Performance Analysis of ACO-based IP Traceback

Jose Anand
Associate Professor/ECE
K C G College of Technology
Chennai, Tamil Nadu, India – 600097

K. Sivachandar
Assistant Professor/ECE
K C G College of Technology
Chennai, Tamil Nadu, India – 600097

ABSTRACT
The Internet has experienced a tremendous expansion in its size and complexity since its commercialization. Internet hosts are threatened by large-scale Distributed Denial-of-Service (DDoS) attacks in the network. DDoS attacks typically rely on compromising a large number of hosts to generate traffic to a single destination node. Thus the severity of DDoS attacks will likely increase to the possible extend, as greater numbers of poorly secured hosts are connected to high-bandwidth Internet connections. To detect and coordinate DDoS attacks in the network usually an Intrusion Detection System (IDS) is used but, this method consumes most of the resources and thereby degrades the network performance. Moreover, the memory-less feature of the routing mechanism makes the operation hard to traceback the source of the DDoS attacks. This paper analyzed the performance of an Ant Colony Optimization (ACO)-based IP traceback method to identify the origin of the attack in the network. The ACO-based IP traceback approach uses flow level information to identify the origin of a DDoS attack. The ACO-based IP traceback method is implemented using NS-2 simulation on various network scenarios consisting of 8 nodes, 10 nodes, and 14 nodes. The results of the experimental and simulation studies demonstrate the effectiveness and efficiency of the proposed system.

Keywords
Ant Colony Optimization, attacks, DDoS, Internet, IP traceback, pheromone.

1. INTRODUCTION
Intrusion detection is a process of monitoring the assorted computer networks and systems for violations of security and this process can be automatically carried out with the help of an IDS. An IDS is a critical component for secure information management. IDS plays a vital role in detecting and disrupting various attacks in the network before cooperating with the software. A computer system should provide confidentiality, integrity and assurance against DDoS attacks. Due to increased connectivity and the vast spectrum of financial possibilities that are opening up, more and more systems are subjected to attack by intruders. Because of the increasing dependence in which companies and government agencies have their own computer networks, the importance of protecting these systems from various attacks is critical. A single intrusion of a computer network can result in the loss of unauthorized utilization or modification of large amount of data and cause users to question the reliability of all of the information on the network.

IDSs can be categorized into three different types namely, network-based intrusion detection system, router-based intrusion detection system, and host-based intrusion detection system [15]. Network-based intrusion detection system, operate at the gateway of a network and examines all the incoming packets. Router-based intrusion detection system is installed on the routers to prevent intruders from entering into the network. Finally, the host-based intrusion detection system receives the necessary audit data required from the host's operating system and analyzes the generated events to keep the local system secure. IDS detect DDoS attacks either by using a priori knowledge of the types of known attacks or by recognizing deviations from normal system behaviors. DDoS attack aims at denying or degrading legitimate users access to a service or network resource, or at bringing down the servers offering such services.

IP networks are vulnerable to source address into packet headers [17]. DDoS block legitimate access by either exhausting victim server's resources or saturating stub networks access links to the Internet. By masquerading as a different host, an attacker can hide its actual identity and location, rendering source-based packet filtering less effective. Many popular attacks use IP Spoofing and requires the ability to forge source addresses. DDoS attacking tools spoof IP addresses by randomizing the 32-bit source-address field in the IP header [7], which conceals attacking sources and dilutes localities in attaching traffic. Some of the features of DDoS attacks in the Internet based network are observed as follows:

- Public networks are used as springboards
- Attackers are well camouflaged
- IP Packets sent by the attackers are flooded

Many companies use firewall and IDS to secure the network access from various attacks. If an attack like DDoS is initiated by an attacker with a massive amount of traffic, that can be detected with the help of IDS and the malicious traffic is blocked by properly coordinating the firewall security. The DDoS attacks not only posed a huge threat to the Internet, but also the devastation caused by them becomes more and more severe. DDoS attackers use spoofed source IP addresses in most of the cases, which makes it difficult to defend against the attack in the network. The topic on IP traceback gains sustained attention within the research scholars and several approaches have been proposed for path identification and trace the source IP address. Many researches are going on to solve the IP traceback problem. Most of the studies aimed to identify the exact origin of the attack in the network.

The rest of this paper is organized as follows. We discuss related work in Section 2. Section 3 describes the need of IP traceback in the network. Section 4 illustrates the operation of Ant Colony Optimization algorithm in a detailed way. Section 5 depicts the proposed ACO based design and section 6 gives the details of the experimental setup and evaluations. Section 7 briefs with the conclusion and future direction in the field of IP traceback on the Internet based network.

2. RELATED WORK

In this section, we review the prior work on IP traceback. Chao et al [3] studied the effectiveness of log-based IP traceback in tracing a single packet under the environment where not all the Autonomous Systems (AS) supports log-based IP traceback. Also, a scheme to conduct the single packet traceback process in AS-level partial deployment scenario was proposed and evaluated the performance of single packet IP traceback in AS-level partial deployment scenario based on the proposed scheme through simulation. Bellovin [1] introduced ICMP traceback, in which routers select packets with low probability, and then send ICMP packets including the contents of sampled packets and local path information to the same destination as the selected packets. Burch et al [2] proposed controlled flooding, and the victims reconstruct attack paths by selectively flooding network links and monitoring the change of incoming traffic. Also, mentioned an idea of IP traceback based on marking packets, either probabilistically or deterministically, with the identification information (e.g., IP addresses) of routers they pass through. Ruiliang et al [14] proposed a novel attack mitigation scheme named, Attack Diagnosis (AD) that combines the concepts of Pushback and packet marking. The AD architecture is in-line with the ideal DDoS attack countermeasure paradigm, in which attack detection is performed near the victim host and attack mitigation is executed close to the attack sources. Also, an extension of AD called Parallel Attack Diagnosis (PAD) is proposed, that is capable of throttling traffic coming from a large number of attack sources simultaneously.

Wei et al [22] proposed a Recursion Nearness approach to characterize the IP address features, which is also helpful to the other security applications, such as firewall log analysis, traffics filter rule management, intrusion investigation and IP traceback. Also, the IP address is characterized with some fields, such as TTL, MAC, TCP/IP stack implementation finger printing and IP geographic location mapping. Experimental analysis shows that the method can characterize IP features more exact and efficiently. Stefan et al [18] described a technique for tracing anonymous packet flooding attacks in the Internet back toward their source. Also a general purpose traceback mechanism based on probabilistic packet marking in the network is proposed. The approach allows a victim to identify the network path traversed by attack traffic without requiring interactive operational support from Internet Service Providers (ISP). Moreover, this traceback can be performed by "post mortem" after an attack has completed. Dalia et al [5] proposed a scheme to detect the flooding agents by considering all the possible kinds of IP spoofing. The proposed scheme is based on the TCP SYN-SYN/ACK protocol pair with the consideration of packet header information. The Counting Bloom Filter is used to classify all the incoming SYN/ACK packets to the sub-network into two streams, to make the scheme generally applicable and the Cumulative Sum algorithm is applied to avoid the dependence of detection on sites and access patterns.

Wei et al [21] proposed a simple and efficient method to detect and defend against TCP SYN flooding attacks under different IP spoofing types, including subnet spoofing. The method makes use of a storage-efficient data structure and a change-point detection method to distinguish complete three-way TCP handshakes from incomplete ones. Experimental analysis shows the proposed method is both efficient and effective in defending against TCP-based flooding attacks under different IP spoofing types. Hikmat [10] proposed an efficient method to defend against IP spoofing. To improve the scalability, the Implicit Token Scheme (ITS) is added with a component using Bloom filters, which is a simple method and saves a substantial amount of router memory, and also does not impose large strain on routers. The efficiency of the method is demonstrated through simulations by using real-world Internet data. Vamsi et al [19] proposed a novel protocol, Fast Autonomous System Traceback (FAST) to traceback to the attack originating AS. The multifold advantages of FAST include reconstruction requires just around 5 to 10 packets and reconstruction takes just a few seconds. Also the performance through extensive simulations over the datasets obtained from trace route is validated. Dean et al [6] proposed the utilization of probabilistic packet marking for IP trace back. The 16-bit identification field in the IP header of each packet will be marked with partial path information. So the entire information regarding the path can be recovered after receiving many packets. But this approach is vulnerable to packet markings forgery.

Yan et al [24] presented a signature and verification based IP Spoofing prevention method named Automatic Peer-to-Peer based Anti-Spoofing (APPA) method. APPA has two levels named Intra-AS level and Inter-AS gateway level. In an Intra-AS level, the end host tags a one-time key into each outgoing packet and the gateway verifies the key. In an Inter-AS level, the gateway at the AS border tags a periodically changed key into the leaving packet and the gateway at border of the destination AS verifies and removes the key. Chu et al [4] proposed a method to monitor the traffic pattern in order to alleviate DDoS attacks. In this method, a bandwidth allocation policy is adopted to assign normal users to a high priority queue and suspected attackers to a low priority queue. The simulations conducted in the network simulator shows that its effectiveness in blocking attack traffic while maintaining constant flows for legitimate traffic. Voravud et al [20] proposed a new methodology for verifying the vulnerability of firewall configurations to IP spoofing attack and for synthesizing IP spoofing-free configurations. The methodology is based on graph theory which provides a simple and intuitive approach to the vulnerability analysis of the attack.

3. IP TRACEBACK

The objective of IP traceback is to identify the path of an IP packet to its origin node in the network. The need for IP traceback is to deal with DDoS attacks, where the source IP address in the Internet based network is spoofed by attackers [16]. The IP traceback techniques in the Internet based network can be classified into two types, named; proactive IP traceback technique and reactive IP traceback technique. The proactive IP traceback technique measures the various records and exchange the tracing information in the same way as packets are routed through the network. The victim uses the traceback data for reconstructing the attacked path and to get the attacker identification. The reactive IP traceback technique initiates the traceback process in response to an attack. In this, the attack has to be active until the trace is identified. In order to identify the attackers, two methods of IP traceback are used named; the Probabilistic Packet Marking (PPM) [8] and the Deterministic Packet Marking (DPM) [23]. Both of these approaches need routers to inject marks into individual packets. Furthermore, the PPM approach can only operate in a local range of the Internet, where the defender has the authority to manage.

4. ANT COLONY OPTIMIZATION

Ants are capable of finding the shortest path from a food source to their nest, and are adaptive to changes in the environment for finding a new shortest path route, once the old path is no longer feasible [12]. For finding the shortest path, on the way ants deposit pheromone, a kind of cumarin which ants are able to smell and by which they marks the route taken. The concentration of pheromone on a certain path is an indication and with time the concentration of pheromone decreases due to diffusion effects [11]. The pheromone is used to probabilistically sample the search space and this behavior explains how the ants can be used to find the shortest path. Figure 1 shows the Ant Colony Optimization (ACO) meta-heuristic with two routes from the source (nest) to destination (food). At the intersection (node B) ants select different branch to reach the destination. Since the lower route is shorter than the upper one, the ants which take the lower path will reach the destination (food) first.

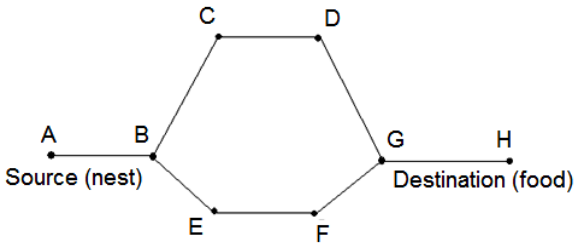

Figure 1 ANT Colony based route selection

The pheromone concentration on the shorter path will be higher after a short time than on the longer path, because the ants using the shorter path will increase the pheromone concentration faster. That is more and more ants will deposit pheromone over the path. Thus the shortest path will be identified. The Ant Colony Optimization algorithm is given below:

Algorithm: Ant Colony Optimization

> ***input****: an instance x of a Combinatorial Optimization problem*
> ***while*** *termination conditions not met* ***do***
> ***Schedule Activities***
> *Ant based Solution Construction( )*
> *Pheromone Update( )*
> *Daemon Actions( )*
> ***end Schedule Activities***
> S_{best} ← *best solution in the population of solutions*
> ***end while***
> ***output****:* S_{best}*, candidate to optimal solution for x*

The basic ingredient of ACO algorithm is a constructive heuristic for probabilistically constructing solutions taken from a finite set of solution compounds v = { v_1, v_2, ……., v_n }. The process of constructing solutions can be regarded as the path on the construction graph G = (V, E), whose vertexes are the solution components V and the set E are the links or connections. The allowed path on G are implicitly defined by the solution construction mechanism and the choice of solution component $N(s^p)$ at each construction step done probabilistically with respect to the pheromone value T_i, and the set of all pheromone tail parameters is denoted by T. The probabilities for choosing the next solution component called transition probabilities is given by

$$P(V_i|S^p = \frac{T_i{\propto}\,.\,\eta(v_i)^{\beta})}{\sum_{vj\,\epsilon\,N_{(sp)}}\,T_i{\propto}\,.\,\eta(v_i)^{\beta})}\,,\forall\, v_i \in N_{(sp)} \quad (1)$$

where, η is the weighting function that depends on the current partial solution and assigns a heuristic value $\eta(v_j)$ to each feasible solution component v_j ε $N(s^p)$. α and β are positive parameters that determines relation between pheromone information and heuristic information. Normally the interval is taken as 0.05. Pheromone update has an evaporation phase that uniformly decreases its values which is used to avoid a rapid convergence of the algorithm towards a sub-optimal region. Daemon actions can be used to implement centralized actions which cannot be performed by a single ant. This decides to deposit extra pheromone on the solution components that belong to the best solution.

5. ACO-BASED DESIGN

If a node identifies a DDoS attack, it has to analyze the packets of the DDoS attack and find out the source IP address of the attack. For the analysis of the IP traceback mechanism ACO-based design is used [9]. For this analysis the attacked node is considered as the starting point of the IP traceback. Initially each node uses the total amount of packets sent in a particular duration. The flow information is used to determine the path selection of the ants. The probability of ant to choose a path is given in equation 2.

$$P_i(t) = \frac{[\tau i(t)]^{\alpha}\,.\,[\rho i]^{\beta}}{\sum_i\ [\tau i(t)]^{\alpha}\,.\,[\rho i]^{\beta}} \quad ----- \quad (2)$$

where, $P_i(t)$ is the probability of an ant to choose a path at the node i. $\tau i(t)$ is the intensity of pheromone trail of an ant at the node i at time t. ρi is the total number of packets sent in a particular duration at the node i. α and β are constants. The probability of the successive move is determined by the flow information of the neighbor nodes. More ants will choose the path with large amount of flow, as the DDoS attack generates large amount of flow. The movement of the ant will be repeated until the ant reaches a boundary node in the network. The intensity of the pheromone is revised after the completion of all ants. The traceback procedure is executed for a user-defined number of cycles or all ants make the same path. Finally, the traceback path of the ants is the path of the DDoS attack path.

6. PERFORMANCE EVALUATIONS

Performance analysis of the algorithm has been carried on an Intel Core 2 Duo CPU system with 2.10 GHz on a 32-bit Windows 7 Ultimate Operating System. Simulations are carried out using network simulator, NS-2, assuming that the nodes are facilitated with NetFlow providing the flow-level information. Simulations are done for three different scenarios each with different network topologies. The first scenario consists of 8 nodes, the second scenario consists of 10 nodes and the third scenario consists of 14 nodes [13] as shown in figure 2, figure 3 and figure 4 respectively. The 10 nodes network is a standard Abilene network consisting of 13 bi-directional links and the 14 nodes network is a standard NSFNET consisting of 21 bi-directional links. It is considered that, eleven legitimate normal flows in the simulation network and one DDoS attack flow through the network.

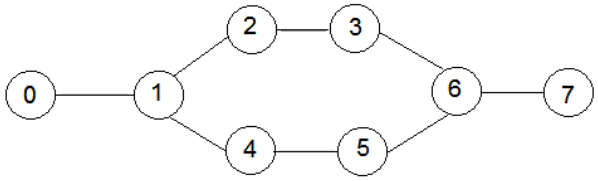

Figure 2 8-nodes Network

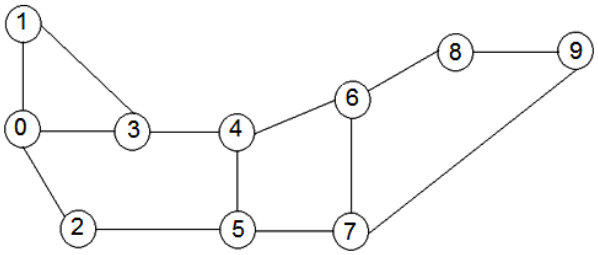

Figure 3 10-nodes (Abilene) Network

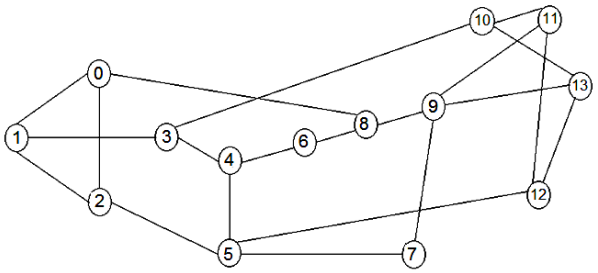

Figure 4 14-nodes NSFNET

For the above three scenarios simulations are done and analyzed the performance of success rate. Success rate defines the positive alert obtained by the traceback when an attack is identified. Figure 5 shows the traceback success rate in the simulation for the three different scenarios. Analysis is made on the deployment rate of attack on the x-axis and success rate on the y-axis. Observation shows that when the deployment rate of the attack is more in the network, higher the chance to identify the attacker with the ACO-based IP traceback algorithm.

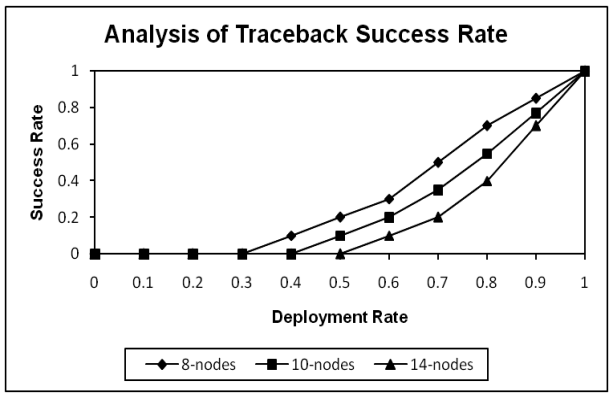

Figure 5 Traceback success rate

7. CONCLUSION AND FUTURE ENHANCEMENTS

To traceback the origin of an Internet attack, strategic importance is given to cyber space security. From the survey it is analyzed that each method has certain features that make it more suitable to implement in one situation than another. DDoS attack is one of the important threats in the Internet and causes many revenue losses to the companies. This paper depicts an ACO-based IP traceback system to identify the attack path. The concept is implemented in a different set of small network topology. This approach uses the flow level information to identify the attacker. Other optimization techniques can also be used to identify the traceback. Further, the implementation and deployment of this concept on a large network can be done to evaluate the scalability of the traceback system.

8. REFERENCES

[1] S. Bellovin, "ICMP traceback messages", Internet draft: Draft-bellovin-itracc-00.txt, March 2000.

[2] H. Burch, and B. Cheswick, "Tracing anonymous packets to their approximate source", Proceedings of the 14th USENIX Systems Administration Conference, New Orleans, USA, December 2000.

[3] Chao Gong, Trinh Le, Turgay Korkmaz, and Kamil Sarac, "Single Packet IP Traceback in AS-level Partial Deployment Scenario", Proceedings of the IEEE Globecom, St. Louis, MO, USA, pp. 1817-1821, Nov. 28th – Dec. 2nd, 2005.

[4] Chu Hsing Lin, Jung Chun Liu, Hsun Chi Huang and Tsung Che Yang, "Using Adaptive Bandwidth Allocation Approach to Defend DDoS Attacks", Proceedings of the International Conference on Multimedia and Ubiquitous Engineering, IEEE Computer Society, pp. 176-181, 24-26 April 2008.

[5] Dalia Nashat, Xiaohong Jiang and Susumu Horiguchi, "Detecting SYN Flooding Agents under any type of IP Spoofing", Proceedings of the IEEE International Conference on e-Business Engineering, IEEE Computer Society, pp. 499-505, 22-24 October 2008.

[6] D. Dean, M. Franklin, and A. Stubblefield, "An Algebraic Approach to IP Traceback," ACM Transactions on Information and System Security (TISSEC), Vol. 5, No. 2, pp. 119-137, May 2002.

[7] S. Dietrich, N. Long, and D. Dittrich, "Analyzing distributed denial of service tools: The shaft ease," Proceedings of USENIX LISA 2000, New Orleans, LA, USA, pp. 329-339, 3-8 Dec. 2000.

[8] M. T. Goodrich, "Probabilistic Packet Marking for Large-Scale IP Traceback", IEEE/ACM Transactions on Networking, Vol. 16, No. 1, pp. 15-24, Feb. 2008.

[9] Gu Hsin Lai, Chia-Mei Chen, Bing-Chiang Jeng, and Willams Chao, "Ant-based IP traceback", International Journal on Expert Systems and Applications, Elsevier, Vol. 34, Issue 4, pp. 3071-3080, May 2008.

[10] Hikmat Farhat, "A Scalable Method to Protect From IP Spoofing", Proceedings of the IEEE International Conference on The Applications of Digital Information and Web Technologies, IEEE press, ISBN: 978-1-4244-2623-2, pp. 569-572, 2008.

[11] Marco Dorigo and Christian Blum, "Ant colony optimization theory: A survey", Journal of Theoretical Computer Science, Elsevier, Vol. 344, Issues 2-3, pp. 243-278, 17th Nov. 2005.

[12] Marco Dorigo and Luca Maria Gambardella, "Ant Colonies for the traveling Salesman problem", University

Libre de Bruxelles, Publication in Bio Systems, (TR/IRIDIA/1996-3), Belgium, 1997.

[13] G. Ramesh, S. Sundara Vadivelu and Jose Anand, "Design of Optimized WDM Networks Using Heuristic Algorithms" International Journal of Advance in Communication Engineering, Vol. 01, No. 2, pp. 93-98, ISSN 0975-6094, July-Dec. 2009.

[14] Ruiliang Chen and Jung-Min Park, "Attack Diagnosis: Throttling Distributed Denial-of-Service Attacks Close to the Attack Sources", Proceedings of 14th International (IEEE) Conference on Computer Communications and Networks (ICCCN 2005), Issue 4, pp. 275-280, Oct. 2005.

[15] Shengrong Bu, Richard Yu F., Xiaoping P. Liu, Peter Mason, and Helen Tang, "Distributed Combined Authentication and Intrusion Detection with Data Fusion in High-Security Mobile Ad hoc Networks", IEEE Transactions on Vehicular Technology, Vol. 60, No. 3, pp. 1025-1036, March 2011.

[16] Shui Yu, Wanlei Zhou, Robin Doss, and Weijia Jia, "Traceback of DDoS Attacks using Entropy Variations", IEEE Transactions on Parallel and Distributed Systems, Vol. 22, No. 3, pp. 412-425, March 2011.

[17] D. Srinath, J. Janet, and Jose Anand, "A Survey of Routing Instability with IP Spoofing on the Internet", Asian Journal of Information Technology, Vol. 9, No. 3, pp. 154-158, 2010.

[18] Stefan Savage, David Wetherall, Anna Karlin and Tom Anderson, "Network Support for IP Traceback", IEEE/ACM Transactions on Networking, Vol. 9, No. 3, pp. 226-237, June 2001.

[19] Vamsi Paruchuri, Arjan Durresi, and Leonard Barolli, "FAST: Fast Autonomous System Traceback", Proceedings of 21st International Conference on Advanced Networking and Applications (AINA'07), IEEE Computer Society, Niagara Falls, Canada, pp. 498-505, 21-23 May 2007.

[20] Voravud Santiraveewan and Yongyuth Permpoontanalarp, "A Graph-based Methodology for Analyzing IP Spoofing Attack", Proceedings of the 18th International Conference on Advanced Information Networking and Application (AINA'04), IEEE Computer Society, Vol. 2, 29-31 March 2004.

[21] Wei Chen, and Dit-Yan Yeung, "Defending Against TCP SYN Flooding Attacks Under Different Types of IP Spoofing", Proceedings of the International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06), IEEE Computer Society, Morne, Mauritius, page 38, 23-29 April 2006.

[22] Wei Ren and Hai Jin, "A Recursion Nearness based method for Characterizing IP Address", Proceedings of the Sixth International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'05), IEEE Computer Society, Dalian, China, pp. 665-669, 5-8 Dec. 2005.

[23] Y. Xiang, W. Zhou, and M. Guo, "Flexible Deterministic Packet Marking: An IP Traceback System to Find the Real Source of Attacks", IEEE Transactions on Parallel and Distributed Systems, Vol. 20, No. 4, pp. 567-580, April 2009.

[24] Yan Shen, Jun Bi, Jianping Wu and Qiang Liu, "A Two-Level Source Address Spoofing Prevention based on Automatic Signature and Verification Mechanism", Proceedings of 13th IEEE Symposium on Computers and Communications (ISCC 2008), Marrakech, Morocco, pp. 392-397, 6-9 July 2008.

AUTHOR'S PROFILE

Jose Anand received his Diploma from the state board of Technical Education, Tamil Nadu in 1995, Bachelor of Engineering Degree from Institution of Engineers (INDIA), Calcutta in 2003, Master of Engineering in Embedded System Technologies from Anna University, Chennai in 2006, Master of Arts in Public Administration from Annamalai University in 2000, and Master of Business Administration from Alagappa University in 2007. He is a member of CSI, IEI, IET, IETE, ISTE, and INS. He received State 3rd Rank in Bachelor of Engineering. He presented forty six papers in National Conferences and eleven papers in International Conferences. He published one paper in National Journal, fourteen papers in International Journal and published twenty three books for various polytechnic subjects in Electrical, Electronics and Computer Science Disciplines.

K. Sivachandar received his Bachelor of Engineering Degree from Anna University Chennai, in 2008, Master of Engineering in Embedded System Technologies from Anna University Coimbatore, in 2010. He presented two papers in National Conferences. His area interest includes sensor area networks, and security.