Detection of Denial of Service Attack on Data Network Using IP Traceback with Entropy Variation

Dr.N.Arumugam

Lecturer (SG), Dept of ECE, Nachimuthu Polytechnic College, Pollachi, Tamilnadu, South India

***Abstract*- In internet data packets are typically routed through different networks until it gets to reach its destination. In the present state of affairs maximum numbers of day to day activities are performed through online using internet. As the easy accessibility of internet anybody share any information to anyone without any prerequisite. The design architecture of internet do not performs any security verification of the originality of each data 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 DoS attack which is characterized by an explicit attempt by attackers to prevent legitimate users of a service from using that service include attempts to flood a network, thereby preventing legitimate network traffic .It attempts to disrupt connections between two machines, thereby preventing access to a service either to prevent a particular individual from accessing a service or to disrupt service to a specific system. A number of detection techniques are proposed by the research community to detect the origin of the attack. This article proposed a traceback based technique is used to identify the attack origin.**

Index Terms- DoS attacks, IP trace back, hop count, pheromone intensity, flow level.

1. INTRODUCTION

Internet is a global system of interconnected computer networks that works on the standard Internet Protocol Suite (TCP/IP) to serve billions of users worldwide. It is a network of networks that consists of millions of private, public, academic, business, and government networks, of local to global scope, that are linked by a broad array of electronic and optical networking technologies. The Internet carries a vast range of information resources and services, such as the inter-linked hypertext documents of the World Wide Web (WWW) and the infrastructure to support electronic mail. The phenomenal growth of the Internet owes much to the simplicity of its design principles, which allow to widely interconnecting heterogeneous systems. However, the penalty of this success has been poor security. The design architecture of Internet permits anybody to send any request to anyone without being authenticated, while the receiver has to process any information that arrives to a provided service. Due to this lack of authentication attackers can create a fake identity, and send malicious traffic on internet. Therefore any systems connected to the Internet are potential targets for attacks since the openness of the Internet makes them accessible to attack traffic.

Since internet does not have any form of control for a server to dictate how much traffic it wants to receive and from whom. During the routing process 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 man-in-the-middle attacks, Denial-of-Service (DoS) and Distributed-Denial-of-Service (DDoS) attacks, in the mean time economic and social impact of internet has grown to considerable proportions. One of the major threats to the Internet is source IP address spoofing. In current Internet communication world, validity of the source of IP packet is an important issue. Thus the problems of IP spoofing is definitely alarm the legitimate users of the Internet. According to a study conducted by CSI Computer Crime and Security Survey 2010, the impact of Denial-of-Service (DoS) attacks is around 19.8%. Thus it is mandatory to detect and diffuse the DoS attack.

2. RELATED WORK

More than a decade a number of network attack detection techniques have been developed and applied in detection of DoS network attacks. Most of the approaches required to modify the network infrastructure either encoding the router's information into the specific fields of the IP header or storing a representative amount of the packet content at the routers for the attack detection purpose. And also they require all the routers, along the DoS attacking path, to support the detection mechanism. Alex C. Snoeren et al proposed a hash-based IP traceback technique which generates audit trails for traffic within the network, and can trace the origin of a *single* IP packet delivered by the network [1].Probabilistic packet marking (PPM) is another technique where each packet is marked with partial path information at routers. Each router marks their IP address onto the packet with the probability along the way the packet traversed. When DDOS attack is detected, the victim can reconstruct the whole path after collecting certain amount of packet by using the information of the mark, despite the source address in the IP header. One of the demerits of this method is its computation overhead discussed in [2]. The limitation of PPM was modified and reduces the computational overhead to an acceptable level were discussed in [3]. M. M. Viana et al combine PPM and the concept of winding number. Their work shows that they are able to correctly trace the attacker's router IP address using integral equation discussed in [4]. Deterministic Packet Marking (DPM) is a new approach for IP traceback which is scalable and simple to implement, and introduces no bandwidth and practically no processing overhead. It is backward compatible with equipment which does not implement it. The approach is capable of tracing back attacks, which are composed of just a few packets. In addition, a service provider can implement this scheme without revealing its internal network topology [5]. On Deterministic Packet Marking is another approach to IP Traceback based on marking all packets at ingress interfaces discussed in [6].Flexible Deterministic Packet Marking (FDPM) provides a defense system with the ability to find out the real sources of attacked packets that traverse through the network. FDPM provides innovative features to trace the source of IP packets and can obtain better tracing capability than others. In particular, FDPM adopts a flexible mark length strategy to make it compatible to different network environments. It is also adaptively changes it's marking rate according to the load of the participating router by a flexible flow-based marking scheme [7]. Path identification (Pi) DDoS defense scheme is a deterministic packet marking scheme that allows a DDoS victim to filter out attack packets on a per packet basis with high accuracy after only a few attack packets are receive. Enhancement of the idea called the StackPi marking, a new packet marking scheme based on Pi, and new filtering mechanisms. This scheme almost completely eliminates the effect of a few legacy routers on a path, and performs 2-4 times better than the original Pi scheme in a sparse deployment of Pi-enabled routers [8]. Another attack mitigation scheme which adopts divides and conquers strategy. Attack diagnosis combines the concepts of pushback and packet marking. Its architecture is in line with the ideal DDoS attack countermeasure paradigm. Attack detection is performed near the victim host and packet filtering is executed close to the attack sources. Attack diagnosis is a reactive defense mechanism that is activated by a victim host after an attack is detected. By instructing its upstream routers to mark packets deterministically, the victim can trace back one attack source. Attack diagnosis enabled router close to the source to filter the attack packets. This process isolates one attacker and throttles it, which is repeated until the attack is mitigated [9]. Lih-Chyau Wuu et al proposed an IP traceback method based on the Chinese Remainder Theorem, where routers with the proposed method can interoperate seamlessly with legacy routers and be incrementally deployable. In the proposed method a victim does not need to maintain the network topology while it reconstructs attacking paths [10].

3. IP TRACEBACK WITH ENTROPY VARIATION

This section describes a DoS attack detection method based on IP Traceback techniques with Entropy Variation to identify the origin of the attack. Hence this paper considers another flow based metric as entropy variation to find the attack origin. The following section narrates the entropy variation of the packet flows in the network to detect the origin of DoS attack.

3.1 Entropy Variation

Entropy is an information theoretic concept, which is a measure of randomness of data flows for a given time interval. Similarly in internet sequence of data packets are flows through routers are also in the form of packet flow till it reaches its destination. That is a sequence of data packet travels from the source to the destination end. The flow of data packet may denoted as $< u_i , d_i , t >, i, j \in I, t \in R$, where I is the set of positive integers , R is the set of real numbers ,u_i is the source node ,the d_i is the destination node and t is the time stamp. Figure 3 represents different data flows between the nodes. Data packets entering the node is called as input flow similarly when the data packets leaving the node is called as output flow. These data packets flows are also denoted as transit flow. It is represented as L.

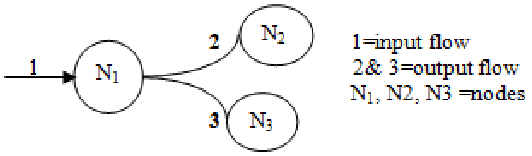


Figure 1: Data flow in a network topology

Thus it is understood that the set U represents the incoming flows to a node and D represents flow of data packets which are leaving the node.

$U = \{u_i, i \in I\} + \{L\}$ (6)

$D = \{d_i, I \in I\}\}$ (7)

Therefore data packet flow at a node can be defined as:

$f_{ij} (u_i, d_j) = \{< u_i , d_i , t >| u_i \in U, d_j \in D, i,j \in I\}$ --(8)

where f_{ij} (u_i, d_j) denotes the number of packets flow at time t. For a given time interval ΔT the variation of the number of packet flow as below:

$N_{ij} (u_i, d_j, t +\Delta T) = | f_{ij} (u_i, d_j, t +\Delta T) |-| f_{ij} (u_i, d_j, t |$ (9)

If $| f_{ij} (u_i, d_j, t | =0$ *then* $N_{ij} (u_i, d_j, t +\Delta T)$ is the number of packets of flow $f_{ij,}$ which went through the initial node during the time interval ΔT. Thus $N_{ij} (u_i, d_j)$ is represents the packets flow.

And hence the probability of each flow at a node based on the large number theorem as:

$p_{ij}(u_i,d_j)=\frac{N_{ij}\ (ui,dj)}{\sum_{i=1}^{\infty}\sum_{j=1}^{\infty} N_{ij}\ (ui,dj)}$ ---------------------- (10)

where $p_{ij}(u_i.d_j)$ *gives the probability of the flow and the denominator is =0*

Let F be the random variable of the number of flows during the time interval ΔTon a local router, therefore the entropy of data packets are defined as follows:

$H (F) = - \sum_{ij} p_{ij}(u_i.d_j) \log p_{ij}(u_i.d_j)$ (11)

where $H(F)$ denoted as entropy variation.

In a network topology N represents the number of data packets flows and the probability of distribution is

P $\{p_1, p_{2,} \dots\dots\dots, p_N\}$ and hence the expression for the entropy is as follows:

$H (F) = H (p_1, p_{2,} \dots\dots\dots, p_N) = -\sum_{i}^{N} p_i \log p_i$ (12)

According to [15] for a non attack case the data packets flows are stable and hence the entropy variation $H(F)$ is stable with minor fluctuations at the same any attacks indulged automatically the entropy variation $H(F)$ is varied.

3.2 Ant System Based Traceback with Entropy Variation

According to the natural behavior of ant, the explorer ant finds out the shortest path to reach the food source and that path is used by the rest of the ants. Among n numbers of path between the source to the destination maximum ants are follows the shortest path. Similarly in data network the maximum flow of data packets are between the source and destination by shortest path only. While traceback the maximum flow of data from the victim it is easy to identify the source of the origin. Here the flow of packets is identified by the chemical substance pheromone intensity. Since it is a chemical substance naturally the intensity of pheromone may disappear thus this article considers another metric based on packet flow as entropy. Entropy is one of the metric used to calculate the random changes. Under normal condition the flow of data packets are normal at the same it is abnormal when there is a flow variations. Thus this article uses these two concepts to find the attack origin.

4. EXPERIMENTAL RESULTS

To verify the performance of the proposed IP Traceback with Entropy Variation algorithm implemented to find the origin of the attack source. A series of experiments was performed through the network simulator NS-2 using a PC with an Intel Dual core CPU 3.0G, DDR2 1G of RAM and the MS Windows XP operating system. Figure 2 show an experimental topology setup constructed with 9 numbers of nodes. The simulation parameters such as the simulation duration, experimental topology size, traffic type, number of nodes and the routing algorithm are list out in table 1. For the analysis it is assumed that out of 9 numbers of nodes, node 1 is treated as an attacker node and the node 9 is a victim node. During the network operation naturally the victim node may receive frequent request from the attacker. To traceback the origin of the attacker all possible paths are identified by implementing ant system algorithm.

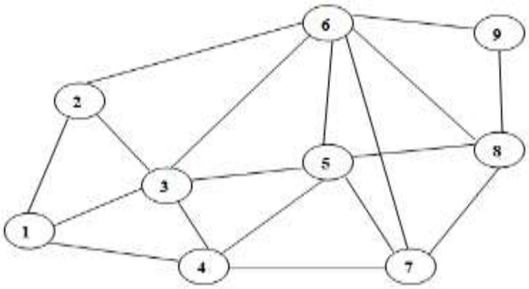
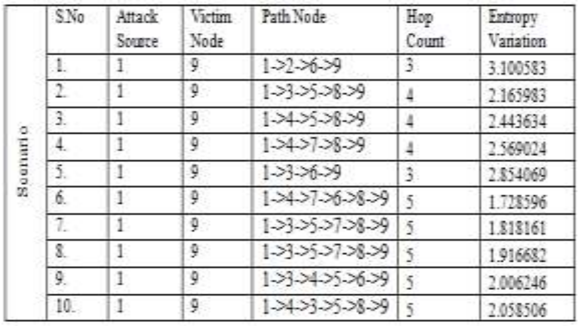


Figure 2: Experimental topology with 9 nodes

S.No.	Parameters	Value
1	Simulation duration	150 seconds
2	Topology	1000m * 1000 m
5	Traffic type	CBR (UDP)
6	Data payload	512 bytes
7	Routing Algorithm	ANT
8	Number of nodes	9

Table1: Simulation Parameters

As per the natural behavior of ant, majority of the ants choose the shortest path to reach the food source. The shortest path is may easily identified by calculating the pheromone intensity. Table 2 shows the details of possible path with entropy variation. From the experimental result it is understood that maximum number of data packets are flooded using the shortest path 1->2->6->9. Since this is a shortest path to reach the victim it is confirmed that node 1 is an attacker node. Figure 3 shows the graphical representation of the scenario discussed.



	S.No	Attack Source	Victim Node	Path Node	Hop Count	Entropy Variation
Scenario	1.	1	9	1->2->6->9	3	3.100583
	2.	1	9	1->3->5->8->9	4	2.165983
	3.	1	9	1->4->5->8->9	4	2.443634
	4.	1	9	1->4->7->8->9	4	2.569024
	5.	1	9	1->3->6->9	3	2.854069
	6.	1	9	1->4->7->6->8->9	5	1.728596
	7.	1	9	1->3->5->7->8->9	5	1.818161
	8.	1	9	1->3->5->7->8->9	5	1.916682
	9.	1	9	1->3->4->5->6->9	5	2.006246
	10.	1	9	1->4->3->5->8->9	5	2.058506

Table 2: Details of possible path with entropy variation of each path.

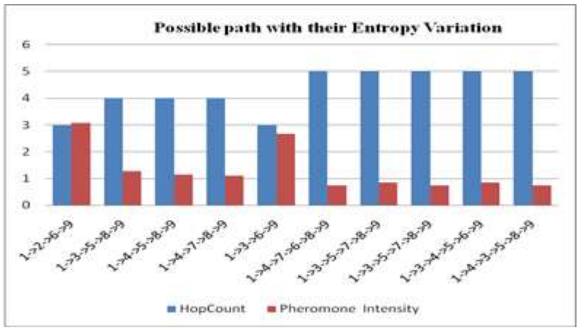


Figure 3: Possible path with their pheromone intensity

5. CONCLUSION

In this paper the proposed method is developed by using the concepts ant system based IP traceback approach with information entropy. According to the ant algorithm more numbers of ants are routed to reach the food source by shortest path. Since ants are using the shortest path to reach their food, naturally in data network data packets are reached the destination by shortest path only. From the experimental result the origin of the attacker is identified by considered the maximum trail of pheromone intensity. To enhance the result further data flow variation (entropy variation) of each path is also considered. From the experimental results it is concluded that the path with maximum pheromone intensity and maximum variation in entropy is the attack path. The simulation results confirmed that maximum pheromone intensity and maximum variation in entropy values are in the shortest path only. Hence, this paper concludes that the proposed solution is an efficient method to find out the DoS attack origin in the networks.

REFERENCES

[1] Alex C. Snoeren et al,“ Hash-Based IP Traceback”, BBN Technologies, SIGCOMM’01, August 27-31, 2001, San Diego, California, USA .

[2] Savage, S., Wetherall, D., Karlin, A., & Anderson, T. (2001). Network support for IP traceback. IEEE/ACM Transactions on Networking, 9(3), 226–237.

[3] D. Q. Li, P. R. Su, and D. G. Feng, "Notes on packet marking for IP traceback," Ruan Jian Xue

Bao/Journal of Software, vol. 15, pp. 250-258, 2004.

[4] M. M. Viana, R. Rios, R. M. De Castro Andrade, and J. N. De Souza, "An innovative approach to identify the IP address in denial-of-service (DoS) attacks based on Cauchy's integral theorem," International Journal of Network Management, vol. 19, pp. 339-354, 2009.

[5] A. Belenky and N. Ansari, "IP traceback with deterministic packet marking," IEEE Communications Letters, vol. 7, pp. 162-164, 2003.

[6] Andrey Belenky, Nirwan Ansari, "On deterministic packet marking," Elsevier, 2006.

[7] 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, pp. 567-580, May 2009.

[8] 8 A. Yaar, A. Perrig, and D. Song, "StackPi: New packet marking and filtering mechanisms for DDoS and IP spoofing defense," IEEE Journal on Selected Areas in Communications, vol. 24, pp. 1853-1863, 2006.

[9] R. Chen, J. M. Park, and R. Marchany, "A divide-and-conquer strategy for thwarting distributed denial-of- service attacks," IEEE Transactions on Parallel and Distributed Systems, vol.18, pp. 577-588, May 2007.

[10] Lih-Chyau Wuu et al, "IP Traceback Based on Chinese Remainder Theorem", Journal of Information Science and Engineering 27, 1985-1999 (2011).

[11] Salah Zidi et al, "Ant Colony with Dynamic Local Search for the Time Scheduling of Transport Networks", International Journal of Computers, Communications & Control, Vol. I (2006), No. 4, pp. 110-125

[12] Marco Dorigo et al, "The Ant System", IEEE Transactions on Systems, Man, and Cybernetics–Part B, Vol.26, No.1, 1996, pp.1-13.

[13] Gu Hsin Lai et al,National Sun Yat-Sen University, Taiwan " Ant-based IP traceback", Elsevier, Expert Systems with Applications 34 (2008).

[14] Marco Dorigo and Thomas St¨utzle, "Ant Colony Optimization: Overview and Recent Advances", Springer Science +Business Media, LLC 2010.

[15] Shui Yu, Member IEEE et al, "Traceback of DDoS Attacks Using Entropy Variations", IEEEE transactions on parallel and distributed systems, vol. 22, no. 3, March 2011.