ISSN 2395-695X (Print)
ISSN 2395-695X (Online)

International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)

Vol. 2, Special Issue 10, March 2016

EFFICIENT PACKET MAKING FOR LARGE – SCALE IP TRACE BACK

Thirumoorthy.M,
PG Student,
Priyadarshini Engineering College, Vaniyambadi, Vellore-635751.
m.thiru86@gmail.com
Mala.V,
Assistant Professor,
Priyadarshini Engineering College, Vaniyambadi, Vellore-635751.
malarajinikanth10@gmail.com

Abstract *- As a network security systems refers to the long known attackers may use forged source IP address to conceal their real locations. To capture the spoofers, a number of IP traceback mechanisms have been proposed. However, due to the challenges of deployment, there has been not a widely adopted IP traceback solution, at least at the Internet level. As a result, the mist on the locations of spoofers has never been dissipated till now. 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. In this way, PIT can find the spoofers without any deployment requirement. This paper illustrates the causes, collection, and the statistical results on path backscatter, demonstrates the processes and effectiveness of PIT, and shows the captured locations of spoofers through applying PIT on the path backscatter data set. These results can help further reveal IP spoofing, which has been studied for long but never well understood. Though PIT cannot work in all the spoofing attacks, it may be the most useful mechanism to trace spoofers before an Internet-level traceback system has been deployed in real.*

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 Vulnerable hosts are usually those that are either running no antivirus software or out-of-date antivirus software, or those that have not been properly patched.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 fortracebacking is also necessary.

II. LITERATURE SURVEY

DDoS attack detection metrics are mainly separated into two categories: the signature-based metric and anomaly-based metric. The signaturebased 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.

III. EXISTING SYSTEM

In the Existing system IP trackback approaches can be classified into five main categories: packet marking, ICMP trackback, logging on the router, link testing, overlay, and hybrid tracing.

Existing trace back mechanisms are either not widely supported by current commodity routers or will introduce considerable overhead to the routers generation especially in high-performance networks.

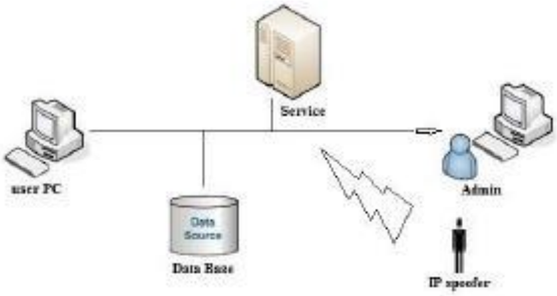


Figure 3.1 Existing System Architecture

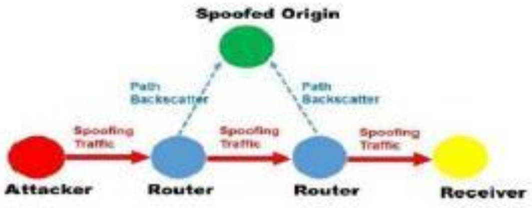
Internet Control Message Protocol (ICMP):

We also implemented a ICMP trace back generates addition ICMP messages to a collector or the destination. The ICMP messages can be used to reconstruct the attacking path. For example, if iTrace is enabled, routers generate ICMP samples to destinations with certain probability. The shortcoming of ICMP trace back is considerable additional traffic will be generated to consume the already stressed bandwidth resource.

Moreover, when the attack is against the bandwidth of the victim, the increased traffic will make the attack more serious. ICMP generation can be performed by the processor, but significant overhead will be introduced to the processor.

IV. PROPOSED SYSTEM

In the proposed system are users and applications passive IP trace back (PIT) that bypasses the deployment difficulties of IP trace back 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. we proposed Passive IP Trace back (PIT) which tracks spoofers based on path backscatter messages and public available information.



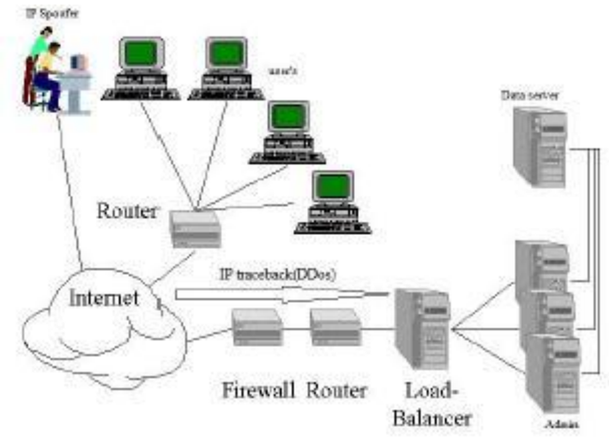


Figure 3.2: Proposed System Architecture

Passive IP Trace back (PIT)

PIT is used to perform IP trace back; it is very different from existing IP trace back mechanisms. PIT is inspired by a number of IP spoofing observation activities. Thus, the related work is composed by two parts. The first briefly introduces existing IP trace back mechanisms, and the second introduces the IP spoofing observation activities.

V. PROPOSED MATHEMATICAL MODEL:

Let S is the Whole System Consists: S= {V, E, P, G}. Where,

1. V is the set of all the network nodes.

2. E is the set of all the links between the nodes in the network.

3. P is path function which defines the path between the two nodes.

4. Let G is a graph.

Suppose, G (V, E) from each path backscatter, the node u, which generates the packet and the original destination v, Where u and v are two nodes in the network. i.e. u∈V and v ∈ V of the spoofing packet can be got. We denote the location of the spoofer, i.e., the nearest router or the origin by s, Where, s€V.

Procedure:

1. For each path backscatter message, at first we check whether it belongs to the classes i.e. dataset or source list. If yes, the reflector should be near the attacker.
2. We simply use the source AS of the message as the location of the spoofer. If the message does not belong to the types, it is mapped into an AS tuple.

3. We determine whether the AS tuple can accurately locate the source AS of the attacker based on our proposed mechanisms. Then if the AS tuple can accurately locate the source AS of the message, the source AS of the spoofer is just this AS.
4. Then we also use the source AS as the location of the spoofer.
We assume some Probability for Accurate Locating on Loop-Free for spoofer based on the Loop-free assumption, to accurately locate the attacker from a path backscatter message (v, s).
There are three conditions:
1) LF-C1: the degree of the attacker sis 1;
2) LF-C2: v is not s;
3) LF-C3: u is s.

Based on the Assumption I, the probability of LF − C1 is equal to the ratio of the network nodes whose degree is 1. To estimate our assumptions of probability, we introduce the power law of degree distribution from,

$$f_d \propto d^{O}$$

Where f_d is the frequency of degree d, and O is the out degree exponent. Transform it to

$$f_d = \lambda d^{O} + b_d$$

Where λ and b_d are two constants. Then,

$$f_1 = \lambda + b_d.$$

Based on the Assumption II, the probability of LF − C2 is simply (N − 1)/N.

Based on the Assumption III, the probability of LF −C3 is equal to 1/(1+len(path(u, v)).

Because s and u are random chosen, the expectation of len(path (u,v)) is the effective diameter of the network δef.
i.e. δef=1+len(path((u,v)).

Based on our three assumptions, these conditions are mutually independent. Thus, the expectation of the probability of accurate locating the attacker is

$$E(P_{LF-accurate}) = \frac{N-1}{N} * \frac{\lambda + b_d}{1 + \delta_{ef}}$$

This form gives some insight on the probability of accurate locating of spoofer. If the power-law becomes stronger, λ will get larger and δefwill get smaller. Then the probability of accurate locating will be larger.

VI. CONCLUSION AND FUTURE WORK

In this project we have presented a new technique, “backscatter analysis,” for estimating denial-of-service attack activity in the Internet. Using this technique, we have observed widespread DoS attacks in the Internet, distributed among many different domains and ISPs. The size and length of the attacks we observe are heavytailed, with a small number of long attacks constituting a significant fraction of the overall attack volume. Moreover, we see a surprising number of attacks directed at a few foreign countries, at home machines, and towards particular Internet services. We try to dissipate the mist on the locations of spoofers based on investigating the path backscatter messages. In this, we proposed Passive IP Traceback (PIT) which tracks spoofers based on path backscatter messages and public available information. We illustrate causes, collection, and statistical results on path backscatter. We specified how to apply PIT when the topology and routing are both known, or the routing is unknown, or neither of them are known. We presented two effective algorithms to apply PIT in large scale networks and proofed their correctness. We proved that, the effectiveness of PIT based on deduction and simulation. We showed the captured locations of spoofers through applying PIT on the path backscatter dataset.

REFERENCES

[1]. S. M. Bellovin, “Security problems in the TCP/IP protocol suite,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 19, no. 2, pp. 32–48, Apr. 1989.

[2]. ICANN Security and Stability Advisory Committee, “Distributed denial of service (DDOS) attacks,” SSAC, Tech. Rep. SSAC Advisory SAC008, Mar. 2006.

[3]. C. Labovitz, “Bots, DDoS and ground truth,” presented at the 50th NANOG, Oct. 2010.

[4]. *The UCSD Network Telescope*. [Online]. Available: http://www.caida.org/projects/network_telescope/

[5]. S. Savage, D. Wetherall, A. Karlin, and T. Anderson, “Practical network support for IP traceback,” in *Proc. Conf. Appl., Technol., Archit., Protocols Comput. Commun. (SIGCOMM)*, 2000, pp. 295–306.

[6]. S. Bellovin. *ICMP Traceback Messages*. [Online]. Available: http://tools.ietf.org/html/draft-ietf-itrace-04, accessed Feb. 2003.

[7]. M. T. Goodrich, “Efficient packet marking for large-scale IP traceback,” in *Proc. 9th ACM Conf. Comput. Commun. Secur. (CCS)*, 2002, pp. 117–126.

[8]. D. X. Song and A. Perrig, “Advanced and authenticated marking schemes for IP traceback,” in *Proc. IEEE 20th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, vol. 2. Apr. 2001, pp. 878–886.

[9]. A. Yaar, A. Perrig, and D. Song, “FIT: Fast internet traceback,” in *Proc. IEEE 24th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, vol. 2. Mar. 2005, pp. 1395–1406.

[10]. M. Adler, “Trade-offs in probabilistic packet marking for IP traceback,” *J. ACM*, vol. 52, no. 2, pp. 217–244, Mar. 2005.

[11]. A. Belenky and N. Ansari, “IP traceback with deterministic packet marking,” *IEEE Commun. Lett.*, vol. 7, no. 4, pp. 162–164, Apr. 2003.

[12]. Y. Xiang, W. Zhou, and M. Guo, “Flexible deterministic packet marking: An IP traceback system to find the real source of attacks,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no. 4, pp. 567–580, Apr. 2009.

[13]. S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, “The internet topology zoo,” *IEEE J. Sel. Areas Commun.*, vol. 29, no. 9, pp. 1765–1775, Oct. 2011.

[14]. L. Gao, “On inferring autonomous system relationships in the internet,” *IEEE/ACM Trans. Netw.*, vol. 9, no. 6, pp. 733–745, Dec. 2001.

[15]. X. Dimitropoulos*et al.*, “AS relationships: Inference and validation,” *ACM SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 1, pp. 29–40, Jan. 2007.