An implementation of a hierarchical IP traceback architecture

Masafumi OE
masa@fumi.org

Youki KADOBAYASHI
youki-k@is.aist-nara.ac.jp

Suguru YAMAGUCHI
suguru@is.aist-nara.ac.jp

Graduate School of Information Science,
Nara Institute Science and Technology,
Ikoma, 630–0192, Japan.

Abstract

The IP traceback technique detects sources of attack nodes and the paths traversed by anonymous DDoS (Distributed Denial of Service) flows with spoofed source addresses. We propose a hierarchical IP traceback architecture, which decomposes the Internet-wide traceback procedure into inter-domain traceback and intradomain traceback. Our proposed method is different from existing approaches in that our method is independent from a single IP traceback mechanism, and domain decomposition is based on existing operational models of the Internet. Moreover, it has the capability of being used for not only the IPv4 network, but also the IPv6 network.

1. Introduction

Distributed Denial of Service (DDoS) attacks are one of the threats on the Internet. In DDoS attacks, the attack nodes are widely set up in the Internet, and transmit a large number of packets to the victim's node. These packets consume network resources and server resources, and obstruct network services like World-Wide-Web in the victim's node. In order to keep this damage to a minimum, the technology that identifies attack nodes and the paths of attack packets is required to be as fast as possible.

However, the source address of the IP packet for DDoS attack is spoofed to a random address. Therefore, investigating an attack node using TRACEROUTE, which depends on a source address, is not effective. For this reason, tracking the attack flow is done by hand using a filtering function, a DDoS attack detection function, (which each router has), etc. The Internet, as an international communications infrastructure, is constructed by various organizations connecting each other with various policies, such as ISPs, companies, research institutes, universities, etc. When tracking attack flows, a lot of time is wasted getting cooperation between organizations on attack paths. Tracking by hand is a big barrier to the reduction of time. IP traceback methods are proposed as solution to this issue. IP traceback detects the attack paths as shown in Fig. 1 and specifes the true origin of an attack flow with a spoofed source address.

In this paper, we propose a hierarchical IP traceback architecture, which decomposes Internet-wide traceback procedure into interdomain traceback and intradomain traceback. Our proposed method is different from existing approaches in that our method is independent from single IP traceback mechanisms, and domain decomposition is based on the existing operational model of the Internet.

The rest of this paper is organized as follows: In section 2, we describe related work. Section 3 outlines our proposed technique. In Section 4 we describe the implementation of our proposal. Finally, we summarize our findings and future work in Section 5.

2. Related work

We briefly introduce three typical proposed methods for IP traceback.

2.1. Link testing method

This method specifies the IP address of a router that forwards the attack flow by the filtering function and the monitoring function in the router. The attack path is clarified by repeating the search from the victim's node to the attacker

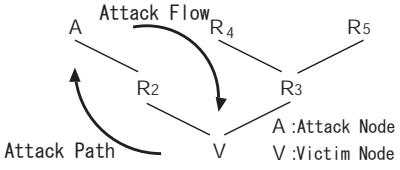

Figure 1. IP traceback : detecting attack flows

node on each router. Additional equipment and materials for the traceback are not required. However, this method can traceback only while the victim's node is under DDoS attack and the attack flow is active on the Internet.

2.2. Hash-based method

Snoeren et al [2] have proposed this method. Every router logs all transmission packets into storage with a hash function for compression of data. This method has a high capability for locating the attacking host. However, this method requires 0.1 % of interface bandwidth on each interface for recoding. Large amounts of cost and hardware infrastructure are required to store recoding data.

2.3. Passive detection method

This method was proposed by Bellovin[1]. Every router sends the router's information, that is, IP address, MAC address, next hop router's IP address and so on into the passive detection packet. The victim hosts can identify the attack path by collecting these packets. Installation cost is low because this method does not require the storage area of the packet and the processing of each packet. However, the passive detection packets increase traffic on the Internet.

2.4. IP traceback difficulty over the Internet

In this chapter, we describe the analytical result of the existing IP traceback methods and the reason it is difficult to operate IP traceback over the Internet. We believe two issues exist to implement IP traceback on the Internet,

1.Organizations connected with the Internet as Autonomous System(AS) does the management and operation independently for each AS. Therefore, emergency action for traceback requires the cooperation between two or more AS over a large amount of time.

2. We expect the attacker to analyze the weak points of current IP traceback methods and develop anti-IP traceback attacks as a countermeasure.

From the two points above, we know that it is difficult to use only one kind of IP traceback method for the Internet. This paper proposes the operational architecture of an IP traceback method that can solve these issues.

On designing the architecture, we modeled our method after the routing architecture in the Internet. The Internet is not operated with one routing protocol to manage the network routing tables. It is a hierarchy of EGP (Exterior Gateway Protocol) and IGP (Interior Gateway Protocol) according to the scale of the network. EGP is used for routing between ASes. In the EGP operation, the routing information has been exchanged through interconnections based on the agreement between ASes. Applying the mechanism of the above-mentioned path control to the traceback mechanism was attempted. This is "Hierarchical Architecture for IP traceback" is based on the idea of "Hierarchy" in the routing system.

3. Hierarchical architecture for IP traceback

The proposed architecture is made up of a construct with three components: eIP traceback, iIP traceback, and ITM network. Each control area of eIP and iIP traceback is the same as each control area of EGP and IGP, which are the routing control protocols shown in Fig. 2. ITM network is used for the association of eIP and iIP. We describe the purpose of each component as follows:

By definition, "Exterior IP (eIP) traceback architecture" designates each AS that the attack flow passed. eIP traceback should have the capability for finding attack paths within 30 minutes. It has been shown that the initial attacking stage ends within 30 minutes, according to the report of CAIDA[CAIDA]. However, in existing research, there is no IP traceback method known that can specify each AS that the attack flow passed. We propose the "IP Option Traceback" for eIP traceback and describe it in section[**?**].

"Interior IP(iIP) traceback architecture" designates the router's IP address that the attack flow has passed in the AS. We researched an existing IP traceback and defined 3 parameters that are needed to make relationships between eIP and iIP. These are the packet dumps of the attack flow, the timestamp, and the recording node(AS/IP). In this paper, we only discuss how to use modified IP traceback as iIP traceback. The technologies of each existing IP traceback are not discussed.

The "IP traceback Manager(ITM)" that exists on each AS exchanges information for the eIP/iIP traceback operation. ITMs are connected with each other and use ITM Protocol(ITMP).

The IP traceback process using this proposed technique is described as follows:

1. Victim's administrator requests a countermeasure to the attack flow to the administrator on a connected network (AS). The administrator of the AS does the monitoring and records the attack flow to the victim's node.
2. eIP traceback, in cooperation with ITM network, calculates the attack path of ASes that the attack flow passed by using this record.
3. ASes in the attack path executes the countermeasure: filtering and bandwidth shaping of the attack flow for easing immediate damage at the victim's site.
4. iIP traceback is executed on each AS in which the attack node has specified the true IP address of the attack node.
5. The DoS attack ends because the attack node is shut off from the network.

Each AS can select a different method of IP traceback because iIP traceback can be executed independently in

each AS. We can switch to another method when vulnerability of an IP traceback that is in operation has been discovered.

3.1. IP option traceback

We propose "IP option traceback(IP-OPT)" as eIP traceback that uses the IP option header of IPv4 and the destination option header of IPv6. We considered the influence of Internet traffic in the passive detection packet that used the IP option. We make a mathematical model and analyze our "IP Option traceback." For more details, please refer to [3][4][5]. In this paper, the outline of IP-OPT is described.

IP-OPT has two components, TOG(Traceback Option Generator) and Packet Monitor(PM). TOG generates the IP traceback option packet constructed with information to construct the attack path by the passive detection method. Also, TOG constructs the attack path from the tracing packets.

PM is set up on each BGP area border router on the AS, and has the following two functions:

Packet selection: The packet output from the interface on the router is selected at probability P, and the copy of the selected packet is recorded.

Cooperation with TOG: PM receives and uses probability P from TOG and sends the selected packet to TOG.

Next, the function of TOG is described. TOG is constructed with the following function.

1. Generation of the IP traceback option: TOG generates IP traceback option from the selected packet and the parameter (AS number and HASH information), and sends it.
2. The option management: TOG generates the "AS message key X" and "key identification number Z", and stores them.
3. The association of PM: TOG sends probability P to each PM and receives selected packets from each PM.
4. Verification of packet: TOG verifies the HMAC certified data and constructs of the attack path.
5. The association between ITMs: For construction of the attack path, TOG exchanges an attack path or IP options with neighboring TOGs via the ITM network. The TOG requests the execution of iIP traceback at each AS including attack nodes via the ITM network.

The traceback information that TOG generates from parameters is recorded in the "IP destination option header " for IPv6 and the " IP option header " for IPv4[7][6]. It has the following parameters:

HMAC tag number (16bit): An HMAC algorithm identifier used in HMAC (Keyed-Hash Message Authentication Code) [8].

Key identification number (64bit): This is the Key identification number Z to AS message key X that is used with HMAC.

MAC data (algorithm dependence and variable-length): HMAC authentication data H generated from message key X to AS number and AS.

When we construct attack paths, the victim records the IP options O from attack flows. Each AS's TOG verifies the AS's MAC data $HMAC_{AS}$, which is calculated from the AS message key X_n and identification number Z, and MAC data in O. We find out the attack path from the concentration of ASes that are verified O.

3.2. ITM and module API

This section describes ITM and ITMP that are used to exchange data between ITMs. ITM network is constructed with a peering connection of ITM to neighboring ITMs as well as the peering of EGP Fig. 3. Thus, the ITM network topology is the same as the AS the network topology.

ITMP has an authentication phase and a connected phase. The authentication phase authenticates ITM and exchanges neighbor information (AS number), supporting eIP/iIP traceback under neighbor ASes, and so on. After the authentication phase, comes the connected phase. In the connected phase, each eIP/iIP traceback can transmit data to each other.

In our implementation, eIP and iIP traceback systems under AS are connected with ITM through ITM API (Fig. 4).

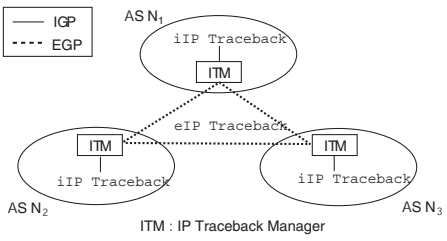

Figure 2. Hierarchical architecture for IP traceback

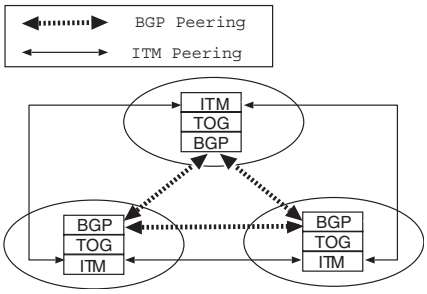

Figure 3. ITM network : interconnection between ITMs

We define ITM API from the analysis of the existing implementations, PAFFI as iIP traceback and IP-OPT as eIP traceback. PAFFI, developed by Yokogawa Electric is based on Hash-based IP traceback and uses NP(Network Processor). ITM API is the access to information of neighboring ITM (kind of the AS number, the connectivity, the state, and the supported IP traceback), the data transfers between modules of neighboring ITMs, and status reports of each module.

4. Feasibility of deployment

In the achievement of this proposal, ITM operation is required in core AS groups that CAIDA selects from the operation result of Skitter[9].

The authors believe that the feasibility of the ITM network is high. ISPs should do the countermeasure to Denial of Service attack as soon as possible because it results in the least amount of damage that the guest receives.

However, ISPs now use link inspection methods for IP traceback. The cost for this is high and the ISP is spending a lot of money, time and human resources, to stop attack flows. This proposal can reduce the cost because ITM automates the cooperation between ASes. Moreover, each AS can select a technique of IP traceback depending on the operation scale and budget of ISP. Therefore, the feasibility of the ITM network is high because there is the advantage that the AS operates ITM.

5. Conclusion and future work

We proposed a hierarchical IP traceback architecture and described the feasibility of this proposal on the Internet. The IP traceback technology is designed as one of the countermeasures to the DDoS attack. However, IP traceback requires the cooperation between two or more ASes over a long preiod of time. The attacker can then analyze the weak points of the contermeasure for developing an anti-IP traceback attack. We believe that it is difficult to use only one kind of IP traceback method for the Internet.

This proposal was split into eIP traceback and iIP traceback according to the relation between EGP and IGP in the routing control architecture. eIP traceback detects ASes that the attack flow has passed. iIP traceback detects the router's IP address passed by the attack flow. We described an implementation architecture for "IP option traceback" that was proposed as eIP traceback. We investigated existing IP traceback and described implementation architecture of the ITM network which was used to make an association between eIP and iIP traceback.

We plan to brush up our implementation and carry out experiments on StarBED[10] which has 500 physical nodes (5000 VM-simulated node) and is a fully programmable Internet simulator. We will then present the experimental results of this simulation.

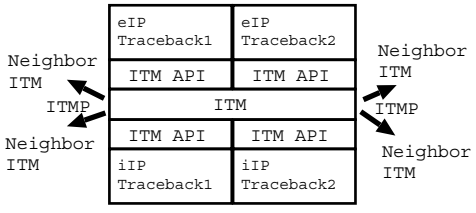

Figure 4. Structure of ITM

References

[1] S.M.Bellovin, M.D Leech, and T.Taylor, "ICMP traceback messages, "Internet-Draft, draft-ietf-itrace-01.txt, Oct.2001 ! %

[2] A.C.Snoeren, C.Partridge, L.A.Sanches, C.E.Jones, F.Tchakountio, S.T.Kent, and W.T.Stayer, "Hash based IP traceback," in Proceedings of SIGCOMM '01, San Diego, USA, Aug.2001.

[3] M.Oe, "A hierarchical architecture for IP Traceback", Proc.54th IETF, ippt BoF, Yokohama, Japan, http://iplab.aist-nara.ac.jp/research/itrace/ippt-naist-ietf54.pdf, Jul.2002.

[4] M.Oe, Y.Kadobayashi and S.Yamaguchi, "A hierarchical architecture for IP Traceback (in Japanese)", IEICE trans. commun., vol.J85-B, no.8, pp.1313-1322, Aug.2002.

[5] Y.Sawai, M.Oe, K.Iida, and Y.Kadobayashi, "Performance evaluation of intra-domain IP traceback," to be presented at ICT'03, Tahiti, Feb.2003.

[6] J.Postel, "Internet protocol," RFC791 http://www.ietf.org/rfc/rfc791.txt , 1981.

[7] S.Deering, and R.Hinden, "Internet protocol, version 6 (IPv6) specification," RFC2460 http://www.ietf.org/rfc/rfc2402.txt, 1998.

[8] H.Krawczyk, M.Bellare, R.Canetti, "HMAC: Keyed-hashing for message authentication," RFC 2104 http://www.ietf.org/rfc/rfc2104.txt, 1997.

[9] D. Moore, G. M. Voelker, and S. Savage, "Inferring internet Denial-of-Service activity," in Proceedings of 10th USENIX Security Symposium '01, Washington, D.C, USA, Aug. 2001.

[10] Telecommunications Advancement Organization of Japan(TAO), "Hokuriku IT open laboratory," http://www.hokuriku-it.tao.go.jp/english/, Sep. 2002.