IOSR Journal of Engineering (IOSRJEN)
e-ISSN: 2250-3021, p-ISSN: 2278-8719, www.iosrjen.org
Volume 2, Issue 10 (October 2012), PP 09-16

A Dynamic Method to Detect IP Spoofing on Data Network Using Ant Algorithm

N.Arumugam[1], Dr.C.Venkatesh[2]
[1](Research Scholar, Anna University, Chennai, Tamilnadu, India,
[2](Dean, Faculty of Engineering. EBET Group of Institutions, Kankayam, Tamilnadu, India, Member IEEE)

Abstract—*A data packet is typically forwarded from one router to another through networks that constitute the internetwork until it gets to its destination node. At the same time routers in the Internet do not perform any security verification of the source IP address contained in the packets. The lack of such verification opens the door for a variety of network security vulnerabilities like denial-of-service (DoS) attacks, man-in-the-middle attacks etc. One of the major threats to the Internet is source IP address spoofing. To avoid the IP spoofing a number of prevention approaches are proposed by the research community. In this paper an ant-based traceback is proposed to detect the IP spoofing. The proposed traceback approach uses flow level information to identify the spoofing request. To validate the detection method further, this paper considers the number of hop needs to reach the destination end. Using a mapping between IP addresses and their flow level with hop-counts, the server can distinguish spoofed IP packets from legitimate ones. The simulations results show that this approach discards almost 90% of spoofed IP request.*

Keywords—*IP spoofing, IP trace back, Ant algorithm, hop count, pheromone intensity, flow level*

I. INTRODUCTION

Packet forwarding in the Internet is based only on the destination IP address contained in the IP packet. This permits forging of the source IP address, commonly referred to as IP spoofing [1]. IP spoofing is a boon for miscreants. Perhaps the most well-known misuse of IP spoofing is in launching denial-of-service (DoS) attacks on critical Infrastructure such as Web and DNS servers, as evidenced by backscatter analysis [2], [3]. Another avenue made possible by spoofing is that of illegal content distribution. UDP-based peer-to-peer (p2p) applications that exploit IP spoofing to mask the identity of the sender already exist [4], [5].Present approaches to curb IP spoofing researchers have taken two distinct approaches: *router-based* and *victim-based*. The router-based approach makes improvements to the routing infrastructure, while the victim based approach enhances the resilience of Internet servers against attacks. The router-based approach performs either off-line analysis of flooding traffic or on-line filtering of DDoS traffic inside routers. But the victim-based prevention methods, which detects and discards spoofed traffic without any router support. Compared to the router-based approach, the victim based approach has the advantage of being immediately deployable. More importantly, a potential victim has a much stronger incentive to deploy defense mechanisms than network service providers. The current victim-based approach protects Internet servers using sophisticated resource management schemes. These schemes provide more accurate resource accounting, and fine-grained service isolation and differentiation [6].

II. SPOOFED PACKETS DETECTION METHODS

A variety of methods are deployed in determining whether a received packet has spoofed source IP address or not. In Internet, when a node receiving a packet can determine whether the packet is spoofed by either an active or passive ways. The term active mean the host must perform some network action but the passive method doesn't require such action. However an active method may be used to validate cases where the passive method indicates the packet was spoofed. Among different methods this paper considers both IP trace back and hop count based detection method.

2.1 TRACEBACK TECHNIQUES

Since the late 1999 research on IP trace back has been active to detection of DDOS attacks. Several approaches have been proposed to trace IP packets to their origins. IP trace back is usually performed at the network layer, with the help of routers and gateways. The trace back techniques can trace packet paths and help in identifying the perpetrators of the DoS attacks with a high probability. These can be useful forensic tools in law enforcement but do nothing to prevent the occurrence of IP spoofing. Among the spoofing prevention

techniques, many focus on shielding the destination from IP spoofing. Their shortcoming lies in the observation that they fail to protect the Internet routing fabric from being misused in forwarding spoofed packets. The rest of the spoofing prevention techniques possess the ideal goal of preventing spoofing near its source [7]-[11].

2.2 TTL METHODS

When IP packets are routed across the Internet, the time-to-live (TTL) field is decremented. This field in the IP packet header is used to prevent packets from being routed endlessly when the destination host cannot be located in a fixed number of hops. It is also used by some networked devices to prevent packets from being sent beyond a host's network subnet. The TTL is a useful value for detecting spoofed packets. Its use is based on several assumptions, which, from our network observations, appear to be true. When a packet is sent between two hosts, as long as the same route is taken, the number of hops will be the same. This means that the initial TTL will be decremented by the same amount. Packets sent near in time to each other will take the same route to the destination. Routes change infrequently. When routes change, they do not result in a significant change in the number of hops [12].

The objective of this work is to use both the concepts of trace back and hop count of the packet while routing from source to destination on internet. The trace back approach is used to finding out the origin of the spoofing attack using the existing traffic flow information. Furthermore, to strengthen the spoofing prevention hop count value of the packet between the source and destination are also validated. An ant-based trace back algorithm is using for finding the traffic flow information as the trace for ants finding the attack path. The hop-count information is indirectly reflected in the TTL field of the IP header, since each intermediate router decrements the TTL value by one before forwarding a packet to the next hop. The difference between the initial TTL (at the source) and the final TTL value (at the destination) is the hop-count between the source and the destination.

III. PROPOSED METHODOLOGY

This paper proposes an optimistic method that validates incoming request before it reach the destination without using any cryptographic methodology. The fundamental idea is to utilize inherent network information that each packet carries. The inherent network information this paper use here is the flow information and the number of hops of a packet takes to reach its destination. This proposed method uses an ant-based traceback algorithm to find the traffic flow information and hop count value, Since an attacker can forge any field in the IP header, he cannot forged the number of hops an IP packet takes to reach its destination, which is solely determined by the Internet routing infrastructure. The hop-count information is indirectly reflected in the Time-to-Live (TTL) field of the IP header, since each intermediate router decrements the TTL value by one before forwarding a packet to the next hop. Figure.1. shows pictorial representation of the proposed methodology.

3.1 ANT ALGORITHM

Ethnologists states that animals like ants could manage to establish shortest route paths from their colony to feeding sources and back. It was found that the medium used to communicate information among individuals regarding paths, and used to decide where to go, consists of *pheromone trails*. A moving ant lays some pheromone (in varying quantities) on the ground, thus marking the path by a trail of this substance. While an isolated ant moves essentially at random, an ant encountering a previously laid trail can detect it and decide with high probability to follow it, thus reinforcing the trail with its own pheromone. The collective behavior that emerges is a form of *autocatalytic* behavior1 where the more the ants following a trail, the more attractive that trail becomes for being followed. The process is thus characterized by a positive feedback loop, where the probability with which an ant chooses a path increases with the number of ants that previously chose the same path [13]. The idea is that if at a given point an ant has to choose among different paths, those which were heavily chosen by preceding ants (that is, those with a high trail level) are chosen with higher probability. Furthermore high trail levels are the same with shortest paths.

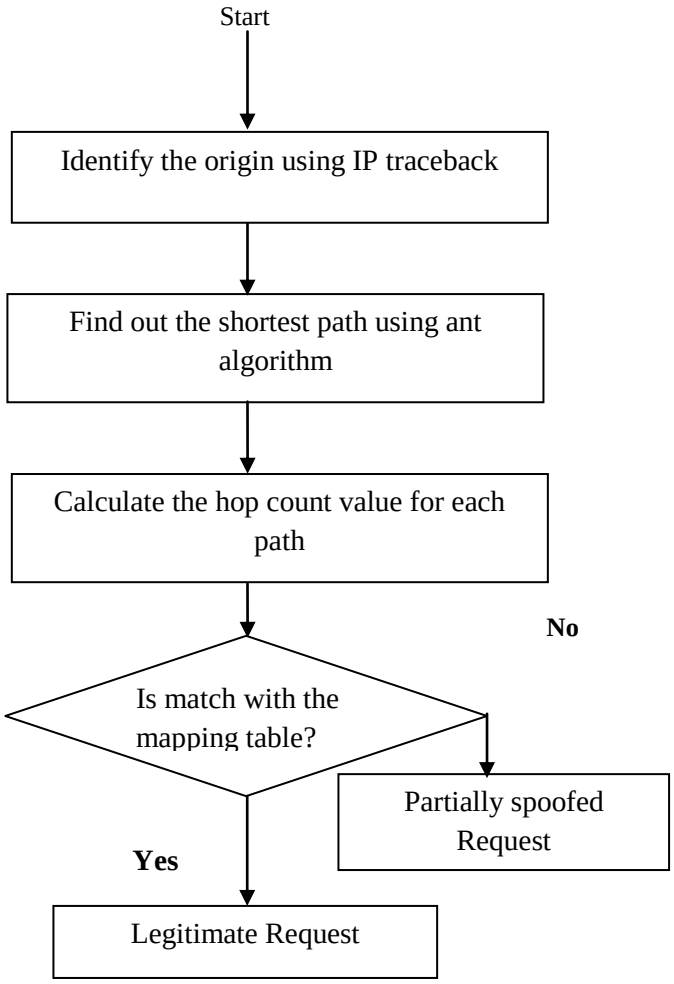


Fig1: Flow chart to detect IP spoofing using ant algorithm

3.2 ANT BASED IP TRACE BACK

Basically, the attack path reconstruction process involves interrogating the routing packets received at the victim in order to find the immediate upstream node and then systematically repeating the interrogation process at each intermediate upstream node until the attack source is reached. The path reconstruction problem could be solved using the ant-based IP traceback. Figure shows the IP trace back of all possible paths from the source node 3 to the destination node D. Basically the ants lay a pheromone trail along the route they select between the source node (the food source) and the destination (e.g., paths 3－2－1, 3－6－5－4 and 3－6－5－1 in Fig. 2), and the relative probability

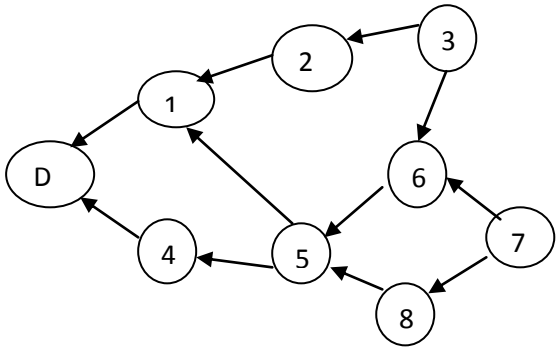


Fig 2: IP trace back of all possible paths

of each path being the actual path is given by the intensity of the pheromone along the corresponding trail. As in nature, the isolated ants in the ant algorithm scheme move essentially at random. However, upon encountering a previously laid trail, the ants decide with a high probability to trace it. As a result, the pheromone intensity of this path progressively increases, and thus the likelihood of the path representing the actual path also increases. The proposed solution could take the victim host as the starting point and perform IP trace back. It is assumed that the legitimate request might reach the victim node in a shortest path [14].
The description of the ant-based IP traceback is as follows:

Step.1: Construct network topology
Step 2: Determine all possible paths between two network nodes (Source node to Destination node)
Step 3: Find out the shortest path

The shortest path searching process is done with the exploitation policy as in the equation (1) chooses the arc with the greatest pheromone intensity and visibility, while the exploration policy as in the equation(2) is a random decision rule. Thus, an ant located at node *i* choose the next node *j* in accordance with the following rule:

$$j=\begin{cases}\arg\max\{[\tau_{ij}(t)^{\alpha}][\eta_{ij}(t)^{\beta}]\} & if\ q \le q^{o}\\ S \quad otherwise\end{cases} \quad \text{---- (1)}$$

$$S=p_{ij}(t)=\begin{cases}\dfrac{[\tau_{ij}(t)]^{\alpha}[\eta_{ij}(t)]^{\beta}}{\sum[\tau_{ij}(t)]^{\alpha}[\eta_{ij}(t)]^{\beta}} \\ 0 \quad otherwise\end{cases} \quad \text{---------- (2)}$$

Where $\tau_{ij}(t)$ =the pheromone intensity of trail between router *i* and router *j* at time

$\eta_{ij}(t)$ = the number of routing packets between router *i* and router *j* between time (t-1) and time (t)

α is the weighting factor of pheromone, β is the weighting factor of visibility.
Ant colony updates the probability density function of feasible attack paths and chooses the right one.

3.3 HOP COUNT COMPUTATIONS

The number of hops a packet takes to reach its destination reflected in the Time-to-Live (TTL) field of the IP header and also each intermediate router decrements the TTL value by one before forwarding a packet to the next hop. Since hop count information is not directly stored in the IP header and it might compute from the final TTL value. TTL indicates the time in which a packet can exist on the network [15].It is defined to prevent a packet from circling on the network and it is decremented by one when passing through one router. Hence it is possible to calculate hop count from the TTL value. TTL is an 8-bit field in the IP header, originally introduced to specify the maximum lifetime of each packet in the Internet.

There are two methods to measure the hop count from a host. One is an active measurement and the other is a passive measurement. The first method is to use ICMP ECHO packets. In most cases this gives an accurate hop count. However applying this method to thousand hosts is not realistic because sending lots of ICMP packets is not recommended as a measuring method. The second method is simply to subtract the TTL of a received IP packet from its initial value. This can be done without sending any sample packets and therefore is ideal of measuring the hop counts of many hosts. However in order to use this method the initial TTL values should be known in advance.

Hop count = (initial TTL) - (TTL)

3.4 PROBLEM WITH INITIAL VALUES OF TTL

According to RFC 1700the recommended initial TTL value is 64. However this rule is often ignored on the real internet. Swiss Academic & Research Network (SWITCH) has researched initial TTL values of different OS (Operating Systems). As a result there are six initial TTL values: 30,32,60,64,128 and 255. The packet whose initial TTL value 255 and the initial TTL value 128 can be distinguished from other easily. However it is more difficult to assess packets whose TTL values are less than 60 or 64.The same problem occurs to the packets with TTL value less than 30.The popular OS like Microsoft Windows, Linux and Free BSD are using 32 and 64 as initial values. Hence the following formula is used to convert TTL to hop count,

Hop Count =

32-TTL	TTL<=32
64-TTL	TTL<=62
128-TTL	TTL<=128
255-TTL	TTL<=255

IV. DIFFERENT PHASES OF PROPOSED METHODOLOGY

To identify the bogus request the following modules are to be implemented. The figure 3 shows the block diagram approach of different phases of detection procedures.

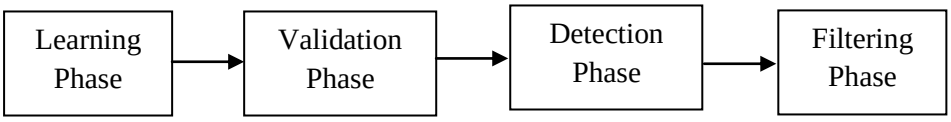


Fig. 3. Different Phases of Fake IP Detection

4.1 LEARNING PHASE:

During this phase the inherent information carries by the request are registered in a database called as mapping table is placed before the server. It contains the successfully connected source IP and the respective hop counts with flow density. It is assumed that during normal connection there is no attacks are happening, hence the respective source IP address its hop count and flow density values can registered in the mapping table. This record which is later used to verify each incoming packet and filter-out the spoofed ones. The learning phase continues for a sufficient time to allow most of the database to be filled up.

4.2 VALIDATION PHASE:

The validation process is simply comparing the request details with the accumulated details of the mapping table. The mapping table consists of IP address and its appropriate hop count, flow level value. This basic technique is straightforward but the implementation is not a simple one. An obvious solution is to collect the IP address and the relevant hop count value when there is no DDoS attack. There are many problems with this approach. First, during a DDoS attack the victim sees a large number of previously unseen address and all of them are considered spoofed because they are not in the database. Second due to frequent routing changes the hop count value most likely will change leads to false positive. A better approach to building the database is to add the IP address, hop count value and flow level for each TCP session separately after the TCP handshake is completed. This guarantees the integrity of the tuple (IP address, hop count, flow level) will used to detect the spoofing of IP address.

4.3 DETECTION PHASE:

After the learning phase, the system begins to perform the detection procedure to filter out the most fake IP request. The spoofing of IP packet is based on verification (Source IP and the relevant hop count with flow level) with the mapping table. When the detail of requested IP is matched with the mapping table, the request is legitimate and allowed to permit otherwise the request is a fake and the request is dropped. A request from an IP whose details are not in the mapping table is assumed as a legitimate request. At the same time the details of the requested IP is registered in the check list database with a probability P. The value of P is set to high (close to 1) initially. In future when the same source IP make a request and successfully completed the verification then the probability value P is decreased. Finally the details are registered in the mapping table when the probability reached zero.

4.4 FILTERING PHASE:

Using the techniques and criteria discussed above, filtering procedure of IPHP is described below. Any packet received by the edge router is authorized to permit the request according to the following rules:

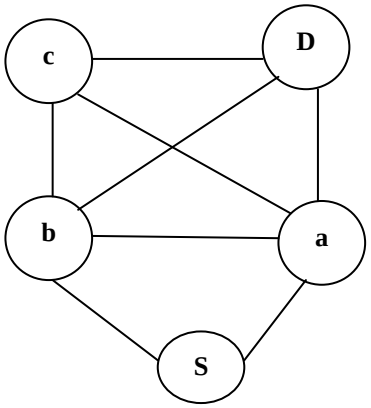
1) If the IP identification details are matched with the records in the mapping table. Permission is granted to access the server.
2) If the source IP address of the packet exists in the mapping table, but the hop count value does not match, then it is a partially spoofed packet and the request is accepted with a probability.

3) If the source IP address of the packet exists in the mapping table, but the flow density value does not match, this packet is considered to be a partially spoofed packet and the request is accepted with a probability.
4) If the source IP address does not appear in the mapping table, then the packet is considered as anew request, accepted with a probability p.

V. RESULTS AND DISCUSSION

In this section, a series of NS2 simulations was performed using a PC with an Intel Dual core CPU 3.0G, DDR2 1G of RAM and the MS Windows XP operating system. Figure 4 show the first experimental topology where 5 numbers of nodes S, D, 1, 2 and 3 were networked in which node S act as source node and D as the destination node.

5.1 Experimental Study#1



S=Source node **D**= Destination node

Fig4: Experimental topology with 5 nodes

Since node S considered as a source node and the node D as s destination, the possible path between the node S and D where identified using ant algorithm. According to the ant system optimization by a colony of cooperating agent, ants follow a path between the source to destination with all possible paths with equal probability. This process continues until all of the ants will eventually choose the shortest path.

Table1: Experimental value for 5 nodes

S.No	Possible Path	Hop Count	Pheromone Intensity
1.	s-> a->d	1	2.029463
2.	s->a>d	1	2.029463
3.	s->a->b->d	2	1.63101
4.	s->a->c->d	2	1.63101
5.	s->b->a->d	2	1.911162
6.	s->a->c->d	2	1.911162
7.	s->a->b->c->d	3	1.427916
8.	s->a->c>b->d	3	1.309615
9.	s->b->a->c->d	3	1.309615
10.	s->b->c->a->d	3	1.512709

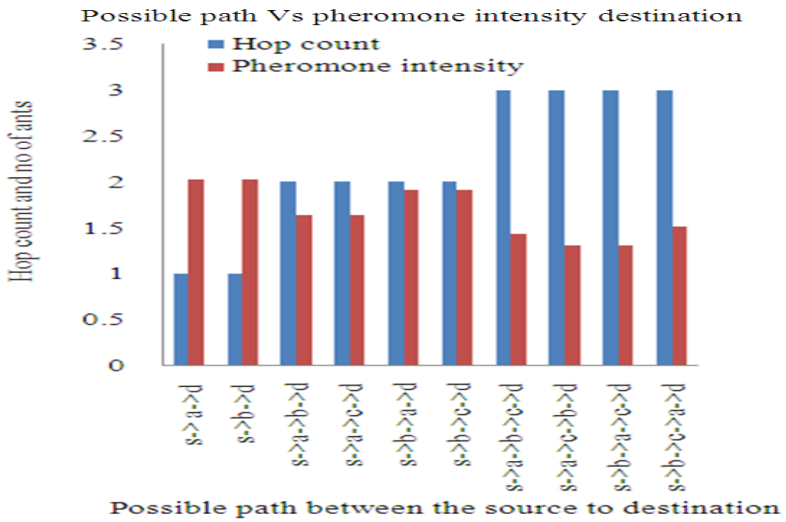


Fig 5: Possible path between source to destination

The idea is that if at a given point an ant has to choose among different paths, those which were heavily chosen by preceding ants are chosen with higher probability based on with a high trail level. Furthermore high trail levels are synonymous with shortest paths. It is understood that the isolated ant would reach the destination in a shortest way. The shortest path is identified by the isolated ant based on the maximum pheromone intensity. Hence it is clear that the shortest path may not have fake request. Experimental values are tabulated as in the table 1 with possible path, hop count and pheromone intensity. From the tabulation it is understood that the legitimate request has minimum hop value and maximum pheromone intensity value. Figure 5 show the possible path among source and destination with pheromone intensity.

5.2 Evaluation of Filtering Accuracy

To evaluate the filtering accuracy under each aggregation method a look up table is constructed based on the destination node IP addresses and evaluate. It is assume that the attacker generates packets by randomly selecting source IP addresses among legitimate clients. It is further assume that the attacker knows the general hop-count distribution for each web server and uses it to randomly generate a hop-count for each spoofed packet. To measure the filtering accuracy of the spoofed request of this proposed method, the term the percentages of false positives and false negatives were used. False positives are those legitimate requests that are incorrectly identified as spoofed. False negatives are spoofed IP addresses that go undetected by HCF. A good aggregation method should minimize both.

5.3 Experiments and results

The initial collection period should be long enough to ensure good filtering accuracy even at the very beginning, and the duration should depend on the amount of daily traffic the server is receiving. HCF will continue adding new entries to the mapping table when requests with previously unseen legitimate IP addresses are sighted. Thus, over time, the IP2HC mapping table will capture the correct mapping between IP address and hop-count for all clients of a server. This ensures that spoofed IP traffic can be detected, and then discarded with little collateral damage during a DDoS attack. From an experiment for different Hop Count values, the mapping table is activated and observed for the false positive and false negative prediction percentage. This is shown in a figure6.

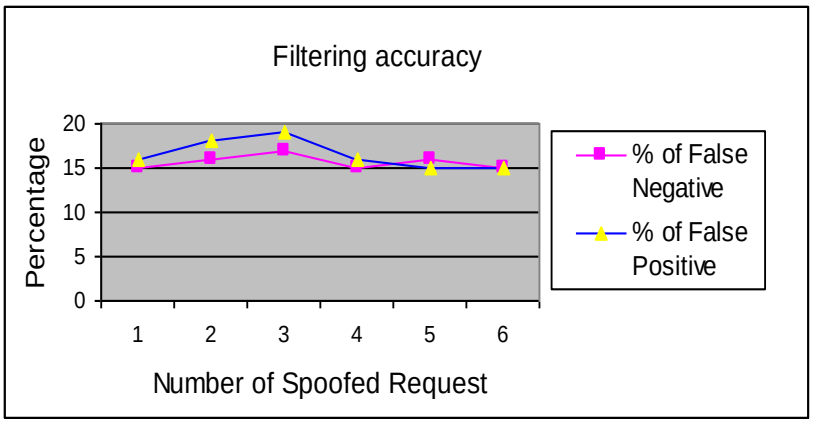


Fig6: Filtering accuracy of the spoofed packets

VI. CONCLUSION

The proposed IP traceback analysis method is an extended of Ant algorithm. This method examines all possible way to reach destination node during it learning stage. Among the different path to reach the destination, it is understood that the legitimate request might prepare shortest path. This shortest path is identified using the pheromone intensity. To strengthen the spoofing identification an additional metric of hop count value is also considered. Thus the legitimate request is validated and permitted to access the destination node based on the metrics hop count and flow level. The simulations results show that this approach discards almost 90% of spoofed IP request.

REFERENCES

[1] R. Beverly and S. Bauer, "The spoofer project: Inferring the extent of Internet source address filtering on the Internet," USENIX Workshop on Steps to Reduce Unwanted Traffic in the Internet (SRUTI), 2005.
[2] D. Moore, G. Voelker, and S. Savage, "Inferring internet denial-of service activity," in USENIX Security Symposium, 2001.
[3] D. Moore, C. Shannon, D. Brown, G. M. Voelker, and S. Savage, "Inferring internet denial-of-service activity," *ACM Transactions on Computer Systems (TOCS), May 2006.*
[4] J. Connelly, "SUMI: A fast anonymous file transfer program website," http://sumi.berlios.de/wiki/Main Page.
[5] RODI,"Rodi website," http://rodi.sourceforge.net/.
[6] Haining Wang, Member, IEEE, Cheng Jin, and Kang G. Shin, Fellow, IEEE," Defense Against Spoofed IP Traffic Using Hop-Count Filtering", *IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 15, NO. 1, FEBRUARY 2007.*
[7] S. Bellovin, M. Leech, and T. Taylor, "ICMP traceback messages," *IETF Draft, Oct. 2001.*
[8] D. Dean, M. Franklin, and A. Stubblefield, "An algebraic approach to IP traceback." *ACM Transactions on Information and System Security, 2002.*
[8] M. Goodrich, "Efficient packet marking for large-scale IP traceback," in *ACM Conference on Computer and Communications Security (CCS), 2001.*
[9] J. Li, M. Sung, J. Xu, and L. Li, "Large-scale IP traceback in high speed Internet: Practical techniques and theoretical foundations," *in IEEE Symposium on Security and Privacy, May 2004.*
[10] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Practical network support for IP traceback," in *ACM SIGCOMM, Aug. 2000.*
[11] A. Snoeren, C. Partridge, L. Sanchez, C. Jones, F. Tchakountio, S. Kent, and W. Strayer, "Hash-based IP traceback," in *ACM SIGCOMM, 2001.*
[12] Steven J. Templeton, Karl E. Levitt Department of Computer Science U.C. Davis. "Detecting Spoofed Packets"
[13] Dorigo, M., Maniezzo, V., & Colorni, A. (1996). The ant system: Optimization by a colony of cooperating agents. *IEEE/ACM Transactions on System, Man and Cybernetics-Part B, 26(1), 1–13.*
[14] Gu Hsin Lai, Chia-Mei Chen, Bing-Chiang Jeng, Willams Chao, Department of Information Management, National Sun Yat-Sen University, Taiwan ," Ant-based IP traceback",*Elsevier,2008.*
[15] W.Richard Stevens, "TCP/IP Illustrated" *Addison Wesley Longman, Inc.,*