IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-3, Mar 2019 E-ISSN: 2347-2693

Mitigating the Effect of Distributed Denial of Service Attack

Lavlish Goyal[1*], Nitika[2]

[1]Information Technology, ABV-Indian Institute of Information Technology, Gwalior, India
[2]Computer Science, PPIMT, Hisar, India

[*]*Corresponding Author: lgoel678@gmail.com, Tel.: +91-89494-51919*

DOI: https://doi.org/10.26438/ijcse/v7i3.10321035 | Available online at: www.ijcseonline.org

Accepted: 05/Mar/2019, Published: 31/Mar/2019

***Abstract*—** As the WORLD WIDE WEB (WWW) has been drawing in more business and transaction to be actualized over the Internet, it also has attracted web server attackers having malicious intends. Among different attacks possible against the web server, Distributed Denial of Service (DDoS) attack is the most harmful kind of thread, as it is hard to detect and dispose of completely. The use of spoofed IP addresses to launch DDoS makes it even much harder to identify the source of an attack. In DDoS, attack traffic resembles to legitimate traffic and difficult to differentiate. Still, there are some features which can be used to differentiate normal request from the malicious ones. The source IP distribution is one of the features. In the proposed mechanism a data structure is maintained at victim server which is used to decide priority among different requests during DDoS based on previous visit history and source IP distribution of that server. In the descending order of priority, the server serves the request of users at the time of DDoS.

Keywords— DDoS; IP spoofing; Source distribution; Priority; Defence

I. INTRODUCTION

As the main aim of the internet is openness and scalability but due to this architecture, security of the internet is very poor. The architecture of the internet includes principals like resource sharing, simple core & complex edge routers, multipath routing, fast core networks & slow edge networks and decentralized internet management [1] which contains vulnerabilities. Internet Protocol (IP) is designed to ease the attachment of new host to the network, and there is little support for verification of IP packet header fields. Due to this, it is possible to fake details in the IP packet header. Packets are delivered to their intended destination and it's up to destination server to check the authenticity of packets and decide whether to accept or reject packets from that source. As the validity of requests is checked by the server, this creates an opportunity for a class of attack known as the denial of service attack. Denial of service attack (DoS) tries to deny access of legitimate request of shared services and resources. When there are large numbers of systems involved in DoS attack than it is Distributed Denial of service attack (DDoS).

In the DDoS attack, the first step involves to find vulnerable systems on the internet and convert them into zombies by installing attacking tools. In second stage, attacker sends command to these zombies to attack on specific victim via secure channels. In DDoS, the original attacker doesn't directly involve in the attack and directs zombies to attack the victim.

Among different attacks possible against web servers, Distributed Denial of Service (DDoS) attack is the most harmful kind of thread, as it is hard to detect and dispose of completely. In TCP SYN DDoS [2] attack, attacker initiates the attack from different places and sends TCP SYN requests from those places. In this type of attack, usually, the source address is spoofed such that response i.e. TCP SYN-ACK from victim server never reaches to clients due to the spoofed source address. But the victim server still keeps some resource for this open connection and wait for ACK from source address till some holding time. At this time server has many open connection holding resources and waiting for ACK. Due to resources held by those malicious open connection, remain no resources for legitimate connections. At this point server becomes inaccessible to legitimate users and service is denied to them. The main problem with these type of attack is that malicious request resembles legitimate requests and very hard to differentiate. Due to this DDoS attacks becomes more harmful and effective.

In DDoS attack, traffic resembles that of legitimate access and difficult to differentiate these two. But there are still some features which can be used to differentiate normal traffic from malicious one [3]. In these features, we will concentrate on source IP distribution of particular server and access

frequency of each IP. By analysing source IP distribution at the time of the DDoS attack to non-attack period, we can differentiate legitimate requests from malicious ones.

Most of the web servers except some like google/yahoo serve the specific community of users and provide specific services to them. Those users have a common interest or they reside geographically nearby. Generally, IP address assignment also depends upon the specific country. We can analyses IP distribution of web server based on the request received from different countries as most of the source IP address range is assigned to ISP of that country. For example, most of the users of www.google.co.in are Indians and if a lot of request received from outside India then it is probably a DDoS. The user IP address distribution of each web server is very specific and it's very hard to mimic by the attacker. At the time of the DDoS attack, attacker use spoofed IP address and seems that web server is receiving the request from all over the world. Some users of web server are those who visit regularly and some are those who visit rarely. So, at the time of DDoS, it is more appropriate to provide priority to those who are more frequent than others or not visited at all. This priority is decided based on previous visit history of IP address.

II. RELATED WORK

Different techniques have been proposed to stop spoofing of IP address in the IP packet header. The process of identifying the genuine source of the packet is called IP trace back.

Ingress/egress filtering [4]: is deployed at edge router to eliminate spoofing of IP address. In this schema edge router examines source IP address with respect to a set of a valid IP address for each packet received. If the source IP address doesn't match than packet is discarded by edge router. In ingress/egress filtering IP spoofing is still possible, the attacker can use any IP address from a valid set of IP address to spoof.

Probabilistic Trace back Schema [5]: In this schema routers marks packet probabilistically. When packets transmission takes place between source- destination pair and packet travel through router than router embed their IP address into packet probabilistically. When the victim server receives a sufficient number of the marked packet than it can reconstruct path followed by packets.

History-based mechanism [6]: Victim server keeps a history of all visitor during normal access and when the load at server reaches to a threshold then this defence mechanism becomes active. Only those sources are allowed to establish a connection which has a history with victim server.

Hop-count filtering [7]: during non-attack period victim server records information regarding the particular source and hop distance between the source and itself. Once the attack is detected at victim than this defence mechanism check hop count for all source IP addresses. If there is a difference between recorded and current hop count than packet is considered as malicious and all subsequent packets are discarded from that source.

III. EXPERIMENTATION

For analysis we have taken 7 days of HTTP logs of the NASA Kennedy Space Centre WWW server in Florida, US provided by Internet Traffic Achieve [8] from 1st July to 7th July 1995. From 7 days of traffic traces, first 6 days of traffic traces are of normal access and 7th-day traffic trace is merged with TCP SYN DDoS attacking tool generated traces.

From traffic traces, it is clear that source distribution during DDoS attack is much different than that of normal access. Source distribution during DDoS is much simpler as compared to various interest user community during a non-attack period. First, we calculate the appearance frequency of legitimate sources during the first six days of traffic traces. The high-frequency users are those who access web server regularly and requests data from the server. It is also possible that these users are behind some proxy or NATed network within some organization or university. During inspected period distinct sources who accessed the web server, only 14% of them have accessed server only once. These could be the curious users who just want to take a look at the website. But at the time of DDoS attack, this behaviour is different from normal behaviour. At the time of DDoS attack around 82% are those IP sources who have accessed server only once. During DDoS, the attacker uses spoofed IP addresses due to which it is unlikely that the same IP appears more than once.

IP address space allocation is managed by The Internet Assigned Numbers Authority (IANA) which provides right to five regional Internet registries (RIRs) to provide space to Internet Service Providers (ISPs) and other entities. To understand the user community IP distribution of a web server it is important to understand regional and local information contained within 32 bit IP address. Generally, MSB bits of IP address contains regional information, middle bits about locality and LSB bits contains information about the specific host. For example, if first 8 bit of address range from 96 to 99 than this IP address resides in North America and belongs to ARIN (American Registry for Internet Numbers) ISP. By analysing traffic traces it is clear that IP addresses during attack time are randomized and even some of them belong to reserved IP space of IANA.

As most of the users have a commonality of interest or resides close to each other. If we aggregate there IP address

on some IP subnet address than we may get many sources in that subnet address, also known as clustering. So, during normal flow, many IP sources may decrease quickly as we reduce subnet address bits but this is not applicable for DDoS attack, as IP addresses are spoofed and randomized. To differentiate legitimate flow from attack, we count number of distinct clusters at various IP prefix levels.

From traffic traces of normal and attack period, we counted numbers of clusters to differentiate these two. Cluster ratio for prefix n is defined as numbers of sources at prefix n divided by numbers of sources at prefix 32. The result from the analysis is shown in figures:

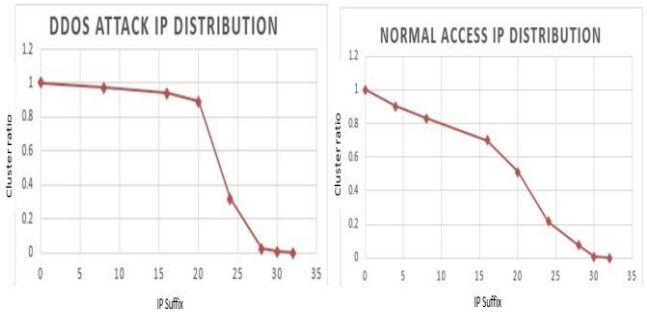

Fig.1.IP source distribution at normal and attack period

From Fig.1 it is clear that during normal access as we increase IP host suffix from 0 to 32, the number of clusters decreases gradually. As most of source IP addresses reside in the proximity of each other. But during DDoS behavior is different from legitimate access. As we increase IP suffix from 0 to 20, numbers of cluster remain same and after more increase in suffix numbers of clusters dropped quickly. This is due to fact that source IP's are distributed over all IP address space and their geographical location is not close to each other).

IV. DEFENCE MECHANISM

At the time of DDoS attack, a large number of request is received by the server from both legitimate users and attacker. The aim of any defence mechanism is to provide access to legitimate users and decline request from the attacker. In this proposed mechanism access request from particular, IP address will get priority over other IP addresses based on access pattern history maintained at the web server.

Any defence mechanism also needs to consider that not all IP addresses are static as a large portion of IP address space is dynamic in nature, means after each reconnection user will get a new IP address. These changes in IP address is in LSB part which provides host information means net/subnet address remains same. For these type of IP, address priority is given to net/subnet addresses which are more frequent than others.

In the proposed mechanism a tree-shaped data structure is maintained at server side to keep track of frequent users. The data structure is maintained as four level 256-ary tree to cover entire IP address space conveniently. Each node is a table of 256 records in the tree and contains 2 fields: one which keeps count of numbers of the request and another for one pointer pointing to the node of next level in the tree. A table contains numbers of requests from IP addresses having 0-bit,8-bit, 16-bit or 24-bit prefix based on the level of the tree. The root node at 0th level of tree contains count of request from addresses 0.*.*.*,1.*.*.*,….255.*.*.*. etc. At root node 120th entry contains the requests count of IP addresses with prefix 119.*.*.*. If the count reaches to a threshold value then pointer points to sub-node at 1-level which contains a table to keep track of request from IP addresses with prefix 119.0*.*, 119.1.*.*…etc. At 1-level if requests count reaches to threshold then another node is added which contains numbers of the request from 24-bit prefix addresses, for example, request from 119.117.0.*.

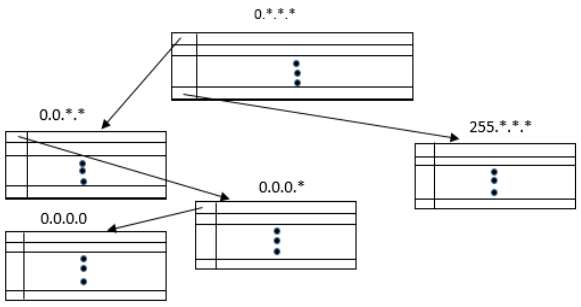

Fig.2. Data structure maintained at victim server

The level 3 of tree contains full IP address which crossed the threshold of level 2 for example 119.117.0.1. The IP addresses which are at level 3 have highest priority over another IP address to access the server. IP address prefixes at level 2 will get priority over level 1 prefixes and they will get priority over level 0 IP prefixes. IP addresses which are at level 0 will get priority over IP addresses which have never visited the site before the attack. To calculate priority over other IP addresses following formula is used:

$$P(I)=a0*(I0/T0) +a1*(I1/T1) +a2*(I2/T2) +a3*(I3/T3)$$

Where P(I) is the priority of IP address 'I' over other IP addresses and a0,a1, a2, a3 are the coefficients for 0,1,2,3 level respectively and satisfy condition of a0<<a1<<a2<<a3.These coefficients show the significance of that level and their value can be adjusted as par requirement. Another term 'Ii/Ti' where i=0,..3 is the ratio of numbers of visits of IP address at ith level of tree and total numbers of visits from all IP addresses at that level.

From priority formula, it is clear that as IP address moves to upper level of tree and more frequent than others will get the highest priority to access the server.

V. ANALYSIS OF PROPOSED MECHANISM

By applying proposed defence mechanism to traffic traces, around 39% of attacking IP addresses were able to pass and the remaining 61% were discarded at level 0. Only 5.1% of attacking IP addresses were present at level 1.But in case of legitimate IP addresses around 85% of those IP addresses were able to pass level 0, around 71% level 1, around 56% level 2 and around 40% level 3.

Those 40% legitimate IP addresses at level 3 will get the highest priority over other IP addresses. Among 40% legitimate IP addresses, those which are more frequent than others will get more priority from the server. The 56% legitimate IP addresses which are at level 2 will get priority over level 1 and level 0 because their subnet addresses were more frequent than level 1 and level 0 subnets.

VI. CONCLUSION

The source IP distribution is the one of feature which is unique to every server and hard to mimic by an attacker. Source distribution during DDoS is much simpler as compared to various interest user community during a non-attack period. In the proposed defence mechanism, priority is calculated based on the previous visit history of IP addresses. Based on the priority of subnet address, the request from particular IP addresses is served. The proposed defence mechanism is activated when the load at the server reaches a threshold value and it can be coupled with existing defence mechanisms of the server.

VII. FUTURE WORK

For our analysis, we have taken numbers of visits from different IP addresses as the only criteria to assign them to different levels of the tree. Also, we can consider other criteria such as numbers of packets sent in each request, time spends during each visit etc. to assign them to the different level of the tree. More accurate results may be obtained by considering previously discussed criteria. Also, we may take n-level tree where n>4 as a data structure to get information about more frequent subnets.

REFERENCES

[1] Peng, Tao, Christopher Leckie, and Kotagiri Ramamohanarao. "Survey of network-based defense mechanisms countering the DoS and DDoS problems." ACM Computing Surveys (CSUR) 39.1 (2007): 3.

[2] http://tools.ietf.org/html/rfc4987

[3] Quyen Le; Zhanikeev, M.; Tanaka, Y., "Methods of Distinguishing Flash Crowds from Spoofed DoS Attacks," *Next Generation Internet Networks, 3rd EuroNGI Conference on* , vol., no., pp.167,173, 21-23 May 2007.

[4] Ferguson, Paul. "Network ingress filtering: Defeating denial of service attacks which employ IP source address spoofing." (2000).

[5] K. Park, and H. Lee, On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack, in Proc. IEEE INFOCOM 2001, pp. 338347.

[6] T. Peng, C. Leckie, and K. Ramamohanarao, Protection from distributed denial of service attacks using history-based IP filtering, ICC '03. May , Vol.1, pp: 482- 486, 2002.

[7] Jin, Cheng, Haining Wang, and Kang G. Shin. "Hop-count filtering: an effective defense against spoofed DDoS traffic." Proceedings of the 10th ACM conference on Computer and communications security. ACM, 2003.

[8] WorldCup 98 and NASA Web server logs are available on the Internet Traffic Archive: http://ita.ee.lbl.gov/html/traces.html

Authors Profile

Mr. Lavlish Goyal pursed Master of Techanology from IIIT Gwalior, India in 2014. He is currently working in MNC as lead engineer in multimedia domain from 4.5 years.On his behaf,MNC has filled 8 patents in various countries including india and USA.

Ms Nitika pursed Bachelor of Techanology from JCD Sirsa and Master of Techanology from PPIMT hisar, India. She is currently teaching computer science to rural students.

