2009

DDoS defense using TCP_IP header analysis and proactive tests

Zhen Ye
Hefei University of Technology, yezhen1952@yahoo.com.cn

Weiwei Shi
Hefei University of Technology, hfut_shiww@yahoo.com.cn

Dayong Ye
University of Wollongong, dayong@uow.edu.au

DDoS defense using TCP_IP header analysis and proactive tests

Abstract

To defend against distributed denial of service (DDoS) attacks, one critical issue is to effectively isolate the attack traffic from the normal ones. A novel DDoS defense scheme based on TCP_IP Header Analysis and Proactive Tests (THAPT) is hereby proposed. Unlike most of the previous DDoS defense schemes that are passive in nature, the proposal uses proactive tests to identify and isolate the malicious traffic. Simulation results validate the effectiveness of our proposed scheme.

Disciplines

Physical Sciences and Mathematics

Publication Details

Z. Ye, W. Shi & D. Ye, "DDoS defense using TCP_IP header analysis and proactive tests," in International Conference on Information Technology and Computer Science, 2009. ITCS 2009, 2009, pp. 548-552.

DDoS Defense Using TCP_IP Header Analysis and Proactive Tests

Zhen YE/ Weiwei SHI
School of Computer and Information, Hefei University of Technology
Hefei, Anhui, P.R.China, 230009
e-mail: yezhen1952/ hfut_shiww @yahoo.com.cn

Dayong YE
School of Computer Science and Software Engineering
University of Wollongong
NSW 2522 AU
e-mail: dy721@uow.edu.au

***Abstract*—To defend against distributed denial of service (DDoS) attacks, one critical issue is to effectively isolate the attack traffic from the normal ones. A novel DDoS defense scheme based on TCP_IP Header Analysis and Proactive Tests (THAPT) is hereby proposed. Unlike most of the previous DDoS defense schemes that are passive in nature, the proposal uses proactive tests to identify and isolate the malicious traffic. Simulation results validate the effectiveness of our proposed scheme.**

Keywords- DDoS defense, TCP_IP header, proactive test

I. INTRODUCTION

Generally, there are four broad categories of defense against DoS attacks [1]: attack prevention, attack detection, attack source identification, attack reaction. Next we will give a mainly introduction of each defense technique and simply analyze their advantages and limitations one by one.

A. Prevention

Attack prevention aims to stop attacks before they can reach their target. This approach assumes the source address of attack traffic is spoofed, so it normally comprises a variety of packet filtering schemes which are deployed in routers to make sure only valid traffic can pass through. This greatly reduces the chance of DDoS attacks occurring. However, it is not easy to specify a filtering rule that can differentiate spoofed traffic from legitimate traffic accurately. Router-based Packet Filtering (RPF) [2] extends ingress filtering to the core of the Internet. It is based on the principle that for each link in the core of the Internet, there is only a limited set of source addresses from which traffic on the link could have originated. RPF is effective against randomly spoofed DoS attacks. However, the filtering granularity of RPF is low. The aim of the Source Address Validity Enforcement Protocol [3] is to provide routers with information about the range of source IP addresses that should be expected at each interface. It overcomes the asymmetries of Internet routing by updating the incoming tables on each router periodically, but it needs to change the routing protocol.

B. Detection

The detection can be divided into two categories according to the ways used in detection. The first category is based on the special features of DoS attacks. MULTOPS [4] assumes that packet rates between two hosts are proportional during normal operation. SYN detection [5] methods detect DoS attacks by monitoring statistical changes. Cheng et al. [6] proposed to use *spectral analysis* to identify DoS attack flows using the number of packet arrivals in a fixed interval as the signal. The assumption of the Kolmogorov test [7] is that multiple attack sources use the same DoS attack tool, resulting the traffic is highly correlated. All above detection techniques are based on one or more assumptions, of which most are not strong. So attackers can evade detection by changing their attack patterns. Another category is anomaly-based detection, which models the behavior of normal traffic, and then reports any anomalies**.** Lightweight Intrusion detection System [8] using the idea of an Artificial Immune System.

C. Source Identification

Source identification is necessary, as we known that blocking the attack traffic at its source is an ideal response when DoS attack is undergoing. There are three kinds of Source Identification. The first kind is IP Traceback by Active Interaction. Link-testing traceback [9] infers the attack path by flooding all links with large bursts of traffic and observing how this perturbs the attack traffic. The common shortcoming for these schemes is only suitable for identifying attack paths within one ISP's network. The second kind is Probabilistic IP Traceback Schemes. Song and Perrig [10] have improved the efficiency and security of the PPM scheme by introducing a new hashing scheme and an authentication scheme. In ICMP *"*traceback*"* scheme [11], when a router receives a packet to a destination, the router generates an ICMP traceback message with low probability. Adjusted Probabilistic Packet Marking [12] was proposed to overcome a problem that is the further the router the less possible it is to receive a marked packet from that router. This kind of source identification is Vulnerable to marking spoofing. The third kind is Hash-Based IP Traceback. Snoeren et al. [13] proposed *hash-based IP traceback,* to trace individual packets by recording every packet at the router which it passed through. However, the success of traceback depends on the number of tracking routers installed, and the area covered by these routers.

978-0-7695-3688-0/09 $25.00 © 2009 IEEE
DOI 10.1109/ITCS.2009.248

D. Reaction

According to the distance to the victim, we divide the location that could be employed the reaction scheme into three types: Victim end reaction, Intermediate network reaction, and Source end reaction. History-based IP Filtering [14] proposed to filter bandwidth attack traffic according to the history. The challenge of victim end reaction is how to differentiate attack traffic from legitimate. There are three types of intermediate network reaction schemes: pushback and controller-agent schemes and secure overlay services. Mahajan et al. [15] provided a scheme in which routers learn a congestion signature that can differentiate legitimate traffic from malicious traffic based on the volume of traffic to the target from different links; the aim of an agent-controller model [16] is to filter attack traffic at the edge routers of one ISP domain. An architecture called secure overlay service [17] was proposed to secure the communication between the confirmed users and the victim. The basic assumption for above schemes is that there are a limited number of attack paths, so how to deal with distributed non-spoofed attacks becomes a challenge. The ultimate goal for DoS attack defense is to filter attack traffic at the source. D-WARD [18] collects flow statistics to defend DoS attack at the source. But how to detect an attack before attack traffic aggregation is a technical challenge to source end reaction.

II. Related Work

Considering the following situation, DDoS attacks are launched from a large army of compromised hosts, most attacks do not spoof source address and each host can behave like a "legitimate" source, attack traffic comes from many geographically distributed links, but the overall effect is a powerful DDoS attack. This attack can evade most of the existing detection mechanism and threaten the victim directly, so we can just employ the reaction scheme to effectively minimize attack damage at the Victim end. Most commercial DoS attack solutions belong to Victim end reaction mechanisms which are easy to implement by deploying at the target or routers close to the target. To ensure good performance and accommodate as many normal users as possible, it is critical to differentiate traffic and then treat them differently. Unfortunately, it is rather difficult to give an accurate classification as DoS attack traffic can mimic any type of legitimate traffic. Discrimination based on packet headers is vulnerable to IP spoofing; Discrimination based on packet contents may be thwarted by the increasing use of end-to-end encryption. Common types of DDoS attacks are listed as follows: TCP SYN Flood; UDP Flood; Ping of Death; Smurf; Teardrop and Land Attacks. We differentiate these attack into Connection oriented and Connectionless oriented, thus we treat them differently using different means. In this paper, a novel DDoS defense scheme based on TCP_IP Header Analysis and Proactive Tests is hereby proposed. Analyzing the TCP_IP Header against the well defined rules is designed to defend the Connectionless oriented attack, such as UDP flood attack. While the proactive test based differentiation technique was proposed to handle Connection oriented (TCP) attack solely because TCP protocol has the built-in congestion control and reliable transmission mechanism that we can use to test every TCP flow to distinguish whether it is a legitimate flow. We identify malicious traffic from their behaviors, such as aggressiveness that is the salient feature of DDoS traffic [19].

The rest of the paper is structured as follows. Section Ⅲ presents in detail our proposal. Then we present in Section Ⅳ some simulation results to validate our scheme. Finally, conclusion is presented in Section Ⅴ.

III. Design of THAPT

The THAPT model's flowchart in Fig. 1, which can be explained as follows:

1) Upon the arrival of a new incoming packet, the receiver first determines which path to go. If it is TCP flow, go to 6, otherwise go to 2.
2) Consider whether the analyzed period is within the considering period (Δt) or not, if not, the packet is admitted, otherwise go to 3.
3) If packet within Δt , calculate the value of VDR.
4) If V>v or D>d or R>r, the packet should be further examined by the packet rules, otherwise admit the packet.
5) Consider whether packet is matched with designed packet rules or not, if yes, the packet should be dropped and then the packet rules and Δt will be update, otherwise the packet is admitted.
6) The receiver determines whether the current packet belongs to a new flow or not by checking the tuple of (source IP address, source port number, destination IP address, destination port number). If it is, go to 7, otherwise go to 9.
7) If it is the first packet of a flow, the receiver examines whether the number of admitted flows is less than the maximum flow count (a threshold set by the receiver to ensure proper provisioning of quality of service) or not , If not , the packet is dropped.
8) Otherwise, the new packet is admitted after updating the flow table maintained by the receiver, resulting in an increment of the flow count by 1, and initialization of several counters, such as the number of successful tests and the number of failure tests.
9) The receiver then checks the behavior history of the flow. If the number of failure tests is no less than a threshold, f, the packet will be dropped.
10) For the flow whose behavior is not so bad in the past, our scheme further examines whether the flow has passed a certain number of tests, h. The receiver will admit directly any packet of flows having passed h tests successfully.
11) For other flows whose pass_num is less than h, we further check the current state of the flow. If the flow is under a test, go to 12, otherwise go to 13.
12) Now the proactive test will be enforced by the receiver by manipulating the reverse ACK rate. The aim of the test is to make the current rate shall not exceed one half

of its previous one. If the flow conforms to that constraint, the flow passes the current test and its pass_num is incremented by 1. Otherwise, the flow fails one test, resulting in an increment of the fail_num by 1 and the possibility of dropping.

13) In the case that the flow is not in the state of being tested, the rate of the flow is compared with that of the fair share of each flow. The result of the comparison is used to determine the test probability for that flow. If the flow needs more tests, then go to 12. Or it will be admit because its sending rate is low.

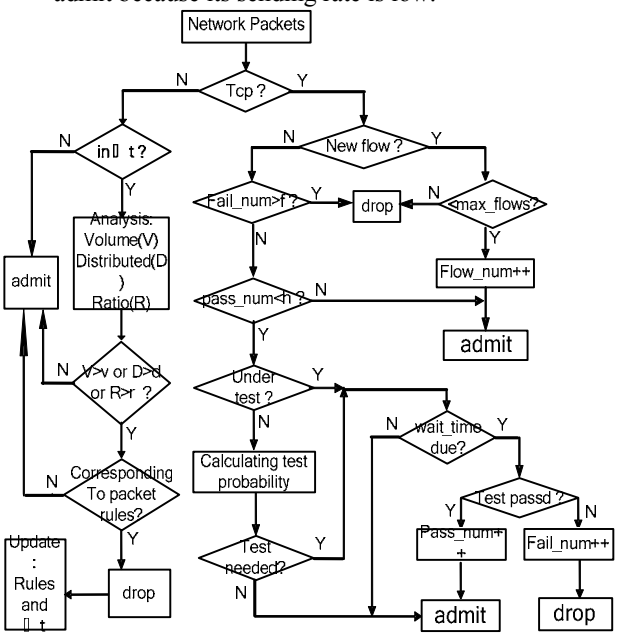

Fig. 1. Flowchart of the traffic differentiation

A. TCP_IP Header Analysis.

- Update the Δt

Δt is the defined time period, it is used to decide how long the each analysis will go on. The initialization value is 0.1 second, so most the packets will be admitted when there is no attack. Once attack happens, some packets will be dropped, resulting in an increment of Δt and more packets will be tested by the well defined rules.

- VDR analysis

VDR Calculating occurs per Δt .

1) Volume Measurement Analysis equation is

$$Volume = \left(\frac{Total \quad number \quad of \quad packet}{\Delta t} \right).$$

2) Distributed Measurement Analysis equation is

$$Distributed = \left(\frac{Total \quad number \quad of \quad distinct \quad packet}{\Delta t} \right).$$

3) Ratio Measurement Analysis equation is

$$Ratio = \left(\frac{Total \quad number \quad of \quad packet \quad incoming}{Total \quad number \quad of \quad packet \quad outcoming} \right).$$

We select VDR analysis on the traffic because they have different strengths and limitations: Volume analysis is suitable for the attacks that have constant attack rate, which is easy to detect. However, its abilities will be reduced in the face of variable rate attack and it is difficult to distinguish the difference between Flash Crowds and real DDoS traffic by this method; Distributed analysis function well when it is used in DoS attributed attack such as Smurf Attack or ICMP Flood. Attack commonly combines between IP Spoofing and Reflectors for the purpose of the best attributed of source IP address distribution. Distribution of source IP address during the attack is easy to recognize, but it is difficult to measure volume of the attack. If the attack concerns about the quantity, not distribution, such as TCP SYN Flood, we cannot catch such attack; Ratio analysis plays well its role for the attack that creates high difference of attack ratio such as TCP SYN Flood and Smurf Attack. However, this way cannot catch the some attack such as Land Attack or Fragile Attack. Therefore, combining advantages of all three methods may increase the efficiency of detection.

The separate threshold of V, D and R is v, d and r. Their value should be defined base on the given environment and extensive simulations.

- Packet Rules

First, Let us analyze the primary attack signature of mentioned DDoS attacks:

1) The TCP SYN flood exploits a vulnerability of the TCP three-way handshake;
2) The Smurf attack uses ICMP echo request packets directed to IP broadcast addresses from remote locations to generate denial of service attacks;
3) ICMP and UDP Flooding sent out ICMP and UDP packets continually;
4) Land attack is caused by sending a packet to a machine with the source address as same as the destination address;
5) Teardrop exploits an overlapping IP fragment bug, making certain operating systems improperly reassemble the overlapping IP fragments with a result of rebooting the machine;
6) Ping-of-death attack sends a large ICMP ping packet that is exceed 65535 byte, which can cause certain operating systems to crash, freeze, or reboot due to buffer overflow;

Then, further discussion is followed: DoS attack can be launched in two forms [20]. The first form aims to crash a system by sending one or more carefully crafted packets that exploit software vulnerability in the target system. The second form is to use massive volumes of useless traffic to occupy all the resources that could service legitimate traffic. While it is possible to prevent the first form of attack by patching known vulnerabilities, the second form of attack cannot be so easily prevented. The targets can be attacked simply because they are connected to the public Internet. So

we just put emphasize on the UDP and ICMP flooding attack here.

Now, base on the analysis above, packet rules are made as followed:

1) IP 1 = {drop | IP.SRC_IP = IP.DST_IP}.

It is used to drop the packets whose source address is the same as the destination one.

2) IP2 = {drop | IP.FLAG_MF = 1}.

This can drop those packets have fragmented packet.

3) ICMP = {drop | ICMP.TYPE = 8}.

It is used to drop ICMP packets that have their type equal to 8 which is a echo request packet.

4) UDP = {drop | UDP.SRC_PORT = UDP.DST_PORT}.

It is used to drop UDP packet that the source port is equal to destination port.

5) IP3 = {drop | IP.SRC_IP = SIP}.

SIP is none initially, while it is selected dynamically to filtering the malicious packets by the analysis of VDR. When the value of Volume or Distributed or Ratio Analysis is high, the most related source address will be selected as SIP.

B. Proactive Tests

- The threshold f and h

It is used to prevent our scheme from falsely identifying the behavior of a flow. A low value of f may exacerbate packet dropping. In case of a false identification, subsequent packets from an innocent flow will be blocked. Selecting a too high value is unwise, either. A high f delays the packet dropping decision, and thus subsequent packets of a malicious stream may still consume system resources. Similarly, some tradeoff has to be made to determine a proper value of h.

- Calculate the rate of a flow

It is calculated according to the following formula, num_ pkt*sz_pkt/t, where t is the time interval (window), num_pkt is the number of packets received during this period, and sz_pkt is the packet size. We update the starting time of a flow once it passes a test. In so doing, we can effectively thwart a low-rate DoS attack which sends a burst of attack packets to incite congestion and keeps silence for a much longer period to significantly lower its average rate in order to escape detection and filtering.

- Calculate the test probability

It is fair that a flow with less bandwidth consumption should be tested by less numbers. The test probability for a high-rate flow is 1/ (pass_num+1). At the very beginning, pass_num is 0 for all flows. Therefore, as long as a high-rate flow has not passed a test, its chance of being tested is 100%. As the number of successful tests of a flow increases, its test probability reduces. The test probability for the less resource-consumption flow is 1/max (m, 2*h), where m is the total number of flows. For the normal case, m is far greater than 2h; thus, p=1/m. We use the max() function to address the case that only a few flows exist in the system and ensure that the test probability for a low-rate flow is at most 1/2 of that of a high-rate one.

IV. SIMULATIONS

To test the effectiveness of our proposed traffic differentiation, we set up a simulation scenario including two sources: Normal source and Attack source which contains TCP Syn, UDP and ICMP flood attack source as shown in Fig.2. These flows pass through the same bottleneck link. The difference is that one simulation uses a Normal sink to accept packets, and the other uses the THAPT sink. We set f to 3, h to 6 and increase the value of Δt by 1 second when it will be updated. The simulation results are shown in Fig. 3.

Fig.3 shows the throughput of the traffic using the Normal sink and the THAPT sink that we proposed, in which the traffic throughput of THAPT sink drops drastically after 2 minutes, then after some minutes, the traffic becomes stabilized. In contrast, using the Normal sink as the receiver, the attacker may keep the highest throughput during its lifetime. The result demonstrates the effectiveness of our proposed traffic differentiation.

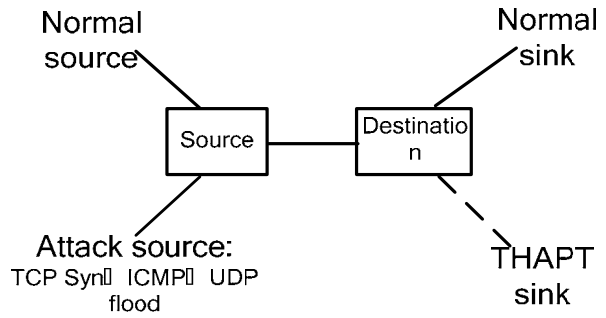

Fig.2. Simulation for comparison study of the effectiveness of THAPT

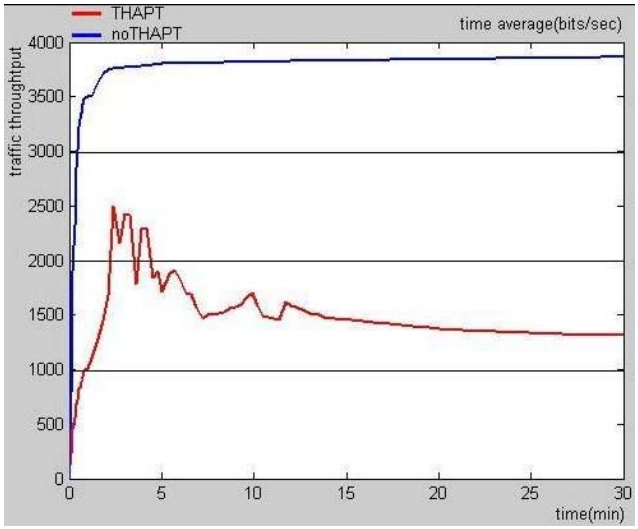

Fig. 3. Traffic throughput of Normol and THAPT Sink

V. CONCLUSION

In this paper, we first analyzed four categories of defense against DDoS attacks. After that, a novel DDoS defense scheme has been presented. The salient benefits of our proposal mainly lie in its capability of defending malicious flows. Preliminary simulation results have validated our design.

REFERENCES

[1] Peng, T., Leckie, C., and Ramamohanarao, K. 2007. Survey of network-based defense mechanisms countering the DoS and DDoS problems. ACM Comput. Surv. 39, 1, Article 3 (April 2007).

[2] PARK, K. AND LEE, H. 2001a. On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack. In Proceedings of IEEE INFOCOM 2001. 338–347.

[3] LI, J., MIRKOVIC, J., WANG, M., REITHER, P., AND ZHANG, L. 2002. Save: Source address validity enforcement protocol. In Proceedings of IEEE INFOCOM 2002. 1557–1566.

[4] GIL, T. M. AND POLETTO, M. 2001. MULTOPS: A data-structure for bandwidth attack detection. In Proceedings of the 10th USENIX Security Symposium.

[5] WANG, H., ZHANG, D., AND SHIN, K. G. 2002. Detecting SYN flooding attacks. In Proceedings of IEEE INFOCOM 2002. 1530–1539.

[6] CHENG, C.-M., KUNG, H. T., AND TAN, K.-S. 2002. Use of spectral analysis in defense against DoS attacks. In Proceedings of IEEE GLOBECOM 2002. 2143–2148.

[7] KULKARNI, A., BUSH, S., AND EVANS, S. 2001. Detecting distributed denial-of-service attacks using Kolmogorov complexity metrics. Tech. rep. 2001CRD176. GE Research & Development Center. Schectades, NY.

[8] FORREST, S. AND HOFMEYR, S. 1999. Architecture for an artificial immune system. Evolution. Computat. J. 7, 1, 45–68.

[9] BURCH, H. AND CHESWICK, B. 2000. Tracing anonymous packets to their approximate source. In Proceedings of the 14th Systems Administration Conference (New Orleans, LA).

[10] SONG, D. X. AND PERRIG, A. 2001. Advanced and authenticated marking schemes for IP traceback. In Proceedings of IEEE INFOCOM 2001. 878–886.

[11] BELLOVIN, S. 2000. The ICMP traceback message. IETF Internet Draft. Internet Engineering Task Force (IETF). Go online to www.ietf.org

[12] PENG, T., LECKIE, C., AND RAMAMOHANARAO, K. 2002a. Adjusted probabilistic packet marking for IP traceback. In Proceedings of the Second IFIP Networking Conference (Networking 2002). (Pisa, Italy). 697–708.

[13] SNOEREN, A. C., PARTRIDGE, C., SANCHEZ, L. A., JONES, C. E., TCHAKOUNTIO, F.,KENT, S. T., AND STRAYER,W. T. 2001. Hash-based IP traceback. In Proceedings of the 2001 ACMSIGCOMM Conference (San Diego, CA). 3–14.

[14] KARGL, F., MAIER, J., ANDWEBER, M. 2001. Protecting web servers from distributed denial of service attacks. In Proceedings of the 10th International World Wide Web Conference. 130–143.

[15] MAHAJAN, R., BELLOVIN, S. M., FLOYD, S., IOANNIDIS, J., PAXSON, V., AND SHENKER, S. 2002. Controlling high bandwidth aggregates in the network. ACM Comput. Commun. Rev. 32, 3 (Jul.), 62–73.

[16] TUPAKULA, U. AND VARADHARAJAN, V. 2003. A practical method to counteract denial of service attacks. In Proceedings of the Twenty-Sixth Australasian Computer Science Conference (ACSC2003) (Adelaide, Australia). 275–284.

[17] KEROMYTIS, A. D., MISRA, V., AND RUBENSTEIN, D. 2002. SOS: Secure overlay services. In Proceedings of the 2002 ACM SIGCOMM Conference. 61–72.

[18] LI, J., MIRKOVIC, J., WANG, M., REITHER, P., AND ZHANG, L. 2002. Save: Source address validity enforcement protocol. In Proceedings of IEEE INFOCOM 2002. 1557–1566.

[19] Gao, Z., AND Ansari, N. 2006. Differentiating Malicious DDoS Attack Traffic from Normal TCP Flows by Proactive Tests. IEEE COMMUNICATIONS LETTERS, VOL. 10, NO. 11.

[20] Limwiwatkul, L., AND Rungsawang, A. 2004. Distributed denial of service detection using TCP_IP header and traffic measurement analysis. International Symposium on Communications and Information Technologes 2004 (ISCIT 2004) Sapporo Japan. October 26- 29. 2004