Detection of Denial-Of-Service Attack on Data NetworkUsing IP Trace Back with Entropy Variation

[1]Shubha Jain, [2]Vijatashwa Awasthi, [3]Ajay Gaur, [4]Amit Sabharwal, [5]Waseed Mohamad Khan

Department of Computer Applications, Axis Institute of Higher Education, Kanpur, Uttar Pradesh, India

Abstract

In internet data packets are typically routed through various networks until they reach their destination. Currently, the majority of daily activities are conducted online using the internet. Due to the ease of internet access, anyone can share information with anyone else without any prerequisites. The design architecture of the internet does not perform security verification of the originality of each data packet. This lack of verification opens the door to various network security vulnerabilities, such as denial-of-service (DoS) attacks and man-in-the-middle attacks. A major threat to the internet is the DoS attack, characterized by explicit attempts by attackers to prevent legitimate users from accessing a service. This can include efforts to flood a network, thereby blocking legitimate network traffic, or disrupting connections between machines to prevent access to a service. Various detection techniques have been proposed by the research community to identify the origin of these attacks. This article proposes a traceback-based technique to identify the source of the attack.

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

1. INTRODUCTION

With billions of users worldwide, the Internet serves as a global network of interconnected computer networks that make use of the standard Internet Protocol Suite (TCP/IP). It is an intricate network system made up of multiple public, private, academic, corporate, and government networks. These networks range in size from small to large and are connected by a variety of optical and electrical networking technologies. The World Wide Web (WWW), which is made up of interconnected hypertext texts, and the infrastructure required for electronic mail are only two examples of the many information resources and services that may be accessed via the Internet.

The phenomenal growth of the Internet is largely attributed to the simplicity of its design principles, which allow for the extensive interconnection of heterogeneous systems. However, this success has come with a significant drawback: poor security. The Internet's design architecture allows anyone to send requests to any other user without authentication, requiring the receiver to process all incoming information. The absence of proper authentication allows malicious actors to fabricate identities and transmit harmful data across the internet. As a result, any device connected to the internet becomes susceptible to potential attacks due to its vulnerability to malicious traffic. The internet lacks a mechanism for servers to regulate the volume and source of incoming traffic During the routing process, routers fail to perform security checks on the source IP addresses, which puts networks at risk for attacks such as Distributed Denial-of-Service (DDoS) and Man-in-the-Middle (DoS). Because of the internet's growing social and economic influence, source IP address spoofing is a serious risk. It is essential in today's internet-based communication environment to confirm the legitimacy of an IP packet's source. For this reason, stopping IP spoofing is crucial for users who are authorized to use the internet. The 2010 CSI Computer Crime and Security Survey estimates that DoS assaults have an approximate 19.8% impact. For this reason, detecting and preventing DoS attacks is essential.

2. RELATED WORK

Many methods for detecting network attacks have been developed and used for more than ten years to identify denial-of-service (DoS) attacks. The majority of techniques necessitated altering the network infrastructure, either by encoding the router's information into particular IP header fields or by storing a portion of the packet content at the routers in order to identify attacks. Additionally, they need the detection technique to be supported by every router along the DoS attack path. A hash-based IP traceback technique was suggested by Alex C. Snoeren et al. It can track the origin of a single IP packet that the network delivers and creates audit trails for network traffic.

[1Another method is called probabilistic packet marking (PPM), in which routers mark each packet with a partial path. Every router appends their IP address to the packet together with the likelihood of each step the packet took. Even if the source address is not included in the IP header, the victim of a denial-of-service attack can reconstruct the entire path after gathering a specific number of packets using the mark's information. The calculation expense of this approach is one of its drawbacks.

PPM's limitations were adjusted, and [3] covered how to lower the computational overhead to a manageable level. PPM and the idea of winding number are combined by M. M. Viana et al. Their research demonstrates that they can use the integral equation covered in [4] to accurately trace the attacker's router IP address. PPM's limitations were adjusted, and [3] covered how to lower the computational overhead to a manageable level. PPM and the idea of winding number are combined by M. M. Viana et al. Their research demonstrates that they can use the integral equation covered in [4] to accurately trace the attacker's router IP address. Deterministic Packet Marking (DPM), a revolutionary approach to IP traceback, is scalable, user-friendly, and requires essentially no processing cost or bandwidth. It is backwards compatible with non-implementing devices as well. This technique can be used to track down attacks that comprise a small number of packets. Moreover, a service provider can use this strategy without having to reveal its internal network structure. [5]. On Deterministic Packet Marking is an additional technique for IP traceback that is described in [6]. It is predicated on marking each packet at ingress interfaces. As packet attacks travel over a network, a security system can determine their true sources by using Flexible Deterministic Packet Marking (FDPM). [5[6] describes On Deterministic Packet tagging, a further technique for IP traceback that is based on tagging each packet at input interfaces. As packet attacks travel over a network, a security system can determine their true sources by using Flexible Deterministic Packet Marking (FDPM).

With its cutting-edge characteristics, FDPM is able to track the origin of IP packets more effectively than other solutions. FDPM specifically uses a flexible mark length approach to provide compatibility with various network environments. Additionally, it uses a flexible flow-based marking technique to adaptively change its marking rate dependent on the load of the participating router. [7].

After receiving a small number of assault packets, a DDoS victim can efficiently filter out attack packets on a per packet basis with great precision thanks to the Path identification (Pi) DDoS protection technique, which is a deterministic packet marking scheme. The Stack Pi marking, a unique packet marking technique inspired by Pi, and creative filtering mechanisms are an advancement on this idea. When Pi-enabled routers are dispersed, this method offers 2-4 times better performance than the original Pi scheme and greatly mitigates the impact of a small number of out-of-date routers along a path. [8]. Using a divide and conquer tactic is an alternate method of preventing attacks. Pushback and packet marking are integrated into the attack diagnosis procedure. This method's structure is in line with the best possible paradigm for defending against DDoS attacks. Attacks are detected close to the host that is being attacked, and packets are filtered close to the attack sources. The victim host initiates attack diagnostics as a reactive defense mechanism when it detects an attack. The target can identify the source of the attack by giving its upstream routers instructions to mark packets deterministically. Thus, the router close to the source is able to filter out the assault packets. The process entails separating specific assailants.

An IP traceback methodology based on the Chinese Remainder Theorem was presented by Lih-Chyau Wuu and colleagues. This technique allows routers using it to coexist peacefully with older routers and be gradually deployed. This method reconstructs the attackers' pathways, relieving the targeted victim of the task of maintaining network topology. [10].

3. IP TRACEBACK WITH ENTROPY VARIATION

This section explains a method for detecting denial-of-service attacks (DoS) that uses IP Traceback techniques along with Entropy Variation to pinpoint the attack's origin. In order to identify the source of the attack, this paper takes into account an additional flow-based statistic called entropy variation. The entropy variation of packet flows in the network to identify the DoS attack's source is described in the section that follows.

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 < ui , di , t >,i, j Є I, t Є R , where I is the set of positive integers , R is the set of real numbers ,ui is the source node ,the di 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.

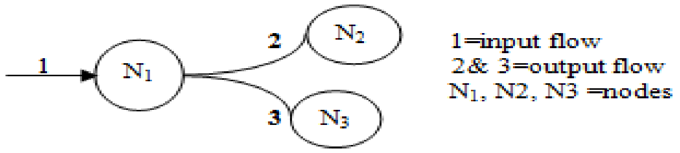


Figure 1: Data flow in a network topology

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 € I} + {L} (6)

D = {d_i, I € I}} (7)

Therefore, data packet flow at a node can be definedas:

f_{ij} (u_i, d_j) = {< u_i , d_i , t >| u_i € U, d_j € D,i,j € I} --(8) where f_{ij} (u_i, d_j) denotes the number of packets flow attime t. For a given time interval ΔT the variation ofthe number of packet flow as below:

N_{ij} (u_i, d_j, t +ΔT) = | f_{ij} (u_i, d_j, t +ΔT) |-| 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 +ΔT) 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: 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, ………, pN} and hence the expression for the entropy is as follows:

H (F) = H (p_1, p_2, ………, pN) = $-\sum^{N} p_i \log p_i$ (12) i

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 Trace back with Entropy Variation

As per its natural way, the explorer ant finds the shortest route to the food supply, which the other ants then take advantage of. Most ants take the shortest way between the source and the destination, out of all the possible paths. Similar to this, in a data network, the shortest path between a source and a destination allows for the most data packet flow. It is simpler to identify the origin when the victim's maximum data flow is traced back. The intensity of the chemical substance pheromone controls the packet flow. Since pheromone intensity can naturally decline, this article presents entropy, another packet flow-based statistic.

One of the metrics used to determine the random changes is entropy. The data packet flow is normal under typical circumstances, but it is abnormal when there is a change in the flow. In order to determine the source of the attack, this article applies these two ideas.

4. Experiment Result

To confirm the Entropy Variable algorithm's and IP track back's performance. NS-2 network simulator running on a PC with an Intel dual core 3.0 GHz processor, DDR2 1 GB RAM, and MS Windows XP. An experimental topology setup with nine nodes is depicted in Figure 2. The characteristics of the simulation, including its duration, the size of the experimental topology, the kind of traffic, the number of nodes, and the routing.

Table 1 contains a list of algorithms. It is assumed for the sake of the analysis that, among the nine nodes, node 1 is an attacker node and node 9 is a victim node. Naturally, the vulnerable node may receive repeated requests from the attacker while the network is operating. In order to determine the attacker's origin, every potential route is discovered and put into practice.

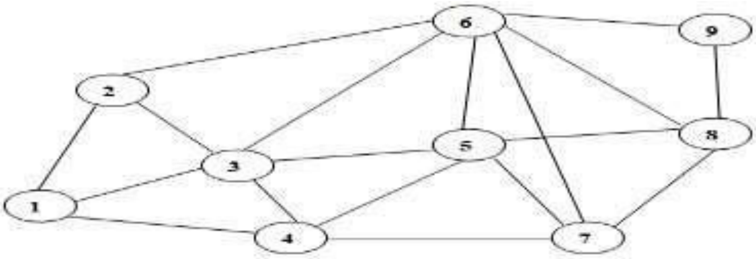
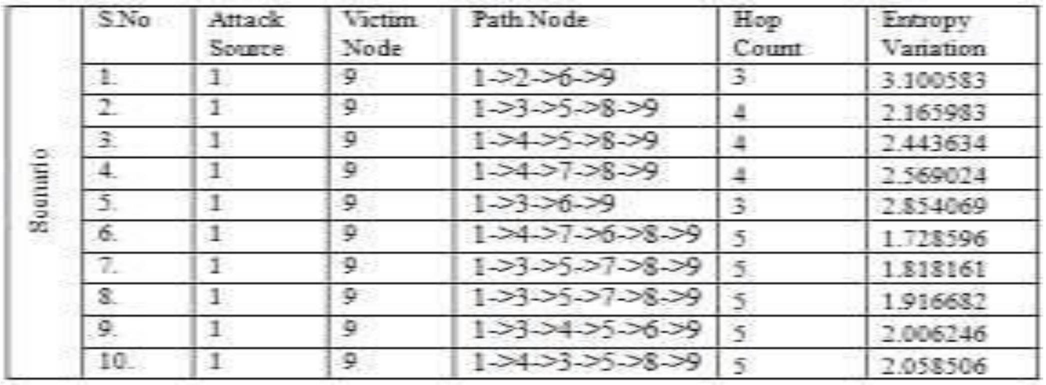


Figure 2: Experimental topology with 9 nodes

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

Most ants take the shortest route to the food source in accordance with their natural habit. Pheromone intensity calculations make it simple to determine the shortest path. The features of a potential path with entropy fluctuation are displayed in Table 2. Based on the experimental results, it can be inferred that the shortest path (1->2->6->9) floods the highest number of data packets. It is established that node 1 is an attacker node because this is the fastest route to get to the victim. The graphical depiction of the scenario under discussion is displayed in Figure 3.



	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.

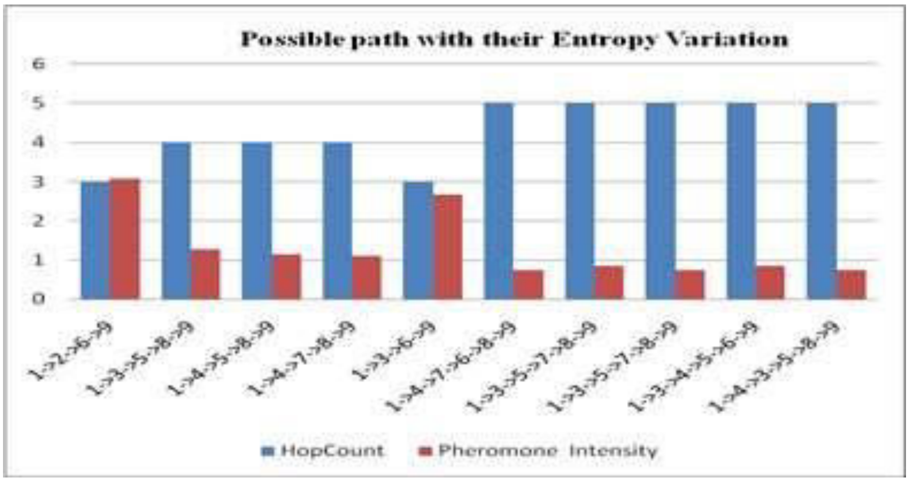


Figure 3: Possible path with their pheromoneintensity

5. Conclusion

This research develops the proposed strategy by combining information entropy and an ant system-based IP traceback approach. The ant algorithm routes a greater number of ants along the quickest path to the food source. In a data network, data packets naturally follow the quickest path to reach their destination, just as ants do to get to their meal. The greatest trail of pheromone intensity is used to determine the attacker's origin based on the experimental results. Further data flow fluctuation (entropy variation) of each path is also taken into consideration to improve the outcome. Based on the experimental findings, the attack path is determined to be the one with the greatest pheromone intensity and the greatest variance in entropy.

REFERENCE

[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 RecentAdvances”, Springer Science +Business Media, LLC 2010.