write anything in the footer.
- The title should appear in single column on the first page in 14 Font size, below which the name of the author(s), in bold, should be provided centrally aligned in 12 font size. The affiliations of all the authors and their E-mail IDs should be provided in the footer section of the first column, as shown in the template.
- To avoid unnecessary errors, the authors are strongly advised to use the "spell-check" and "grammar-check" functions of the word processor.
- The complete template has been prepared, which can be used for paper structuring and formatting, and is available at http://www.bvicam.ac.in/BIJIT/Downloads/Template_For_Full_Paper_BIJIT.pdf.
- The structure of the paper should be based on the following details:-

Essential Title Page Information

- **Title:** Title should be Concise and informative. Avoid abbreviations and formulae to the extent possible.
- **Authors' Names** and **Affiliations:** Present the authors' affiliation addresses (where the actual work was done) in the footer section of the first column. Indicate all affiliations with a lower-case superscript letter immediately after the author's name

and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and e-mail address of each author.

- **Corresponding Author:** Clearly indicate who will handle correspondence at all stages of refereeing and publication. Ensure that phone numbers (with country and area code) are provided, in addition to the e-mail address and the complete postal address.

Abstract

A concise abstract not exceeding 200 words is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. References and non-standard or uncommon abbreviations should be avoided. As a last paragraph of the abstract, 05 to 10 Index Terms, in alphabetic order, under the heading Index Terms ***(Index Terms - …….)*** must be provided.

NOMENCLATURE

Define all the abbreviations that are used in the paper and present a list of abbreviations with their definition in Nomenclature section. Ensure consistency of abbreviations throughout the article. Do not use any abbreviation in the paper, which has not been defined and listed in Nomenclature section.

Subdivision - numbered sections

Divide paper into numbered Sections as 1, 2, 3, ….. and its heading should be written in CAPITAL LETTERS, bold faced. The subsections should be numbered as 1.1 (then 1.1.1, 1.1.2, ...), 1.2, etc. and its heading should be written in Title Case, bold faced and should appear in separate line. The Abstract, Nomenclature, Appendix, Acknowledgement and References will not be included in section numbering. In fact, section numbering will start from Introduction and will continue till Conclusion. All headings of sections and subsections should be left aligned.

INTRODUCTION

State the objectives of the work and provide an adequate background, with a detailed literature survey or a summary of the results.

Theory/Calculation

A Theory Section should extend, not repeat the information discussed in Introduction. In contrast, a Calculation Section represents a practical development from a theoretical basis.

RESULT

Results should be clear and concise.

DISCUSSION

This section should explore the importance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate.

CONCLUSION AND FUTURE SCOPE

The main conclusions of the study may be presented in a short Conclusion Section. In this section, the author(s) should also briefly discuss the limitations of the research and Future Scope for improvement.

APPENDIX

If there are multiple appendices, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similar nomenclature should be followed for tables and figures: Table A.1; Fig. A.1, etc.

ACKNOWLEDGEMENT

If desired, authors may provide acknowledgements at the end of the article, before the references. The organizations / individuals who provided help during the research (e.g. providing language help, writing assistance, proof reading the article, sponsoring the research, etc.) may be acknowledged here.

REFERENCES

Citation in text

Please ensure that every reference cited in the text is also present in the reference list (and vice versa). The references in the reference list should follow the standard IEEE reference style of the journal and citation of a reference.

Web references

As a minimum, the full URL should be given and the date when the reference was last accessed. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list, as well.

Reference style

Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given. Example: '..... as demonstrated [3,6]. Barnaby and Jones [8] obtained a different result ....'

List: Number the references (numbers in square brackets) in the list, according to the order in which they appear in the text.

Two sample examples, for writing reference list, are given hereunder:-

Reference to a journal publication:

[1] I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread-spectrum watermarking for multimedia", *IEEE Transactions on Image Processing*, Vol. 6, No. 12, pp. 64 – 69, December 1997.

Reference to a book:

[2] J. G. Proakis and D. G. Manolakis – Digital Signal Processing – Principles, Algorithms and Applications; Third Edition; Prentice Hall of India, 2003.

Mathematical Formulae

Present formulae using Equation editor in the line of normal text. Number consecutively any equations that have to be referred in the text

Captions and Numbering for Figure and Tables

Ensure that each figure / table has been numbered and captioned. Supply captions separately, *not attached to the figure*. A caption should comprise a brief title and a description of the illustration. Figures and tables should be numbered separately, but consecutively in accordance with their appearance in the text.

3. Style for Illustrations

All line drawings, images, photos, figures, etc. will be published in black and white, in Hard Copy of BIJIT. Authors will need to ensure that the letters, lines, etc. will remain legible, even after reducing the line drawings, images, photos, figures, etc. to a two-column width, as much as 4:1 from the original. However, in Soft Copy of the journal, line drawings, images, photos, figures, etc. may be published in colour, if requested. For this, authors will need to submit two types of Camera Ready Copy (CRC), after final acceptance of their paper, one for Hard Copy (compatible to black and white printing) and another for Soft Copy (compatible to colour printing).

4. **Referees**

Please submit, with the paper, the names, addresses, contact numbers and e-mail addresses of three potential referees. Note that the editor has sole right to decide whether or not the suggested reviewers are to be used.

5. Copy Right

Copyright of all accepted papers will belong to BIJIT and the author(s) must affirm that accepted Papers for publication in **BIJIT** must not be re-published elsewhere without the written consent of the editor. To comply with this policy, authors will be required to submit a signed copy of Copyright Transfer Form, available at http://bvicam.ac.in/bijit/Downloads/BIJIT-Copyright-Agreement.pdf, after acceptance of their paper, before the same is published.

6. Final Proof of the Paper

One set of page proofs (as PDF files) will be sent by e-mail to the corresponding author or a link will be provided in the e-mail so that the authors can download the files themselves. These PDF proofs can be annotated; for this you need to download Adobe Reader version 7 (or higher) available free from http://get.adobe.com/reader. If authors do not wish to use the PDF annotations function, they may list the corrections and return them to BIJIT in an e-mail. Please list corrections quoting line number. If, for any reason, this is not possible, then mark the corrections and any other comments on a printout of the proof and then scan the pages having corrections and e-mail them back, within 05 days. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the paper that has been accepted for publication will not be considered at this stage without prior permission. It is important to ensure that all corrections are sent back to us in one communication: please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely authors' responsibility. Note that BIJIT will proceed with the publication of paper, if no response is received within 05 days.

BVICAM'S International Journal of Information Technology

(A Biannual Publication)

Subscription Rates

Category	I Year		3 Years	
	India	**Abroad**	**India**	**Abroad**
Companies	Rs. 400	US $ 45	Rs. 1000	US $ 120
Institution	Rs. 300	US $ 40	Rs. 750	US $ 100
Individuals	Rs. 250	US $ 30	Rs. 600	US $ 075
Students	Rs. 150	US $ 25	Rs. 375	US $ 050
Single Copy	Rs. 250	US $ 25	-	-

Subscription Order Form

Please find attached herewith Demand Draft No._____ dated _____

For Rs._____ drawn on _____Bank in favor of **Director, "Bharati Vidyapeeth's Institute of Computer Applications and Management, New Delhi"** for a period of 01 Year / 03 Years

Subscription Details

Name and Designation _____

Organization _____

Mailing Address _____

_____ PIN/ZIP _____

Phone (with STD/ISD Code)_____FAX_____

E-Mail (in Capital Letters)_____

Date: **Signature**

Place: (with official seal)

Filled in Subscription Order Form along with the required Demand Draft should be sent to the following address:-

Prof. M. N. Hoda
Chief Editor – BIJIT,
Director, Bharati Vidyapeeth's
Institute of Computer Applications & Management
A-4, Paschim Vihar, Rohtak Road, New Delhi-110063 (INDIA).
Tel.: 91 – 11 – 25275055 Fax: 91 – 11 – 25255056 E-Mail: bijit@bvicam.ac.in
Visit us at: www.bvicam.ac.in

Announcement & Call for Papers

Organized by

Bharati Vidyapeeth's Institute of Computer Applications & Management

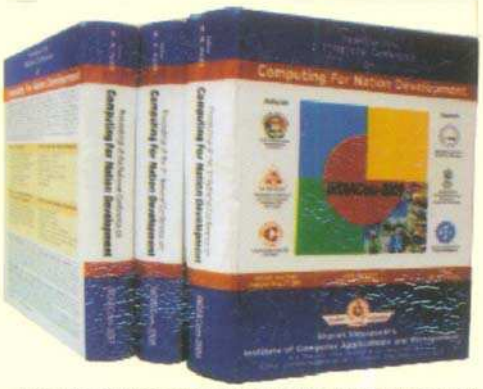
A-4, Paschim Vihar, Rohtak Road, New Delhi-63

Jointly with

GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY

The Institution of Electronics and Telecommunication Engineers (IETE), Delhi Centre

Computer Society of India (CSI), Delhi Chapter



(Copies of the proceedings of past *INDIAComs*)

Correspondence
All correspondences related to the conference may be sent to the address:

Prof. M. N. Hoda
Chief Convener, *INDIACom* - **2010**
Director, Bharati Vidyapeeth's
Institute of Computer Applications and Management
A-4, Paschim Vihar, Rohtak Road, New Delhi-63.
Tel./Fax: 011-25275055, 09212022066 (Mobile)
E-Mails: conference@bvicam.ac.in
indiacom2010@gmail.com

For further details, visit us at: http://www.bvicam.ac.in

INDIACom-2010

4th National Conference on

Computing For Nation Development

(25th-26th February, 2010)

Information and communication technologies play a dramatic impact on effectiveness, efficiency, growth and development in various areas such as education, health-care & modernization. Foreseeing the importance and impact of the above and encouraged by the resounding success met with the previous three editions of the INDIAComs; INDIACom-2009, INDIACom-2008 and INDIACom-2007; we hereby announce **INDIACom – 2010**, which aims to develop a strategic plan for balanced growth of our economy through IT in critical areas like E-Governance, E-Commerce, Disaster Management, GIS, Nano-Technology, Intellectual Property Rights, AI and Expert Systems, Networking, Software Engineering and other Emerging Technologies.

The **INDIACom - 2010** intends to bring eminent academicians, scientists, researchers, industrialists, technocrats, government representatives, social visionaries and experts from all strata of society, under one roof, to explore the new horizons of innovative technology to identify opportunities using IT and defining the path forward. This new path will envision to eliminate isolation, discourage redundant efforts and promote scientific progress aimed to accelerate India's overall growth to prominence on the International front. The INDIACom - 2010 will feature regular paper presentation sessions, invited talks, key note addresses, panel discussions and poster exhibitions. More than 500 papers have been received from over 850 authors from all over country. Eminent speakers from Academia, Industry and Government have already confirmed to participate in INDIACom -2010. Our previous editions of Pre-Conference Proceedings have widely been appreciated from all academic circles. As earlier, this year also, we will publish both soft and hard copies of the Pre-Conference Proceedings with ISSN and ISBN serials. Maximum benefits from this event can be derived by participating in huge number and together making it a grand success. Further details are available at our website **www.bvicam.ac.in/indiacom**.

Registration Fee :

	Early Bird on or before 05th December, 2009 (in Rs.)		After 05th December, 2009 (in Rs.)		Spot Registration (only in Cash)	
Category of Delegates	***CSI/IETE Members**	**General**	***CSI/IETE Members**	**General**	***CSI/IETE Members**	**General**
Students# (Delegate only)	600.00	800.00	800.00	1000.00	1000.00	1200.00
Teachers/Research Scholars	2000.00	2300.00	2300.00	2600.00	2600.00	3000.00
Industry	3000.00	3500.00	3500.00	4000.00	4000.00	4500.00

10% discount will be given on three or more registrations from one organization in General Category only.

NSC-2010

3rd National Students' Convention on

Computing For Nation Development

(February 27, 2010)

Bharati Vidyapeeth's CSI Students' Branch is also organizing 3rd National Students' Convention (NSC-2010) on the same theme of "Computing For Nation Development" on 27th February, 2010. Further details are available on the website **www.bvicam.ac.in/nsc**.