International Journal On Engineering Technology and Sciences – IJETS™
ISSN (P): 2349-3968, ISSN (O): 2349-3976
Volume 2 - Issue 5, May 2015

A Novel Approach to Denial-of-Service Attack Detection with Tracebacking

Jasheeda P
M.tech. Scholar
Department of CSE
Kathir College of Engineering
Coimbatore
jashi108@gmail.com

Faisal E
M.tech. Scholar
Department of CSE
Kathir College of Engineering
Coimbatore
faisalecse@gmail.com

T.K.P.Rajagopal
Associate Professor
Department of CSE
Kathir College of Engineering
Coimbatore
tkprgrg@gmail.com

Abstract— *The reliability and availability of network services are being threatened by the growing number of Denial-of-Service (DoS) attacks. This paper proposes a multivariate correlation analysis approach to investigate and detect the Dos attack. The proposed system applies the idea of Multivariate Correlation Analysis (MCA) to network traffic characterization and employs the principal of anomaly-based detection in attack recognition. One major difficulty to defend against Distributed Denial-of-service attack is that attackers often use fake, or spoofed IP addresses as the IP source address. To capture the spoofers, this paper proposes passive IP traceback (PIT) that bypasses the deployment difficulties of IP traceback techniques. PIT investigates Internet Control Message Protocol error messages (named path backscatter) triggered by spoofing traffic, and tracks the spoofers based on public available information (e.g., topology). In this way, PIT can find the spoofers without any deployment requirement.*

Index Terms—Denial of Service, Mahalanobis distance, Multivariate correlation analysis, Traceback, Triangle Area Map

I. INTRODUCTION

The distributed denial of service (DDoS) attack is a serious threat to the security of cyberspace. It typically exhausts bandwidth, processing capacity, or memory of a targeted machine or network. To launch a DoS attack, malicious users first build a network of computers that they will use to produce the volume of traffic needed to deny services to computer users. To create this attack network, attackers discover vulnerable sites or hosts on the network. Vulnerable hosts are usually those that are either running no antivirus software or out-of-date antivirus software, or those that have not been properly patched. Vulnerable hosts are then exploited by attackers who use their vulnerability to gain access to these hosts. The next step for the intruder is to install new programs (known as attack tools) on the compromised hosts of the attack network. The hosts that are running these attack tools are known as zombies, and they can carry out any attack under the control of the attacker. Many zombies together form what we call an army.

DoS attack detection is essential to the protection of online services. Network-based detection mechanisms are widely used. Network–based detection systems[1] are classified into misuse-based detection systems and anomaly-based detection systems[2].Due to various drawbacks of misuse-based detection systems, anomaly based detection systems are widely used. Since spoofed packets are used for DoS attack, it is difficult to find out the route of attack. An effective method for tracebacking is also necessary.

II. RELATED WORKS

DDoS attack detection metrics are mainly separated into two categories: the signature-based metric and anomaly-based metric. The signature-based metric depends on technology that deploys a pre defined set of attack signatures such as patterns or strings as signatures to match incoming packets. The anomaly-based detection metric typically models the normal network (traffic) behavior and deploys it to compare differences with incoming network behavior. Anomaly based network intrusion detection techniques are a valuable

technology to protect target systems and networks against malicious activities [2]. Anomaly-based detectors attempt to estimate the ‘‘normal’’ behavior of the system to be protected, and generate an anomaly alarm whenever the deviation between a given observation at an instant and the normal behavior exceeds a predefined threshold. Another possibility is to model the ‘‘abnormal’’ behavior of the system and to raise an alarm when the difference between the observed behavior and the expected one falls below a given limit.

Tan et al. [3] proposed a system which applies the idea of Multivariate Correlation Analysis (MCA) to network traffic characterization and employs the principal of anomaly-based detection in attack recognition. This makes the solution capable of detecting known and unknown DoS attacks effectively by learning the patterns of legitimate network traffic only. Furthermore, a triangle area technique is proposed to enhance and speed up the process of MCA. Traffics are monitored at destination. Anomaly based detectors, sample by sample detection, multivariate correlation based method along with Triangle Area Map generation are used to recognize the malicious users.

Security community does not have effective and efficient trace back methods to locate attackers as it is easy for attackers to disguise themselves by taking advantages of the vulnerabilities of the World Wide Web, such as the dynamic, stateless, and anonymous nature of the Internet. The memory less feature of the Internet routing mechanisms makes it extremely hard to trace back to the source of these attacks. As a result, there is no effective and efficient method to deal with this issue so far. Number distributions of packet flows, which will be out of control of attackers once the attack is launched, and found that the similarity of attack flows are much higher than the similarity among legitimate flows. An approach, based on *ICMP messaging* [4], is to have each router X decide, with some probability q (typically = 1/20000 is mentioned), for each packet P to send an additional ICMP packet to the destination, which identifies X and some content of P. The main idea of this approach is that during a DDOS, a sufficient amount of attack packets will trigger ICMP messages from the routers in the attack tree T so that the victim can identify the leaves of T from these messages. The main drawback of this approach is that it causes additional network traffic even when no DDOS is present. Also it is not efficient, for identifying the n leaf nodes in the attack tree T.

Some researchers have advocated a logging approach to the IP traceback problem. In a logging solution, we either ask routers to log the packets they process or we augment the data packets themselves to contain a full log of all the routers they have encountered on their way to their destinations. Stone [5] and Baba and Matsuda [6] advocate logging of packet information at the routers, and Snoeren *et al.* [7] propose the logging of message digests of packets at the routers. The drawback with these approaches is that they require additional storage at the routers.

Michael T. Goodrich [8] proposed IP traceback based on the probabilistic packet marking paradigm called *randomize-and-link*, uses large checksum *cords* to “link” message fragments in a way that is highly scalable, for the checksums serve both as associative addresses and data integrity verifiers.
The DPM method [9] requires all the internet routers to be updated for packet marking. The DPM mechanism poses an extra ordinary challenge on storage for packet logging for routers. DPM require update on the existing routing software which is extremely hard to achieve on the internet. The DPM tries to spare space of a packet with the packet’s initial router information. Therefore the receiver can identify the source location of the packets once it has sufficient information of the marks. The major problem of DPM is that it involves modification of the current routing software and it may require large amount of marks for packet reconstruction.

III. PROPOSED SYSTEM

Architecture for proposed denial of service (DOS) attack detection and tracebacking mechanism is shown in fig.1,fig 2. This system consists for four major steps. Sample by sample detection mechanism is involved in it.

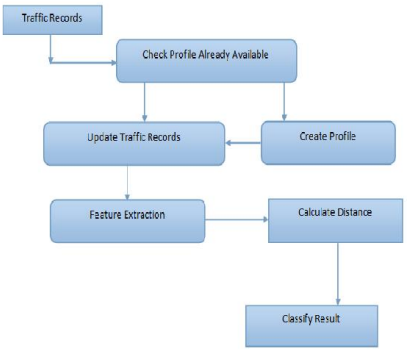

Figure 1.DoS Detection

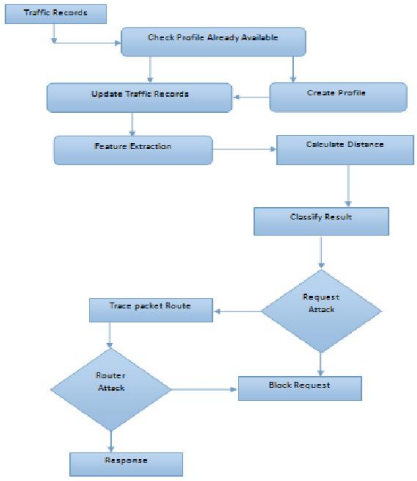

Figure 2.DoS Tracebacking concept

Step 1. Over-Provision Bandwidth to Absorb DDoS Bandwidth Peaks. This is one of the most common measures to alleviate DDoS attacks, but it is also probably the most expensive, especially since DDoS attacks can be ten times or even one hundred times greater than standard Internet traffic levels. An alternative to over-provisioning Internet bandwidth is to use a security service to scale on-demand to absorb and filter DDoS traffic. DDoS protection services are designed to stop massive DDoS attacks without burdening businesses' Internet connections.

Step 2. Monitor Application and Network Traffic The best way to detect when you are under an attack is by monitoring application and network traffic. Then, you can determine if poor application performance is due to service provider outages or a DDoS attack. Monitoring traffic also allows organizations to differentiate legitimate traffic from attacks. Ideally, security administrators should review traffic levels, application performance, anomalous behavior, protocol violations, and Web server error codes. Since DDoS attacks are almost always executed by botnets, application tools should be able to differentiate between standard user and bot traffic. Monitoring application and network traffic provide IT security administrator's instant visibility into DDoS attack status.

Step 3. Detect and Stop Malicious Users There is two primary methods to identify DDoS attack traffic: identify malicious users and identify malicious requests. For application DDoS traffic, often times identifying malicious users can be the most effective way to mitigate attacks.

1. Recognize known attack sources, such as malicious IP addresses that are actively attacking other sites, and identifying anonymous proxies and TOR networks. Known attack sources account for a large percentage of all DDoS attacks. Because malicious sources constantly change, organizations should have an up-to-date list of active attack sources.
2. Identify known bot agents; DDoS attacks are almost always performed by an automated client. Many of these client or bot agents have unique characteristics that differentiate them from regular Web browser agents. Tools that recognize bot agents can immediately stop many types of DDoS sources.
3. Perform validation tests to determine whether the Web visitor is a human or a bot. For example, if the visitor's browser can accept cookies, perform JavaScript calculations or understand HTTP redirects, then it is most likely a real browser and not a bot script.
4. Restrict access by geographic location. For some DDoS attacks, the majority of attack traffic may originate from one country or a specific region of the world. Blocking requests from undesirable countries can be a simple way to stop the vast majority of DDoS attack traffic.

Step 4. Detect and Stop Malicious Requests because application DDoS attacks mimic regular Web application traffic, they can be difficult to detect through typical network DDoS techniques. However, using a combination of application-level controls and anomaly detection, organizations can identify and stop malicious traffic. Measures include:

ISSN (P): 2349-3968, ISSN (O): 2349-3976
Volume 2 - Issue 5, May 2015

1. Detect an excessive number of requests from a single source or user session—automated attack sources almost always request Web pages more rapidly than standard users.
2. Prevent known network and application DDoS attacks—Many types of DDoS attacks rely on simple network techniques like fragmented packets, spoofing, or not completing TCP handshakes. More advanced attacks, typically application-level attacks, attempt to overwhelm server resources. These attacks can be detected through unusual user activity and known application attack signatures.
3. Distinguish the attributes, and the aftermath, of a malicious request. Some DDoS attacks can be detected through known attack patterns or signatures. In addition, the Web requests for many DDoS attacks do not conform to HTTP protocol standards. The Slowloris attack, for example, includes redundant HTTP headers. In addition, DDoS clients may request Web pages that do not exist. Attacks may also generate Web server errors or slow Web server response time

A. Algorithm for Normal profile generation based on triangle-area-map

Require: $X_{TAMlower}^{normal}$ with g elements

1. $\overline{TAM_{lower}^{normal}}$ $TAM_{lower}^{normal} \leftarrow 1/g \sum_{i=1}^{g} TAM_{lower}^{normal,i}$
2. Generate covariance matrix *Cov* for $X_{TAMlower}^{normal}$
3. for i=*1* to *g* do
4. $MD^{normal,i} \leftarrow$ MD $(TAM_{lower}^{normal\,i}, \overline{TAM_{lower}^{normal}})$
5. end for
6. $\mu \leftarrow \frac{1}{g}\sum_{i=1}^{g} MD^{\text{normal, i}}$
7. $\sigma \leftarrow \sqrt{\frac{1}{g-1}\sum_{i=1}^{g}(MD^{normal,i}-\mu)^2}$
8. *Pro*$\leftarrow$ (N (μ, σ$^{2)}$, . $\overline{TAM_{lower}^{normal}}$, *Cov*)
9. return *Pro*

Where $X_{TAMlower}^{normal}$ generated lower triangles of the TAMs of the set of *g* legitimate training traffic records, $TAM_{lower}^{normal,i}$ are the legitimate training traffic records TAM_{lower}^{normal} is the expectation of g legitimate training traffic records, MD is the Mahalanobis distance which is adopted to measure the dissimilarity between traffic records and described through mean μ and the standard deviation σ , *Pro* is the normal profile generated which contains the obtained distribution $N(\mu, \sigma^{2})$ of the normal training traffic records, $\overline{TAM_{lower}^{normal}}$ and *Cov*.

B. Algorithm for attack detection based on Mahalanobis distance

Require: Observed traffic record $x^{observed}$,normal profile *Pro* :(N (μ, σ^2), $\overline{TAM_{lower}^{normal}}$, *Cov*) and parameter α

1. Generate $TAM_{lower}^{observed}$ for the observed traffic record $x^{observed}$
2. $MD^{observed} \leftarrow$ MD $(TAM_{lower}^{observed}, \overline{TAM_{lower}^{normal}})$
3. if $(\mu - \sigma * \alpha) \leq MD^{observed} \leq (\mu + \sigma * \alpha)$ then
4. return Normal
5. else
6. return Attack
7. end if

A specific Threshold which is $(\mu + \sigma * \alpha)$ is used to differentiate attack from the legitimate one. For a normal distribution α is usually ranged from 1 to 3.If the MD between an observed traffic records $x^{observed}$ and the respective normal profile is greater than the threshold, it will be considered as an attack.

C. PIT: TRACKING BASED ON PATH BACKSCATTER

The below algorithm helps us to trace the source of attack.

1. Function GetSuspectSet_LoopFree(G,r,od)
2. $SuspectSet \leftarrow \phi$
3. $c \leftarrow null$
4. P shortest path from r to od
5. For vertex v in P do
6. If $v == r$ then
7. Continue
8. End if
9. $G'' \leftarrow G.remove(v)$
10. If r and od are disconnected in G'' then
11. $c \leftarrow v$
12. break
13. end if
14. end for

15. $SG \leftarrow G.remove(c)$
16. For Vertex v in SG do
17. If v and r are connected in SG then
18. $SuspectSet \leftarrow SuspectSet + v$
19. End if
20. End for
21. Return $SuspectSet$
22. End function

It is possible to get the topology of the network in some traceback scenarios. For example, the router-level topology can be got from trace route, and the AS(Autonomous System)-level topology can be inferred from the BGP data and supplementary means. Besides, a number of ASes make public their topologies. However, the routes of a network are always treated as business secret and are non-public. Above algorithm is based on Loop-Free Assumption [9]. This assumption states there is no loop in the paths. This assumption always holds unless misconfiguration or the routing has not converged.

IV CONCLUSION

This paper provides an effective method to detect DoS attack based on Multivariate Correlation analysis along with proper tracebacking with Passive IP tracebacking algorithm to find the source of attack. This system is able to distinguish both known and unknown DoS attacks from legitimate network traffic. The proposed IP traceback scheme based on information metrics can effectively trace all attacks until their own LANs (zombies).

REFERENCES

[1] Zhiyuan Tan, Aruna Jamdagni, Xiangjian He, Priyadarsi Nanda, Member, and Ren Ping Liu, "A System for Denial of Service Attack Detection based on Multivariate CorrelationAnalysis"*IEEE transactions on parallel and distributed systems*,vol.25, no.2, 2014

[2] P. Garca-Teodoro, J. Daz-Verdejo, G. Maci-Fernndez,and E.Vzquez, "Anomaly-Based Network Intrusion Detection: Techniques,Systems and Challenges," Computers and Security, vol. 28, pp. 18-28, 2009.

[3] Z. Tan, A. Jamdagni, X. He, P. Nanda, and R.P. Liu, "Triangle-Area-Based Multivariate Correlation Analysis for Effective Denial-of- Service Attack Detection," Proc. IEEE 11th Int'l Conf. Trust,Security and Privacy in Computing and Comm., pp. 33-40, 2012.

[4] S. M. Bellovin. ICMP traceback messages. In *Work in Progress, Internet Draft draft-bellovin-itrace-00.txt*, 2000

[5] R. Stone. Centertrack: An IP overlay network for tracking DoS floods.In *Proc.of 9th USENIX Security Symposium*,August 2000.Michael

[6] T. Baba and S. Matsuda. Tracing network attacks to their sources. *IEEEInternet Computing*, 6(2):20–26, 2002.

[7] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio,S. T. Kent, and W. T. Strayer. Hash-based IP traceback. In *Proc. Of the ACM SIGCOMM 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, 2001

[8] Michael T. Goodrich: Probabilistic Packet Marking for Large- Scale IP Traceback, IEEE/ACM transactions on networking,vol.

[9] Guang Yao, Jun Bi x, no.x,2007, Passive IP Traceback: Disclosing the Locations of IP Spoofers From Path Backscatter IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 10,No3.March 2015