DYNAMIC TRUST AND SECURITY MANAGEMENT PROTOCOL FOR DELAY TOLERANT NETWORKS

Mrs. Suvarna L. Kattimani[1], Mr. Jaeerahmad N. Indikar[2], Dr Suvarna Nandyal[3]

[1]Assistant Professor, [2]PG Scholar, [3]Professor, HOD, Dept of CSE,
BLDEA'S Dr Halkatti College of Engineering & Technology, Vijayapur, Karnataka, (India)

ABSTRACT

Delay tolerant networks (DTNs) are depicted by high frequent disconnection, end-to-end latency, and opportunistic communication over unreliable wireless links. To avoid these anomalies we propose a dynamic trust & security management protocol for secure routing optimization in DTN environments in the presence of well-behaved, selfish and malicious nodes. Which will be a novel model-based methodology for the analysis of our trust protocol and validate it via simulation, we address dynamic trust and security management using the information centric networks (ICN) architecture, i.e., determining and applying the best operational settings at runtime in response to dynamically changing network conditions to minimize trust bias and to maximize the routing application performance. The results demonstrate that our protocol is able to deal with selfish behaviours and is resilient against trust-related attacks comparison to Bayesian trust-based routing protocols, Dynamic trust and security management without ICN architecture and with using ICN architecture .Furthermore, our trust and security based routing protocol can effectively trade off message overhead and message delay for a significant gain in delivery ratio.

Keywords: DTN, DTSMP, ICN

I. INTRODUCTION

Mobile network typically consist of many heterogeneous nodes performing end-to-end wireless communications to achieve the system functionality. There are various types of mobile networks, including delay/disruption tolerant networks (DTNs) [9] ,mobile ad-hoc networks (MANETs) [11], Internet of things (IoT) systems [5] ,mobile wireless sensor networks (WSNs) [4] etc. The key features of mobile networks are low dependency on infrastructure, no centralized entity needed for managing the network (distributed control), and change of network topology, population size, etc (dynamic). Because of these main features, mobile networks have been widely deployed in many applications. For example, conference attendees can set up an ad-hoc network using their laptops for discussion instant messaging. In war situations, a soldier can dynamically assemble and manage a mobile network consisting of group members to achieve a critical mission assigned. In zoology research, sensors can be attached to wild animals to form a delay tolerant WSN in order to track animal behaviors.

Trust management in mobile wireless network is always been challenging because of frequently changing network environment. This will cause delay tolerance networks (DTN) a high latency, frequent disconnection over unreliable wireless links. Many researchers worked and designed and validate the Trust management for delay tolerant networks (DTN).

The contribution of the paper related to the some of the existing work in trust management for DTNs which are summarized as follows

1. We have combined the social trust and quality of service which are derived from social network and communication network respectfully. We have used the two social trust metrics called "unselfishness" and "healthiness" to find the both malicious and socially selfish nodes in the DTN environment.
2. We address the issue of the trust based DTN routing through dynamic trust and security management protocol by adjusting trust protocol setting dynamically for the changing DTN environment.
3. We deploy trust and security management protocols for delay-tolerant, self-contained message forwarding applications based on the information-centric networks (ICN) architecture.
4. We perform comparative analysis of trust and security management protocol with respective the Bayesian trust-based routing protocols and Dynamic trust and security management protocol using ICN.

II. RELATED WORK

2.1 Computational Trust Models

There are many computational trust models being proposed in the literature, including Bayesian [14], weighted summation [3], game theory based [7], fuzzy logic based [8], routing algebra based [19], graph based [18],belief based, flow based [6] , and information theory-based [10] models. Below we survey and contrast our work with the first three computational trust models which have been used most frequently in the literature.

2.1.1 Bayesian Models

In Bayesian trust models, the evidence of trust is considered as a stochastic process. First, a prior distribution of the trust value is assumed. Then, the evidence is observed and can be used as the likelihood to calculate the posterior distribution following Bayes' Theorem. After new evidence is observed, the previous posterior distribution obtained can be used as a new prior distribution to calculate the next posterior distribution iteratively. The new evidence could be from direct observations or indirect recommendations. Direct observations may be used to update the numbers of positive and negative interaction experiences, whereas indirect recommendations may be discounted by the confidence [12] or belief [15] of the trustor toward the recommenders. Since this is an iterative computing process, it is desirable if both the prior and posterior distributions follow the same distribution and only the parameters are updated iteratively after new evidence is observed. Therefore, conjugate prior distributions, like Beta distribution [14] and Dirichlet distribution [14] are usually used as the prior distribution to build trust models.

2.1.2 Weighted Summation Models

One of most popular and straightforward computational trust models is the weighted summation or average model [2,3]. Models in this category aggregate trust using a weighed calculation on information collected from different sources (e.g., direct observation vs. indirect observation [3], past experience vs. recent experience, etc.). The weight parameters are determined by factors such as the trustworthiness of the information provider, the rate of trust decay, etc. For example, eBay [16] employs this model to calculate the feedback score. The advantages of this kind of models are, first it is simple and easy to understand, and second the linear calculation is easy to implement and efficient. However, it is a challenge to find the best weight parameters to achieve an accurate trust evaluation. Our dissertation research considers weighted summation as one of the many possible ways for trust formation and it seeks the best trust composition and formation to maximize application performance.

2.1.3 Game Theory Models

Game theory based trust models [7] usually use incentives to stimulate the cooperation between nodes, such that the system can reach a stable state where the overall utility is maximized. However, these models only consider selfish nodes and cannot deal with malicious nodes that intend to disrupt the system functionality. Staab, et al. [13] proposed a trust model by considering a game between normal nodes and attackers, given the knowledge of the strategies that attackers will use in each system configuration. Their model can be used to find the optimal parameters for an evidence based trust model to maximize the expected utility. However, in reality, it is difficult to obtain a complete set of attacker strategies and the attacker behavior may change dynamically. In our dissertation research we do not make assumptions of the attacker strategies. Rather, we design dynamic trust management protocols that can learn from past experiences and adapt to changing environment conditions to maximize application performance and enhance operation agility.

2.1.4 Information Theory Models

In information theory models [17], trust is considered as a measure of certainty of whether the trustee will perform an action in the trustor’s point of view. Depending on the way of aggregating trust, there are two trust models: entropy-based and probability based. In the entropy-based trust model, trust is calculated as the entropy of information (recommendations) from others. In the probability-based model, trust is obtained by aggregating recommendations using conditional probability. Similar to Bayesian trust management, information theory models do not have direct trust vs. indirect trust as design parameters and only address trust aggregation protocol design. In our dynamic trust management, we consider the design of trust composition, trust propagation, trust aggregation and trust formation protocols.

2.2 Information centric networks (ICN) architecture.

Information-Centric Networking (ICN) has emerged as a promising candidate for the architecture of the Future Internet. Inspired by the fact that the Internet is increasingly used for information dissemination, rather than for pair-wise communication between end hosts, ICN aims to reflect current and future needs better than the existing Internet architecture. By naming information at the network layer, ICN favours the deployment of in-network caching deployment of in-network caching (or storage, more generally) and multicast mechanisms, thus facilitating the efficient and timely delivery of information to the users. However, there is more to ICN than information distribution, with related research initiatives employing information-awareness as the means for addressing a series of additional limitations in the current Internet architecture, for example, mobility management and security enforcement, so as to fulfil the entire spectrum of Future Internet requirements and objectives.

Survey papers exist for research in the Future Internet area (e.g., [27] and [28]), due to their broad coverage they treat ICN architectures and related research efforts either sketchily or incompletely. The aim of this survey is to focus on ICN and cover the state-of-the-art evenly, broadly, and at some depth. Compared to other ICN surveys (e.g. [29] and [30]) the present survey covers in more detail and depth the most representative and mature ICN architectures and approaches, instead of a subset. In addition to describing the goals and basic concepts of the various research projects on ICN, it identifies the core functionalities of all ICN architectures and highlights their similarities and differences in how these functionalities are implemented. Furthermore, it provides a critical analysis of the main unresolved research challenges in ICN that require further attention by the community.

III. SYSTEM MODEL

We design a DTN environment with no centralized trusted authority. Nodes communicate through multiple hops. When a node encounters another node, they exchange encounter histories certified by encounter tickets so as to prevent black hole attacks to DTN routing. We differentiate socially selfish nodes from malicious nodes. A selfish node acts for its own interests including interests to its friends, groups, or communities. So it may drop packets arbitrarily just to save energy but it may decide to forward a packet if it has good social ties with the source, current carrier or destination node. We consider a friendship matrix to represent the social ties among nodes. Each node keeps a friend list in its local storage. A similar concept to the friendship relationship is proposed in [26], where familiar strangers are identified based on colocation information in urban transport environments for media sharing. Our work is different from [26] in that rather than by frequent colocation instances, friendship is established by the existence of common friends. Energy spent for maintaining friend lists and executing matching operations is negligible because energy spent for computation is very small compared with that for DTN communication and matching operations are performed only when there is a change to the friend lists. When a node becomes selfish, it will only forward messages when it is a friend of the source, current carrier, or the destination node, while a well-behaved node performs altruistically regardless of the social ties. A malicious node aims to break the basic DTN routing functionality. In addition to dropping packets, a malicious node can perform the following trust-related attacks:

1. **Self-promoting attacks:** it can promote its importance (by providing good recommendations for itself) so as to attract packets routing through it (and being dropped).

2. **Bad-mouthing attacks:** it can ruin the reputation of well-behaved nodes (by providing bad recommendations against good nodes) so as to decrease the chance of packets routing through good nodes.

3. **Ballot stuffing**: it can boost the reputation of bad nodes (by providing good recommendations for them) so as to increase the chance of packets routing through malicious nodes (and being dropped).

A malicious attacker can perform random attacks to evade detection. We introduce a random attack probability P(rand) to reflect random attack behavior. When P(rand) <1, the malicious attacker is a reckless attacker; when P(rand) < 1 it is a random attacker.

A collaborative attack means that the malicious nodes in the system boost their allies and focus on particular victims in the system to victimize. Ballot stuffing and bad-mouthing attacks are a form of collaborative attacks to the trust system to boost the reputation of malicious nodes and to ruin the reputation of (and thus to victimize) good nodes. We mitigate collaborative attacks with an application-level trust optimization design by setting a trust recommender threshold Trec to filter out less trustworthy recommenders, and a trust carrier threshold Tf to select trustworthy carriers for message forwarding. These two thresholds are dynamically changed in response to environment changes.

A created for the Application Developers, Enabling them to build and test the system. Many organizations look at System Design primarily as the preparation of the system component specifications; however, constructing the various system components is only one of a set of major steps in successfully building a system. The preparation of the environment needed to build the system, the testing of the system, and the migration and preparation of the data that will ultimately be used by the system are equally important. In addition to designing the technical solution, System Design is the time to initiate focused planning efforts for both the testing and data preparation activities.

Software design is an important activity in SDLC (Software Development Life Cycle). This is the third activity that emphasize on the requirements that are analyzed in Requirement Analysis Phase by building models that provides initial glimpse of what the real system looks like. Some of the software development activities are.

- Planning
- Implementation
- Testing and Documentation
- Deployment

IV. DYNAMIC TRUST AND SECURITY MANAGEMENT PROTOCOL (DTSMP)

4.1 DTSMP

Our trust and security protocol considers trust composition, trust aggregation, trust formation and application-level trust optimization designs. Figure 1 shows a flowchart of our trust management protocol execution. For trust composition design (described in the top part of Figure 1), we consider two types of trust properties:

- *QoS trust*: QoS trust [22] is evaluated through the communication network by the capability of a node to deliver messages to the destination node. We consider "connectivity" and "energy" to measure the QoS trust level of a node. The connectivity QoS trust is about the ability of a node to encounter other nodes due to its movement patterns. The energy QoS trust is about the battery energy of a node to perform the basic routing function.
- *Social trust*: Social trust [22, 25] is based on honesty or integrity in social relationships and friendship in social ties. We consider "healthiness" and social "unselfishness" to measure the social trust level of a node. The healthiness social trust is the belief of whether a node is malicious. The unselfishness social trust is the belief of whether a node is socially selfish. While social ties cover more than just friendship, we consider friendship as a major factor for determining a node's socially selfish behavior.

The selection of trust properties is application driven. In DTN routing, message delivery ratio and message delay are two important factors. We consider "healthiness", "unselfishness", and "energy" in order to achieve high message delivery ratio, and we consider "connectivity" to achieve low message delay.

We define a node's trust level as a real number in the range of [0, 1], with 1 indicating complete trust, 0.5 ignorance, and 0 complete distrust. We consider a trust formation design (described in the middle part of Figure 1) by which the trust value of node *j* evaluated by node *i* at time *t*, denoted as $T_{i,j}$(t) is computed by a weighted average of healthiness, unselfishness, connectivity, and energy as follows:

$$T_{i,j}(t) = \sum_{X}^{all} W^{X} \times T_{i,j}^{X}(t)$$

where X represents a trust property explored ðX ¼ healthiness, unselfishness, connectivity or energy), Tx is node i's trust in trust property X toward node j, and wX is the weight associated with trust property X with the sum equal to 1. wX is application-dependent. However, it is not related to the application priority [22] but dependent on the operational profile of an application [23]. We aim to identify the best weight ratio under which the application performance (secure routing) is maximized, given an operational profile [23] as input. Before this can be achieved, however, one must address the accuracy issue of trust aggregation. That is, for each QoS or social trust property X, we must devise and validate the trust aggregation protocol executed by a trustor node to assess X of a trustee node such that the trust value computed is accurate with respect to actual status of the trustee node in X. This is achieved by devising a trust propagation protocol (described in the middle part of Fig. 1) with tunable parameters which can be adjusted based on each trust property.

The design principles of DTN-ICN are described below:

1. We design the service abstraction that is provided to applications by defining an information model, as well as a service model, that is exposed to them. We utilise existing DTN and ICN solutions as a basis for this common abstraction, providing an object-level graph-based information abstraction. Information is split into several items or objects and each such object is associated with a context (also known as scoping). Scope represents sets of information. Both information objects a nd scopes are represented as directed acyclic graphs (DAG) manipulated through a set of publish/subscribe operations. While we expect applications to natively utilise this common information-centric interface of the architectural framework, we also foresee interfaces being defined that allow, or example, socket emulation [21] that would enable backward compatibility.
2. We functionally decompose the network components using PURSUIT ICN and existing DTN (Bundle Protocol [20]), into three core functions, namely rendezvous, topology management and forwarding. The functional decomposition also addresses the interaction with the underlying networks, such as satellite, cellular, WiFi or optical networks. This is accomplished mainly through the topology management function, which manages the resources available in the form of links, spectrum, wavelength but also storage and computational capability.

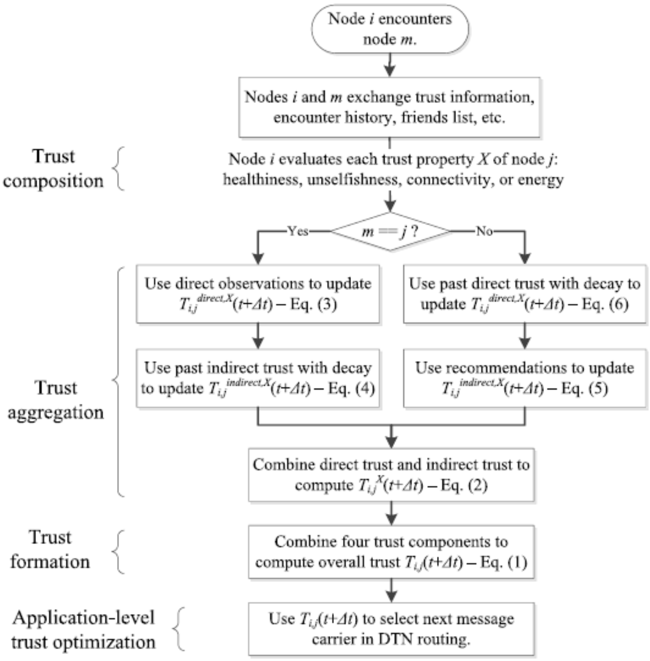

Node i encounters node m.
Nodes i and m exchange trust information, encounter history, friends list, etc.
Trust composition
Node i evaluates each trust property X of node j: healthiness, unselfishness, connectivity, or energy
$m == j$?
Yes
No
Trust aggregation
Use direct observations to update $T_{i,j}^{direct,X}(t+\Delta t)$ – Eq. (3)
Use past direct trust with decay to update $T_{i,j}^{direct,X}(t+\Delta t)$ – Eq. (6)
Use past indirect trust with decay to update $T_{i,j}^{indirect,X}(t+\Delta t)$ – Eq. (4)
Use recommendations to update $T_{i,j}^{indirect,X}(t+\Delta t)$ – Eq. (5)
Combine direct trust and indirect trust to compute $T_{i,j}^{X}(t+\Delta t)$ – Eq. (2)
Trust formation
Combine four trust components to compute overall trust $T_{i,j}(t+\Delta t)$ – Eq. (1)
Application-level trust optimization
Use $T_{i,j}(t+\Delta t)$ to select next message carrier in DTN routing.

Figure 1: A Flowchart for Trust Protocol Execution.

3. Based on our decomposition, we define the interfaces between the core components of our architectural framework, e.g., for initiating discovery requests, assembling network resources for store-and-forward operations or forwarding information objects over paths that were assembled through the topology

management function. These interfaces are realised through various dissemination strategies that enable traversal across the various connectivity options, e.g., over challenged and opportunistic network environments (using DTN), IP-based backhauls (IP being used as a 'framing' (link layer) based backhauls (IP being used as a 'framing' (link layer) protocol) or using native ICN for high speed optical links.

V. RESULTS

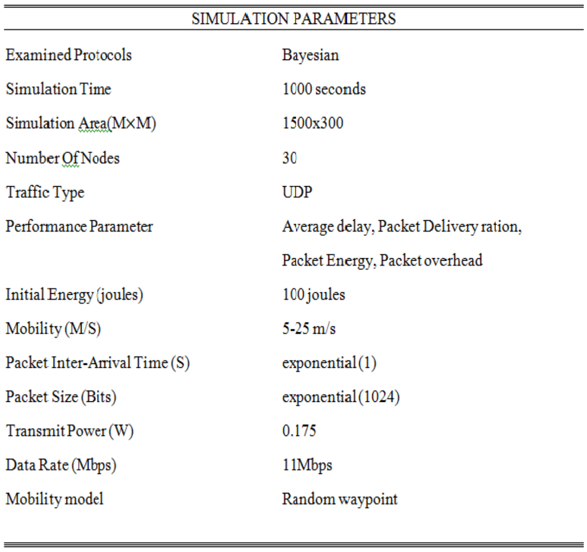
5.1 Simulation Setup

As in our paper we going to compare the three protocols Bayesian, DTM and DTM-ICN for each protocol the simulation setup is explained as follows.

Bayesian Trust protocol: The simulation parameter for Bayesian is as shown in Table 5.1. A single scenario comprising of 30 mobile nodes moving at a variable speed from 5 meter per seconds to 25 meter per second. The number of node can be select explicitly and even the mobility (m/s) can be set explicitly, Simulation time was taken 1000 seconds. Simulation area taken is 1500 x 300 meters. Packet Inter-Arrival Time (sec) is taken exponential (1) and packet size (bits) is exponential (1024). The data rates of mobile nodes are 11 Mbps with the default transmitting power of 0.175 watts. Random way point mobility is selected.

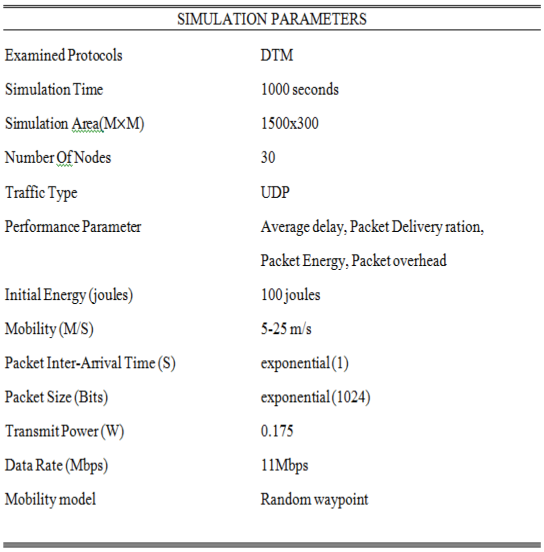
DTM Protocol: The simulation parameter for DTM is as shown in Table 5.2. A single scenario comprising of 30 mobile nodes moving at a variable speed from 5 meter per seconds to 25 meter per second. The number of node can be select explicitly and even the mobility (m/s) can be set explicitly, Simulation time was taken 1000 seconds. Simulation area taken is 1500 x 300 meters. Packet Inter-Arrival Time (sec) is taken exponential (1) and packet size (bits) is exponential (1024). The data rates of mobile nodes are 11 Mbps with the default transmitting power of 0.175 watts. Random way point mobility is selected.

DTM-ICN: DTM Protocol: The simulation parameter for DTM-ICN is as shown in Table 5.3.A single scenario comprising of 32 mobile nodes moving at a variable speed from 5 meter per seconds to 25 meter per second where 2 nodes act as servers in topology. The number of node can be select explicitly and even the mobility (m/s) can be set explicitly, Simulation time was taken 1000 seconds. Simulation area taken is 1500 x 300 meters. Packet Inter-Arrival Time (sec) is taken exponential (1) and packet size (bits) is exponential (1024). The data rates of mobile nodes are 11 Mbps with the default transmitting power of 0.075 watts. Random way point mobility is selected.



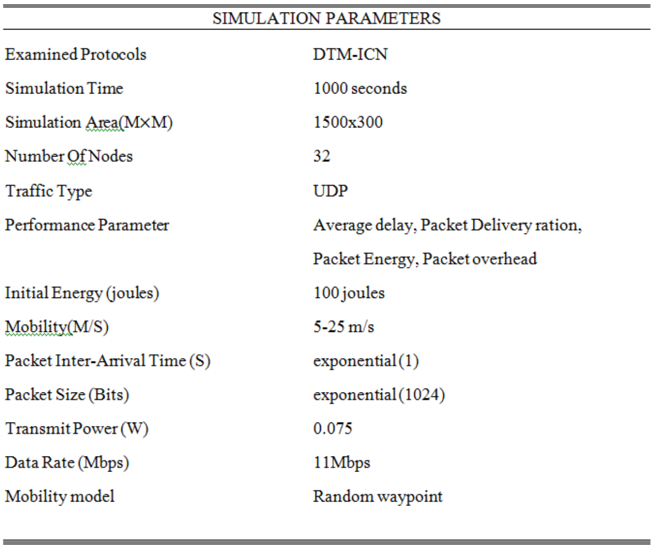
SIMULATION PARAMETERS	
Examined Protocols	Bayesian
Simulation Time	1000 seconds
Simulation Area(M×M)	1500x300
Number Of Nodes	30
Traffic Type	UDP
Performance Parameter	Average delay, Packet Delivery ration, Packet Energy, Packet overhead
Initial Energy (joules)	100 joules
Mobility (M/S)	5-25 m/s
Packet Inter-Arrival Time (S)	exponential (1)
Packet Size (Bits)	exponential (1024)
Transmit Power (W)	0.175
Data Rate (Mbps)	11Mbps
Mobility model	Random waypoint

Table 5.1 Bayesian Simulation Parameters



SIMULATION PARAMETERS	
Examined Protocols	DTM
Simulation Time	1000 seconds
Simulation Area(M×M)	1500x300
Number Of Nodes	30
Traffic Type	UDP
Performance Parameter	Average delay, Packet Delivery ration, Packet Energy, Packet overhead
Initial Energy (joules)	100 joules
Mobility (M/S)	5-25 m/s
Packet Inter-Arrival Time (S)	exponential (1)
Packet Size (Bits)	exponential (1024)
Transmit Power (W)	0.175
Data Rate (Mbps)	11Mbps
Mobility model	Random waypoint

Table 2.5 DTM Simulation Parameters



SIMULATION PARAMETERS	
Examined Protocols	DTM-ICN
Simulation Time	1000 seconds
Simulation Area(M×M)	1500x300
Number Of Nodes	32
Traffic Type	UDP
Performance Parameter	Average delay, Packet Delivery ration, Packet Energy, Packet overhead
Initial Energy (joules)	100 joules
Mobility(M/S)	5-25 m/s
Packet Inter-Arrival Time (S)	exponential (1)
Packet Size (Bits)	exponential (1024)
Transmit Power (W)	0.075
Data Rate (Mbps)	11Mbps
Mobility model	Random waypoint

Table 5.3 DTM-ICN Simulation Parameter

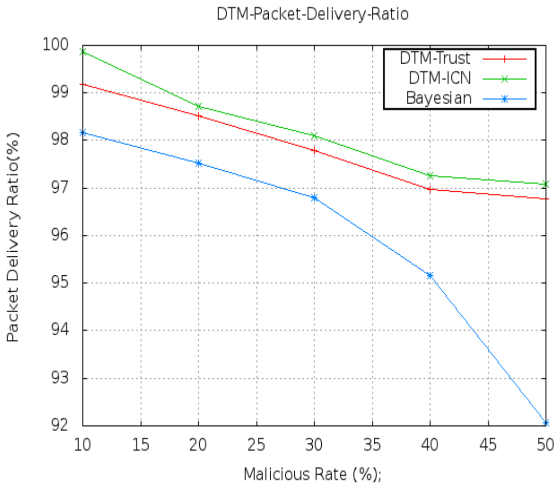
5.2 Comparative Analysis

we conduct a comparative analysis, contrasting our trust and security-based protocol operating under the best settings identified with Bayesian trust-based routing [12, 15], Dynamic trust and security management protocol without ICN architecture and with INC architecture. Bayesian trust-based routing relies on the use of trust information maintained by a Bayesian based trust management system (such as a Beta reputation system [12, 15]) to make routing decisions. In a Bayesian trust management system, the trust value is assessed using the Bayes estimator, updated by both direct observations and indirect recommendations. The direct observations are directly used to update the number of positive and negative observations, whereas the recommendations are discounted by the confidence [12] or belief [15] of the trustor toward the recommender. Under Bayesian trust-based routing, a node is chosen as the message carrier only if its trust value is in the top Ω percentile and higher than the message carrier trust threshold T_f.

Figure 2 compares the packet delivery ratio of Bayesian, DTM and DTM-ICN.The results demonstrate that our trust-based secure routing protocol designed to maximize delivery ratio, As compare to our protocol to Bayesian trust-based protocol and DTM protocols have less performance degradation in message delivery ratio.

Figure 3 compare the Packet Average delay of Bayesian, DTM and DTM-ICN.The results demonstrate that our trust-based secure routing protocol designed to minimize the Average delay, As compare to our protocol to Bayesian trust-based protocol and DTM protocols have less performance degradation Average delay.

Figure 4 compare the Packet Overhead of Bayesian, DTM and DTM-ICN.The results demonstrate that our trust-based secure routing protocol designed to minimize the packet overhead, As compare to Bayesian and DTM protocols.



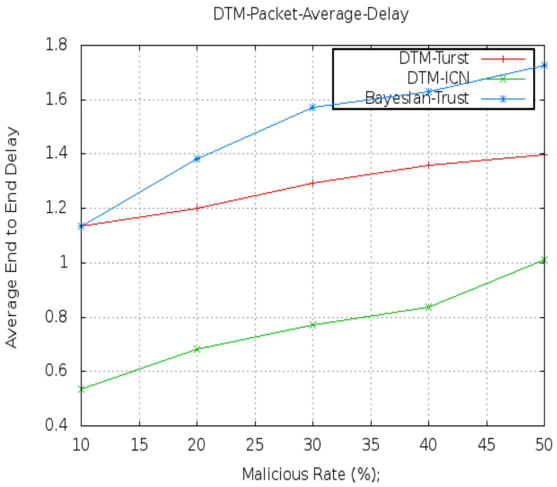


Figure 2 : Packet Delivery Ratio **Figure 3: Packet Average Delay**

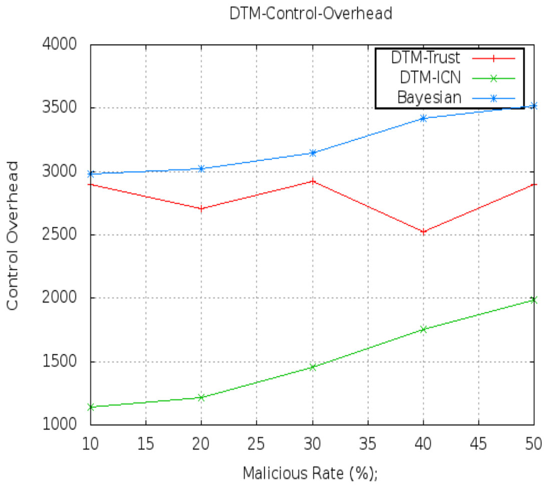


Figure 4: Packet Control Overhead

VI. CONCLUSION AND FUTURE WORK

In this paper, we designed and validated a trust and security management protocol using Information Centric-Network (ICN) architecture for DTNs and applied it to secure routing to demonstrate its utility. Our trust management protocol combines QoS trust with social trust to obtain a composite trust metric. We demonstrated how the results obtained at design time can facilitate dynamic trust management for DTN routing in response to dynamically changing conditions at runtime. We performed a comparative analysis of trust-based secure routing running on top of our trust management protocol with Bayesian trust-based routing and DTM routing in DTNs. Our results backed by simulation validation demonstrate that our trust-based secure routing protocol outperforms Bayesian trust-based routing. Our protocol approaches the ideal performance of epidemic routing in delivery ratio and message delay without incurring high message or protocol maintenance overhead.

There are several future research areas including (a) exploring other trust-based DTN applications with which we could further demonstrate the utility of our dynamic trust management protocol design; (b) designing trust management for DTNs considering social communities and performing comparative analysis with more recent works such as [2, 3].

REFERENCES

[1] "The ns-3 Network Simulator," Nov. 2011, http://www.nsnam.org/.

[2] E. Aivaloglou, and S. Gritzalis, "Trust–Based Data Disclosure in Sensor Networks," IEEE International Conference on Communications, 2009, pp. 1-6.

[3] E. Aivaloglou, and S. Gritzalis, "Hybrid Trust and Reputation Management for Sensor Networks," Wireless Networks, vol. 16, no. 5, July 2010, pp. 1493-1510.

[4] J. N. Al-Karaki, and A. E. Kamal, "Routing Rechniques in Wireless Sensor Networks: A Survey," IEEE Wireless Communications, vol. 11, no. 6, Dec. 2004, pp.6-28.

[5] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A Survey," Computer Networks, vol. 54, no. 15, Oct. 2010, pp. 2787-2805.

[6] E. Ayday, H. Lee, and F. Fekri, "Trust Management and Adversary Detection for Delay Tolerant Networks," Military Communications Conference, 2010, pp. 1788-1793.

[7] S. Braynov, and T. Sandholm, "Contracting with Uncertain Level of Trust,"Computational Intelligence, vol. 18, no. 4, 2002, pp. 501-514.

[8] J. Carbo, J. M. Molina, and J. Davila, "Trust Management Through Fuzzy Reputation," International Journal of Cooperative Information Systems, vol. 12, no. 1,2003, pp. 135-155.

[9] V. Cerf, S. Burleigh, A. Hooke, L. Torgerson, R. Durst, K. Scott, K. Fall, and H. Weiss, "Delay-Tolerant Networking Architecture," RFC 4838, IETF, 2007.

[10] T. Chen, F. Wu, and S. Zhong, "FITS: A Finite-Time Reputation System for Cooperation in Wireless Ad Hoc Networks," IEEE Transactions on Computers, vol. 60, no. 7, July 2011, pp. 1045-1056.

[11] I. Chlamtac, M. Conti, and J. J.-N. Liu, "Mobile Ad Hoc Networking: Imperatives and Challenges," Ad Hoc Networks, vol. 1, no. 1, 2003, pp. 12-64.

[12] M. K. Denko, T. Sun, and I. Woungang, "Trust Management in Ubiquitous Computing: A Bayesian Approach," Computer Communications, vol. 34, no. 3, 2011, pp. 398-406.

[13] T. E. Eugen Staab, "Tuning Evidence-Based Trust Models," International Conference on Computational Science and Engineering, Vancouver, Canada, August 2009, pp. 92-99.

[14] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, "Reputation-Based Framework for High Integrity Sensor Networks," ACM Transactions on Sensor Networks, vol. 4, no. 3, May 2008, pp. 1-37.

[15] A. Josang, and R. Ismail, "The Beta Reputation System," Bled Electronic Commerce Conference, Bled, Slovenia, June 17-19 2002, pp. 1-14.

[16] P. Resnick, and R. Zeckhauser, "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System," Advances in Applied Microeconomics, vol. 11, no. 12, 2002, pp. 127-157.

[17] Y. L. Sun, W. Yu, Z. Han, and K. J. R. Liu, "Information Theoretic Framework of Trust Modeling and Evaluation for Ad Hoc Networks," IEEE Journal on Selected Areas in Communications, vol. 24, no. 2, Feb. 2006, pp. 305-317.

[18] G. The odorakopoulos, and J. S. Baras, "On Trust Models and Trust Evaluation Metrics for Ad Hoc Networks," IEEE Journal on Selected Areas in Communications, vol. 24, no. 2, Feb. 2006, pp. 318-328.

[19] C. Zhang, X. Zhu, Y. Song, and Y. Fang, "A Formal Study of Trust-Based Routing in Wireless Ad Hoc Networks," IEEE Conference on Computer Communications, March 2010, pp. 1-9.

[20] K. Scott, S. Burleigh, ”Bundle Protocol Specification”, IETF FC 5050, experimen-tal, November 2007, http://www.ietf.org/rfc/rfc5050.txt.

[21] G. Xylomenos, B. Cici, Design and Evaluation of a Socket Emulator for Pub-lish/Subscribe Networks, Proc. of the Future Internet Symposium, 2010.

[22] J. H. Cho, A. Swami, and I. R. Chen, “A Survey on Trust Management for Mobile Ad Hoc Networks,” *IEEE Communications Surveys & Tutorials,* vol. 13, no. 4, 2011, pp. 562-583.

[23] J. D. Musa, “Operational Profiles in Software-Reliability Engineering,” *IEEE Software,* vol. 10, no. 2, March 1993, pp. 14-32.

[24] I. Psaras, L. Wood, and R. Tafazolli, *Delay-/Disruption-Tolerant Networking: State of the Art and Future Challenges*, Dept. of El. Eng., University of Surrey, 2009.

[25] S. Trifunovic, F. Legendre, and C. Anastasiades, “Social Trust in Opportunistic Networks,” *IEEE Conference on Computer Communications Workshops*, San Diego, CA, USA, March 2010, pp. 1-6.

[26] L. McNamara, C. Mascolo, and L. Capra, “Media Sharing Based on Colocation Prediction in Urban Transport,” Proc. 14th Ann. Int’l Conf. Mobile Computing and Networking, 2008.

[27] P. Stuckmann and R. Zimmermann, “European research on future Internet design,” IEEE Wireless Commun., vol. 16, no. 5, pp. 14–22, October 2009.

[28] J. Pan, S. Paul, and R. Jain, “A survey of the research on future Internet architectures,” IEEE Commun. Mag., vol. 49, no. 7, pp. 26–36, July 2011.

[29] J. Choi, J. Han, E. Cho, T. Kwon, and Y. Choi, “Survey on content-oriented networking for efficient content delivery,” IEEE Commun. Mag., vol. 49, no. 3, pp. 121–127, March 2011.

[30] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman,“A survey of information-centric networking,” IEEE Commun. Mag., vol. 50, no. 7, pp. 26–36, July 2012

AUTOMATIC DETECTION OF FAKE PROFILES

[1]Sanju, [2] Dinesh

[1,2]Computer Science & Engineering Department, M.D.U, (India)

ABSTRACT

This paper presents the study of various methods for detection of fake profiles.In this paper a study of various papers is done, and in the reviewed paper we explain the algorithm and methods for detecting fake profiles for security purpose. The main part of this paper covers the security assessment of security on social networking sites.

Keywords: Objective, Problem Statement, Scope, Conclusion, Survey

I. INTRODUCTION

A social networking site is a website where each user has a profile and can keep in contact with friends, share their updates, meet new people who have the same interests. These Online Social Networks (OSN) uses web2.0 technology, which allows users to interact with each other.

These social networking sites are growing rapidly and changing the way people keep in contact with each other. The online communities bring people with same interests together which makes users easier to make new friends.There are no feasible solution exist to control these problems. In this project, we came up with a framework with which automatic detection of fake profiles is possible and is efficient framework uses classification techniques like Support Vector Machine, Nave Bayes and Decision trees to classify the profiles into fake or genuine classes. As, this is an automatic detection method,it can be applied easily by online social networks which has millions of profiles whose profiles can not be examined manually.. These social networking sites are growing rapidly and changing the way people keep in contact with each other. The online communities bring people with same interests together which makes users easier to make new friends. In the present generation, the social life of everyone has become associated with the online social networks. These sites have made a drastic change in the way we pursue our social life. Adding new friends and keeping in contact with them and their up- dates has become easier.

The online social networks have impact on the science, education, grassroots organizing, employment, business, etc. Researchers have been studying these online social networks to see the impact they make on the people. Teachers can reach the students easily through this making a friendly environment for the students to study

II OBJECTIVE

In todays online social networks there have been a lot of problems like fake profiles,online impersonation, etc. Till date, no one has come up with a feasible solution to these problems.In this project we intend to give a framework with which the automatic detection of fake profiles can be done so that the social life of people become secured and by usingthis automatic detection technique we can make it easier for the sites to manage the huge number of profiles, which cant be done manually.

III. LITERATURE SURVEY

Fake profiles are the profiles which are not genuine i.e. they are profiles of persons who claim to be someone they are not, doing some malicious and undesirable activity,causing problems to the social network and fellow users.

Social Engineering in terms of security means the art of stealing confidential information from people or gaining access to some computer system mostly not by using technical skills but by manipulating people themselves in divulging information. The hacker doesnt need to come face to face with the user to do this.

The social engineering techniques are like Pretexting, Diversion theft, phishing, baiting, quid pro quo, tailgating, etc. Social bots are semi-automatic or automatic computer programs that replicate the human behavior in OSN. These are used mostly by hackers now-a-days to attack online social networks. These are mostly used for advertising,campaigning purposes and to steal users personal data in a large scale.

These social bots communicate with each other and are controlled by a program called botmaster. The botmaster may or may not have inputs from a human attacker. The social bots look like human profiles with a randomly chosen.

IV. CLASSIFICATION

Classification is the process of learning a target function f that maps each records,x consisting of set of attributes to one of the predefined class labels, y. A classification technique is a approach of building classification models from an input data set. This technique uses a learning algorithm to identify a model that best fits the relationship between the attribute set and class label of the training set. The model generated by the learning algorithm should both fit the input data correctly and correctly predict the class labels of the test set with as high accuracy as possible. The key objective of the learning algorithm is to build the model with good generality capability

V. SCOPE

The proposed framework shows the sequence of processes that need
to be followed for continues detection of fake profiles with active leaning from the feedback of the result given by the classification algorithm. This framework can easily be implemented by the social networkingcompany.By using method and parameters fake profiles detection becomes easy.As a result of this cyber crime may be reduced.

VI. CONCLUSION

From the above study we conclude that we an detect the fake profiles on social networking sites

VII. ACKNOWLEDGMENT

I show my thanks to all the departments' personals and sponsors who give us an opportunity to present and express my paper on this level. We wish to place on my record my deep sense of gratitude to all reference papers authors for them valuable help through their papers, books, websites etc.

REFERENCES

[1] T. Stein, E. Chen, and K. Mangla. Facebook immune system. In Proceedings of the 4th Workshop on Social Network Systems, SNS, volume 11, page 8,2011.

[2] Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu. The socialbot network: when bots socialize for fame and money. In Proceedings of the 27th Annual Computer Security Applications Conference, pages 93{102. ACM, 2011.

[3] C. Wagner, S. Mitter, C. K□orner, and M. Strohmaier. When social bots attack: Modeling susceptibility of users in online social networks. InProceedings of the WWW, volume 12, 2012.

[4] G. Kontaxis, I. Polakis, S. Ioannidis, and E.P. Markatos. Detecting social network profile cloning. In Pervasive Computing and Communications Workshops (PERCOM Workshops),2011 IEEE International Conference on, pages 295{300. IEEE, 2011.

[5] A. Wang. Detecting spam bots in online social networking sites: a machine learning approach.Data and Applications Security and Privacy XXIV, pages335{342, 2010.

[6] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B.Y. Zhao. Detecting and characterizing social spam campaigns. In Proceedings of the 10th annual conference on Internet measurement, pages 35{47. ACM, 2010.

[7] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia. Who is tweeting on twitter: human, bot, or cyborg? In Proceedings of the 26th Annual Computer Security Applications Conference, pages 21{30. ACM, 2010.

[8] S. Krasser, Y. Tang, J. Gould, D. Alperovitch, and P. Judge. Identifying image spam based on header and _le properties using c4. 5 decision trees and support vector machine learning.In Information Assurance and Security Workshop, 2007. IAW'07. IEEE SMC, pages 255{261.IEEE, 2007.

[9] G.K. Gupta. Introduction to Data Mining with Case Studies. Prentice Hall India, 2008.

[10] Rajan Chattamvelli. Data Mining Methods. Narosa, 2010.

[11] Spies create fake facebook account in nato chief's name to steal personal details, http://in.news.yahoo.com/spies-create-fake-facebook-account-nato-chiefs-name-114824955.html.

[12] Man arrested for uploading obscene images of woman colleague, http://www.ndtv.com/article/andhra-pradesh/man-arrested-for-uploading obscene-images-of-woman-colleague-173266.

[13] How obamas internet campaign changed politics, /bits.blogs.nytimes.com/2008/11/07/how-obamas-internet-campaign-changed-politics.

[14] S. Nagaraja, A. Houmansadr, P. Piyawongwisal, V. Singh, P. Agarwal, and N. Borisov. Stegobot:A covert social network botnet. In Information Hiding, pages 299{313. Springer, 2011.

[15] M. Huber, M. Mulazzani, and E. Weippl. Who on earth is mr. cypher: Automated friend injection attacks on social networking sites. Security and Privacy{Silver Linings in the Cloud, pages 80{89, 2010. Y. Boshmaf, I. Muslukhov, K. Beznosov, and M. Ripeanu. The socialbot network: when bots socialize for fame and money. In Proceedings of the 27th Annual Computer Security Applications Conference, pages 93{102. ACM, 2011.

[16] C. Wagner, S. Mitter, C. K□orner, and M. Strohmaier. When social bots attack: Modeling susceptibility of users in online social networks. In Proceedings of the WWW, volume 12, 2012.

[17] G. Kontaxis, I. Polakis, S. Ioannidis, and E.P. Markatos. Detecting social network profile cloning. In Pervasive Computing and Communications Workshops (PERCOM Workshops),2011 IEEE International Conference on, pages 295{300. IEEE, 2011.

[18] A. Wang. Detecting spam bots in online social networking sites: a machine learning approach. Data and .

[19] H. Gao, J. Hu, C. Wilson, Z. Li, Y. Chen, and B.Y. Zhao. Detecting and characterizing social spam campaigns. In Proceedings of the 10th annual conference on Internet measurement, pages 35{47. ACM, 2010.

[20] Z. Chu, S. Gianvecchio, H. Wang, and S. Jajodia. Who is tweeting on twitter: human, bot, or cyborg? In Proceedings of the 26th Annual Computer Security Applications Conference, pages 21{30. ACM, 2010.

[21] S. Krasser, Y. Tang, J. Gould, D. Alperovitch, and P. Judge. Identifying image spam

ONTOLOGY ORIENTED DIAGNOSTIC MODEL FOR MEDICINE BASED ON RELATION REFINEMENT

Kavita[1], Ankush[2], Sachin[3]

[1,2]Computer Science & Engineering Department, M.D.U, (India)

ABSTRACT

We define the diagnosis in Traditional Chinese Medicine (TCM) as discovering the fuzzy relations between symptoms and syndromes. An Ontology-oriented Diagnosis System (ODS) is created to address the knowledge-based diagnosis based on a well-defined ontology of syndromes. The ontology transforms the implicit relationships among syndromes into a machine-interpretable model. The clinical data used for feature selection is collected .

Keywords: Introduction, OODS, Problem, Conclusion, Model

I. INTRODUCTION

Ontology Oriented is an emerging Knowledge Engineering paradigm. It aims at the discovering, documenting and maintaining a set of requirements. Knowledge Engineering (KE) is believed to promote better achievement of Requirement Document and thus to improve the quality of software systems and the efficiency of software development. As a result, Ontology Oriented (OO) is gaining increasing attention from academic organizations as well as from industrial companies' research firms.

Ontology-Oriented consists of a series of development discipline, covering a wide span of development activities in the software development lifecycle. OOD focuses on the systematic, identification, modularization, composition and analysis of reusable requirement which are evident at the requirement engineering stage. Hundreds of papers regarding OO have been published [23]. Researchers in OOD communities believe that identifying and capturing reusable requirement early on, at the requirement engineering stage, will benefit downstream development activities such as architecture design and implementation [6]. The identified reusable requirement may offer valuable insight at the architecture design and implementation stages. They often eventually correspond to reusable requirement in architecture, and then in code. As a result, pedigrees of reusable requirement throughout the entire software development lifecycle will be established, improving the traceability of a wide range of requirements in a software system and facilitating the system's resolvability. Moreover, identifying reusable requirement at the requirement engineering stage may help to reveal the scope of each requirement, detect potential conflicts between reusable requirements and support trade-off negotiation.

However, there is limited evidence that early identification of requirements is a productive software engineering practice. First of all, although a great amount of literature on OOD has been published worldwide, none of these papers, to the best of our knowledge, addresses this question. The Knowledge of early identification and management of reusable requirement may outweigh the benefits. Intuitively, an argument can be made that this is especially the case when requirements are not fully developed – when there is a large amount of uncertainty and volatility. Secondly, most proposed Algorithm approaches in the literature are supported with small scale, simplified, and sometimes artificial examples. Therefore, there is no convincing evidence that proposed Algorithm approaches are feasible and productive when applied to larger scale real world projects. A careful analysis of a real world software requirement document could provide some insight into the value or lack of value of the proposed SRR approaches.

Moreover, most proposed OOD approaches in the literature aim at identifying and thus capturing reusable requirement from well structured, formal (or at least semi-formal) knowledge engineering requirement documents that are organized. Only a few Knowledge reuse approaches deal with reusable requirement in less structured requirement documents such as informal software requirement specifications. There is not sufficient evidence to show that identifying and capturing reusable requirement is feasible for less structured and informal requirement documents.

II. OODS (ONTOLOGY-ORIENTED DIAGNOSIS SYSTEM)

Applying mathematical models and information technologies to medical intelligence has long been a hot spot in the academic research domains and real-life health care applications. Plenty of the efforts in this field allow researchers and medical practitioners to identify required information more efficiently, discover new substances or relationships, and integrate different sources of information more easily.

Traditional Medicine, known as TM, that is an ancient and unique branch of medical science, which probably covers broad range of practices such as herb medicine, acupuncture, attracts much attention on how to defeat the inconsistency of therapeutic patterns and vagueness of medical terms in TM in order to improve the user experience of TM diagnosis through the traditional methods and therapy.The basic theories of TM are based on the ancient philosophy of holistic understanding of the universe and the human body, which are commonly associated with the flow and the balancing of complementary opposites, cough and cold, or five elements. Based on the interacting forces in the human body, the forms of many pathological conditions in TM usually differ from that of identifiable diseases in terminology of the Western medicine. “Syndrome” (also called pattern) refers to a pattern of disharmony or functional disturbance within the functional entities that the TM model of the body is composed of. A “disease” in TM refers to disease entity or disease category, focusing on the macroscopic classification of specific manifestations. In TM diagnosis, the therapy is determined mainly according to the pattern rather than the disease. Two patients with the same disease but different patterns will receive different therapy; vice versa patients with similar patterns might receive similar therapy even if their diseases are different. Hence, the concept of “syndrome differentiation” is of great importance in deciding the diagnosis for patients. But the difficulty of differentiating syndromes is rooted in the unclear definition of the range of syndromes, which means the determination of specific syndromes, the similarity among syndrome patterns, and this information are usually hidden in the literature texts or the doctors’ experience.

III.PROPOSED MODEL FOR OODS

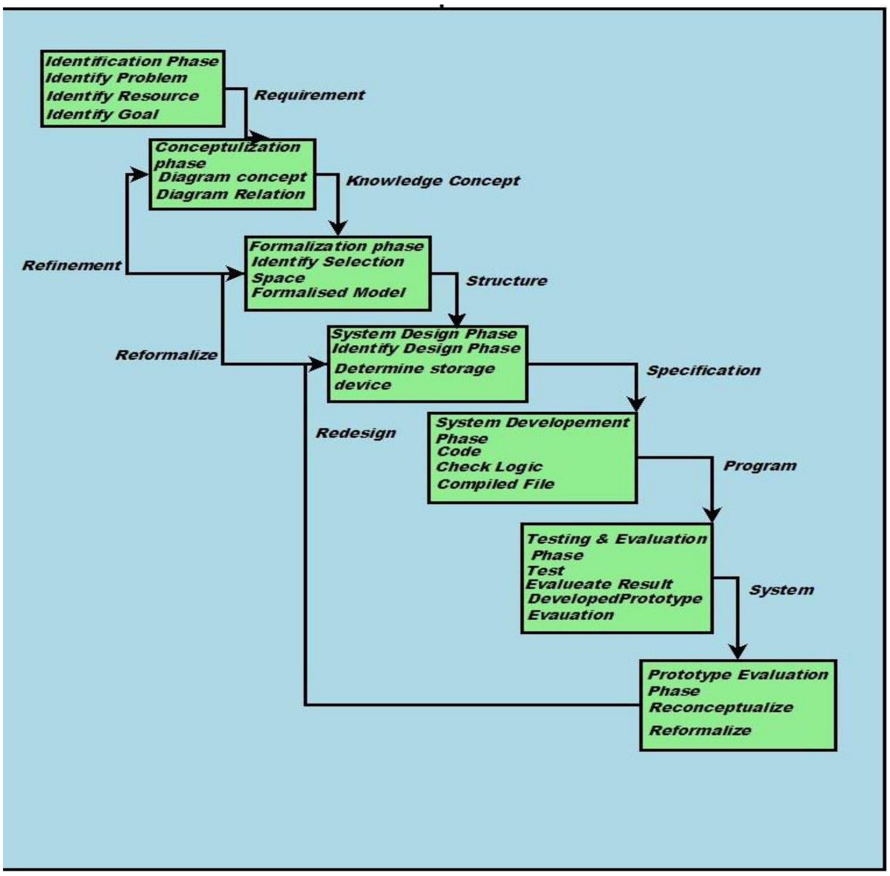


Fig.3.1: Ontology-Oriented Diagnostic model

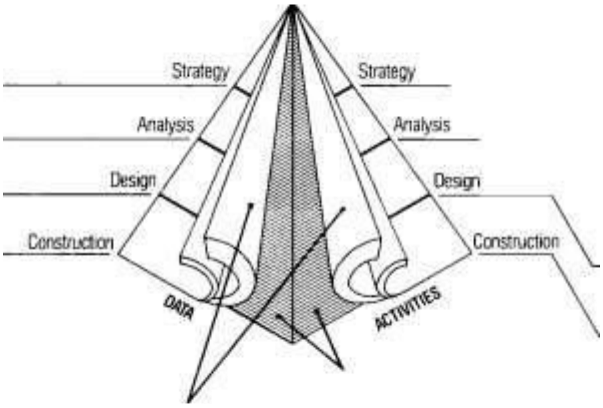


Fig3.2: Basic Model for Knowledge Based System Development

IV. PROBLEM OUTLINE

An alternative source of health care, TM is interpreted as to have intangible connections between human and nature rather than anatomical parts, which leads to complex semantic inclusion relations. We have done some researches on expressing biological facts into ontological statements, by constructing domain ontology as RDF models. A medical diagnosis problem can be formed as a probabilistic relation between a group of symptoms and a diagnosis.

Let be the set of symptoms to be queried, in which is the cardinality of, namely, the number of symptoms. Each is a professional description of a pathological symptom, such as cough and fever. The diagnostic result is an integrated answer of some disease and at least one syndrome, meaning that the patient catches a disease named with a collection of syndromes.

The difficulties in determining this relation between and lie on the obscure descriptions of each clinic case and uncertain (expertise-dependent) differentiation of syndromes. As an outsider of medical science, we are suggested by the medical experts that a relation table should be defined in advance, in which there are two kinds of relations and Represents the correspondence between a specific disease and a set of symptoms and represents the correspondence between a specific syndrome between a set of symptoms, and a table of these correspondence relations form a feature selection source for medical diagnosis. For example, we could obtain a minimum associated set of symptoms for each disease or syndrome such as by analyzing prior clinical data, which is formatted as (disease or syndrome, related symptoms). The process of clustering and filtering data is a variant of set intersection problem. We put our considerations on the performance and efficiency of the algorithm.

In order to rectify the data biases of clinical data and facilitate knowledge-based diagnosis, we define a knowledge base which takes the role of the conceptual model. In the procedure of automatic diagnosis, the user-queried symptoms are compared with the relation generated from the preprocessed relation table and the knowledge base. By calculating the similarity, the diagnostic result will be given as the combined answer of a disease and a group of syndromes.

The process of Chinese medical diagnosis can be concluded as an expertise-dependent case search based on observed pathological patterns, where expertise is consisted of practical experiences and theoretical principles. In this paper, we integrate both practical experiences and theoretical principles together to construct a complete computation model for reliable medical diagnosis.

V. RESEARCH OBJECTIVE

Objective 1: A feature set extraction method based on a revised adaptive set intersection algorithm for symptoms and syndromes.

Objective 2: A medical diagnosis can be formed as a probabilistic relation between a group of symptoms and a diagnosis.

VI. RESEARCH METHODOLOGY

KADS Methodology for developing AI (Knowledge Based) Systems. KADS traditionally has employed a step-wise Life Cycle Model (LCM), consisting of analysis, design, implementation, installation, use and maintenance.

Knowledge-based DSS is a category of DSS built using an expert system . These systems have their own expertise based on knowledge on many aspects of the problem: In the application domain, the definitions of problems within that domain or related to the domain and the necessary skill and methods proposed to solve them The knowledge of the system is often coded as a set of rules by one or more human experts: this kind of systems are referred to as rule-based expert systems.

Along with the development of Expert and Decision Support Systems, in recent years in bioinformatics a new type of tools, called Workflow Management Systems (WFMS) have begun to spread out. WFMSs provide a simple way to build and run a custom experiment and designed and tasted software using the most common bioinformatics resources, like online databases, software and algorithms.

The most used and famous WFMS for bioinformatics is Taverna it is able to automatically integrate tools on databases available both locally and on the web in order to build workflows of complex tasks; to run the workflows and to show results in different formats. The system works by means of a Graphical user Interface.

VII. ALGORITHM BASED ON ONTOLOGY

In this algorithm there is multiple set of symptoms and syndrome diagnose. Due to the generation development some diseases are not identified and there is some cases of symptoms and syndromes are enhanced. To refine these syndromes and symptoms this algorithm is designed. In this algorithm took a number of clinical cases and match the symptoms and syndromes. Hare through this algorithm is refined symptoms and syndromes resulted as compared to traditional medicines.

```
Input: multi-set
Output:
if   then
      return  null;
End
else if   then
       return  . get (0);
End
 = generateRatioNum (. size());
Let  largest elements of the multi-set ;
Let ;
Let  element of ;
Mark green all the sets  such that  = , and remove all copies of  from ;
Mark red all the sets  such that ;
Mark white all the sets  such that  < , and remove   from ;
While   do
    if    sets are red then
           if   sets are green then


           end
```

```
            Take  (white sets) of the green sets and mark them white;
            For each remaining green set , insert in  the first element of  which is , and change the
            mark of  to red;
            Let ;
            change the mark to green for all the sets which have as a representative in , and
    remove  from ;
        end
        Let  be the next white set;
        if   then
            Mark  in green;
        end
        else
            Insert in  the first element of  which is strictly larger than , and mark  in red;
        end
    End
```

VIII. CONCLUSION

The ontology assisted diagnostic system interprets the correspondence between symptoms and syndromes in an integrated method of minimum set mapping and ontology refinement, instead of static rules which are difficult to conclude. Web users could access the online user interface and fetch a diagnostic result according to the specific input symptoms.

IX. ACKNOWLEDGEMENT

I thanks to all who support me to work in this work.

REFERENCES

[1]. Potolea, R., Cacoveanu, S., Lemnaru, C., "Meta-learning Framework for Prediction Strategy Evaluation", LectureNotes in Business Information Processing, 2011, Volume 73, Part 3, 280- 295.

[2]. Potolea, R., Barbantan, I., Lemnaru, C.,"A Hierarchical Approach for the Offline Handwritten Signature Recognition", Lecture Notes in Business Information Processing, 2011, Volume 73, Part 3, 264-279.

[3]. Halalai, R., Lemnaru, C., Potolea, R., "Distributed community detection in social networks with genetic algorithms", Proceedings of the 2010 IEEE 6th International Conference on Intelligent Computer Communication and Processing, pp.35-41.

[4]. Vidrighin, B.C., Potolea, R., "Unified Strategy for Feature Selection and Data Imputation", Proceedings of SYNASC 2009, pp. 413-419.

[5]. C. VidrighinBratu, C. Savin, R. Potolea, "A Hybrid Algorithm for Medical Diagnosis". Proceedings of EUROCON 2007, pp. 668-673.

[6]. Y. J. Chen, Y. M. Chen and H. C. Chu, “Enabling collaborative product design through distributed engineering knowledge management,” Computer in Industry, vol. 59, no. 4, pp.395-409, 2008.

[7]. C. K. Mok, K. S. Chin and H. Lan, “An Internet-based intelligent design system for injection moulds,” Robotics and Computer-Integrated Manufacturing, vol. 24, no. 24, pp. 1-15, 2008.

[8]. D. L. Mcguinness and F. V. Harmelen, OWL Web Ontology Language Overview, World Wide Web.

[9]. E. Antezana, M. Kuiper, and V. Mironov, “Biological knowledge management: the emerging role of the Semantic Web technologies,” Briefings in Bioinformatics, vol. 10, no. 4, pp. 392–407, 2009.

[10]. G. Sun, The Basic Theory of Chinese Medicine, China Press of Traditional Chinese Medicine, 2005.

[11]. I. Salomie, M. Dinsoreanu, C. Rat, S. L. Suciu, "Efficient Ontology Processing Using Hierarchical Data Models Representation", The 5th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP 2009), ISBN 978-1-4244-5007-7, pp. 211-215

[12]. T. Cioara, I. Anghel, I. Salomie, M. Dinsoreanu, "A Context–based Semantically Enhanced Information Retrieval Model", The 5th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP 2009), ISBN 978-1-4244-5007-7, pp.291-298.

[13]. Y. J. Chen, Y. M. Chen, H. C. Chu and H. Y. Kao, “On technology for functional requirement-based reference design retrieval in engineering knowledge management,” Decision Support Systems, vol. 44, no. 4, pp. 798-816, 2007.

APPLICATION OF MULTI-OBJECTIVE GENETIC ALGORITHM FOR SOLVING OPTIMAL POWER FLOW PROBLEM

Saket Gupta[1], Laxmi Srivastava[2]

Department of Electrical Engineering MITS, Gwalior, (India)

ABSTRACT

This paper presentsmulti-objective genetic algorithm for solving the optimal power flow (OPF) problem. The proposed method is employed for optimal adjustment of the power system control variables which involve continuous variables of the OPF problem namely active power generation at the PV buses except at the slack bus, voltage magnitude at PV buses, tap settings of transformer and shunt VAR compensation. Solution of multi-objective optimization problem providesa number of trade-off solutions. The decision maker has an option to choose a solution among the different trade-off solutions provided in the Pareto-optimal front.The proposed method is tested in standard IEEE 30-bus test system with different objective such as fuel cost minimization, voltage stability enhancement and transmission losses minimization. The numerical results clearly show that the proposed method is capable to produce true and well distributed Pareto-optimal solutions for multi-objective OPF problem.

Keywords*: *Multi-Objective Genetic Algorithm, Multi-Objective Optimization, Optimal Power Flow,Pareto-Optimal Front.

I. INTRODUCTION

The concept of the optimal power flow (OPF) was first proposed by Carpenters [1] in the early 1960's based on the economic dispatch problem. Optimal Power Flow problem is one of the fundamental issues of power system operation, designed and planning. The main purpose of an OPF algorithm is to find steady state operation point which minimizes objective function, while satisfying various operating constraints [2].

Early, several conventional optimization techniques where apply to solve OPF problem such as linear programming (LP), quadratic programming (QP), nonlinear programming (NPL), Mixed Integer Programming (MIP), interior point method (IP) and Newton-based method. Generally, most of these approaches have been applied to solve OPF problem assuming convex, analytic, differentiable and linear. But unfortunately, OPF problem is a highly non-linear and a multi-modal optimization problem, i.e. there exist more than one local optimum. Hence, conventional optimization techniques are not suitable for such a problem and conventional optimization methods that make use of derivatives and gradients are in general not able to locate or identify the global optimum [3]. Hence, it becomes essential to develop optimization techniques that are able to overcome these drawbacks and handling such difficulties. Complex constrained optimization problems have been solved by many evolutionary computational optimization techniques in the recent years. These techniques have been successfully applied to non-convex, non-smooth and non-differentiable optimization problems. Some of the set echniquesare genetic algorithm, simulated annealing, particle swarm optimization (PSO), evolutionary programming, hybrid evolutionary programming (HEP), chaotic ant swarm optimization (CASO), Bacteria foraging optimization (BFO), Teaching-Learning-Based Optimization (TLBO) [4, 5].

Genetic algorithm was first introduced based on Darwin's principle of evolution. GA is random search algorithms based on the principles of genetic variation and natural selection and is considered to offer a high probability of finding the global or near global optimum solution of difficult optimization problems. GA combine solution evaluation with stochastic genetic operator namely, selection, crossover and mutation to obtain near optimality. An optimization problemtreats simultaneously more than one objective function is called as multi-objective optimization problem. Multi-objective GAis an extension of classical GA. The main difference between a conventional GA and Multi-Objective Genetic Algorithm (MOGA) lies in the fitness assignment to an individual. The rest of algorithm issame as that in a classical GA.

The main aim of this paper is to apply the MOGA to solve the OPF problem.Multi-Objective Genetic Algorithm produces multiple solutions in one single simulation run for solving a multi-objective optimization problem. Genetic Algorithm toolbox of matlab has been used for solving Multi-Objectiveoptimal power flow(MO-OPF) problem.

II. PROBLEM FORMULATION

The main objective of OPF problem solution is to optimize a selected objective function such as fuel cost minimization, voltage stability enhancement and transmission losses minimizationvia optimal adjustment of the power system control variables, while at the same time satisfying various equality and inequality constraints. The problem can be described as follows [6]:

Mathematically,

Min F(x,u)(1)

Subject to: g(x,u) = 0(2)

h (x,u) ≤ 0 (3)

Where x is the vector of dependent variables or state variables; u is the vector of independent variables or control variables;F is the objective function to be optimized;g is the equality constraints representing nonlinear load flow Equations;h is the inequality constraints representing system operating constraints.

a. State Variables

In eq (1) – (3), x is the vector of dependent variables in a power system network that includes:

1. Slack bus generated active power P_{G_1}.

2. Load (PQ) bus voltageV_L.

3. Generator reactive power output Q_G .

4. Transmission line loading (line flow) S_l.

Hence, x can be expressed as:

$$\mathbf{x^T = [P_{G_1},\ V_{L_1} \dots V_{L_{NL}},\ Q_{G_1} \dots Q_{G_{NG}},\ S_{l_1} \dots S_{l_{nl}}]}\ \mathbf{(4)}$$

Where *NL*,*NG* and *nl* are denote the number of load buses,the number of generatorsunitand the number of transmission lines, respectively.

b. Control Variables

In eq (1) – (3), u denotes the independent or control variables of a power system network that includes:

1. Generator active power output P_G except at slack bus P_{G_1}.

2. Generator bus voltageV_G.

3. Transformer taps setting$T$.

4. Shunt VAR compensationQ_C.

Hence, u can be expressed as:

$$u^T = [P_{G_2} \dots P_{G_{NG}}, V_{G_1} \dots V_{G_{NG}}, T_1 \dots T_{NT}, Q_{C_1} \dots Q_{C_{NC}}] \quad (5)$$

Where *NG, NT* and *NC* are denote the number of generators unit, the number of regulating transformers and the number of shunt VAR compensators, respectively.

c. Objective function

In this paper, three different objective functions are considered. The objective functions are as follows:

1. Minimization of total fuel cost

In this case, the objective function$F_1(x,u)$represents the total fuel cost, and it can be expressed asfollows [6]:

$$F_1(x,u) = \sum_{i=1}^{NG} f_i \ (\$/h) \quad (6)$$

Where,f_i is thetotal fuel cost of the ithgenerating unit.

The fuel cost characteristics is represented by quadratic functions as:

$$f_i = a_i + b_i P_{G_i} + c_i P_{G_i}^2 (\$/h) \quad (7)$$

Where a_i, $b_i$and $c_i$are the fuelcost coefficients of the ithgenerating unit and $P_{G_i}$is real power output of the ith generator.

2. Voltage stability enhancement

Voltage stability is one of the important issues in power system planning and operation. The static approach for voltage stability analysis involves determination of an index known as voltage collapse proximity indicator. This index is an approximate measure of closeness of the system operating point to voltage collapse. There are different type methods of determining the voltage collapse proximity indicator. One such method is the L-index of the load buses in theSystem proposed in[6]. It is based on power flow analysis and its value ranges from 0 (no load condition) to 1 (voltage collapse). The bus with the largest*L*-index value will be the most vulnerable bus in the system. The *L*-index determine for a power system is briefly discussed below [6, 7].

For a power system with *NB, NG* and *NL* buses representing thetotal number of buses, the total number of generator bus (or *PV* buses) and the total number of load buses (or *PQ* bus), respectively, we can separate buses into two parts: *PQ*buses at the head and *PV*buses at the tail as follows

$$\begin{bmatrix} I_L \\ I_G \end{bmatrix} = [Y_{bus}] \begin{bmatrix} V_L \\ V_G \end{bmatrix} = \begin{bmatrix} Y_{LL} & Y_{LG} \\ Y_{GL} & Y_{GG} \end{bmatrix} \begin{bmatrix} V_L \\ V_G \end{bmatrix} \quad (8)$$

Where,Y_{LL}, Y_{LG},Y_{GL} and Y_{GG} are sub matrix of Y_{bus}.The following hybrid system of equations can be written:

$$\begin{bmatrix} V_L \\ I_G \end{bmatrix} = [H] \begin{bmatrix} I_L \\ V_G \end{bmatrix} = \begin{bmatrix} H_{LL} & H_{LG} \\ H_{GL} & H_{GG} \end{bmatrix} \begin{bmatrix} I_L \\ V_G \end{bmatrix} \quad (9)$$

Where H matrix is generated by the partial inversion ofY_{bus} ,$H_{LL}$$H_{LG}$, $H_{GL}$and $H_{GG}$are sub matrix of H,V_G,I_G,V_L and I_L are voltage and current vector of generator buses and load buses, respectively

The matrix H is given by:

$$[\mathrm{H}] = \begin{bmatrix} \mathrm{Z_{LL}} & -\mathrm{Z_{LL}Y_{LG}} \\ \mathrm{Y_{GL}Z_{LL}} & \mathrm{Y_{GG}} - \mathrm{Y_{GL}Z_{LL}Y_{LG}} \end{bmatrix} \mathrm{Z_{LL}} = \mathrm{Y_{LL}^{-1}} \qquad (10)$$

Therefore, the L-index denoted by $\mathrm{L_j}$of bus j is represented as follows

$$L_j = \left| 1 - \sum_{i=1}^{NG} H_{LG_{ji}} \frac{V_i}{V_j} \right| \qquad j = 1, 2, 3 \dots \dots \dots . NL \qquad (11)$$

For stable situations the condition $L_j \leq 1$ must not be violated for any of the buses j. Hence, a power system L-index describing the voltage stability of the complete subsystem is given by

$$\mathrm{F_2(x,u)} = L_{max} = max\,(L_j), \;\; j = 1 \dots \dots . NL(12)$$

The lower value ofL_{max} system ismore stable.

3. Minimization of total power losses

This objective is to minimize power transmission loss in the system. The power loss is a non- linear function of bus voltages. Total power loss in the transmission system can be mathematically represented as follows [6]:

$$\mathrm{F_3(x,u)} \;\; = \sum_{k=1}^{NT} \mathrm{G_k} \left| \mathrm{v_i^2} + \mathrm{v_j^2} + 2|\mathrm{V_i}||\mathrm{V_j}|\cos(\delta_i - \delta_j) \right| \qquad (13)$$

Where,G_k is the conductance of *kth* line connected between *ith* and *jth* buses: *NT* is the number of transmission lines: V_i is the voltage magnitude at bus *i*: V_j is the magnitude at bus j: δ_i is the voltage angles at bus *i*: δ_j is the voltage angles at bus *j.*

d. Constraints

OPF constraints can be classified into equality and inequality constraints, as detailed in given below:

A. . Equality Constraints

The equality constraints g represented by (2), are typical load flow equations which are defined as follows:

- Real Power Constraints

$$\mathrm{P_{Gi}} - \mathrm{P_{Di}} - \mathrm{V_i} \sum_{j=1}^{NB} \mathrm{V_j} [\mathrm{G_{ij}} \cos(\theta_{ij}) + \mathrm{B_{ij}} \sin(\theta_{ij})] = 0 \qquad (14)$$

- Reactive Power Constraints

$$\mathrm{Q_{Gi}} - \mathrm{Q_{Di}} - \mathrm{V_i} \sum_{j=1}^{NB} \mathrm{V_j} [\mathrm{G_{ij}} \sin(\theta_{ij}) - \mathrm{B_{ij}} \cos(\theta_{ij})] = 0 \qquad (15)$$

Where, $\theta_{ij} = \theta_i - \theta_j$,$V_i$ and $V_j$are the voltage magnitudes at bus *i* and bus *j*respectively,*NB* is the number of buses, P_{Gi} is the active power generation at bus *i*,Q_{Gi} is the reactive power generation at bus *i*, $P_{Di}$is the active load demand at bus *i*, $Q_{Di}$is the reactive load demand at bus *i*, G_{ij} andB_{ij} are the elements of the admittance matrix (Y_{ij}− G_{ij} + jB_{ij}) representing the conductance and susceptance between bus *i* and bus *j*, respectively.

B. Inequality Constraints

The inequality constraints h represented are the power system operating limits includes:

i. Generator Constraints

For all generators including the slack: voltage, active and reactive outputs ought to be restricted by their lower and upper limits as follows:

$$\mathbf{V_{G_i}^{min} \leq V_{G_i} \leq V_{G_i}^{max} \quad i = 1 \ldots\ldots NG}$$

(16)

$$\mathbf{P_{G_i}^{min} \leq P_{G_i} \leq P_{G_i}^{max} \;\; i = 1 \ldots\ldots NG}$$

(17)

$$\mathbf{Q_{G_i}^{min} \leq Q_{G_i} \leq Q_{G_i}^{max} i = 1 \ldots\ldots NG}$$

(18)

ii. Transformer Constraints

Transformer taps have minimum and maximum setting limits as follows:

$$\boldsymbol{T_i^{min} \leq T_i \leq T_{G_i}^{max} \quad i = 1 \ldots\ldots NT}$$

(19)

iii. Shunt VAR compensator constraints

Shunt VAR constraints must be restricted by their lower and upper limits as follows

$$\boldsymbol{Q_{C_i}^{min} \leq Q_{C_i} \leq Q_{C_i}^{max} \quad i = 1 \ldots\ldots NG}$$

(20)

iv. Security Constraints

These contain the constraints of voltage magnitude at load buses and transmission line loadings. Voltage magnitude of each load bus must be prohibited within its lower and upper operating limits. Line flow through each transmission line ought to be restricted by its capacity limits. These constraints can be mathematically formulated as follows:

$$\mathbf{V_{L_i}^{min} \leq V_{L_i} \leq V_{L_i}^{max} \qquad i = 1 \ldots\ldots NL}$$

(21)

$$\mathbf{S_{l_i} \leq S_{l_i}^{max} \qquad i = 1 \ldots\ldots nl}$$

(22)

III. MULTI-OBJECTIVE GENETIC ALGORITHM

An optimization problem treats simultaneously more than one objective function is called as multi-objective optimization problem. Multi-Objective optimization Problem (MOP) can be mathematically presented as [8, 9]:

Min

$$\mathbf{[\,F(x) \;=\; f_1(x), \ldots\ldots f_n(x)]} \qquad \textbf{(23)}$$

$$\textbf{Subject to:}\begin{cases} \mathbf{g_j \;=\; 0 \quad j = 1, 2, \ldots\ldots\ldots M} \\ \mathbf{h_k \;\leq\; 0 \quad k = 1, 2, \ldots\ldots\ldots N} \end{cases} \textbf{(24)}$$

Where *F(x)* consists of *n* conflicting objective functions, *x* is the decision vector, g_j is the *jth* equality constraint and h_k is the *kth* inequality constraint.

For a multi-objective optimization problem, any two solutions $x_1$and $x_2$can have any one of two possibilities, where one dominates other or not. In a minimization problem, without loss of generality, solution $x_1$dominates x_2 if the following conditions are satisfied.

1. $\forall_i \in \{1, 2, \ldots \ldots N\}$: $f_i(\, x_1) \leq f_i(\, x_2)$(25)
2. $\exists_j \in \{1, 2, \ldots \ldots N\}$: $f_j(\, x_1) < f_j(\, x_2)$(26)

If any one of the above conditions is violated, then the solution $x_1$does not dominate x_2. If $x_2$dominates by the solution x_1,$x_1$is called as the non-dominated solution. A solution is said to be Pareto optimal if it is not dominated by any other solution in the solution space. A Pareto optimalsolution cannot be refined with respect to any objective without worsening at least one other objective. The set of possible feasible non-dominated solutions in *X*is referred to as the Pareto optimal set, and for a given Pareto optimal set, the corresponding objective function values in the objective space is called the Pareto front. For several optimization problems,all Pareto optimal solutions are enormous (maybe infinite). The main goal of a multi-objective optimization algorithm is to identify solutions in the Pareto optimal set. However, searching the all Pareto optimal set, for many multi-objective problems, is practically impossible due to its size.

MOGA [9, 10]was the first multi-objective GA that explicitly used Pareto based ranking and niching techniques together to encourage the search toward the true Pareto front while maintaining diversity in the population. Once fitness has been assigned, selection can be performed and genetic operators are applied as usual. To a solution *i*, a rank equal to one plus the number of solutions η_i that dominate solution i is assigned:

$$r_i = 1 + \eta_i (27)$$

The rank one is assigned to non-dominated solutions since no solution would dominate a non-dominated solution in a population. After ranking, raw fitness is assigned to each solution based on its rank by sorting the ranks in ascending order of magnitude. Then, a raw fitness is assigned to each solution by linear mapping function. Thereafter, solutions of each rank are considered at a time and their averaged raw fitness is called assigned fitness. Thus the mapping and averaging procedure ensures that the better ranked solutions have a higher assigned fitness. In order to maintain diversity in the population, niching among solutions of each rank are introduced. The niche count is calculated with following equation [9, 10, 11]:

$$nc_i = \sum_{j=1}^{\mu(r_i)} Sh(d_{ij}) \qquad (28)$$

Where,$\mu(r_i)$ is the number of solutions in a rankand $Sh(d_{ij})$is the sharing function value of two solution *i* and *j*. The sharing function$Sh(d_{ij})$ is calculated by using objective function as distance metric as:

$$Sh(d_{ij}) = \left\{ \begin{matrix} 1 - \left(\frac{d_{ij}}{\sigma_{share}}\right)^{\alpha} & \text{if } d \leq \sigma_{share} \\ 0 & \text{otherwise} \end{matrix} \right\} \qquad (29)$$

The parameter *d* is the distance between any two solutions in the population and is $\sigma_{share}$the sharing parameter which signifies the maximum distance between any two solutions before they can be considered to be in the same niche. The above function takes a value in [0, 1] depending on the values of $\sigma_{share}$and d_{ij}. If $\alpha = 1$is used, the effect linearly reduces from one to zero.

The normalized distance between any two solutions can be calculated as follows:

$$d_{ij} = \sqrt{\sum_{k=1}^{M}\left(\frac{f_k^{(i)} - f_k^{(j)}}{f_k^{max} - f_k^{min}}\right)^2} \qquad (30)$$

Wheref_k^{max} and f_k^{min} are the maximum and minimum objective function value of the *kth*objective.

In MOGA, the shared fitness is calculated by dividing the fitness of a solution by its niche count. Even though all solutions of any particular rank have the identical fitness, the shared fitness value of each solution residing in less crowded region has a better shared fitness which produces a large selection pressure for poorly represented solutions in any rank. The fitness of the solution is reduced by dividing the assigned fitness by the niche count. In order to keep the average fitness of the solutions in a rank same as that before sharing, the fitness values are scaled. This rule is continued until all ranks are processed. This paper, tournament selection, BLX- α crossover and non-uniform mutation operators are applied to create a new population [12].

Best Compromise Solution

Having obtained the Pareto optimal set, choosing a best compromise solution is important in decision making process. In this paper, fuzzy membership approach is used to find a best compromise solution. Due to imprecise nature of the decision maker'sjudgement the *ith* objective function F_i of individual *k* is represented by a membership function u_i^kdefined as

$$u_i^k = \begin{cases} 1 & F_i \geq F_i^{min} \\ \dfrac{F_i^{max} - F_i}{F_i^{max} - F_i^{min}} & F_i^{min} < F_i < F_i^{max} \\ 0 & F_i^{max} \leq \quad F_i \end{cases} \qquad (31)$$

Where $F_i^{min}$and F_i^{max} are the minimum and maximum value of *ith* objective function among all non-dominated solutions, respectively. For each non-dominated solution *k*, the normalized membership function u^k is calculated as

$$\mu^k = \frac{\sum_{i=1}^{NO}\mu_i^{ko}}{\sum_{k=1}^{M}\sum_{i=1}^{NO}\mu_i^k} \qquad (32)$$

Where, *M* is thetotal number of non-dominated solutions; *NO* is the number of objectives. Finally, the best compromise solution is the one achieving the maximum member ship function (μ^k).

IV. Numerical Results

The proposed MOGA algorithm is tested on the standard IEEE 30-bus test system [13]. This system consists of six generators at buses 1, 2, 5, 8, 11 and 13, four transformers with off-nominal tap ratio at lines 6–9, 6–10, 4–12 and 27–28 and addition, buses 10, 12, 15, 17, 20, 21, 23, 24 and 29 were selected as shunt VAR compensation buses for reactive power control. The complete system data with minimum and maximum limits of control variables are given in [13]. Bus 1 was taken as the slack bus. The proposed algorithm has been applied to solve the OPF problem for several cases with different objective functions. Before MOGA is applied to OPF problem, following parameters need to be defined. The number of population NP = 30 andthe number of variable = 24.

A. Case 1: Fuel cost Vs Transmission line Losses

In this case, two competing objectives, i.e. fuel cost and transmission linelosses were considered. This multi-objective optimization problem was solved by the proposed algorithm. The Pareto optimal solution obtained with the help of proposed MOGA algorithm is shown in Fig. 1. Pareto optimal solution, it is clear that the proposed MOGA method is giving well distributed solutions. The best compromise solution was found with the help of fuzzy membership approach. The best solution for minimum fuel cost and minimum loss and the compromise solutionare given in Table 1.

B. Case 2: Fuel cost Vs L-index

In this case, L-index is considered in place of transmission line losses. The L-index of a bus indicates the proximity of voltage collapse condition of that bus. It varies zero (no load case) to one (voltage collapse). These two competing objective functions were optimized by the proposed MOGA method. The Pareto optimal solution for this case is shown in Fig.2. The best compromise solution for minimum fuel cost and minimum L-index are given in Table 1.

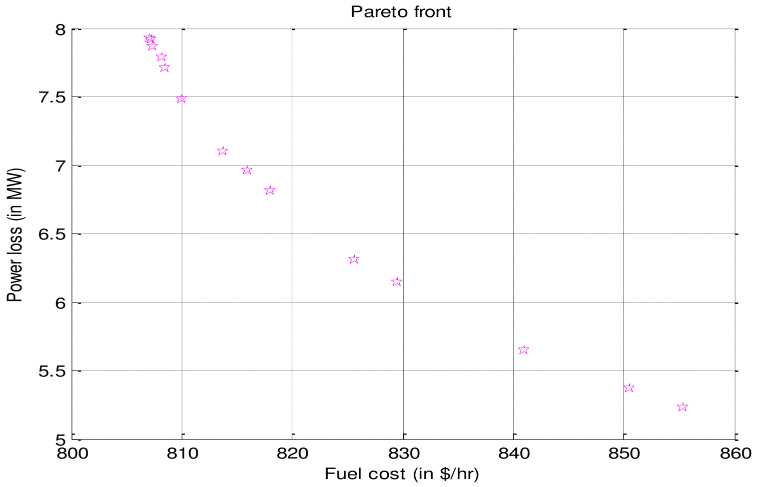


Fig.1. Pareto Optimal Solutions for Case 1

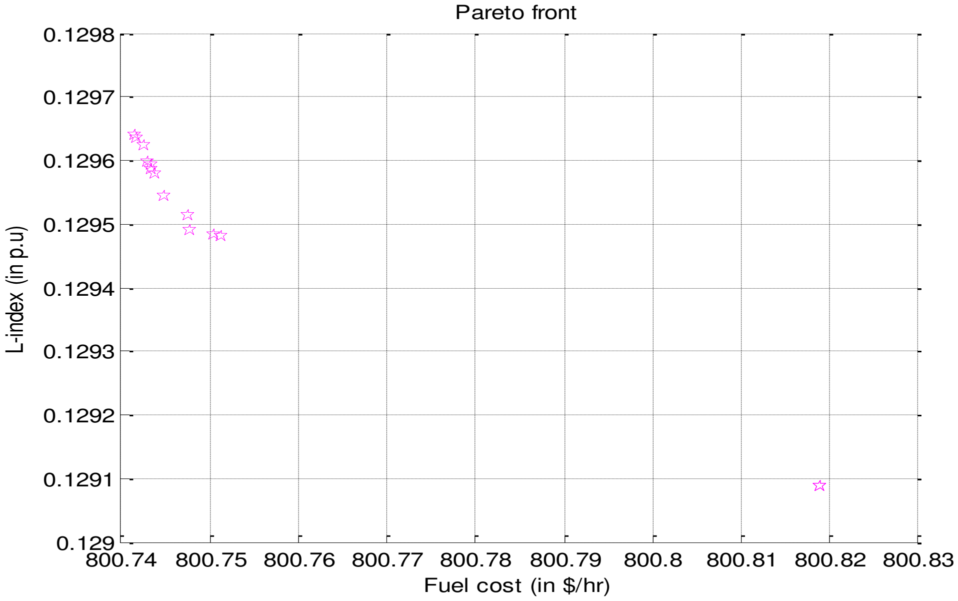


Fig.2. Pareto Optimal Solutions for Case 2

Table 1 Simulation Results for IEEE-30 Bus system

Control variable (p.u.)	Initial	Best Cost	Best Losses	Best Comp.	Best Cost	Best L-Index	Best Comp.
P2	.8000	.4897	.5895	.5675	.4894	.4928	.4900
P5	.5000	.2342	.3962	.3586	.2149	.2155	.2156
P8	.2000	.2409	.2877	.2742	.2022	.1967	.2012
P11	.2000	.1844	.2829	.2719	.1227	.1259	.1223
P13	.2000	.2030	.2361	.2371	.1201	.1237	.1201
V1	1.0500	1.0609	1.0508	1.0588	1.0806	1.0788	1.0806
V2	1.0400	1.0485	1.0207	1.0127	1.0190	1.0270	1.0190
V5	1.0100	1.0210	1.0162	1.0156	1.0266	1.0290	1.0265
V8	1.0100	1.0381	1.0349	1.0283	1.0456	1.0469	1.0456
V11	1.0500	1.0751	1.0433	1.0487	1.0664	1.0704	1.0664
V13	1.0500	1.0396	1.0504	1.0412	1.0723	1.0716	1.0723
T11	10780	.9986	1.0158	1.0120	1.0050	1.0036	1.0051
T12	1.0690	1.0007	1.0260	1.0172	.9925	.9906	.9926
T15	1.0320	1.0276	1.0233	1.0215	1.0110	1.0138	1.0110
T36	1.0680	.9793	.9836	.9851	.9701	.9732	.9700
Qc10	0.0	.0193	.0360	.0357	.0421	.0433	.0419
Qc12	0.0	.0224	.0304	.0239	.0497	.0499	.0498
Qc15	0.0	.0408	.0179	.0257	.0345	.0397	.0343
Qc17	0.0	.0285	.0229	.0232	.0409	.0408	.0409
Qc20	0.0	.0209	.0236	.0304	.0369	.0389	.0369
Qc21	0.0	.0154	.0254	.0276	.0500	.0500	.0500
Qc23	0.0	.0195	.0296	.0310	.0238	.0252	.0239
Qc24	0.0	.0423	.0329	.0260	.0500	.0500	.0500
Qc29	0.0	.0340	.0323	.0308	.0411	.0497	.0456
Fuel Cost($/h)	902.00	807.13	855.37	**841.05**	800.74	800.81	**800.75**
L_{max}	0.1772	-	-	-	.1296	.1291	**.1295**
P_{LOSS}	5.8423	7.930	5.228	**5.6482**	-	-	-

*Bold values represent the best values of the objective functions chosen and best comp.Indicate best compromise solution.

Table 2

The best compromise solution for case-1 using different multi objective algorithms.

Algorithms	Fuel cost ($/h)	Losses (MW)

MOSPEA [14]	847.01	5.666
NSGA-II [15]	823.88	5.7699
MOGA	841.05	5.6482

Table 3

The best compromise solution for case-2 using different multi objective algorithms.

Algorithms	Fuel cost ($/h)	L-index (p.u)
MOSPEA [14]	809.79	.1146
MOTLBO [16]	803.63	.1020
MOGA	800.75	.1295

V. CONCLUSIONS

In this paper a multi-objective genetic algorithm (MOGA) has been proposed to solve the multi-objective optimal power flow(MO-OPF) problem with many constraints in IEEE 30-bus system. The proposed approach successfully applied to solve various types of optimal power flow (OPF) problems with different objective function like fuel cost minimization, voltage stability enhancement and transmission losses minimization.The proposed approach results are compared with the results reported in the literature.The numerical results show that the proposed technique is efficient and outperforms for solving MO-OPF problem.

VI. ACKNOWLEDGEMENT

The authors sincerely acknowledge the financial support received from University Grants Commission (UGC), New Delhi, India under Major Research Project received vide F. No. 41-657/2012[SR] dated 26-07-2012 and the Director, Madhav Institute of Technology &Science, Gwalior, India to carry out this research work.

REFERENCES

[1] J. Carpentier, Contribution à l'étude du Dispatching Economique, Bull. Soc. Fran-caise Electriciens (1962) 431–447.

[2] Dommel, H. W. and W. F. Tinney (1968). Optimal power flow solutions. IEEE Trans on Power Apparatus and System, PAS-87(10), 1866–1876.

[3] James A. Momoh, M. E. El-Hawary and RamababuAdapa "A Review of Selected Optimal Power Flow Literature''IEEE Transactions on Power Systems, Vol. 14, No. 1, February 1999 .

[4] M.R. AlRashidi and M.E. El-Hawary, Applications of computational intelligence techniques for solving the revived optimal power flow problem, Electric Power Systems Research 79 (2009) 694–702.

[5] P.K. Roy , S.P. Ghoshal and S.S. Thakur. Biogeography based optimization for multi-constraint optimal power flow with emission and non-smooth cost function, Expert Systems with Applications 37 (2010) 8221–8228, Expert Systems with Applications 37 (2010) 8221–8228

[6] M. RezaeiAdaryani, A. Karami. Artificial bee colony algorithm for solving multi-objective optimal power flow problem, Electrical Power and Energy Systems 53 (2013) 219–230.

[7] H.R.E.H. Bouchekara, M.A. Abido, M. Boucherma, Optimal power flow using Teaching-Learning-BasedOptimization technique Electric Power Systems Research 114 (2014) 49–59.

[8] C. Coello,1999. –"A comprehensive survey of evolutionary based Multi-objective optimization technique" *Knowl. Inf. Syst.*,vol.1, no.3,pp.129 156.

[9] M. A. Abido, "A Niched Pareto Genetic Algorithm for Multi-objective Environmental EconomicDispatch," *Iniernational Journal of Electrical Power andEnergv Systems,* Vol. *25,* No. 2, February 2003,pp. 79-105.

[10] M. A. Abido, "Environmental/Economic Power Dispatch Using Multi-objective Evolutionary Algorithms" *IEEE Trans on Pawer Systems,* Vol. 18, No. 4, November 2003, pp. 1529-1 537.

[11] M. A. Abido, "A novel multi-objective Evolutionary Algorithms for environmental /economic power dispatch," *Electric Power Systems Research,* Vol. *65,* No.1, April 2003,pp. 71-81.

[12] Eshelman LJ, Schaffer JD. In: Whitley D, editor. Real coded genetic algorithms and interval schemata; 1993. p. 187–202.

[13] K. Lee, Y. Park, and J. Ortiz, "A United Approach to Optimal Real and Reactive Power Dispatch," IEEE Trans. on Power Apparatus andSysrems, Vol. 104, No. 5, 1985, pp. 1147-1 153.

[14] Kumari MS, Maheswarapu S. Enhanced genetic algorithm based computation technique for multi-objective optimal power flow solution. Int J Electr Power Energy Syst 2010;32(6):736–42.

[15] Sivasubramani S, Swarup KS. Multi-objective harmony search algorithm for optimal power flow problem. Int J Electr Power Energy Syst 2011;33(3): 745–52.

[16] Nayak MR, Nayak CK, Rout PK. Application of multi-objective teaching learning based optimization algorithm to optimal power flow problem. ProcTechnol 2012;6:255–64.

INNOVATION- CHALLENGE FOR ENTREPRENEURS AND NGO’S

Shubhi Khare[1], Dr. Pradeep Chaurasia[2], Bhola Nath Jaiswal[3], Anurag Singh Parihar[4]

[1]Student M.Com, [3]Research Scholar PHD, [4]Student LLB, SRKC Satna, Affiliated by APSU Rewa, Madhya Pradesh, (India)

[2]Asst. Professor, Department Of Management Studies, AKS University, Madhya Pradesh, India,

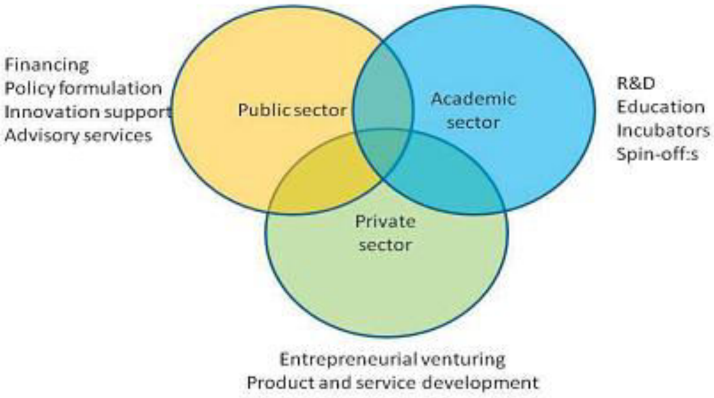
ABSTRACT

Innovation is change that unlocks new value. Innovation distinguishes between a leader and a follower and we can consider private companies as leaders and NGO as followers and vice versa as both help each other with the resources required by them. Entrepreneurial activity and innovation have been seen as an indispensable factor behind societal development and prosperity. Almost all political parties agree upon the necessity of increased entrepreneurial venturing in society. Entrepreneurship and innovation from policymakers are emphasizing high-tech, high growth, individualist ventures. Entrepreneurial methods can help NGOs maintain and expand their membership and funding sources, through continuous innovation in the services provided to meet their target constituencies’ needs. There are certain points that focused on this concept are social entrepreneur brings innovation and entrepreneur are those who combine innovation with growth and risk.

Keypoints: Innovation, Ngo, Risk Taking, Growth, Alliances.

I. INTRODUCTION

What entrepreneurial skills can be useful for NGOs and how NGOs can help entrepreneurs develop successful business models that integrate a human rights approach? Communication and education about human rights and corporate social responsibility are necessary to develop a better interaction between the business community and the NGOs, and with this innovation can be possible in different sectors and development would be done faster. Entrepreneurial ventures may play important role in the development. Given below are some sectors and innovation ideas under:-



II. ENTREPRENEURS AND NGOs

- Entrepreneurs are actors that combine innovation with rapid growth and risk taking.
- A social entrepreneur brings innovative solutions to persistent social problems and to provide solution for the same in the products, the processes, the positioning and the paradigms of the organization.
- Social entrepreneurship articulates the fuelling of a sustainable entrepreneurial spirit and the search for solutions to establish a balance between economic profits and the respect for social, environmental, cultural, and spiritual ecosystems.
- As social entrepreneurs, NGOs can contribute to more equitable and rights-based approaches to globalization by developing models that balance profit and non-profit objectives.
- NGOs can advance respect for human rights and corporate social responsibility by engaging large corporations as well as small and medium enterprises in entrepreneurial initiatives that benefit social and humanitarian causes.
- Entrepreneurship has different implications for NGOs depending on the size and scope of the organization.
- Entrepreneurial methods can help NGOs maintain and expand their membership and funding sources, through continuous innovation in the services provided to meet their target constituencies' needs.
- Entrepreneurship can become a shared value for both NGOs and for Private Sector.

III. DIFFICULTIES FACED BY ENTREPRENEURS

- There is a tension between how funding is allocated and the goals of recipient organizations.
- The private sector and foundations generally wish to fund outcomes rather than generic processes, despite the fact that the work of NGOs entail substantial administrative costs.
- The private sector functions with short term quantifiable goals, whereas NGOs function with long range plans and goals that are not as easily measurable.
- Corporations often operate according to their own norms and processes when they give funds to NGOs.
- NGOs may experience risks related to their autonomy and integrity when accepting funds from the private sector.

IV. PROPOSALS FOR ALLIANCES

- A consistent interaction between NGOs and the business world could involve big business sponsorship of NGO activities, as well as collaboration between young companies to form entrepreneurial initiative.
- An entrepreneurial alliances between the private sector and NGOs requires mutual respect and that each party identify its helpful resources.
- The private sector could provide free legal and accounting advice to NGOs to get benefited with positive gains.
- Established companies could help create "brand awareness" of different social causes in collaboration with NGOs and support them with communications and websites.
- Financial advisors could help NGOs identify diverse sources of funding to enhance the sustainability of their activities.

- NGOs could help entrepreneurs with important insights realities on the socioeconomic, cultural and political sectors.
- The direct allocation of private sector funds to NGO projects and programmes would increase transparency and effectiveness.

V. POINTS OF DIVERGENCE

- Participants expressed clearly divergent views on:
- The political role of business, with some stating that businesses are separate from politics and shouldn't assume responsibility for political issues and others stating that businesses are political actors that influence political decisions and decision-makers.
- The potential of private sector contributions to NGOs, with some stating that companies can only be useful to NGOs by giving them funds, and others stating that companies have more to offer NGOs than just charity.

VI. PROJECTS RECENTLY UNDERTAKEN

- Camp Fosters Future Entrepreneurs.
- Business Incubator Nurtures Start-Ups.
- Railroads Containers Become Facility For Homeless.
- Bakery Advances Students Education.
- Texas State University- Project ROW.

VII. CONCLUSION

If you want something new you have to stop doing something old. The above projects undertaken by leaders create and implement community empowerment projects around the globe. The experience not only transforms lives, but also helps to develop the kind of talent and perspective that are essential for innovative challenges. Innovation is requirement of every sector and the combination of risk taking and growth helps entrepreneurs and NGOs both to develop strong relation and come with some innovation that can change the society and livelihood.

REFERENCE

[1] Aktion Afrika Hilfe (AAH). 2006. "Updates on AAH programmes in Kyangwali." Internal report. September 29, 2006.

[2] Barr, Abigail and Marcel Fafchamps. 2006. "A Client-Community Assessment of the NGO Sector in Uganda." *Journal of Development Studies*, forthcoming

[3] INNOVATION AND ENTREPRENEURSHIP IN TODAY'S SCENARIO. International Journal of Marketing, Financial Services & Management Research Vol.1 Issue 8, August 2012, ISSN 2277 3622

[4] Batra Promod, Batra Vijay, Outside the Box- Great Ideas that transformed Business, published by Promod batra Vijay batra and Associates, New Delhi

[5] Hisrich D Robert, Peters P Michael, Shepherd A Dean, Entrepreneurship, sixth edition (2007), Tata McGraw-hill publishers, New Delhi

[6] Oats David, A Guide to Entrepreneurship, second edition (2007), Jaico publishing house, Mumbai

[7] http://www.gujaratchamber.org/ - The Gujarat Chamber of Commerce and Industry is a local chamber of commerce in Ahmedabad.

[8] http://www.biztradeshows.com/cii-chennai/ - The Chennai wing of Confederation of Indian Industry.

[9] http://www.fapcci.in/ - The Federation of Andhra Pradesh Chambers of Commerce is a local chamber of commerce.

[10] http://msme.gov.in/ and http://www.laghu-udyog.com/ - The Ministry of Micro, Small And Medium Enterprises lists various government schemes for entrepreneurs.

[11] http://www.techno-preneur.net/ - The Technology Innovation Management and Entrepreneurship Information Service assists techno-preneurs in finding technologies, projects, funding options and information on policy environment, incentive schemes and industrial infrastructure available in the country, covering central and state governments

INFORMATION SECURITY RISK MANAGEMENT

[1]Usha, [2] Ankush Goyal

[1,2]*Computer Science & Engineering Department, M.D.U, (India)*

ABSTRACT

Optimizing risk to information to protect the enterprise as well as to satisfy government and industry mandates is a core function of most information security departments. Risk management is the discipline that is focused on assessing, mitigating, monitoring a optimizing risks to information. Risk assessments and analyses are critical sub-processes within risk management and are used to generate data that drive organizational decisions to accomplish this objective

***Keywords*: Risk Management, Security, Vendor, Introduction, Impact**

I. INTRODUCTION

Information risk management is the activity directed towards assessing, mitigating (to an acceptable level) and monitoring of risks associated with information. The principle goal of an organization's risk management process should be to protect the organization and its ability to perform their mission, not just its IT assets [70]. Peter Drucker once said "the diffusion of technology and commoditization of information transforms it into a resource equal in importance to the traditionally important resources of land, labor and capital" [14]. The exponential growth and availability of information after the Internet boom of 1990's goes to show the accuracy of his foresight. In today's world, the fortunes of most organizations are tied with the information they possess and the sophistication with which they are able to manage it. Most of these risk management methodologies, while providing a structured and systematic process for risk management, either lack specific guidance on which risk assessment methods to use or provide for a weak approach. This does not satisfy the rigorous data needs of business leaders as well as audit needs of compliance auditors. This was clearly identified as a significant issue in the recent RSA report based on discussions with top risk management leaders in Global 1000 companies [61]: "Risk should be managed to an acceptable level, based on the enterprise's risk appetite with decision-making guided by a risk assessment model. A structured, consistent and repeatable process for making the risk/reward calculation helps to ensure that it is done consistently across the organization".

II. CHARACTERSTICS IN METHODOLOGY IN RISK MANAGEMENT

A comprehensive definition of the characteristics desired from a risk management system in one place is missing from current literature on this topic. We propose the following criteria based on our research (these criteria are articulated in our paper [64]):

1. It must manage risks to an acceptable level based on enterprise's risk appetite [61]. 2. It must provide decision-support [61]. Security investments are expensive and risk is one criterion that is used to address the economics around it.

3. It must be a continual process [35]. Risk management is not conducted at a point in time; it should be considered throughout the lifecycle of systems development.

4. It must be aligned with an organization's business objectives [70].

As the amount as well as complexity of information resources within organizations is increasing at an exponential rate, we also consider the following characteristics as desirable traits: 5. It must be adaptive. Since an organization's risk profile, threats and vulnerabilities change frequently, it is important that risk management should be adaptive to these changes.

6. It must be scalable to accommodate for this increasing complexity while not impacting the window desired to conduct the assessment activities.

7. It must ensure compliance with government and industry mandates.

8. It must produce consistent results irrespective of who conducts the responsibilities associated with risk management

III. IMPACT OF SECURITY ENHANCEMENT

Based on the risk assessment analytics, a risk assessor provides recommendation on how controls need to be adjusted or whether new controls need to be added. However, decision makers want to measure the impact of these security enhancements. For e.g. increasing the strictness of configuration of a control might mean that the end user sees increased response times; for a decision maker, it is critical to understand whether increasing that strictness and the inconvenience caused to the end user as a result is worth it or not in terms of prevention of security threats. While this area has been researched in other disciplines such as microprocessor simulation [78], it remains unaddressed within the domain of information security.

IV. SECURITY THREATS

The risk profile of an organization changes on a very dynamic basis because new threats come into existence on an almost continuous basis. Thus any approaches to deal with the threats have to be dynamic as well. This issue has not been dealt with in existing research, either methodologically or architecturally. One essential aspect of being able to manage the information security risk to the enterprise is configuring security controls appropriately to ensure that the organization is protected against the threats impacting it. However, despite this critical need, there is a significant opportunity in current approaches that are used for this purpose. They are initially configured during the installation phase and then changed only on an event driven basis. These events could be things like an incident, or observation from logs or recommendations from a risk assessment exercise. There are significant issues with this approach: these changes are ad-hoc and either happen after the fact (i.e. the loss to the enterprise has already happened at that point) or are not dynamic in nature (it makes sense to manage security configuration as soon as the security controls start sensing that the nature of threats around it has started changing).

V. VENDOR SECURITY

A policy is typically a high level articulation of management's intent. As such, it does not provide more granular direction and measurable metrics, which would make the task of adhering to it easier for rest of the enterprise. [25] demonstrates the effective way of writing security policies. Standards are used for this purpose. A standard is refinement of the policy to a more granular level and provides the requirements that need to be met for adherence to the policy. Figure 5.3.2 shows a sample vendor information security standard. It is based on the controls and control objectives provided by ISO 27001 [36]. This standard can be used as a starting point if one doesn't exist already for the enterprise. Note that just the creation of policy and standard is not going to be sufficient unless it is followed up by extensive propagation through the enterprise. This needs to be accomplished through training. Our recommendation is to make it mandatory for all key stakeholders.

VI. CONCLUSION

A survey of current literature as well as prevalent risk management practices in enterprise environments indicates that there are some significant limitations in current risk management approaches. Although control selection and management is a crucial part of risk assessment process of these methodologies, no formalized methods exist that help manage these aspects. In addition, the area of managing risks due to vendors of the enterprise as well as a requirements engineering framework for determining an appropriate GRC strategy remain unaddressed as well.:

VII. ACKNOWLEDGMENT

I show my thanks to all the departments' personals and sponsors who give us an opportunity to present and express my paper on this level. We wish to place on my record my deep sense of gratitude to all reference papers authors for them valuable help through their papers, books, websites etc.

REFERENCES

[1]. R. Anderson, "Why Information Security is Hard- An EconomicPerspective", 17th Annual Computer Security ApplicationsConference, Dec. 2001.

[2]. Jonathan D. Andrews, "Erosion of Trust - E-commerce and theLoss of Privacy", Information Systems Control Journal, Vol. 3, 2001, pp. 46-49.

[3]. Georges Ataya,"Risk-aware Decision Making for New ITInvestment", Information Systems Control Journal, Vol. 2, 2003, pp. 12-14.

[4]. Rudy Bakalov,"Risk Management Strategies for OffshoreApplications and Systems Development", Information Systems Control Journal, Vo. 5, 2004, pp. 36-38.

[5]. S. P. Bennett and M. P. Kailay, "An application of qualitative risk analysis to computer security for the commercial sector", Eighth Annual IEEE ComputerSecurity Applications Conference, Nov.-4Dec. 1992, pp.64-73.

[6]. Nicholas A. Benvenuto & David Brand, "Outsourcing: A RiskManagement Perspective", Information Systems Control Journal,Vol. 5, 2005, pp. 35-40.

[7]. B. Blakley, E. McDermott and D. Geer, "Information Security isInformation Risk Management", Proceedings of the2001workshop on New security paradigms, 2001, pp. 97-104.

[8]. PaulBrooke,"RiskAssessmentStrategies",NetworkComputing,30thofOctober,(http://www.networkcomputing.com/1121/1121f32.html?ls=NCJ_1121bt).

[9]. S. A. Butler,"Security Attribute Evaluconference on Software engineering, ACM, May 2002, pp. 232-240.

[10]. K. Campbell, L. A. Gordon, M. P. Loeb and L. Zhou, "The Economic Cost of Publicly Announced Information Security Breaches: Empirical Evidence from the Stock Market", Journal of

Computer Security, Vol. 11, 2003, pp. 431-448.

[11]. H. Cavusoglu, B. Mishra and S. Raghunthan,“The Effect of Internet Security Breach Announcements on Market Value of Breached Firms and Internet Security Developers”, International Journal of Electronic Commerce, Volume 9, Issue 1, 2004, pp. 70-104.

[12]. F. Cohen, “A Cost Analysis of Typical Computer Viruses andDefenses”, Computers & Security, Vol. 10, 1991, pp. 239-250.

[13]. Center for Internet Security (http://www.cis.org)

[14]. Drucker, Peter, 1988. ‘The Coming of New Organization’, HarvardBusiness Review.

[15]. Peter Drucker,“The Practice of Management”, Butterworth-Heinemann, 2007.

[16]. CriminalTake ControlofCheckFree Web Site(http://pcworld.about.com/od/networkin1/Criminals-Take-Control-of-Chec.htm)

[17]. COBIT 4.1, ISACA (http://www.isaca.org/).

a. Enterprises Risk Managemen Integrated Framework(http://www.coso.org/).

[18]. Covington, Prahlad Fogla, Zhiyuan Zhan, and Mustaque Ahamad. A context aware secur A context aware security architecture for emerging applications. In Proceedings of 18th Annual Computer Security Applications Conference (ACSAC), pages 249-258, as Vegas, NV, December 2002.

[19]. Risk Management: Implementation principles and Inventories forRisk Management/Risk Assessment methods and tools, European Network and Information Security Services.

[20]. G. Eschellbeck,“Active Security- A Proactive Approach for Computer Security Systems, Journal of Network and Computer Applications, No. 23, 2000, pp.109-130.

[21]. F. Farahmand, W. J. Malik, S. B. Navathe and P. H. Enslow,“Security Tailored to the Needs of Business”, Proceeding of theACM CCS BIZSEC, October 2003.

[22]. F. Farahmand, S. B. Navathe and P. H. Enslow, Electronic Commerce and Security A Management Perspective, ISS/INFORMS Seventh Annual Conference on Information Systems and Technology, San Jose, 2002.

a. Shared Assessments, http://www.sharedassessments.org.

[23]. Todd Fitzgerald,“Ten Steps to Effective Web-Based Security Policy Development and Distribution”, in Information Security Handbook, Harold Tipton and Mickey Krause Eds., Auerbach Publications, Boca Raton, FL, 2005.

[24]. Todd Fitzgerald,“Building Management Commitment Through Security Councils”, Information Systems Security, May/June 2005.

[25]. D. E. Geer, “Making Choices to Show ROI”, Secure Business Quarterly, Vol. 1, Issue 2, 2001, pp. 1-3. K. Ghosh and T. M. Swaminatha, Software Security and Privacy Risks in Mobile E-Commerce, Communications of the ACM, Feb.2001, Vol. 44, No. 2, pp. 51-57.

[26]. L. A. Gordon and M. P. Loeb, “Return on Information Security Investments”, Strategic Finance, Nov. 2002.

[27]. John Hagerty,“The Governance, Risk Ma Risk Management, and Compliance Spending Report, 2008-2009: Inside the $32B GRC Market”, http://www.amrresearch.com.

DATA MINING USING PARALLEL COMPUTING FOR E-GOVERNACE INFORMATION SECURITY

Agnivesh Kumar Agnihotri[1], Pinki Sharma[2], Omanand Tiwari[3]

[1]Research Scholar, (M.Tech-Cse), Shri Ram Group Of Institution Jabalpur, RGPV, Bhopal M.P. (India)

[1,2]Research Scholar, M.PHILL, SRKC Satna APSU, Rewa M.P. (India)

ABSTRACT

In this paper, a scalable framework aimed at providing a platform for developing and using high performance data mining applications on heterogeneous platforms. The framework incorporates a software infrastructure and a library of high performance kernels. The exponential increase in the generation and collection of data has led us in a new era of data analysis and information extraction and focuses on the security of the information. In which the three methods is being used first K-means algorithm, decision tree and the artificial neural network.

After using the data mining concept there is use of more mature association model for information system. The information security is more concerning concept for various large scale organizations. As information system provides many facilities to us but the information security is another concern of any computer system. The information system is "Double Edge Sword" because it provides many facilities but if there is no security there will be a great loss and inconvenience. The security of any information system is interrupted by virus, hacker, and leakage of secrets, system failure and many other things.

Conventional systems based on general-purpose processors are unable to keep pace with the heavy computational requirements of data mining techniques. High performance coprocessors like GPUs and FPGAs have the potential to handle large computational workloads.

Data mining is widely used in various domains and has significant applications. But the data mining is mostly used in many fields differently like E-governance, information security, spatial mining, web mining and also with parallel computing. This paper will present the combination of three concepts which are E-government and information security by using parallel computing plate form. Using Parallel Computing concept will makes the E-Governance system much faster and the Information Security will make it more reliable.

In this paper the CUDA, (the Compute Unified Device Architecture), is a parallel computing platform and programming model created by NVIDIA and implemented by the graphics processing units (GPUS) that they produce. CUDA gives program developers direct access to the virtual instruction set and memory of the parallel computational elements in CUDA GPUs. Using CUDA, the GPUs can be used for general purpose processing (i.e., not exclusively graphics); this approach is known as GPGPU. Unlike CPUs, however, GPUs have a parallel throughput architecture that emphasizes executing many concurrent threads slowly, rather than executing a single thread very quickly.

***Keywords:* CUDA, GPU, Double Edge Sword, Parallel Computing, K-Mean Algorithm, Apriori, Decision Tree, Security, Information Extraction.**

I. INTRODUCTION

1.1 Data Mining in Information Security for E Government

E-government security is considered one of the crucial factors for achieving an advanced stage of e-government. As the number of e-government services introduced to the user increases, a higher level of e-government security is required. Since an underdeveloped country whose development can be rapid through proper E-Government implementation. Presently, it is in infnt ancy stage. One of the major failure factors identified at this stage is the improper security consideration. This dissertation also contributes in proposing a cost effective security framework for underdeveloped country.

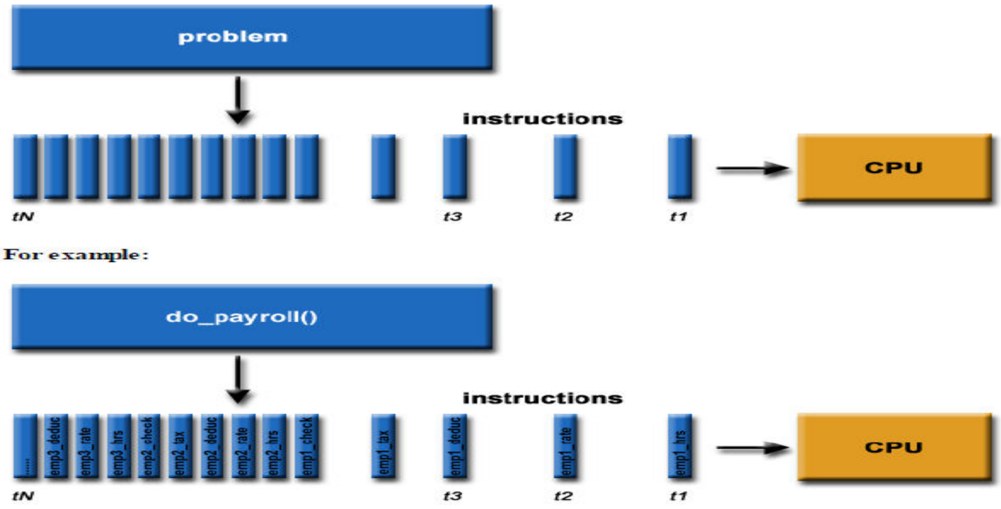
In twenty-first Century, information technology has rapidly permeated into every field of human society. Form the Large national or international fields to the small families or individuals, more and more people use of information technology to provide convenient, fast and efficient work and business. The expanding demands of All walks of life prompted the investment and the scale of informatization construction from the underlying physical layer to the top of the application system constantly strengthen intensify. According to the United States of FBI survey, economic losses caused by the network security are more than $170 billion in USA per year. 75% companies reported that the financial losses were caused by results of computer system security problems. From CN CERT and the China Internet Network Information Center's annual reports, in the first half of 2010, CN CERT had received 4780 network security incident reports, increase of 105%. In the past year, the service fee expenditure for processing safety events totaling up to 153 billion Yuan.

1.2 Parallel Computing

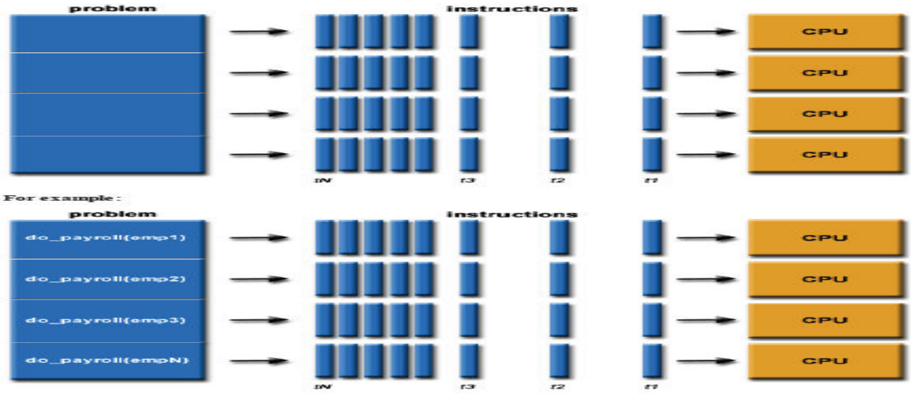
HISTORY OF GPU COMPUTING:- Graphics chips started a fixed-function graphics pipelines. Over the years, these graphics chips became increasingly programmable, which led NVIDIA to introduce the first GPU or Graphics Processing Unit. In the 1999-2000 timeframe, computer scientists in particular, along with researchers in fields such as medical imaging and electromagnetic started using GPUs for accelerating a range of scientific applications. This was the advent of the movement called GPGPU or General Purpose computing on GPUs. The challenge was that GPGPU required using graphics programming languages like OpenGL and Cg to program the GPU. Developers had to make their scientific applications look like graphics applications and map them into problems that drew triangles and polygons. This limited the accessibility of tremendous performance of GPUs for science.

NVIDIA realized the potential to bring this performance to the larger scientific community and decided to invest in modifying the GPU to make it fully programmable for scientific applications and added support for high-level languages like C, C++, and Fortran. This led to the CUDA parallel computing platform for the GPU.

To be run on a single computer having a single Central Processing Unit (CPU). A problem is broken into a discrete series of instructions. Instructions are executed one after another. Only one instruction may execute at any moment in time.



In the simplest sense, *parallel computing is* the simultaneous use of multiple compute resources to solve a computational problem. To be run using multiple CPUs. A problem is broken into discrete parts that can be solved concurrently. Each part is further broken down to a series of instructions. Instructions from each part execute simultaneously on different CPUs.



II. K-MEANS ALGORITHM

K-means algorithm (Lloyd, 1982) is a simple and effective statistical clustering technology, it gives specific classes number K, and put N objects into the K classes, to make the maximum similarity within objects in any class, and to make the minimum similarity among class. The algorithm first need to make an initial judgment that is to select the initial class number and the initial cluster center, and then, each sample is placed in the similar class. Similarity measure can be defined in many different ways. The most commonly used similarity metric is simple Euclidean distance. After all samples are placed into the appropriate class, the class center was update through the calculation of each new class of average. The process is repeated until a certain iteration of generating class center no longer change.

III. DECISION TREE

The decision tree algorithm was first proposed as ID3 algorithm by Quinlan, Later there reappeared many kinds of decision tree algorithm such as ID4, ID5, C4.5, CART, CLOUDS, PUBLIC, SLIQ, RAINFOREST, SPRINT, and ScalParC and so on. It is a common structure to supervise learning, mainly used for data classification. First, we should select portion of the samples to create a decision tree from the training set, and the remainder of the training samples are used to inspect the accuracy of the tree established. If the decision tree can correctly classify the remaining samples, the process will be end. If some sample's classification is error, this sample is added to the training set and create a new tree. In this way, we design a tree which can classify all training samples correctly.

IV. ARTIFICIAL NEURAL NETWORK

Artificial neural network was born in 1950, and Rosenblatt put the single-layer perception application in pattern classification. Its principle is the human brain thinking system's simple structure simulation. It's a multilayered network that is made up of a number of neuronal connections, and can imitate the human brain function of neuron. It is also the adaptive function estimator that does not rely on the model. It does not need any model to realize arbitrary function relation. Its advantage is capable of parallel processing, and has the learning ability, adaptability and strong fault tolerant ability.

V. THE FORMAL DESCRIPTION FOR CORRELATION ANALYSIS MODEL

Correlation analysis model is adapted to find out the meaningful connection hiding in the large data sets. The connection found can be represented by association rules or frequent itemsets form. We need to deal with two critical issues in the data correlation analysis. First, finding in computing mode from the large object data concentration may be costly, second, some models found may be false, because they may have occurred by chance.

5.1 Data's Two Elements Expression

The hypothesis that certain things and the Item set included in these things in the set of relations as shown in

Table 1

TID	i1	i2	i3	i4	i5	i6
1	1	1	0	0	0	0
2	1	0	1	1	1	0
3	0	1	1	1	0	1
4	1	1	1	1	0	0
5	1	1	1	0	0	1

If I={i1,i2,…,id} is the collection of all items, and T={t1,t2,…,tn} is the collection of all things. Each thing of ti contains item set is the subset of I, that is ti being contained in I. Item set's an important property is the support count, that is the number of things containing specific item sets, With σ(X) expressed as:

$$\sigma(X) - |\{t_i | x \subseteq t_i, t_i \in T\}|$$

5.2 The Support Degree and Confidence Degree

Support degree is an important measure, because the rules which support degree is very low may

Sometimes occur. From the data analysis point of view, most of low support degree of rules will be meaningless, because it was more important to improve the research and find out countermeasures to focus on more important things than on the things which does not often happen. Therefore, when we carry out data analysis, we can use the support threshold interval to delete those meaningless rules. In addition, support degree also has desired properties, can be used for association rules effectively discover. The confidence degree measures the reasoning reliability by rules or mathematical model. For a given rule X→Y, the higher the confidence degree, the more likely that Y will appear in the things which contain X. The confidence degree can also estimate the Y defined in X under the conditional probability. If the association rules are expressed as shaped like X→Y implication expression, where X and Y are disjoint sets, namely X∩Y=ø. The strength of association rule can be measured by support degree and confidence degree. Support degree determines the rules that can be used for a given data set frequency, while the confidence degree determines the frequent degree that Y in X contains things appeared. Support degree (s) and confidence degree (c) these two metric representation just as below:

$$s(X \rightarrow Y) = \frac{\sigma\ (X \cup Y)}{N}$$

$$c(X \rightarrow Y) = \frac{\sigma\ (X \cup Y)}{\sigma\ (X)}$$

Support degree measures the importance (or range) of association rule, confidence degree measures the accuracy of association rules. At the same time satisfying the minimum support threshold (min-support) and minimum confidence threshold (min-confidence) rules called the strong rule. The problem of association rule mining is the strong rule that is satisfied for min-support and min-confidence at the same time when mining in transaction database.

5.3 Algorithm optimization

Mining association rules in a primitive method is to calculate for each possible rules support degree and confidence degree. But this method is costly, and step back, because of the number of rules extracted from the data set up to index level. For example, the total number of rules extracted from a data set containing d is, R=3d-2d+1+1. If there are 6 items of data set, it needs to calculate the 602 rules of the support degree and confidence degree. The optimization strategy of mining algorithm of association rules is to decompose the association rules mining task into two processes. The first is the frequent item sets generation process that is the process finding to meet the minimum support threshold. The second process is the generation rules that are to extract all the high confidence rules from frequent item sets found from the previous step. There are many optimization algorithm in frequent item sets generation process, such as the transcendental principle, apriori algorithm, candidate item sets generation and pruning algorithm.

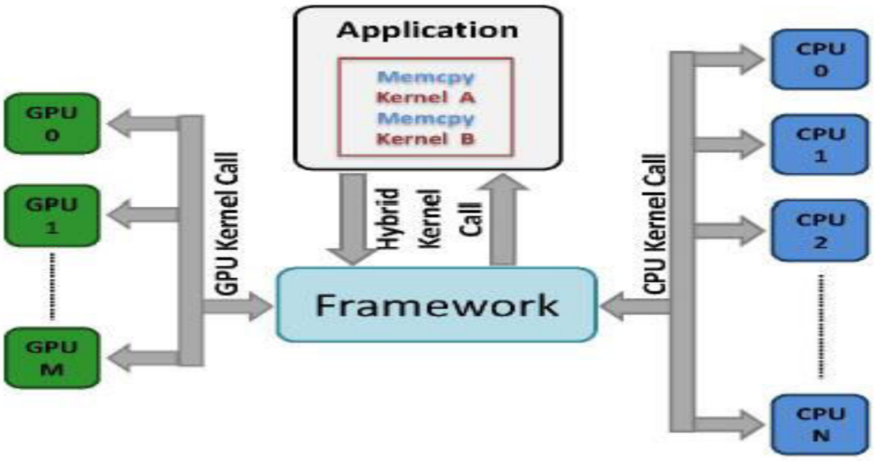
5.4 Association Analysis Application in the Analysis Of Information Security

In this part, we give an IT environment of the government with hundreds of branches as a case. We give an information security risk assessment analysis on years of Information Security Survey and information security events and other research data according to the proposed model relational analysis, and give a step by step sample. This paper describes the data up to the end of 2010 census data completed, and streamlines a part of the data to be example. Exclusion of different industry, the information security census data is up to thousands items, users wanted to investigate the relationship amount some projects according to some empirical, and analyzed

Some valuable information related with information security from the artificial.

5.5 Hybrid Implementation

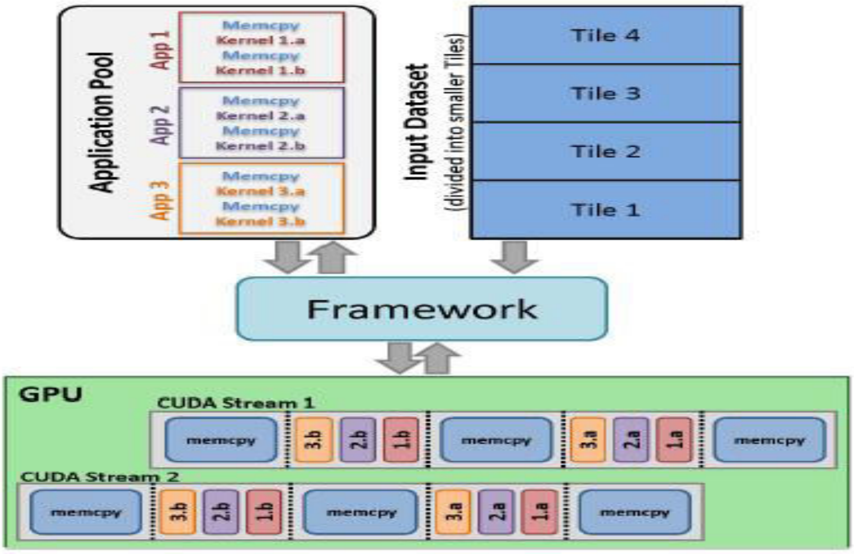
Hybrid implementation refers to harnessing the capabilities of both



GPUs and CPUs simultaneously. Consider a situation when we have a GPU and a multi-core CPU in the system. It would be desirable to distribute the tasks between the GPU and the CPU cores. Since the computational power of GPU is significantly higher than that of the CPUs, the data need to be distributed such that the work remains balanced. Our framework provides the functionality to run an application in the hybrid mode. In this mode, the data will be distributed among the nodes and the corresponding CPU or GPU kernels will be launched. Figure 5 shows how a hybrid kernel call gets broken down into architecture specific kernel calls using the framework.

5.6 Multiple Kernel Optimizations

This optimization is specific to the CUDA implementation. We notice that data mining kernels process huge amounts of data and it is not always possible to fit the entire data in the GPU device memory. As mentioned in Section III-B2, this will require the usage of CUDA Streams to lower the overhead caused by copying data from host memory to the device memory. We further notice in our experiments that kernel execution time is smaller than the time it takes to copy smaller tiles of data to the GPU device memory. This presents us a unique opportunity to leverage the time difference. In practical situations, a number of different data mining algorithms are used on a given dataset. We propose to run kernels from different applications on the dataset while it is in the device memory so as to reduce the overhead of memory copy as much as possible. Figure 6 shows the idea



behind this optimization. As an example, three different applications App1, App2, and App3 are shown in the figure. For each memory transfer call put on a CUDA Stream, one kernel call from each of the applications (1.a, 2.a, and 3.a) is allocated on that particular stream as shown. This can be viewed as a single kernel whose execution time is close to the combined execution time of the same kernels running separately. The kernel execution and host-device memory copy times can be used to predict the number of applications which can be interleaved in the above fashion.

VI. CONCLUSION

As explained earlier the information technology is the double edge sword that's why the security is main issue for it. The developed country are also analyze that information issue is big issue for any country development in modern competitive era. The governing bodies of country or the state will also moves towards the online application which makes the people to easy and quick transaction with government and each other. E-governance is electronic government means everyone in the area will access it so it must be secure that's why no one can suffer from the any security problems. As at the time many people use same server access so the parallel computing will give fast access to the system without any interference and less cost will also beneficial for the governing bodies.

REFERENCES

[1]. High Performance Data Mining Using Data Cubes On Parallel Computers _Sanjay Goil Alok Choudhary ECE Department and CPDC, Northwestern University Technological Institute, 2145 Sheridan Road, Evanston IL-60208.

[2]. High Performance Data Mining Using R on Heterogeneous Platforms, Prabhat Kumar, Berkin Ozisikyilmaz, Wei-Keng Liao, Gokhan Memik.

[3]. Alok Choudhary Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA, fpku649, boz283, wkliao, memik, choudharg@ece.northwestern.edu.

[4]. Parallel Data Mining Techniques on Graphics Processing Unit, with CUDA, Liheng Jian Ying Liu(yingliu@gucas.ac.cn) Shenshen Liang, School of Information Science and Engineering, Graduate University of Chinese Academy of Sciences, Beijing, China

[5]. Data Mining Applications in E-Government Information Security Tongwei yuan*, Peng Chen.

[6]. http://www.infosecurity-us.com/view/11292/. Cyber crimes cost organizations 38 million per year. 27 July 2010.

[7]. http://www.edu.cn. Network black industry challenge network security 04 Jan 2011.

[8]. Litao, Zhang; Security management model research based on system boundary analysis information. (2005)

[9]. Chen. Jinbo; The study on the application of data mining in telecommunication CRM. (2006)

[10]. PangNing Tan, Michael Steinbach, Vipin Kumar; Introduction to Data Mining; Pearson Addison Wesley (2006) 200-238

[11]. J.A.Berry Michael, S.Linoff Gordon; Data Mining Techniques; Wiley Publishing (2011) 245-290.

AN EFFICIENT FDM-KC ALGORITHM FOR SECURE MINING IN HORIZONTALLY DISTRIBUTED DATABASES USING ASSOCIATION RULES

C M Sumana[1], Rajeswari R P[2]

[1]B.E., M.Tech, [2]Assistant Professor, Dept. of Computer Science and Engineering,
RYMEC, Ballari (India)

ABSTRACT

In recent years, Data mining is the most fast growing area in which data mining is the process of finding correlations or patterns among various fields in large relational databases in which it is used to extract important knowledge from large datasets, but sometimes these datasets are split among various parties.We propose a protocol for secure mining of association rules in horizontally distributed databases. The current integral protocol is that of Kantarcioglu and Clifton well known as K&C protocol. This proposed protocol is based on an unsecured distributed version of the Apriori algorithm termed as Fast Distributed Mining (FDM) algorithm of Cheung et al.

The main constituents in this protocol are two novel secure multi-party algorithms one that computes the union of private subsets that each of the interacting players hold and another that tests the whether an element held by one player is included in a subset held by another. This protocol offers enhanced privacy with respect to the protocol. This is simpler and is significantly more efficient in terms of communication rounds, communication cost and computational cost.

Keywords: Privacy Preserving Data Mining, Distributed Computation, Frequent Item sets, Association Rules.

I. INTRODUCTION

Data mining is the extraction of interesting patterns or knowledge from huge amount of data. Today, we have far more information than we must handle from business transactions and scientific data, to satellite pictures, text reports and military intelligence. Thus, Data mining is the computational process of discovering patterns in large data sets. In simple words, it is the process of analyzing data from different perspectives and summarizing it into useful information. A distributed database can reside on network servers on theInternet, on corporate internets or extranets, or on other company networks. Because they store data across multiple computers, distributed databases can improve performance at end-user worksites by allowing transactions to be processed on many machines, instead of being limited to one.

Homogeneous distributed database all sites have identical software and are aware of each other and agree to cooperate in processing user requests. Each site surrenders part of its autonomy in terms of right to change schema or software. A homogeneousDDBMS(Distributed database management system)appears to the user as a single system. The homogeneous system is much easier to design and manage. The following conditions must be satisfied for homogeneous database:

1. The operating system is used, at each location must be same or compatible.

2. The data structures used at each location must be same or compatible.

3. The database application used at each location must be same or compatible.

Data miningis ready for application in the business community because it is supported by three technologies that are now sufficiently mature:

- Massive data collection
- Powerful multiprocessor computers
- Data mining algorithms

The limitations of frequent or rare itemset mining motivated to develop a secure based mining approach, which allows a user to conveniently express his or her perspectives concerning the usefulness of itemsets as secure values and then find itemsets with high utility values higher than a user-specified threshold.

In the literature we have studied the different methods proposed secure mining from large datasets.That goal defines a problem of secure multi-party computation. In such problems, there are M players that hold private inputs, $x1, \dots, xM$, and they wish to securely compute $y = f(x1, \dots, xM)$ for some public function f. If there existed a trusted third party, the players could surrender to him their inputs and he would perform the function evaluation and send to them the resulting output. In the absence of such a trusted third party, it is needed to devise a protocol that the players can run on their own in order to arrive at the required output y. Such a protocol is considered perfectly secure if no player can learn from his view of the protocol more than what he would have learnt in the idealized setting where the computation is carried out by a trusted third party. Thus to overcome this challenges the efficient algorithm is presented in this paper.

The main aim of this proposed protocol is to achieve the following aspects:

- Reducing the number of scans in the original database.
- Distributed databases remain up-to-date and current replication and duplication.
- Minimize memory utilization (Reducing the search space).
- Reducing the total execution and computation time.
- Reducing the resource utilization.
- Increase the performance in terms of time and space complexity.

The proposed protocol improves upon that in terms of simplicity and efficiency as well as privacy. We propose an alternative protocol for the secure computation of the union of private subsets. In particular, our protocol does not depend on commutative encryption and oblivious transfer (what simplifies it significantly and contributes towards much reduced communication and computational costs). While our solution is still not perfectly secure, it leaks excess information only to a small number (three) of possible coalitions, unlike the protocol of that discloses information also to some single players. In addition, we claim that the excess information that our protocol may leak is less sensitive than the excess information leaked by the protocol.

We propose here computes a parameterized family of functions, which we call threshold functions, in which the two extreme cases correspond to the problems of computing the union and intersection of private subsets. Those are in fact general-purpose protocols that can be used in other contexts as well. Another problem of secure multiparty computation that we solve here as part of our discussion is the set inclusion problem namely, the problem where Alice holds a private subset of some ground set, and Bob holds an element in the ground set, and they wish to determine whether Bob's element is within Alice's subset, without revealing to either of them information about the other party's input beyond the above described inclusion.

II. METHODOLOGY

Process Design

Consider D be a transaction database. The database is partitioned horizontally between P1, P2 . . . , Pm players, denoted 1 M. Player Pm holds the partial database Dm that contains Nm = lDm l of the transactions in D, 1≤m≤M . The unified database is D =D_1 U···U DM. An itemset X is a subset of A. Its global support, supp(X), is the number of transactions in D that contain it. Its local support, sup (X), is the number of transactions in D_m that contain it.

Support

The rule X ⇒ Y holds with support s if s% of transactions in D contains X ∪ Y. Rules that have a s greater than a user-specified support is said to have minimum support or threshold support. The support of rule is defined as,sup(X) = no of transactions that contain X / total no of Transactions.

Confidence:

The rule X ⇒ Y holds with confidence c if c% of the transactions in D that contain X also contain Y. Rules that have a c greater than a user-specified confidence is said to have minimum confidence or threshold Confidence. The confidence of a rule is defined as, conf(X =>Y) = sup(X U Y)/ supp(X).

2.1 Apriori Algorithm

Apriori is designed to operate on databases containing transactions. The purpose of the Apriori Algorithm is to find associations between different sets of data. It is sometimes referred to as "Market Basket Analysis". Each set of data has a number of items and is called a transaction. The output of Apriori is sets of rules that tell us how often items are contained in sets of data.

2.2 Algorithm - Fast Distributed Mining (Fdm)

The FDM algorithm proceeds as follows:

(1) Initialization

(2) Candidate Sets Generation

(3) Local Pruning

(4) Unifying the candidate item sets

(5) Computing local supports

(6) Broadcast Mining Results

III. SYSTEM ARCHITECTURE

Data mining consists of various techniques that are applied for secure mining of association rules in horizontally distributed database. In the figure 1 shows two novel secure multiparty algorithmis to provide enhanced privacy, security, and efficiency. In this paper we propose a protocol for secure mining of association rules in horizontally distributed database. This protocol is based on FDM Algorithm which is an unsecured distributed version of the Apriori algorithm. In this protocol two secure multiparty algorithms are involved:

1. Computes the union of private subsets that each interacting players hold.
2. Tests the inclusion of an element held by one player in subset held by another.

Mapper is a database management and processing system. It is a software tool that enables end-users to share computer power in a corporation. Mapper is mapping between database attributes to the java object.Make the given itemsets are to be frequent itemsets in horizontally distributed database by using association rules, privacy preserving techniques and FDM algorithm to provide secure mining.

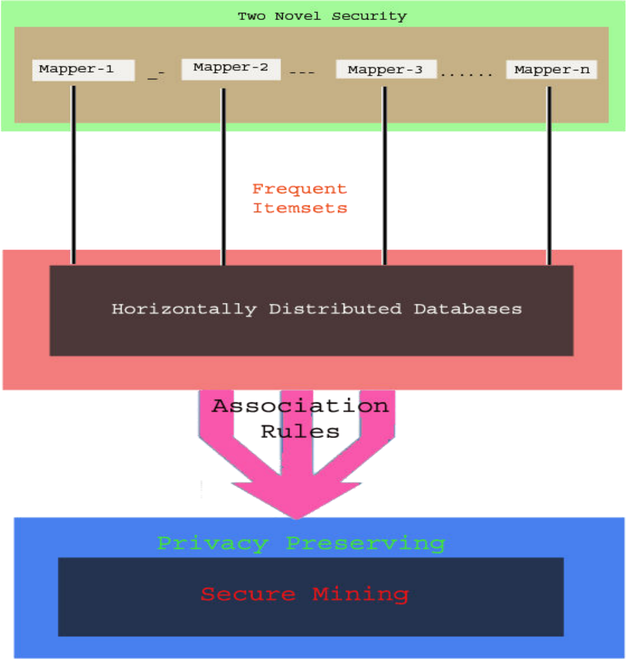


Figure 1: System Architecture for Secure Mining of Association Rules in Horizontally Distributed Data Bases

IV. CONCLUSION

The main problems with the existing methods are the generation a huge set of candidate items and scanning of the original database several times. In this proposed a protocol for secure mining of association rules in horizontally distributed databases that improves significantly upon the current leading protocol in terms ofprivacy and efficiency.In this proposedprotocolhasthepotentialtoovercomeseveral restrictionsinthecurrent data mining model thatcanprevent system'sfunctionalityandlimitations,andmakefurther refinements.

One of the main ingredients in our proposed protocol is a novel secure multi-party protocol for computing the union (or intersection) of private subsets that each of the interacting players hold. Another ingredient is a protocol that tests the inclusion of an element held by one player in a subset held by another. Those protocols exploit the fact that the underlying problem is of interest only when the number of players is greater than two.

Since the algorithm generates few candidate items it takes less time to find the frequent itemsets. And also the memory used is less compared to the existing algorithms. Thus, pruning the itemsets very well at early stages saves the time as well as space.

V. ACKNOWLEDGEMENT

I consider it is a privilege to express my gratitude and respect to all those who guiding me in the progress of my paper.

I wish my grateful thanks to Smt Rajeswari R PM.Tech, project guide, for her invaluable support and guidance.

C M Sumana

REFERENCES

[1]. R. Agrawal and R. Srikant.Fast algorithms for mining association rules in large databases. In VLDB, pages 487–499, 1994.

[2]. R. Agrawal and R. Srikant.Privacy-preserving data mining. In SIGMOD Conference, pages 439–450, 2000.

[3]. D. Beaver, S. Micali, and P. Rogaway. The round complexity of secure protocols. In STOC, pages 503–513, 1990.

[4]. M. Bellare, R. Canetti, and H. Krawczyk. Keying hash functions for message authentication. In Crypto, pages 1–15, 1996.

[5]. A. Ben-David, N. Nisan, and B. Pinkas. FairplayMP - A system for secure multi-party computation. In CCS, pages 257–266, 2008.

[6]. J.C. Benaloh. Secret sharing homomorphisms: Keeping shares of a secret secret. In Crypto, pages 251–260, 1986.

[7]. J. Brickell and V. Shmatikov.Privacy-preserving graph algorithms in the semi-honest model. In ASIACRYPT, pages 236–252, 2005.

[8]. D.W.L. Cheung, J. Han, V.T.Y. Ng, A.W.C. Fu, and Y. Fu. A fast distributed algorithm for mining association rules. In PDIS, pages 31–42, 1996.

[9]. D.W.L Cheung, V.T.Y. Ng, A.W.C. Fu, and Y. Fu. Efficient mining of association rules in distributed databases. IEEE Trans. Knowl. DataEng., 8(6):911–922, 1996.

[10]. T. ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms.IEEE Transactions on Information Theory, 31:469– 472, 1985.

BIOGRAPHY

C M Sumana,is a student pursuing her Master degree in Computer Science and Engineering department at RYMEC,Ballari, Karnataka, India. Her research interests are Computer Science related aspects such as Data minig technology, Java programming language and web technology.

Rajeswari R P, is an assistant professor in the department of Computer Science and Engineering at RYMEC, Ballari, Karnataka, India. She received her Master degree in computer science and engineering. Her research interests are related to Biomedical image processing and big data analytics in health care.

GREEN SYNTHESIS AND MODELING

Gnanasangeetha D[1], Sarala Thambavani D[2]

[1]Assisant ProfessorDepartment of Chemistry, PSNA College of Engineering and Technology, Dindigul, Tamilnadu, (India)

[2]Associate Professor Research and Development Centre, Bharathiar University, Coimbatore, Tamilnadu, (India)

ABSTRACT

An Artificial Neural Network (ANN) representation was developed to anticipate the biosorption proficiency of Zinc oxide nanoparticle deep-seated on activated silica using Emblica officinalis (ZnO-NPs-AS-Eo) for the confiscation of total As (III) from aqueous solution based on 95 data sets obtained in a laboratory batch study. Experimental parameters affecting the biosorption succession such as initial concentration, dosage, pH, contact time and agitation were intended. A contact time of 120 min was drivable to bring about equilibrium. The utmost adsorption capacity of (ZnO-NPs-AS-Eo) in AS (III) removal was found to be 2.96 g/L. The sensitivity analysis confirmed that MSE values decreased as the number of variables used in the ANN model increased. The qualified increase in the performance due to inclusions of V2, adsorbent dosage and V5, agitation speed is larger than the contribution of other variables. The projected ANN model provided pragmatic experimental data with a reasonable correlation coefficient of 0.999 for two operating variables like adsorbent dosage and agitation speed.

Keywords: Agitation, Back propagation, Concentration, Contact time, Dosage, Neurons, pH, Zinc oxide

I. INTRODUCTION

In waste water treatment various technologies are available such as chemical precipitation, ion exchange, electrochemical precipitation, solvent extraction, membrane separation, concentration, evaporation, reverse osmosis, emulsion per traction and adsorption. Among these technologies adsorption is a user friendly technique for the removal of heavy metals. This process includes the selective transfer of solute components in the liquid phase onto the surface or onto the bulk of solid adsorbent materials. In last two decades artificial neural network (ANN) models have been extensively studied in different fields of engineering, finance with a basic objective of achieving human like performance. The neural networks are powerful tools to identify underlying highly complex relationships from input– output data. ANN derived from the biological counterparts and based on the concept that a highly interconnected system of simple processing elements known as nodes or neurons, enables to learn highly complex nonlinear interrelationships existing between input and output variables of the data-set. In ANN model of system feed-forward architecture namely multilayer perception (MLP) is most commonly used. This network consists of at least three layers namely input layer, one or several hidden layers and output layer. Each layer consists of a number of elementary processing units known as neurons. Each neuron in the input is connected to its hidden layer through weights. Also there is connection between hidden and output layers. When an input is introduced to the neural network the synaptic weights between the neurons are simulated and these signals propagate through layers and the output result is formed. The main objective is to form the output by the network in such a way that it should be close to the expected output. The weights between the layers and the neurons are modified in such a way that next time the same input will provide an output that are closer to the expected output. Various algorithms are available for the training of the neural networks. Feed-forward back propagation (BP) algorithm is the most versatile and robust technique which provides the most efficient learning procedure for MLP networks. This algorithm is especially capable of solving predictive problems . Researchers pointed out that increasing the number of hidden layers enables a trade-off between smoothness and closeness-of-fit. The greater number of hidden layers improves the closeness-of-fit while a smaller number of hidden layers improve the smoothness or extrapolation capability of the ANN. Single hidden layer with arbitrarily large quantity of neurons is capable of modeling accurately. It is also observed that two hidden layer networks are better than the single hidden layer network for specific problem . Single hidden layer can solve most of the problems for more input variables and outputs. Recently researchers have successfully modeled a three layer feed forward BP network to predict the removal of Cu(II) from industrial leachate by pumice and Zn(II) from hazelnut shell . The present paper deals with a development of a more general and system-independent neural network based on MLP having a single hidden layer trained with BP and Levenberg-Marquardt (LM) algorithms for the prediction of the percentage removal of As (III) from aqueous solution using five different variables under different operating conditions using two different transfer functions in a single hidden layer. Recently, the use of neural networks has gained popularity for modeling biological wastewater treatment processes . The details of the adsorption study of these adsorbents are reported in our earlier publications and the relevant experimental data are taken for this ANN analysis [1-7].

II. MATERIALS AND METHODS

2.1. Adsorbent Preparation and Characterisation

Aqueous leaf extract of Emblica officinalis was stirred for 30 min to that 1g of Zinc acetate dihydrate was added under vigorous stirring. After 1hr stirring 10 g of activated silica was introduced into the above solution followed by the addition of aqueous NaOH resulted in a white aqueous solution at pH 12. This was then sited in a magnetic stirrer for 2hr. The activated silica supported ZnO nanoparticle were then filtered and washed with double distilled water. The synthesized ZnO-NPs-AS-Eo was maintained at 60ºC for 12 hrs. ZnO-NPs-AS-Eo structure was primed by green synthesis method. A mortar was used to homogeneously ground ZnO-NPs- AS-Eo. The proposed sorbent were stored in air at room temperature. The X-Ray powder diffraction pattern of the as- synthesized sample was recorded on an X-ray diffractometer (XRD, PW 3040/60 Philips X'Pert) using Cu (Kα) radiation (λ =1.5416 A°) operating at 40 kv and 30 mA with 2θ ranging from 10- 90°.The external morphology of the sample were characterized by scanning electron microscope (SEM) (LEO 1530FEGSEM).

2.2. Batch Adsorption Studies

The equilibrium sorption capacity of the sorbent at the corresponding equilibrium conditions was calculated using a mass balance equation as in Eq. (1).

$$Qe = \frac{Ci - Ce}{M} xV \quad \text{———} \quad (1)$$

where Q_e is the amount of the metal uptake by the bioadsorbent (mg/g) in the equilibrium; C_i is initial metal ion concentration in solution (mg/L); C_e is the equilibrium metal ion concentration in solution (mg/L); V is volume of the medium (L); and M is the amount of the bioadsorbent used in the reaction mixture (g). The percent removal (%) of As (III) was calculated using the following equation:

$$Removal\ \% = \frac{Co - Ce}{Co} x100 \quad \text{———} \quad (2)$$

Where C_0 and C_e are the initial and final equilibrium As(III) concentration.

Batch adsorption experiments were conducted in 250 mL glass-stoppered, Erlenmeyer flasks with 20 mL As (III) solution of desired concentration and pH. A weighed amount of adsorbent was added to the solution. The flasks were agitated at a constant speed of 250 rpm until reaching equilibrium. The influence of pH (1.0 , 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0), adsorbent dose (0.5, 1, 1.5. 2, 2.5, 3, 3.5, 4, 4.5 5, 5.5, 6, 6.5,7,7.5,8 g), contact time (10, 20, 30,40, 50, 60, 70, 80, 90,100,110,120,130 min) and initial As(III) concentration (0.005,0.075,0.01,0.02,0.03, 0.04 ,0.05, 0.06, 0.07 ,0.08, 0.09, 0.1N) were evaluated during the present study. Each test lasted for nearly 2 h after which the adsorbent was separated from the solution by centrifugation at 400 rpm for 20 min. The residual As (III) concentration in the adsorbent was then characterized using ED-AX and XRD.

2.3 ANN Structure and its Optimization Procedure

Neural networks can map a set of input patterns onto a corresponding set of output patterns after a series of past process data from a given system have been acquired. Moreover, neural network has a distinctive ability to learn nonlinear functional relationships without the requirement for structural knowledge of the process to be modeled. Among the various ANN models, the one of our interest was the feed forward back propagation network . The feed forward back propagation neural network consisting of forward five neurons corresponding to the five process variables (initial metal ion concentration, pH, time, dosage and agitation) were used in the input layer, twenty in the hidden layer and one in the output layer of the network. The number of neurons per layer should be high enough to allow the network to reproduce the behavior of the system. However, too large of a neuron number can cause data over fitting a situation that can be encountered when correlating experimental data. This is due to the fact that the large number of parameters to be adjusted when using too many neurons might induce the network to memorize the data used in the training while losing one of its more functional characteristics generalization [8-12]. Once the neural network was created it was trained to accurately model the given phenomenon by using the experimental data in MATLAB. The mean square error (MSE) was used as the error function and defined as:

$$MSE = \sum \frac{(y' - y)^2}{n} \quad \text{———} \quad (3)$$

where **y** is the measured values, **ý** the corresponding predicted values and n is the number of samples. Sensitivity tests were conducted to ascertain the relative significance of each of the independent parameters (input neurons) on the removal efficiency (output) in the ANN model. In the sensitivity analysis, each input neuron was in turn eliminated from the model and its influence on prediction of removal efficiency (Qe) was evaluated in terms of correlation coefficient (R^2), and mean square error (MSE).

III. RESULTS AND DISCUSSION

3.1. ANN Model for LM algorithm

Artificial neural network (ANN) models have been used with basic objective underlying highly complex relationships from input–output data of achieving human like performance with accuracy. The main objective is to form the output by the network in such a way that it should be close to the expected output. In the present study, an ANN based model was developed for predicting the As (III) removal efficiency of (ZNO-NPs-AS-*Eo*). The present paper deals with a development of a more general and system-independent neural network based on MLP having a single hidden layer trained with BP and Levenberg-Marquardt (LM) algorithms for the prediction of the percentage removal of As (III) from aqueous solution using five different variables under different operating conditions using two different transfer functions in a single hidden layer. The input layer had five neurons as pH, adsorbent dosage, initial concentration, agitation and contact time while the output layer had the As (III) removal efficiency as the only neuron. In order to determine the optimum number of hidden nodes, a series of topologies was used, in which the number of nodes were varied from 2 to 20. Each topology was repeated three times to avoid random correlation due to random initialization of the weights. The Mean square error (MSE) was used as the error function to measure the performance of the network according to the above equation (3). The MSE was minimum just about 10 neurons. Therefore, the number of neurons in the hidden layer was selected as 10. A regression analysis of the network response between ANN outputs and the corresponding targets performed shows a good agreement between ANN outputs (predicted data) and the corresponding targets (experimental data). The best linear fit was indicated in the Figure 1 with a good correlation coefficient of 0.987.

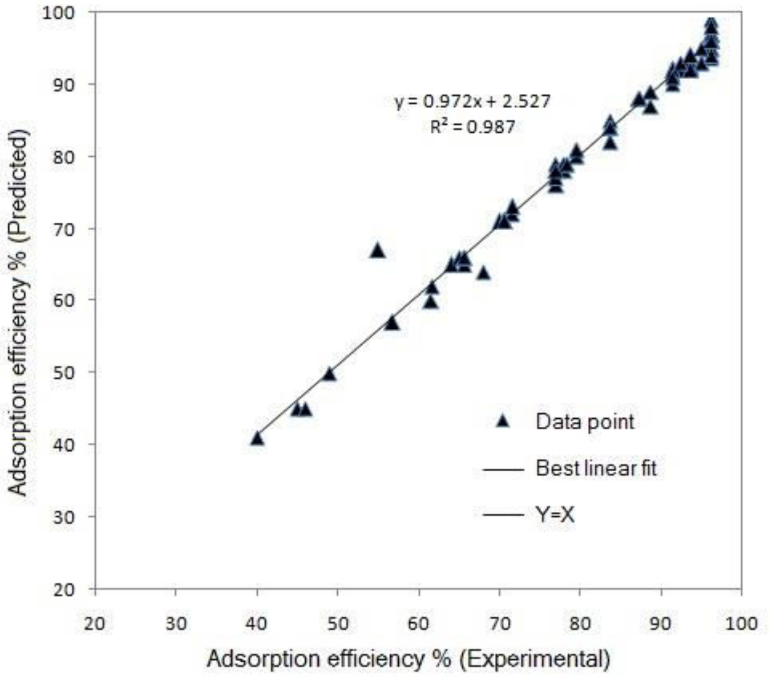


Figure 1 Ann Outputs Plotted Versus the Corresponding Experimental Targets for the Levenberg-Marquardt Algorithm

3.2. Sensitivity Analysis

Table 1 Recital Appraisal of Grouping of Input Variables for Lma With 10 Neurons in The Hidden Layer For Sensitivity Analysis

S.No.	CN	MSE	R^2	IN	Gradient	BLE
1.	**V_1**	**9.06**	**0.917**	**47**	**1.00 x 10^{-5}**	**0.83x +16**
2.	V_2	203.26	0.322	7	2.68 x 10^{-3}	0.35x +53
3.	V_3	110.48	0.449	7	8.48 x 10^{-3}	0.21x +69
4.	V_4	229.17	0.66	10	1.26 x 10^{-4}	0.43x+44
5.	V_5	32.37	0.887	8	7.43 x 10^{-3}	0.77x +20
6.	V_1+V_2	8.38	0.995	56	3.89 x 10^{-3}	0.99x+1
7.	V_1+V_3	71.66	4.530	11	9.34 x 10^{3}	0.36x+61
8.	V_1+V_4	42.63	0.915	09	7.49x10^{3}	0.98x +2.5
9.	V_1+V_5	25.52	0.872	11	1.26 x 10^{3}	0.76x +21
10.	V 2+V_3	35.40	0.873	16	5.23 x 10^{3}	0.75x +22
11.	V_2+V_4	51.27	0.903	17	1.35 x 10^{4}	0.81x +16
12.	**V_2+V_5**	**5.4**	**0.999**	**100**	**7.31 x 10^{4}**	**1x**
13.	V_3+V_4	123.4	0.738	12	5.32 x 10^{3}	0.57 x+37
14.	V_3+V_5	94.12	0.620	13	3.53x 10^{3}	0.35x+57
15.	V_4+V_5	12.04	0.974	23	5.52x 10^{3}	0.94x+5
16.	$V_1+V_2+V_3$	43.67	0.782	13	1.93x 10^{3}	0.6x+35
17.	**$V_1+V_2+V_4$**	**5.71**	**0.985**	**27**	**2.3x 10^{3}**	**0.97x+2.7**
18.	$V_1+V_2+V_5$	8.82	0.963	14	3.68x 10^{3}	0.9x+8.3
19.	$V_1+V_3+V_4$	33.96	0.899	64	1.68x 10^{3}	0.81x+17
20.	$V_1+V_3+V_5$	53.19	0.739	12	7.84x 10^{3}	0.49x+45
21.	$V_1+V_4+V_5$	11.54	0.967	20	6.97x 10^{3}	0.93x+6
22.	$V_2+V_3+V_4$	27.94	0.94	23	5.06x 10^{3}	0.88x+9.9
23.	$V_2+V_3+V_5$	21.95	0.91	33	2.08x 10^{3}	0.8x+16
24.	$V_2+V_4+V_5$	34.32	0.922	14	2.83x 10^{3}	0.85x+14
25.	$V_3+V_4+V_5$	31.43	0.932	26	1.38x 10^{3}	0.86x+11
26.	$V_1+V_2+V_3+V_4$	19.03	0.941	15	1.27x 10^{3}	0.88x+11
27.	$V_1+V_2+V_3+V_5$	25.68	0.895	14	1.66x 10^{3}	0.79x+19
28.	**$V_1+V_2+V_4+V_5$**	**6.94**	**0.98**	**17**	**1.01x 10^{3}**	**96x+3.5**
29.	$V_1+V_3+V_4+V_5$	25.5	0.92	18	2.54x 10^{3}	0.87x+11
30.	$V_2+V_3+V_4+V_5$	2.54	0.94	30	1.89x 10^{3}	0.9x+8.9
31.	V 1+$V_2+V_3+V_4+V_5$	18.55	0.94	21	1.43x 10^{3}	0.89x+9.9

CN-combination; MSE-mean squared error; R^2-correlation coefficient; IN-iteration; BLE- best linear equation; V_1-concentration; V_2-adsorbent dosage; V_3-contact time; V_4-pH; V_5-agitation speed.

In this cram a sensitivity analysis was accomplished to establish the extent of effectiveness of a variable using the projected ANN model. In the analysis recital assessment of various possible combinations of variables were investigated. Therefore performance of the groups of one, two, three, four, and five variables were tested by the optimal ANN structure using the LMA with 10 hidden neurons. The groups of input vectors were defined in this form V_1, initial As (III) ions concentration; V_2, adsorbent dosage; V_3, contact time; V_4, initial pH and V_5, agitation speed. Results of the performance evaluation of 31 combinations are summarized in Table 1. Findings of the sensitivity analysis showed that V_1, initial As (III) ions concentration was found to be the most effective parameter, among those considered in the group of one variable. As shown in Table 1, the MSE value significantly decreased from 71.86 to 5.4 when V_2 + V_5 (dosage & agitation) was used in combination with subsequent group of two variables. The minimum MSE in the group of three variables was determined to be 5.71 using the combination of $V_1 + V_2 + V_4$(As (III) ions concentration ,adsorbent dosage and pH) with a further contribution of V_5 (agitation speed) the MSE decreased up to 6.94 from 25.5 which is the minimum value of the group of four variables. The MSE value significantly deceased 25.5 to 18.55 when V_3 (contact time) was used in combination with other variables in the subsequent group of five variables with reasonable correlation coefficient of 0.94. On the basis of the performance evaluation of combinations of input variables best group performances according to number of parameters are listed in Table 1. The respective MSE values as given in Table 1 show that MSE values decrease as the number of variables in the group increases. Furthermore it can also be concluded that the relative increase in the performance due to inclusions of V_2, adsorbent dosage and V_5, agitation speed is larger than the contribution of other variables. Single hidden layer with arbitrarily 10 neurons is capable of modeling accurately. A single hidden layer solved the problems for 5 input variables and 1 output. This is in agreement with the work reported for adsorption of Lanaset Red G dye on walnut husk [13-17]. These results confirm that the developed ANN model reproduces the adsorption in this system within experimental ranges adopted in the fitting model with best linear correlation of 0.999.

IV. CONCLUSION

The (ZnO-NPs-AS-*Eo*) used as a low-cost adsorbent showed good adsorption performance for removal of As (III) ions from aqueous solutions. Batch adsorption experiments showed that optimal operating initial concentration of 0.05N, pH of 5, an adsorbent dosage of 3 g and agitation speed of 250 rpm and contact time of 30 min was found to be sufficient to achieve equilibrium. The optimal neuron number for the LMA was determined to be 20 hidden neurons with MSE of 5.4 with a tangent sigmoid transfer function (tansig) at hidden layer and a linear transfer function (purelin) at output layer. The proposed ANN model showed a precise and an effective prediction of the experimental data with a satisfactory correlation coefficient of 0.999 for two operating variables like V_2, adsorbent dosage and V_5, agitation speed. The maximum adsorption capacity of the (ZnO-NPs-AS-*Eo*) in As (III) removal was found to be 2.96 g/L. The relative increase in the performance is due to inclusions of V_2, adsorbent dosage and V5, agitation speed. The sensitivity theoretical analysis confirmed that this system was in good agreement with experimental pseudo second order kinetics.

REFERENCES

[1] Bansal, A., Kanuffman, R.J. and Weitz, R.R. Comparing the modeling performance of regression and neural networks as data quality varies: A buisnessvalue approach. J. Manage. Inform. Syst. 10(1), 1993, 11–32.

[2] Barnard, E., Wessels, L.. Extrapolation and interpolation in neural network classifiers. IEEE Control Syst. 12(5), 1992, 50–53.

[3] Celekli, A. and Geyik, F. Artificial neural network (ANN) approach for modelling of removal of Lanaset Red G on Chara contraria. Bioresour. Technol. 102 , 2011, 5634-5638.

[4] Gnanasangeetha, D. and Sarala Thambavani, D. Neural Network Modeling and Sorption of As III with Zinc Oxide Nanoparticle Bounded on Activated Silica using *Ocimum Sanctum*. International Journal of Engineering Sciences & Research Technology. 3(5) , 2014, 206-213.

[5] Haykin, S.Neural Networks—A Comprehensive Foundation, 2nd ed., Prentice-Hall, 1999, USA.

[6] Hornik, K. Approximation capabilities of multilayer feed forward networks. Neural Network. 4, 1991, 251–257.

[7] Imandi, S.B., Karanam, S.K. and Garapati, H.R. Optimization of fermentation medium for the production of lipopeptide using artificial neural network and genetic algorithms. IJNES. 2, 2008, 105–109.

[8] Lee, D.S., Jeon, C.O., Park, J.M. and Chang, K.S. Hybrid neural network modeling of a full-scale industrial wastewater treatment process. Biotechnol Bioeng. 78, 2002, 670–682.

[9] Naiya, T.K., Chowdhuary, P., Bhattacharya, A.K. and Das S.K. Saw dust and neem bark as low-cost natural biosorbent for adsorptive removal of Zn(II) and Cd(II) ions from aqueous solutions. Chem. Eng. J., 148(1), 2009, 68–79.

[10] Pal, M.P., Vaidya, B.K., Desai, K.M., Joshi, R.M., Nene, S.N. and Kulkarni, B.D. Medium optimization for biosurfactant production by Rhodococcus erythropolis MTCC 2794: artificial intelligence versus a statistical approach. J Ind Microbiol Biotechnol. 36, 2009., 747–756.

[11] Plippman R. An introduction to computing with neural nets. IEEE ASSP Mag. 4, 1987, 4–22.

[12] Tamura, S. and Tateishi, M. Capabilities of four layered feedforward neural network: Four layers versus three. IEEE Trans. Neural Networks. 8(2), 1997, 251–255.

[13] Turan, N.G., Mesci, B. and Ozgonenel, O. Artificial neural networks (ANN) approach for modeling Zn(II) adsorption from leachate using a new biosorbent. Chem. Eng. J. 173, 2011, 98–105.

[14] Turan, N.G., Mesci, B. and Ozgonenel, O. The use of artificial neural networks (ANN) for modeling of adsorption of Cu(II) from industrial leachate by pumice. Chem. Eng. J. 171, 2011,1091–1097.

[15] Walczak, S. Developing neural nets currency trading. Artif. Intel. Finance. 2(1), 1995, 27–34.

[16] White, H. Connectionist nonparametric regression: Multilayer feed-forward networks can learn arbitrary mapping. Neural Networks 3, 1990,535–549.

[17] Zhao, H., Hao OJ, McAvoy TJ, Chang CH Modeling nutrient dynamics in sequencing batch reactor. Journal of environmental engineering. 123(4), 1997 , 311-319.

IDENTIFYING DIABETIC PARAMETER IN BASILAR MEMBRANE MECHANICS AND MODELS.

Mrs. Nirmala N. Kamble[1], Dr V R Mankar [2]

[1]Head of Department, Thakur Polytechnic, Electronics Department, Mumbai, Maharashtra (India)
[2] Deputy Secretary, RBTE, Pune (India)

ABSTRACT

The cochlear frequency-place map is believed to be an important determinant of the frequencies that a species can hear. The cochlear frequency-place map is created partially by a stiffness gradient in the basilar membrane (BM) in which stiff regions respond best to high frequencies and more compliant regions respond best to low frequencies.BM mass and stiffness play significant role infrequency –place map. Outer hair cells (OHC) present inside cochlea plays important role for mechanical amplification by introducing active feedback. Mathematical model of active cochlea involves the factor α, the motility factor, which reflects the active feedback mechanism of OHC. Mathematical model of BM shows the presence of stiffness gradient in BM. Both motility factor in cochlear models and stiffness in BM model are affected by prolonged diabetic condition. Diabetic factor in cochlear model is been identified previously. In this paper we present our study of available BM model and identify the diabetic parameter in BM model.

Keywords: Collagen Fibers, Glycation, Ages, Longitudinal Coupling, SEM, TEM.

I. INTRODUCTİON

Studies suggest that the basilar membrane (BM) is structurally designed to support a radial tension (Engström, 1955; Henson and Henson, 1988). The BM is composed of a homogeneous, soft ground substance that is traversed radially by fibers, which extend between the spiral lamina and the spiral ligament.The BM can be divided into two regions based on the arrangement of the fibers: The lateral pectinate zone, where the fibers are grouped into bundles; and the arcuate zone, where the bundles separate into individual fibers(Iurato, 1967). The parallel arrangement of fiber bundles in the pectinate zone suggests that the bundles are under a radial tension (Engström, 1955). Studies also suggest that such a tension is maintained by the spiral ligament. It has been shown that the spiral ligament fibers are anchored to the bony cochlear wall by fibroblasts. (Henson *et al (*1984). Fibroblasts contain fibers composed of contractile proteins and have been shown to create tension (Harris *et al.*, 1981). In the spiral ligament, the configurations in which fibroblasts are arranged and oriented suggest that these cells actively maintain a radial tension in the BM (Henson and Henson, 1988).

Our paper is organized in following manner.

Section I give the details of BM and explain that the fibroblast proteins are basically collagen fibers, the basic building blocks of the BM. Latest Laser-confocal microscopy, high resolution scanning (SEM)and transmission electron microscopy (TEM) reveals BM in detail.

Section II explains that the collagen fibers undergo glycation which can affect its mechanical properties prominently its stiffness which has major effect in BM response.

Section III explains the available mathematical model of BM and its solution.

Section IV identifies the diabetic parameter in BM model and explains the need to study the variation of this parameter.

II. SECTION I: HUMAN BM

The human BM consists of four separate layers: (1) epithelial basement membrane positive for laminin-β2 and collagen IV, (2) BM proper composed of radial fibers expressing collagen II and XI, (3) layer of collagen IV and (4)tympanic covering layer (TCL) expressing collagen IV, fibronectin and integrin[1].

SEM view of the epithelial basement membrane shows it as a carpet-like structure as seen in figure 1. Beneath the basement membrane lies a fibrous layer extending from the tympanic lip of the lamina spiral is to the basilar crest of the spiral ligament which is the proper BM and consisted of various sized radial parallel fiber bundles (10–20 nm in diameter) as seen in figure 2

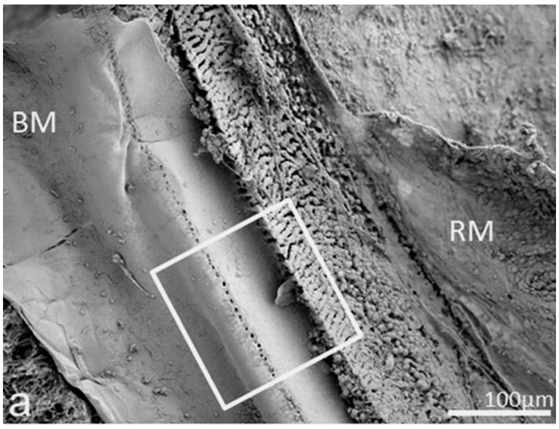


Figure1:SEM view of Basement membrane[1]

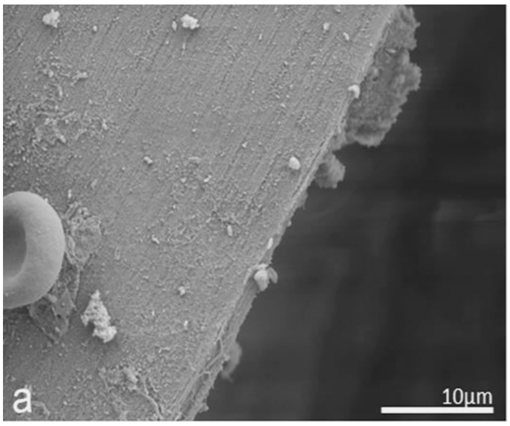


Figure2:SEM view of Proper BM Below Basement membrane [1]

Lateral to the basilar crest, in the spiral ligament, the BM thickened and formed an anchor-like structure from which several fibers emerged. The BM also displayed longitudinally arranged fibrils. The BM merged with the tympanic lip. BM viewed from scala media after rupture of the basement membrane expose collagen fibers as shown in figure 3.

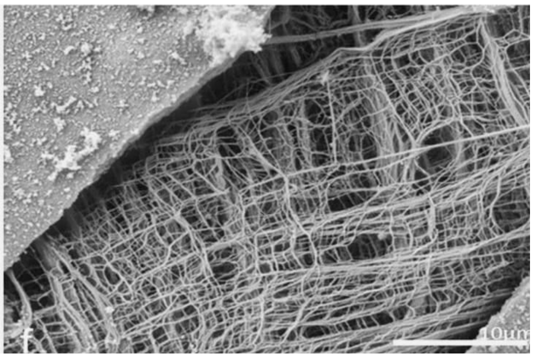


Figure3 SEM view of ruptured proper BM exposing collagen fibers [1]

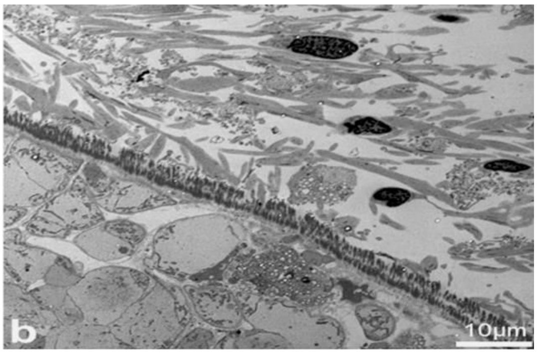


Figure4: TEM view of Proper BM below Basement membrane [1]

Figure 4 shows the SEM view of BM below basement membrane.

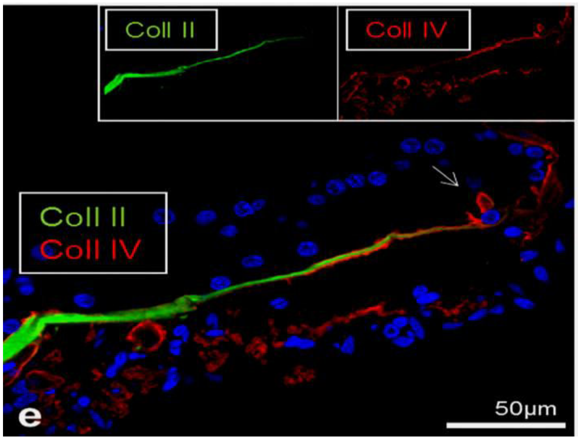


Figure5: TEM View of Proper BM below Basement Membrane [1]

Figure 5 shows Collagen II and IV expression in the BM. Collagen molecules make up 30% of total protein in the body and form the basis of many vital organs (Kadler et al., 2007).The collagen molecule is synthesized as a trimetric molecule containing two α1, and one α2 chains, each of about 1000 amino acids. Upon secretion from the cell, collagen molecules assemble into fibrils and are enzymatically crosslinked (Sweeney et al., 2008).

The BM consists of a filamentous layer consisting primarily of collagen fibers that run radially from medial to lateral i.e., perpendicular to the length of the cochlear duct.

Collagen is regarded as a main source of basilar membrane stiffness and therefore plays a critical role in establishing its tuned resonant frequency map [2].Collagen molecules are packed together in a hexagonal manner to form fibrils, which are covalently bound together with groups of other fibrils to form collagen fibers. Ground substance fills in the spaces between fibers.

III. SECTION II: GLYCATION OF COLLAGEN FIBERS

Glycation and Advanced glycation end products (AGEs)are form in vivo on collagen via non enzymatic reactions that covalently add a sugar moiety onto the protein(Paul and Bailey,1996) and their accumulation is particularly high in long-lived proteins, such as collagen. The low biological turn over of collagen makes it therefore susceptible to interaction with metabolites, primarily glucose.

Several glycation crosslinks have been proposed (Avery and Bailey,2006) but are all present in minute quantities except glucosepane(Sell et al., 2005), which is a lysine–arginine crosslinking AGE which could make a significant change to the biomechanics and biological activity of fibrillar collagen as shown in figure 6.

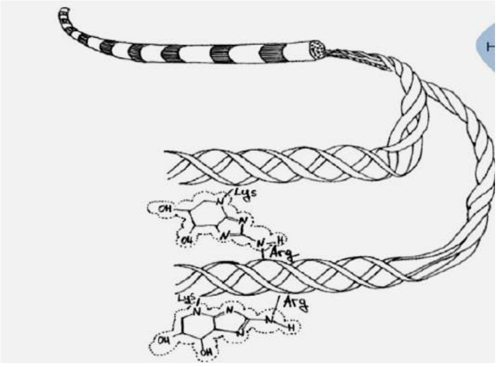


Figure6: Cross Linking in Collagen Fibers [3]

One is the biomechanical effects of non enzymaticinter molecular cross linking: glucose reaction with the amino acidside-chains, and subsequent further reaction to form a crosslink with an adjacent collagen molecule, result in a modification of the physical properties of the collagen (e.g.: elasticity), but the detailed effects of AGEs on collagen nano-mechanics are still unknown. The mechanical effects of AGEs on collagenous tissues are long known and include stiffening, increased failure load and decreased visco elasticity of tendons (Galeski et al., 1977; Andreassen et al., 1981; Daniel sen and Andreassen, 1988; Li et al., 2013).Various studies have shown that already a single non enzymatic crosslink (in addition to the physiological enzymatic ones) is able to drastically change the mechanical properties of collagen tissues, inducing increased stiffness and decreased toughness. Intermolecular crosslinks is likely ≈10 per molecule and thus undoubtedly responsible for the observed stiffening of collagen tissues [3].

IV. SECTION III:BM MODEL

Among various BM models we select the BM model proposed in [4].As per structural mechanics the architecture of the BM is remarkably similar to the architecture of a reinforced slab of concrete [4]. The ground substance may act as the concrete and the fiber bundles can be thought as the reinforcing steel bars as shown in Figure7. It means that the preferential reinforcement of the BM along its width is designed to particularly enhance the, otherwise poor, ability of the BM ground substance to support large radial forces. This is similar to the manner by which reinforcing steel bars greatly increase the tensile strength of concrete along the direction of reinforcement.

At any given location along the cochlea, the BM is modeled as a plate as shown in Figure7. The dimensions of the plate change with position along the cochlea. The width of the plate is oriented along the radial direction(x) of the cochlea. The length of the plate is oriented along the length (y) of the cochlear spiral.

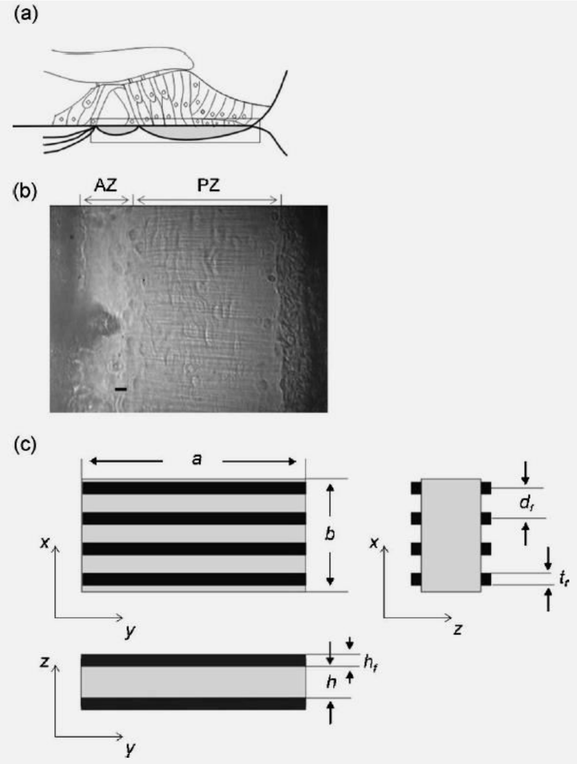


Figure 7 BM Plate Approximations for Model [4]

The BM plate is composed of two components: (a) Anisotropic plate of BM ground substance; and (b) two sets of fiber bundles, which travel along the top and bottom of the plate in order to preferentially reinforce the plate along its width. Due to the preferential reinforcement by the fiber bundles along the *y* direction, the plate is an isotropic, whereby it has different mechanical properties along its width and length.

4.1 Determination of BM Plate Dimensions

At a given position, *x*, along the cochlea, the dimensions of the plate are determined as follows. The width of the plate, *a*, is equal to the width of the BM, *w*(*x*). The length of the plate, *b*, is calculated as

$$b(x) = 5\lambda_c(x), \qquad (1)$$

where $\lambda_c(x)$ is the space constant that describes the amount of BM longitudinal coupling[5]. As shown in the figure 8 the plate length is taken as per longitudinal coupling involved which means relevant length affected by deflection for model estimation

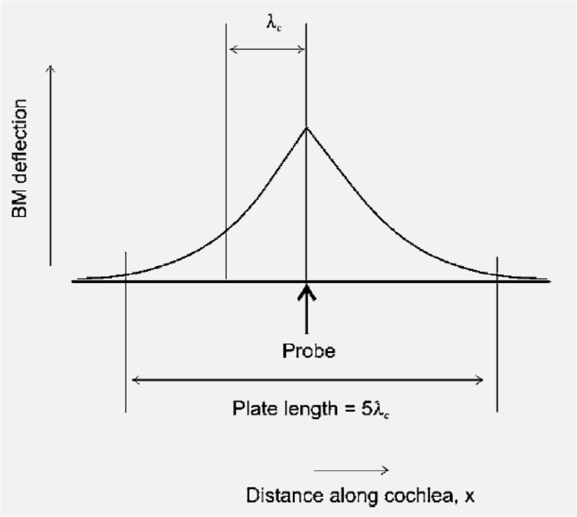


Figure 8: Deflection Profile from Excitation [4]

The effective thickness, *h(x)* of the plate is calculated as the cross-sectional area of the BM divided by its width.

4.2 Determination of BM Plate Material Properties

The deflection equation of a plate is determined by its flexural rigidities, *Dx*, *Dy*, and *H*, which are functions of the plate architecture, the elasticity of the fibers, and the elasticity of the ground substance. Given the preferential reinforcement of the plate along the radial (*x*) direction, the flexural rigidities, *Dx*, *Dy*, and *H*, of the plate can be calculated (Timoshenko and Woinowsky- Kreiger, 1959) to be

$$D_x(x) = \frac{E_g h^3(x)}{12(1-v^2)} \qquad (2)$$

$$D_y(x) = \frac{E_g h^3(x) + E_f I_f(x)}{12(1-v^2) d_f(x)} \qquad (3)$$

$$H(x) = D_x(x) \qquad (4)$$

Where E_g and *Ef*r represent the elasticity moduli of the ground substance and fiber bundles, respectively, *v* is Poisson's ratio of the ground substance, *h(x)* is the thickness of the BM plate, $d_f(x)$ is the spacing between fiber bundles, and

$I_f(x)$ is the moment of inertia of each set of fiber bundles (top and bottom) with respect to the middle axis of the cross section of the plate calculated as

$$I_f(x) = \frac{1}{12} t_f(x)h^3(x)[1-(1-\frac{2h_f(x)}{h(x)})^3]$$

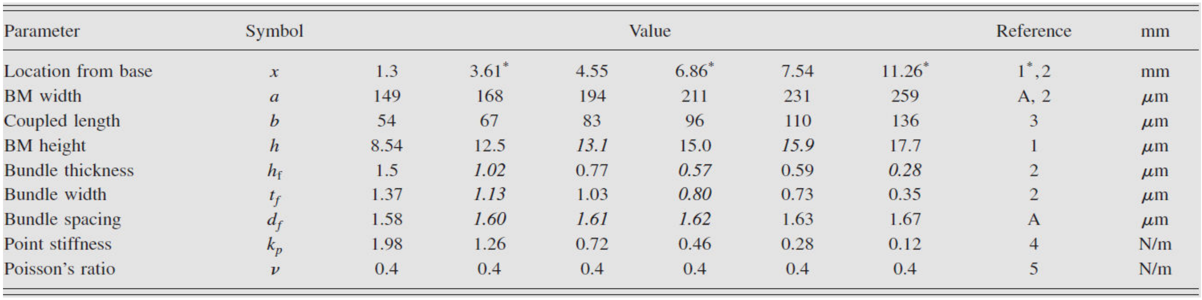
This can be approximated as

$$I_f(x) = \frac{1}{12} t_f(x)h^2(x)h_f(x) \qquad (5)$$

Where $h_f(x)$ is the height of fiber bundles and $t_f(x)$ is the thickness of fiber bundles.

4.3 Parameter Estimation

The model was used to determine BM tension at six locations along the cochlea [4]. Table 1 lists the experimental measurements which were used to calculate the model dimensions and mechanical properties.



Parameter	Symbol	Value						Reference	mm
Location from base	x	1.3	3.61*	4.55	6.86*	7.54	11.26*	1*,2	mm
BM width	a	149	168	194	211	231	259	A, 2	μm
Coupled length	b	54	67	83	96	110	136	3	μm
BM height	h	8.54	12.5	*13.1*	15.0	*15.9*	17.7	1	μm
Bundle thickness	h_f	1.5	*1.02*	0.77	*0.57*	0.59	*0.28*	2	μm
Bundle width	t_f	1.37	*1.13*	1.03	*0.80*	0.73	0.35	2	μm
Bundle spacing	d_f	1.58	*1.60*	*1.61*	*1.62*	1.63	1.67	A	μm
Point stiffness	k_p	1.98	1.26	0.72	0.46	0.28	0.12	4	N/m
Poisson's ratio	ν	0.4	0.4	0.4	0.4	0.4	0.4	5	N/m

Anatomical dimensions of the fiber bundles were obtained in following manner. The thickness and height of the fiber bundles were determined from table 1 [6]. At each location of interest, the average thickness of the fiber bundles was calculated as the mean of the measured thickness of the upper and lower fiber bundles. The spacing of the fiber bundles at locations in the three turns were obtained using digitized images of the BM in excised turns of gerbil cochlea.

The stiffness at the center of the BM was estimated from the measurements of Naidu and Mountain (1998b) as being equal to one-half the stiffness of the OC as measured under the outer hair cell region. The stiffness was experimentally measured at a BM deflection of $z_0$approximately equal to5µm. Measurements of longitudinal coupling of the BM were obtained from Naidu and Mountain (2001). Poisson's ratio of the ground substance of the BM was estimated to be0.4 from measurements on dog lung tissue by Lai-Fook *et al. (*1976).

The ratio of the elasticity of the fibers to the elasticity of the ground substance was estimated to be 100. This ratio value makes the BM weakly an isotropic in the apex, which is consistent with the experimental measurements of Naidu and Mountain (1998a).

4.4 Model Solutions

The general equation describing the deflection, *z*(*x*, *y*),of the anisotropic BM plate in response to a generalized load ,*q*(*x*, *y*), and tension, *Ny*, along the *y* direction is given by

$$D_x \frac{\partial^4 z(x,y)}{\partial x^4} + 2H \frac{\partial^4 z(x,y)}{\partial x^2 \partial y^2} + D_y \frac{\partial^4 z(x,y)}{\partial y^4} - N_y \frac{\partial^2 z(x,y)}{\partial y^2}$$
$$= q(x,y), \qquad \text{(A1)}$$

where the flexural rigidities *Dx*, *Dy*, and *H* are given by Eqs (2)–(4).

The deflection in response to a concentrated force, *P* ,applied at the center of the plate may be then obtained from Eq(A1) in the form of a trigonometric series (Timoshenko and Woinowsky-Krieger, 1959).

$$z(x,y) = \frac{4P}{ab\pi^4} \times \sum_{n=1,3\ldots}^{\infty} \sum_{m=1,3\ldots}^{\infty} \frac{\sin\frac{m\pi}{2}\sin\frac{n\pi}{2}}{\left[D_x\frac{m^4}{b^4} + 2D_x\frac{m^2n^2}{a^2b^2} + D_y\frac{n^4}{a^4} + \frac{N_x m^2}{\pi^2 a^2}\right]} \times \sin\frac{m\pi x}{b}\sin\frac{n\pi y}{a}, \qquad \text{(A2)}$$

where *H* has been substituted by *Dx* using Eq (4). After approximation the deflection profile is then given by the expression

$$z(x,y) = \frac{4Pab}{K\pi^4 D_x}\frac{1}{(1+\alpha)}\sin\frac{\pi x}{a}\sin\frac{\pi y}{b}, \qquad \text{(A3)}$$

where

$$K = \left(\frac{D_y}{D_x}\frac{b^2}{a^2} + 2 + \frac{a^2}{b^2}\right) \qquad \text{(A4)}$$

and

$$\alpha = \frac{N_y b^2}{\pi^2 D_x K}. \qquad \text{(A5)}$$

Two constraints are then applied to calculate the BM tension. First, the physiological stiffness, *kp*, is measured experimentally at the center of the BM at a deflection, *zo* in response to a force, *P*. The measured stiffness must be equal to the stiffness at the center of the plate calculated from the deflection profile using Eq (A3) as follows:

$$\frac{1}{k_p} = \left.\frac{dz(a/2,b/2)}{dP}\right|_{z_p} = \frac{z_o(a/2,b/2)}{P} = \frac{4ab}{K\pi^4 D_x}\frac{1}{(1+\alpha)}. \qquad \text{(A6)}$$

Second, during deflection, the tensile strain must balance the strain due to deflection. The strain, ε_y, generated in the plate by the tensile stress, σ_y

$$\varepsilon_y = \frac{\sigma_y}{E_y} = \frac{N_y(1-\nu^2)}{hE_y}, \qquad \text{(A7)}$$

where *Ey* represents the equivalent elasticity modulus along the *y* direction, which is related to the flexural rigidity *Dy* by

$$D_y = \frac{E_y h^3(x)}{12}. \qquad \text{(A8)}$$

The strain, ε_d due to deflection is given by

$$\varepsilon_d = \frac{1}{2a}\int_0^a \left(\frac{dz(b/2,y)}{dy}\right)^2 dy, \qquad \text{(A9)}$$

where the right hand side of the equation is the ratio of the plate extension produced by deflection to the width *a* of the plate (Timoshenko, 1955). The plate extension is approximated by the difference between the arc length of the deflection profile and *a* Equating the tensile strain and strain due to deflection, we get

$$\frac{N_y(1-\nu^2)}{hE_y} = \frac{1}{2a}\int_0^a \left(\frac{dz(b/2,y)}{dy}\right)^2 dy. \qquad \text{(A10)}$$

Substituting for *z (b*/2, *y)* using Eq. (A3), and substituting for *Ey* using Eq. (A8), Eq. (A10) becomes

$$\boxed{\frac{N_y h^2}{3D_y} = \left[\left(\frac{4Pab}{k\pi^4 D_x}\right)\left(\frac{1}{1+\alpha}\right)\right]^2\left(\frac{\pi}{a}\right)^2. \qquad (A11)}$$

By dividing Eq. (A11) by Eq.(A6) and substituting for *Ny* using Eq. (A5), the value of α is calculated as

$$\boxed{\alpha = \frac{3}{Kh^2} z_o^2 \left(\frac{b}{a}\right)^2 \left(\frac{D_y}{D_x}\right). \qquad (A12)}$$

The tension, *Ny*is then be obtained from α by dividing Eq. (A5) by Eq. (A6)

$$\boxed{N_y = \frac{4a}{\pi^2 b}\frac{\alpha}{(1+\alpha)}k_p. \qquad (A13)}$$

Dx and *Dy* may then be calculated from Eqs. (A5),(3), and (4).

4.5 Model Solution

The model is used to calculate BM tension by simulating the experimental measurements of BM stiffness (Naidu and Mountain, 1998b). An analytical expression is derived for the deflection profile of the plate in response to a concentrated force, which is applied at the center of the BM. The deflection profile is a function of the elasticity of the fibers and ground substance, and the radial tension. Since the ratio of the elasticity of the fibers and ground substance is assumed to be 100, the deflection profile is only a function of two unknown variables, the elasticity of the ground substance and the radial tension. Therefore, only two equations, instead of three, are required to solve the tension of the BM. The two equations are determined as follows. The first equation is obtained by equating the calculated stiffness to the stiffness as measured experimentally. The second equation is derived from the constraint that edges of the BM are not observed to move towards each other when the BM is deflected during the experimental stiffness measurement.

Mathematically, the constraint is equivalent to balancing the strain developed in the BM due to deflection by the strain developed in the BM due to tension. The two equations formulated as described above are solved simultaneously to calculate the tension in the BM.

4.6 Model Results

Figure 9(a) shows the calculated BM tension *Ny*(*x*) plotted as a function of position along the length of the cochlea. The tension decreases by about two orders of magnitude from a value of about 0.76 N/m at the base to about0.001 N/m at the apex. The continuous variation in tension, in units of N/m, with distance *x* along the cochlea is described by a regression fit to the predicted values, which is given by the function

$$\boxed{N_y(x) = 1.51e^{-0.58x},}$$

where *x* is specified in units of mm. The corresponding tensile stress σ_y(x) acting on the BM was calculated by dividing *Ny*(*x*) by the effective height, *h*(*x*), of the BM plate. Figure 9(b) shows the calculated tensile stress as a function of position along the length of the cochlea. The continuous variation in tensile stress, in units of N/m2, with distance *x* along the cochlea is described by a regression fit to the predicted values, which is given by the function

$$\boxed{\sigma_y(x) = 1.671\,66 \times 10^5 e^{-0.65x},}.$$

where *x* is specified in units of mm. The predicted flexural rigidities of the plate, *Dx* and *Dy*, are plotted as a function of position along the cochlea in Fig.4. The continuous variation in *Dx* and *Dy*, in units of Nm, with distance *x* along the cochlea are described by the regression fit to the predicted values, which are given by the functions

$$\boxed{D_x(x) = 3 \times 10^{-11} e^{-0.01x},}$$

$$\boxed{D_y(x) = 3.3 \times 10^{-9} e^{-0.33x}}$$

where *x* is specified in units of mm.

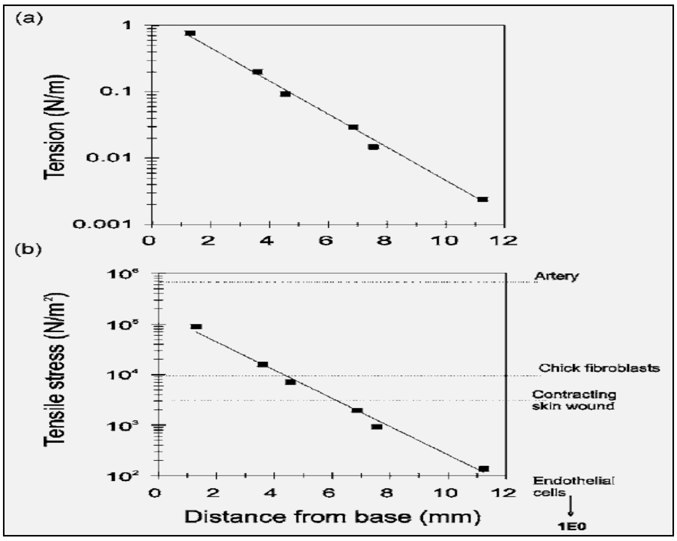


Figure 9: a) BM tension profile and b) BM tensile stress profile [4]

V. SECTION IV: EFFECTS OF DIABETES

Besides the elderly, people who suffer with type II diabetes are particularly badly affected by AGE cross-linking (Andreassen et al., 1981; Schnider and Kohn, 1982). Whilst elder people have abundance of long-lived proteins that have slowly accumulated AGE crosslinks, type II diabetics have abnormally high levels of glucose in their system leaving a surplus which is available to glycate proteins.

The deflection profile equation (A3) in BM mathematical model contains *K* which is function of *Dx* and *Dy* is a diabetic parameter.

As *Dx* and *Dy* are dependent E_f which represent the elasticity moduli of fiber bundles, which is altered by prolonged diabetic condition.

Cochlear active model proposed in [7] involves the motility factor α. Motility factor α varies due to loss of hair cells under prolonged diabetic condition [8]

BM model proposed in [4] involves *K* which varies due to glycation of collagen fibers involved in BM structure

VI. CONCLUSION

After identifying diabetic parameter *K* , we propose that the realistic variations of *K* should be imposed on the mathematical model to investigate the impairments caused by diabetes.

REFERENCES

[1]. Wei Liu, Francesca Atturo, RobiarAldaya, Peter Santi, Sebahattin Cureoglu, Sabrina Obwegeser, Rudolf Glueckert, Kristian P faller, AnneliesSchrott- Fischer, Helge Rask-Andersen, “Macromolecular

organization and fine structure of the human basilar membrane-Relevance for cochlear implantation" Cell Tissue Res(2015) 360:245-262

[2] Alphanso Gautieri, Alberto Radaelli, Markus J. Buehler, Simone Versentini,"Age related dibetes-related nonenzymatic crosslinks in collagen febrils: Candidate amino acids involved in Advanced Glycation End Products " Matrix Biology Journal.

[3] Ram C. Naidu, David C Mountain," Basilar membrane tension calculation for the gerbil cochlea", Hearing Research Centre, Boston University, Massachusetts 02215, USA.

[4] Ram C. Naidu, David C Mountain,"Longitudinal Coupling in the Basilar Membrane", Hearing Research Centre, Boston University, Boston MA 02215, USA

[5] Laura Schweitzer, Carrie Lutz, Michael Hobbs, Sally Weaver ," Anatomical correlates of the passive properties underlying the developmental shift in the frequency map of the mammalian cochlea", Department of Anatomical Science and Neurobiology, University of Louisville School of Medicine, Louisville, KY4092,USA.

[6] Bo Wen , " Modelling the nonlinear cochlea ", University of Pennsylvania.

[7] Dr V.R. Mankar, Nirmala N. Kamble, " Identifying Diabetic Parameters in Cochlear Mechanics and Models".