RESEARCH INVENTY: International Journal of Engineering and Science
ISBN: 2319-6483, ISSN: 2278-4721, Vol. 1, Issue 8 (November 2012), PP 13-19
www.researchinventy.com

Principal Component Analysis in Routing to Identify the Intrusion by Sequential Hypothesis Testing.

K.Santhi M.C.A., M.Phi [1]
Assistant professor
Sri Ramakrishna CAS For Women
Coimbatore., India

P.Bakeyalakshmi [2]
M.Phil C.S (Research Scholar)
Sri Ramakrishna CAS For Women
Coimbatore., India

ABSTRACT: *A Mobile Ad-hoc Network (MANET) has become very important technology in security problem. Recent intrusion detection have emerged an important technique for information security systems. Many researchers have been done to improve the security for intrusion detection in MANET. In security, the unattended replica nodes attack is extremely critical, because they allow the attacker to leverage, or else the attacker may be able to capture the nodes, and then use them to injected fake data, data loss in network communication. In this proposed work a fast and effective mobile replica node detected using the Sequential Probability Ratio Test (SPRT) scheme in Game Theory. This work is to track the problem of replica node attacks in Mobile Ad-hoc Network from intruders and provide the best security. This paper aims to detect the intruders and increase the both speed and predictive accuracy. In this paper, the simulation experiments shown for detecting the claims and increase the trust value by decrease the number of claims.*

KEYWORDS: Manet, Sequential Probability Ratio Test, Game Theory, Intrusion Detection.

I. INTRODUCTION.

A Mobile Ad-hoc Network (MANET) is a collection of mobile nodes that using wireless network to communicate with each other without predefine or static infrastructure, is an unstructured wireless network that can be established temporarily, e.g. application for MANET are include deployment in battle field, small network like universities, Disaster recovery Operations, Civilian environments, Mine site operation [9]. In such a network, the nodes are mobile and can communicate dynamically in an arbitrary manner. The network is characterized by the absence of central administration devices such as source node or destination node. The node plays an important role in the route discovery and route maintenance of the routes from the source to the destination or from a node to another one. It's a great challenge to such a network to provide a security to the nodes from Intruders.

Intrusion detection models were introduced by Denning in 1987 and rather are a new technology. Intrusion detection can be defined as a process of monitoring activities. An intrusion detection system (IDS) collects activity information and then analyzes it to determine whether there are any activities that violate the security rules. There are many mechanism used to find the intruders. Intrusion detection systems can be categorized into two models: Signature-Based Intrusion Detection and Anomaly-Based Intrusion Detection. Signature-Based Intrusion detection uses signatures of the attacks to detect the intrusion. This type of detection monitors the network for finding a match between the network traffic and a known attack pattern. On the other hand, Anomaly-Based Intrusion Detection creates a profile based on the normal behavior of the network [11]. The advantage of the anomaly-based detection is its ability to detect new attacks without any prior knowledge about it.

This paper proposed a anomaly- based neighbor monitoring intrusion detection based on the traffic profile of the node. The proposed approach uses the anomaly-based intrusion detection method. In this scenario, the game theory is used to find the intrusion by sequential probability ratio test which check's for normal and ubnormal node using hypothesis testing. The reported work intends to reduce the number of claims, its lead to increase the speed and accuracy in the network.

This paper is organized in the following way: Section II presents related work for intrusion detection system (IDS). Section III describes the techniques used in existing work. Section IV describes about proposed work and its procedure. Section V describes the simulation and comparison results. Section VI provides conclusion. Section VII describes future work. Section VIII describes reference paper.

II. RELATED WORKS.

The first solution for IDS in MANET was proposed by Anukool Lakhina et al [2]. In their paper a general method to diagnose anomalies. Their method is based on a separation of the high-dimensional space occupied by a set of network traffic measurements into disjoint subspaces corresponding to normal and anomalous network conditions. They show that their separation can be performed effectively by Principal Component Analysis. Using only simple traffic measurements from links, they study volume anomalies and show that the method can: (a) accurately detect when a volume anomaly is occurring; (b) correctly identify the underlying origin-destination (OD) flow which is the source of the anomaly; and (c) accurately estimate the amount of traffic involved in the anomalous OD flow.

In a work reported by Chatzigiannakis et al.,[3] an anomaly detection approach that fuses data gathered from different nodes in a distributed sensor network is proposed and evaluated. The emphasis of this work is placed on the data integrity and accuracy problem caused by compromised or malfunctioning nodes. Their proposed approach utilizes and applies Principal Component Analysis simultaneously on multiple metrics received from various sensors. One of the key features of the proposed approach is that it provides an integrated methodology of taking into consideration and combining effectively correlated sensor data, in a distributed fashion, in order to reveal anomalies that span through a number of neighboring sensors. The efficiency and effectiveness of the proposed approach is demonstrated for a real use case that utilizes meteorological data collected from a distributed set of sensor nodes.

In a work reported by Chong Elkloo et al., [4] security is a critical challenge for creating robust and reliable sensor networks. For example, routing attacks have the ability to disconnect a sensor network from its central base station. In their paper, they present a method for intrusion detection in wireless sensor networks. Their intrusion detection scheme uses a clustering algorithm to build a model of normal traffic behavior, and then uses their model of normal traffic to detect abnormal traffic patterns.

In a work reported by Farid et al., [6] Their paper presents, theoretical overview of intrusion detection and a new approach for intrusion detection based on adaptive Bayesian algorithm. This algorithm correctly classify different types of attack of KDD99 benchmark intrusion detection dataset with high detection accuracy in short response time. The experimental result also shows that, this algorithm maximize the detection rate (DR) and minimized the false positive rate (FPR) for intrusion detection.

In a work reported by Anjum. et al., [1] the intrusion detection community has been concentrating mainly on wired networks. Techniques geared towards wireline networks would not suffice for an environment consisting of Multihop wireless links because of the various differences such as lack of fixed infrastructure, mobility, the ease of listening to wireless transmissions, lack of clear separation between normal and abnormal behavior in ad hoc networks, and consider the signature detection technique and investigate the ability of various routing protocols to facilitate intrusion detection when the attack signatures are completely known. Show that reactive ad-hoc routing protocols suffer from a serious problem due to which it might be difficult to detect intrusions even in the absence of mobility. Mobility makes the problem of detecting intruders harder. Also investigate a relationship between the probability of detecting an intrusion and the number of nodes that must participate in the process of detecting intrusions.

In a work reported by Frullo et al., [12] presents a new approach that combines specification-based and anomaly-based intrusion detection, mitigating the weaknesses of the two approaches while magnifying their strengths. The approach begins with state-machine specifications of network protocols, and augments these state machines with information about statistics that need to be maintained to detect anomalies. They present a specification language in which all of this information can be captured in a succinct manner. Feature selection was a crucial step that required a great deal of expertise and insight in the case of previous anomaly detection approaches.

In a work reported by Hyeon et al., [7] developed Threshold Delay time sequence embedding in that n length subsequences are extracted from the training set and the are assigned a probability given, P(subsequence) = Frequency (subsequence) / total number of subsequences. During prediction the test sequence windows are looked up in stored model and assigned corresponding probabilities. If the window had never occurred in the training data then it is given a zero probability score. Now the sequence of symbols gets converted to a sequence of probabilities. These probabilities need to be combined to get an anomaly score. Different combining methodologies were checked but log average was found to be superior.

In a work reported by Juha et al., [8] discuss about various generalizations of neural PCA (Principal Component Analysis)-type learning algorithms containing nonlinearities using optimization-based approach. Standard PCA arises as an optimal solution to several different information representation problems, and justify that this is essentially due to the fact that the solution is based on the second-order statistics only. If the respective optimization problems are generalized for nonquadratic criteria so that higher-order statistics are taken into account, their solutions will in general be different. Their solutions define in a natural way, several meaningful extensions of PCA and give a solid foundation for them. In their framework the study more closely generalizations of the problems of variance maximization and mean-square error minimization. For these problems, the derive gradient-type neural learning algorithms both for symmetric and hierarchic PCA-type networks. As an important special case, the well-known Sanger's generalized Hebbian algorithm (GHA) is shown to emerge from natural optimization problems.

In a work reported by Peyman et al., [11] the goal of PCA is to reduce the dimensionality of a data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set .The extracted non correlated components are estimated from the eigenvectors of the covariance matrix of the original variables. The objective of the anomaly detection algorithm is to provide an efficient and effective methodology of fusing and combining data of heterogeneous monitors that spread throughout the network, in order to provide a generalized framework, capable of detecting a wide range of classes of anomalies, such as the ones created randomly by faulty nodes or others that result from coordinated compromised nodes. In their work, this is achieved by applying a PCA-based approach simultaneously on one or more monitored metrics.

In a work reported by Sapon et al., [13] introduces a new approach that addresses data contamination problems from attacks in unattended wireless sensor networks. They propose a sliding-window based spatio-temporal correlation analysis called Abnormal Relationships Test (ART)" to effectively detect, respond and immune to inserted spoofed data from both various-ID impersonators and compromised nodes. Also a systematic approach is given to identify the appropriate sliding window size and correlation coefficient threshold. Their study shows that correlation property of observed phenomenon is not always transitive, different phenomenon from same set of nodes at the same or different period of time can have different correlation coefficients.

In a work reported by Sutharshan et al., [14] identifying misbehaviors is an important challenge for monitoring, fault diagnosis and intrusion detection in wireless sensor networks. A key problem is how to minimize the communication overhead and energy consumption in the network when identifying misbehaviors. Their approach to this problem is based on a distributed, cluster-based anomaly detection algorithm. While minimize the communication overhead by clustering the sensor measurements and merging clusters before sending a description of the clusters to the other nodes.

In a work reported by Tu-Liang et al., [15] investigated some intrusion detection techniques using machine learning and proposed a profile based neighbor monitoring intrusion detection method. Further analysis shows that the features collected by each node are too many for wireless devices with limited capacity, and can apply Markov Blanket algorithm to the feature selection of the intrusion detection method. Experimental studies have shown that Markov Blanket algorithm can decrease the number of features dramatically with very similar detection rate.

In a work reported by Xia Wang et al., [16] they proposed Markovian based intrusion detection techniques that underlying hypothesis of this class of technique is that the next event can be predicted by looking at a short history of past events. They use the short memory property of sequences which has been shown to exist across domains. The history window can be fixed or variable. Length of the window is n.

In a work reported by Yonggung et al., [18] They the vulnerabilities of wireless networks and argue that must include intrusion detection in the security architecture for mobile computing environment, and developed an architecture and evaluated a key mechanism in this architecture, anomaly detection for mobile ad-hoc network, through simulation experiments. To build anomaly detection models for mobile wireless networks. Detection based on activities in different network layers may differ in the format and the amount of available audit data as well as the modeling algorithms.

III. Existing Method.

The existing work based on profile based monitoring intrusion detection based on the traffic profile of the node. Feature selection is used to improve its performance to reduce the dimensionality of the network features. This reduction may lead to decrease intrusion detection speed, since the IDS would have fewer features to analyze. Network features such as movement and number of the nodes are also considered in this work [11].

PCA is a classic technique in statistical data analysis, feature extraction and data compression. Goal is to introduce a smaller set of variables in a set of multivariate measurements with less redundancy. The header of the packet has a list of nodes addresses to pass it in source routing. At first, source route discovers the path to the source node. All nodes can listen to the packets in the DSR routing network. Nodes can update the routing information in cache table based-on available paths in packet header for further usages.

IV. Proposed Work.

There are many techniques that can be used for monitoring the nodes and analyzing the results. In the proposed method of approach, a anomaly-based intrusion detection technique is used with sequential hypothesis testing in Game theory.

A.SEQUENTIAL HYPOTHESIS TESTING.

In the proposed technique, Sequential Probability Ratio Test (SPRT) is to find the intrusion by Fast and effective replica node, improves the communication, computation, and storage overheads [9]. And the game-theoretic model is to find the attacker and defender of their respective optimal strategies.

In the principal component analysis each profile includes the features of source address, destination address, route request from node, and send packet. In RREQ and RREP the Sequential Probability Ratio Test is used to find the intrusion. In RREQ the hypothesis testing is used to find the normal and ubnormal node. The RREP is used to find the bit checking [11].

SPRT on every mobile node using a null hypothesis, that the mobile node has not been replicated and an alternate hypothesis that it has been replicated. A two player game model is to interaction between the attacker and the defender, when they derive the optimal attack and defense strategies. And show that the attacker's gain in respective optimal strategies. A game-theoretic model is to claim response and quarantine defense. For an e.g. a network is taken with 60 nodes in that first a path is selected between source node and destination node. In that game theory will find the claims nodes in that path, by using SPRT a new path is selected without any claims. Then the Improved Principal Component Analysis (IPCA) will send a key to the path which without claims. A node in the path will send REQ to another node, that node will REPLY to the REQ node if the key matches between the nodes they transfer the DATA between them. Finally the nodes in the path will send KEY to the one-hob neighbor node, if the key matches the node is Active. When key does not matches means the node is Idea.

V. Simulation Result.

The following metrics to evaluate the performance of our scheme:

(1) Number of Claims: The number of claims required for the base station to decide whether a node has been replicated or not.
(2) False Positive: The error probability that a benign node is misidentified as a replica node.
(3) False Negative: The error probability that a replica node is misidentified as a benign node.

For each execution, we obtain each metric as the average of the results of the SPRTs that are repeated.

A. SIMULATION PROCEDURE.

The proposed scheme is simulated using following performance parameters.

- ✓ **Control Overhead:** It's defined as the number of control packets (RQ/RP packets) needed to establish routes to all destinations from a source.
- ✓ **End- to-End Delay:** It's define as the inter arrival of packet time with total data packet delivery time.

- ✓ **Normalized Average Trust Value:** It compare claim nodes with normal nodes of a network.

B. COMPARISON RESULT.

The simulation is carried out on network simulator NS2.34. The analysis of performance parameters are given in the following scenario.

$$\text{Control Overhead} = \frac{\text{No of Control RREQ} + \text{No of Control RREP}}{\text{Total control packet}}$$

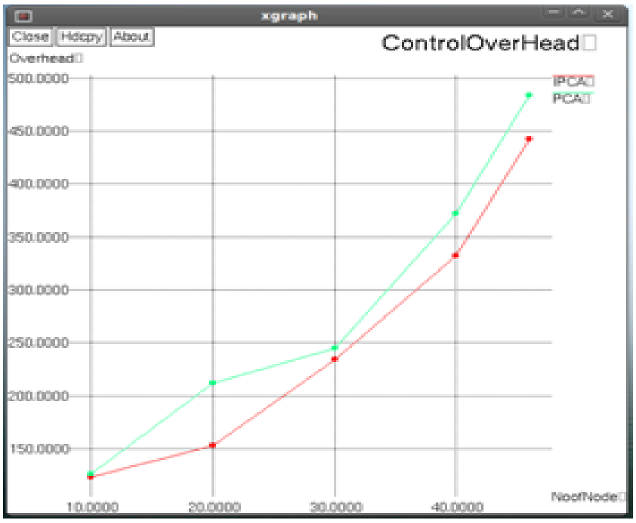

Figure 1. Control Overhead.

The Control Overhead is also reduced in IPCA.

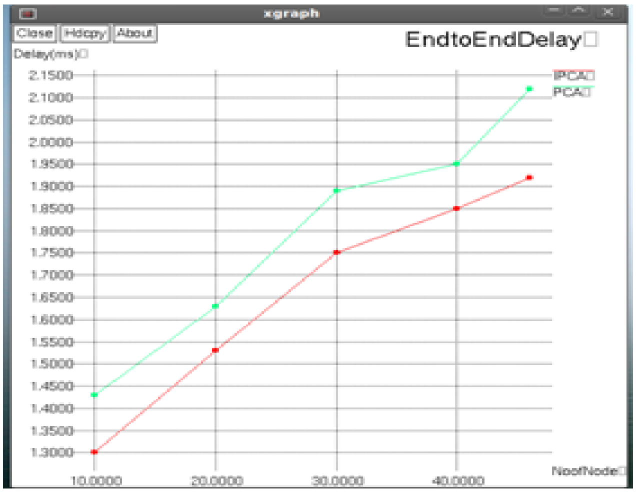

Figure2. End-to-End Delay

$$\text{End-to-End Delay} = \frac{\text{Inter arrival of first Packet time} - \text{Inter arrival of second packet time}}{\text{Total data packets deliver time.}}$$

The End-to-End Delay is also reduced in IPCA.

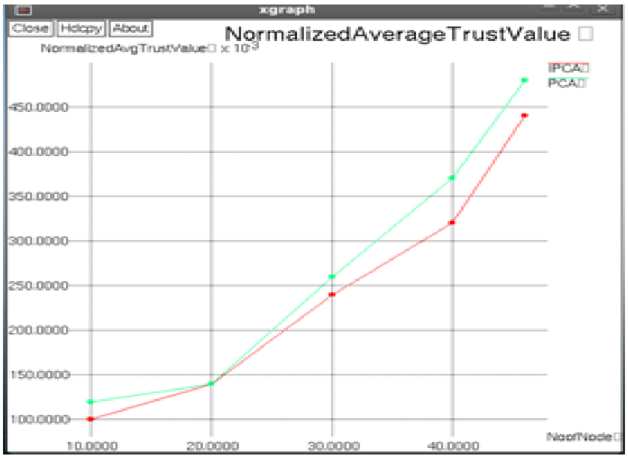

Figure3. Normalized Average trust value.

Normalized Average trust value is compare claim nodes with normal nodes of a network. In this IPCA give better performance when compare with normal PCA.

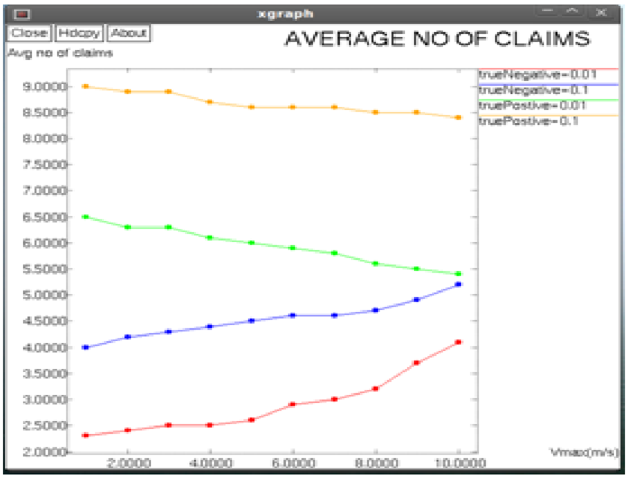

Figure 4. Average number of claims.

The average number of claims is compare with true positive and true negative values; the values are 0.01 and 0.1

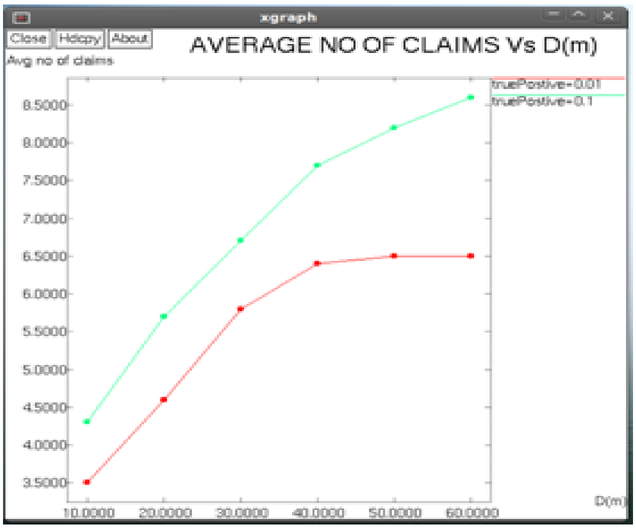

Fig5. Average number of claims (vs.) Distance (m)

The average number of claims (vs.) Distance (m) compare only with true positive values, the values are 0.01 and 0.1

VI. Conclusion.

In this paper, a anomaly based intrusion detection approach in MANET was presented. We have proposed a replica detection scheme for mobile networks based on the SPRT. And we analytically demonstrated the limitations of attacker strategies for providing the secured path without any claims. The results of these simulations show that our technique quickly detects mobile replicas with a small number of location claims.

VII. Future Work.

In this paper, DSR protocol is selected as routing protocol. Ad-hoc On Demand Vector (AODV) can also be selected for the routing. And also do not investigate cooperation amongst the various nodes in the intrusion detection subsystem. Trying to identify attacks based on incomplete information is also not pursued.

VIII. Reference.

[1]. **Anjum.F, D. Subhadrabandhu, S. Sarkar.** “Signature-based intrusion detection for wireless Ad-hoc networks," Proceedings of Vehicular Technology Conference, vol. 3, pp. 2152-2156, USA, Oct. 2003.

[2]. **Anukool Lakhina, Mark Crovella, Christophe Diot,** “Diagnosing Network wide Traffic Anomalies,” International Journal of Networks Security, Vol.12, No.1, PP.42-49, Feb 2004.

[3]. **Chatzigiannakis.V, Papavassilion.S** “ Diagnosing Anomalies and Identifying Faulty Nodes in Sensor Networks”, Sensor Journal, IEEE, Vol.7, Issue.5, PP.637 – 645, May 2007.

[4]. **Chong Eik Loo, Mun Yong Ng, Christopher Leckie, Marimuthu Palaniswami,** “Intrusion Detection for Routing Attacks in Sensor Networks,” International Journal of Distributed Sensor Networks, V.2, PP. 313–332, 2006.

[5]. **Denning. D. E.,** “An Intrusion Detection Model," IEEE Transactions in Software Engineering, vol. 13, no. 2, pp. 222-232, USA, 1987.

[6]. **Farid.D.M, M.Z Rahman,** “Learning intrusion detection based on adaptive Bayesian algorithm," 11th International Conference on Computer and Information Technology (ICCIT2008), pp. 652-656, 2008.

[7]. **Hyeon-Kyn Lee, Kim,J.H,** “An HMM-Based Threshold Model Approach For Gesture Recognition,” Pattern Analysis and Machine Intelligence, IEEE Transaction, Vol.21, Issue.10, PP.961-973, Oct 1999.

[8]. **uhaKarhunen, yrki Joutsensalo, “**Generalizations of Principal Component analysis, Optimization problems, and Neural Networks,” Helsinki University Of Technology, Finland, Vol.8, Issue.4, PP. 549-562, April 2000.

[9]. Jun-**Won Ho, Matthew Wright, Sajal K. Das,** “Fast Detection of Mobile Replica Node Attacks in Wireless Sensor Networks Using Sequential Hypothesis Testing.” IEEE Transaction on Mobile Computing, Vol. 10, No.6, PP.767-782, June 2011.

[10]. **MohapatraP,S.V.Krishnamurthy**, Ad-hoc Network Technologies and Protocols, Springer-Verlag, Business Media Inc., USA, 2005.

[11]. **Peyman Kabir, Mehran Aghaei,**“Feature Analysis for Intrusion Detection in Mobile Ad-hoc Networks,”International Journal of Network Security, Vol.12, No.1, PP.42-49, Jan. 2011.

[12] **Sekar. R, A. Gupta, J. Frullo, T. Shanbhag, A. Ti-wary, H. Yang, S. Zhou,** “Specification-based anomaly detection: A new approach for detecting network intrusions," Proceedings of the 9th ACM Conference on Computer and Communication Security, pp. 265-274, USA, 2002.

[13]. **SaponTanachaiwiwat,AhmedHelmy,** “Correlation Analysis for Alleviating Effects of Inserted Data in Wireless Sensor Networks,” Proceedings of the Second Annual International Conference on Mobile and Ubiquitous Systems, Networking and Services,2005.

[14] **Sutharshan Rajasegara1, Christopher Leckie, Marimuthu Palaniswami, James C. Bezdek** “Distributed Anomaly Detection In Wireless Sensor Networks,” Computer Science Department University of West Florida USA, 2006.

[15]. **Tu-Liang Lin.T, J. Wong,** “Feature Selection in Intrusion Detection System over Mobile Ad-hoc Network,” Technical Report, Computer Science, Iowa State University, USA, 2005.

[16]. **Xia Wang, J. Wong,** “Feature Selection in Intrusion Detection System over Mobile Ad-hoc Network,” Technical Report, Computer Science, Iowa State University, USA, 2005.

[17]. **Ye.N, S. M. Emran, X. Li, Q. Chen,** “Statistical process for computer intrusion detection," Proceeding in DARPA Information Survivability Conference and Explosion (DISCEX'01), vol. 1, pp. 3-14, USA, 2001.

[18]. **Yongguang Zhang, Wenke Lee, Yi-An Huang,** “Intrusion Detection Techniques For Mobile Wireless Networks,” Wireless Networks,Vol. 9, PP.545–556, 2003.

BAKEYALAKSHMI.P, received the B.Sc (Computer Science) degree from the Department of Computer Science at Bharathiar University, Coimbatore, India, the M.Sc (Information Technology) degree from the Department of Computer Science at the Anna University, Erode, India, and she perusing her Master of philosophy in Computer Science at Bharathiar University, Coimbatore, India. Her area of interest is Networking.

K.Santhi, received the B.Sc (Phy) degree from the Department of Physics, M.C.A and M.Phil degree from the Department of computer science, Bharathiar University, Coimbatore, in 1998, 2001 and 2004 respectively. She is currently pursuing the Ph.D degree in Anna University, Coimbatore. She is working closely with Prof Dr.M.Punithavalli in the Sri Ramakrishna College of arts and science for women from 2001 to till date. Her area of interest is Networking.