Advanced Packet Marking Mechanism with Pushback for IP Traceback

Hyung-Woo Lee

Dept. of Software, Hanshin University, Osan, Gyunggi, Korea, 447-791
hwlee@hs.ac.kr
http://netsec.hs.ac.kr

Abstract. Distributed Denial-of-Service(*DDoS*) attack can be done by generating a large volume of traffic through spoofing the IP address of the target system. In response to such attacks, IP traceback technology has been proposed. The method identifies the source of a DDoS attack and restructures the path on the network through which the attacking packet has been transmitted. Existing traceback techniques marked path information on packets or generated separate traceback messages but they increase network load and cannot cope with DDoS attacks actively because they generate traceback information for arbitrary packets without identifying DDoS attacks. Thus this study proposed an improved marking technique that identifies DDoS traffics at routers by applying the pushback function and cope with DDoS attack packets efficiently. According to the result of experiments, the proposed technique reduced network load and improved traceback performance.[1]

1 Introduction

The current TCP/IP system is vulnerable to *DoS (Denial of Service)* attacks such as TCP SYN flooding[1], there have been researches on how to cope with hacking on networks and the Internet[2]. As for techniques to cope with hacking attacks, firewall systems that adopt access control are passive to hacking attacks. IDS(Intrusion Detection System) provides the functions of detecting and blocking abnormal traffic that has reached the victim system, so it is also passive to hacking.

Thus currently available technologies do not provide active functions to cope with hacking such as tracing and confirming the source of DoS hacking attacks. It is because most hacking attacks are carried out by spoofing the IP address of the source system. Thus it is necessary to develop a technology to cope actively with such hacking attacks. Even if the trace-route technique is applied to identify the source address, the technique cannot identify and trace the actual address because the address included in DDoS(Distributed Denial of Service) is spoofed.

Methods of defeating hacking like DDoS attacks are largely divided into passive ones such as vaccines, intrusion detection and tolerance technology, and

[1] This work is supported by the University Basic Research Program of IITA and partially supported by University IT Research Center(ITRC) Project.

M. Jakobsson, M. Yung, J. Zhou (Eds.): ACNS 2004, LNCS 3089, pp. 426–438, 2004.

active ones such as traceback of the origin of attacks. Active methods are again divided into proactive traceback and reactive traceback according to how to detect the origin of hacking attacks.

When a DDoS hacking attack has happened, methods like ingress filtering filter and drop malicious packets at routers on the network, so they are passive to DDoS attacks. An efficient solution is for the victim system to trace back the spoofed actual address of the origin of the DDoS attack. In traceback methods, routers generate information on the traceback path while transmitting packets on the network, and insert it into the packets or deliver it to the IP address of the target of the packets.

If a victim system is hacked, it identifies the spoofed source of the hacking attacks using the generated and collected traceback path information. PPM (probabilistic packet marking)[5,6] and iTrace(ICMP traceback)[7] are this type of traceback methods. A recently proposed Pushback[3] method provides a identification function for packets when a DDoS attack happens and a transmission control function for packets along the packet transmission path. The method provide a control function for DDoS attack traffic but does not provide the function of trace back the source of the attack. It only provides a transmission control function for packets along the packet transmission path, so enhances the overall network performance.

Thus this study proposes a technique to trace back the source IP of spoofed DDoS packets by combining the existing pushback method, which provide a control function against DDoS attacks, with a traceback function. A router performs the functions of identifying/controlling traffic using the pushback technique, and when a DDoS attack happens it sends a pushback message to its upper router and transmits traceback information by marking it on the header of the corresponding packet. Compared to existing traceback techniques, the proposed technique reduced management system / network load and improved traceback performance.

Chapter II reviewed the weaknesses of existing technologies for tracing back the source of hacking attacks and directions for improvement, and Chapter III reviewed the weaknesses of existing pushback techniques. Chapter IV and V proposed a new packet marking technique that adopted a pushback technology to trace back the source of DDoS attacks, and Chapter VI compared and evaluated the performance of the proposed technique.

2 Related Works

2.1 Taceback Mechanisms

The rapidly spreading DDoS attacks generate a number of servers and a lot of subordinate servers (clients), connects to the master server, and carry out DDoS attacks to one or several IP addresses. In that case, *Trinoo* Master communicates with subordinate servers in order to attack one or several IP addresses during a specific period.

Because an attacker can carry out fatal DDoS attacks to victim systems by controlling a large number of servers where attacking tools are installed, such a method can be abused by hackers who mean to disturb the Internet. Up to now, when hacking attacks occur in the Internet, they have been defeated passively using firewall, IDS, scanning and trusted OS-based system security, etc. In particular, existing methods cannot restrict or prevent an attempt at hacking itself, so they are often useless and powerless against attacks paralyzing the Internet. To solve such a problem, active hacking prevention methods were proposed.

Traceback: *an essential technology to cope with hacking and virus actively. Traceback technology traces back the source of hacking attacks real-time and resultantly suppresses hacking attacks fundamentally.*

2.2 Discussion on Existing IP Source Traceback Technologies

Existing IP Traceback methods can be categorized as proactive or reactive tracing. Proactive tracing (such as packet marking and messaging) prepares information for tracing when packets are in transit. Reactive tracing starts tracing after an attack is detected.

Proactive methods sample and transmit packets at a probability of p, and there can be several variations. If a router generating PPM or iTrace messages can adjust probability actively according to the characteristics of entire network traffic rather than sampling at a fixed probability of p, the method may be superior to existing ones in network load, memory, traceback function, etc. In addition, an advanced method can be provided by integrating a traceback module with traditional security structure in order to prevent hackers from restructuring error paths.

PPM Methods[5,6]. In PPM mechanism, a router, an important component of a network, inserts information on packets transmitted through the router into IP packets in order to find the packet transmission route for spoofed packets.

That is, for packets transmitted through the Internet, a router routes them by checking packet header information centering on the IP layer. At that time, the router inserts information on the router address into a writable field of the IP header and sends the packet to the adjacent router.

Information inserted at each router is transmitted to the next router and finally to the target victim system. If a hacking attack occurs later, router information recorded in the packet corresponding to the hacking attack is reconstructed and generates the actual packet transmission path. *However, because all packets are marked with information at each router, transmission rate throughout the entire network will be lowered.* According to how to compose information marked at routers, there are methods such as *node sampling, edge sampling and improved packet marking.*

2.3 Weaknesses and Improvement of Existing PPM Technologies

With PPM technology, a router samples packet information at a probability of p, and marks the message header with its IP address and sends it to the target of the packet. A router samples packets at a probability of p and sends them, but a large number of marked packets are necessary to restructure the path to the source of DDoS attacks. If packets are transmitted without the edge or node information of a specific router, it is impossible to restructure the complete path using marked information. In addition, in order to mark the information of a node or an edge, the algorithm has to select and mark at least eight packets, so the overall efficiency is low.

What is more, existing PPM techniques may not mark hacking traffic in sampling and transmitting packets if the probability of p is satisfied. Because, in such a case, traceback path information is marked on general packets, the spoofed source of the attack cannot be restructured when hacking attacks such as DDoS happen. Thus if a router can adjust the probability rather than fixing it in sampling, the PPM method will have improved performance in network load, memory use and traceback compared to existing methods.

3 Pushback Mechanism

3.1 Pushback Based DDoS Traffic Identification/Control Mechanism

From the viewpoint of a router composing the network, *a hacking attack on the Internet is a kind of congestion.* Thus coping with hacking attacks may be approached from *congestion control* between end systems and relevant technologies. A DDoS attack transmits a large volume of traffic from one or more source hosts to a target host, there should be researches on how to identify and block DDoS traffic in order to cope with hacking attacks on the Internet.

A technology to control DDoS traffic at routers is *ACC (aggregate-based congestion control)* and pushback. Because hacking attacks are extremely diverse, it evaluates traffic based on *congestion signature*, which is corresponding to the congestion characteristic traffic.

ACC: *If traffic shows congestion exceeding a specific bandwidth based on the characteristic of DDoS attack nework traffic, the ACC module judges based on congestion signature that a hacking attack has happened and, working with a filtering module, provides a function to block the transmission of traffic corresponding to the DDoS attack.*

The Fig. 1 shows the structure of *ACC-based identification/control* when a router is congested. As in the figure, the process of identification/control is integrated with a pushback module. The pushback module confirms a DDoS attack and sends a pushback message to its adjacent previous router on the network path. In the figure below, if traffic explodes at link L_0, router R_0 detects (identifies) a high bandwidth. ACC module at R_0 blocks traffic to link L_0 and

send a pushback message to router R_2 and R_3, upper routers on the transmission path. R_2 delivers the pushback message just for congestion control to its upper router R_4, and R_3 to R_7. *However, it cannot trace back the final origin of the attack when a hacking attack occurs.*

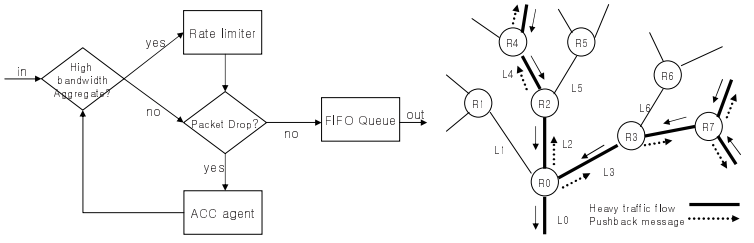

Fig. 1. ACC-based traffic identification/control and pushback mechanism.

3.2 Weaknesses and Improvement of Existing Pushback Technique

A network is defined as a graph $G = (V, E)$ composed of a set of nodes V and a set of edges E. The node set V is again divided into end systems and routers corresponding to internal nodes. Edges are physical connections among nodes in set V. Here, $S \subset V$ is defined as an attacker and $t \in V/S$ as a victim system.

$|S| = 1$ means an attack by a single attacker, and attack path information $P = (s, v_1, v_2, ..., v_d, t)$ means an attack path through which an attacking system s attacked a victim system t using routers on the path d. Let's say the number of packets transmitted is N. If there is a field in packets to mark with router link information $(v, v') \in E$, routers sample the packets at a probability of p. Routers can sample packets at a fixed probability of p and transmit information on edges and distances between routers by including it in the packets.

In existing methods, routers sample packets at a certain static probability of p and transmit router information by marking it on the packets. Probability α_i that a packet is marked at node v_i on the network and not remarked at other routers is computed as follows.

$$\alpha_i = Pr(x_d = (v_{i-1}, v_i)) = p(1-p)^{d-1} (i = 1, 2, ..., d)$$

Thus α_i means a probability that an attack packet is delivered to the victim system without being remarked by other routers. After all, p should be large in order to heighten α_i, which, however, means that routers have to perform marking frequently and consequently the network performance is degraded.

The existing pushback method sends a message to upper routers for the source of attack, however, it cannot trace back the final origin of the attack when a hacking attack occurs. That is, an additional process is necessary for a hacking victim system to trace back the path to the origin of the attack.

4 Packet Marking Based Traceback of DDoS Attacks

4.1 Traceback Structure Using Pushback

The method proposed in this study does not sample and mark at a fixed probability of p but mark packets when abnormal traffic is found by a pushback-based ACC module. Of course, unlike the method used in existing ACC techniques, when abnormal traffic is found, a pushback message is not delivered recursively to the upper router but marking is performed while the pushback message is delivered to the upper router. On receiving the pushback message, the upper router recognizes the characteristic of hacking traffic included in the message, performs marking with two router addresses and sent the message to the target system. The structure proposed in this study is as the figure below.

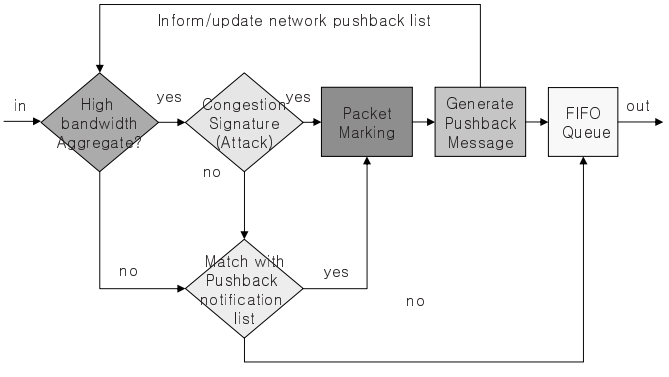

Fig. 2. Proposed Marking-based DDoS source traceback structure.

In the proposed structure, a router checks the traffic bandwidth of a received packet and if the bandwidth exceeds a certain level the router judges whether it is a congestion signature corresponding to an attack pattern. If the traffic bandwidth corresponds to an attack pattern, the router marks the packet, generates a pushback message for the packet and sends it the next router through the output queue of the router. If the traffic does not meet the bandwidth condition, the router check whether there is information coming through a pushback message from neighboring routers and if there is, it marks the packet. If the packet meets none of conditions above, the router regards it as a normal packet and delivers it to the next router.

4.2 Traceback Marking Method Using Pushback

(1) Packet Header Marking Field M_x. Let's say A_x is the IP address of R_x, P_x is IP packet arrived at R_x, and M_x is 24 bits on the header of P_x in

which marking information can be stored. In packet P_x, M_x is composed of 8-bit *TOS(type of service)* field, and 16-bit *ID field*. TOS field has been defined is not used currently. Thus the use of TOS field does not affect the entire network. In TOS field, the first 3 bits are priority bits, and next three bits are minimum delay, maximum performance and reliability fields but not used currently.

Recently, however, TOS field is redefined as Differentiated Service field(DS field) according to *RFC2474*, in which only the first 6 bits are used. Thus this study defines the unused 2 bits out of TOS field as *PF(pushback flag)* and *CF(congestion flag)*. Particularly for CF, RFC2474 defines it as 1 if the network is congested.

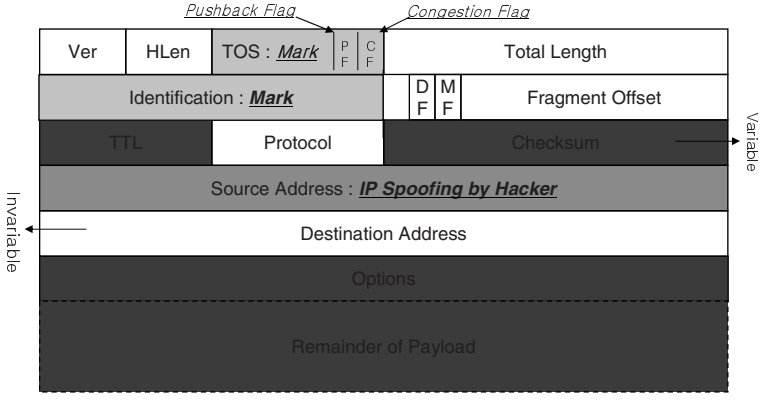

Fig. 3. Packet marking field in the proposed method.

(2) Marking Structure using TTL Information. the IP address A_x of router R_x is marked on 24-bit M_x through the following process. When abnormal traffic happens in the course of pushback for the writable 24 bits of a packet, router R_x marks A_x, which is its own IP address, and A_y, which is the IP address of the next router R_y. To mark the two router addresses within the 24 bits, the router uses address values based on the hash values of the routers, which also provide an authentication function.

TTL(time to live) in all packets is an 8-bit field, which is set at 255 in ordinary packets. The value of TTL field is decreased by 1 at each router until the packet reaches the target.

Currently TTL value is used to secure bandwidth in transmitting packets on the network and to control packets that have failed to reach the target. In previous researches, TTL value was not used but a separate hop counter field was used to calculate the distance that the packet has travelled. This study, however, uses part of TTL value in packets arrived at router R_x for packet marking.

Specifically because the maximum network hop count is 32 in general, the distance of packet transmission can be calculated only with the lower 6 bits out

of the 8 bits of TTL field in packet P_x arrived at router R_x. That is, the router extracts information of the lower 6 bits from the TTL field of packet P_x, names it T_x and stores it in TOS 6-bit field P_x^{TF} of the packet.

$$T_x = TTLofP_x \wedge 00111111$$

T_x value indicates the distance of the packet from the attack system. If the packet with the value is delivered to target system V, it is possible to calculate the distance from router R_x to target system V using the value V and T_v obtained in V in the same way.

(3) Traceback Path Marking at Routers. When informed of the occurrence of abnormal traffic by the ACC-based pushback module proposed above, router R_x performs marking for packet P_x corresponding to congestion signature included in the pushback message.

First of all, because the router received a pushback message, it resets PF field in TOS field as 1. Then it calculates T_x for 8-bit TTL field of packet P_x and stores it in the 6 bits of TOS field. Then the router calculates 8-bit hash value for A_x the address of router R_x and T_x calculated earlier using hash function $H(\cdot)$, and marks the value on P_x^{MF1}, the first 8 bits of ID field. The marked packet is delivered to R_y, the next router on the routing path to the target address.

Now when router R_y checks P_x^{PF} the value of PF field in the packet and finds it is 1, the router applies the hash function to the value obtained by subtracting 1 from P_x^{PF}, which is corresponding to the 6 bits of TOS field in the packet, and router IP address A_x and marks the resulting value on P_x^{MF2}.

$$P_x^{MF1} = H(T_x|A_x), P_x^{MF2} = H(P_x^{TF} - 1|A_y)$$

After marking, the router set CF at 1 and sends the packet to the next router. The next router, finding PF and CF are set at 1, does not perform marking because the packet has been marked by the previous router.

5 Structuring Packet Traceback Path

5.1 DDoS Attack Packet Traceback

For a packet transmitted through the network, victim system V restructures the DDoS attack path. As in the figure below, let's assume that DDoS attacks have been made against S_1, S_2, S_3. For the attack packet, router R_x, R_y and R_z marked 24 bits in the packet header with its own IP information and the information of 6-bit TTL field of the packet. When the DDoS attack occurred, the victim systems perform traceback as follows for packets arrived.

First of all, let's say P_v is a set of packets arrived at victim system V. P_v is a set of packets corresponding to DDoS attacking, and M_v is a set of packets within P_v, which were marked by routers.

To distinguish M_v from packet set P_v, the system selects packets in which PF field P_x^{PF} and CF field P_x^{CF} have been set at 1 as below.

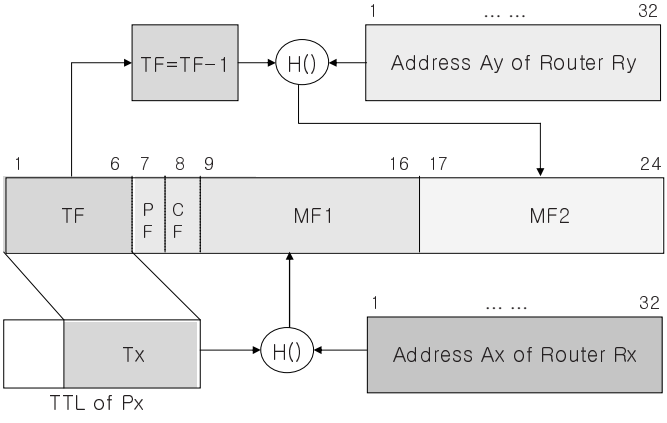

Fig. 4. Packet marking structure in the proposed method

$$M_v = \{P_x | P_x^{PF} == 1 \wedge P_x^{CF} == 1, x \in v\}$$

That is, for packet M_i belonging to packet set M_v in a victim system, its 8-bit TTL value can be defined as $TTLofM_i$. The value is compared with T_{M_i} marked on TOS field, and the network hop count $D(M_i)$, which is the distance since packet M_i was marked, is calculated as follows.

$$D(M_i) = M_i^{TF} - (TTLofM_i \wedge 00111111)$$

If $D(M_i) == 1$, it indicates that the packet was marked at the router just in front of the victim system. The method proposed in this study, however, adopts a pushback technique, it can restructure a traceback path using a packet with $D(M_i) == 2$.

5.2 DDoS Attack Path Reconstruction

Packet M_i satisfying $D(M_i) == 2$ means that the packet was marked by router R_y and R_x two hops apart from the end router in front of the victim system. That is, $D(M_i)$ for packet M_i is 2 because the packet was marked by router R_x, which is 2 hops apart from the router directly connected to the victim system. Thus R_x, 2 hops apart from packet M_i can be identified in the following way.

$$M_i^{MF1} == H(M_i^{TF} | R_x), (R_x \in D(M_i) == 2) and$$
$$M_i^{MF1} == H((TTLofM_i \wedge 00111111) + 2 | R_x), (R_x \in D(M_i) == 2)$$

Of course, packet M_i can prove in the following way that a packet was marked by router R_y 1 hop apart from the victim system.

$$M_i^{MF2} == H(M_i^{TF} - 1 | R_y), (R_y \in D(M_i) == 1) and$$
$$M_i^{MF2} == H((TTLofM_i \wedge 00111111) + 1 | R_y), (R_y \in D(M_i) == 1)$$

Now the victim system can restructure the actual attack path through which packets in DDoS attack packet set P_v were transmitted by repeating the same process for M_j satisfying $D(M_j) == n, (n \geq 3)$. When the proposed method is applied to a network structured as below, DDoS attack path AP to a victim system can be obtained as follows.

$$AP_1 = R_y \rightarrow R_x \rightarrow R_z \rightarrow S1, AP_2 = R_y \rightarrow R_3 \rightarrow R_7 \rightarrow S2,$$
$$AP_3 = R_y \rightarrow R_3 \rightarrow R_7 \rightarrow S3$$

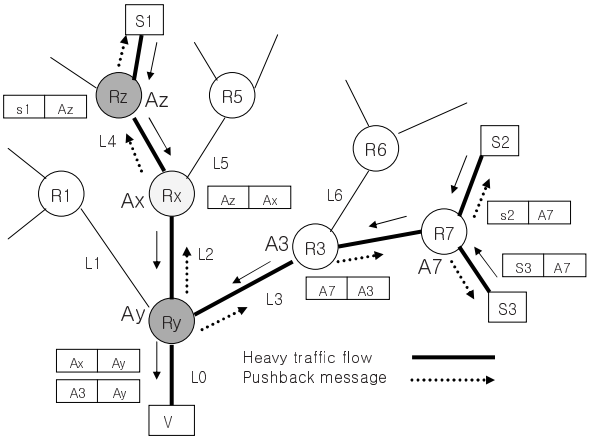

Fig. 5. Attack path traceback using the proposed method.

Through the process, routers could perform not only a monitoring/identification function on network traffic using an ACC module but also a network control function using modified pushback technology. What is more, the proposed method could restructure the source of attackers by providing the function of tracing back spoofed packets adopting improved packet marking technology in order to trace back DDoS hacking paths. Furthermore, it provided a structure for verifying information marked by attackers using a hash technique.

6 Performance Analysis for the Proposed Method

6.1 Experiment Results

In order to evaluate the performance of the proposed method, the author analyzed the performance using ns-2 Simulator in Linux. A network was composed as shown in the figure below and DDoS attacks were simulated against node 0, 1 and 2.

According to the results of the experiment, in existing packet marking methods each router samples and marks at a probability of p to cope with DDoS

attacks. Thus the number of marked packets has increased in proportion to DDoS traffic. In the method proposed in this study, a pushback technique is adopted in marking DDoS traffic and as a result the number of marked packets has decreased by 25.4%.

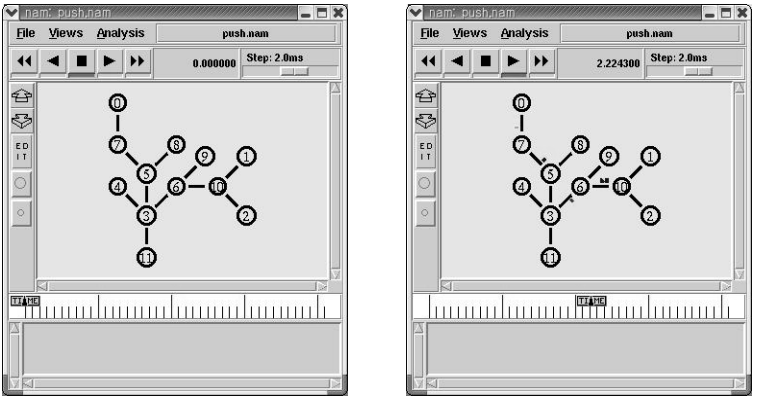

Fig. 6. Attack Architecture and Its Simulation on ns-2.

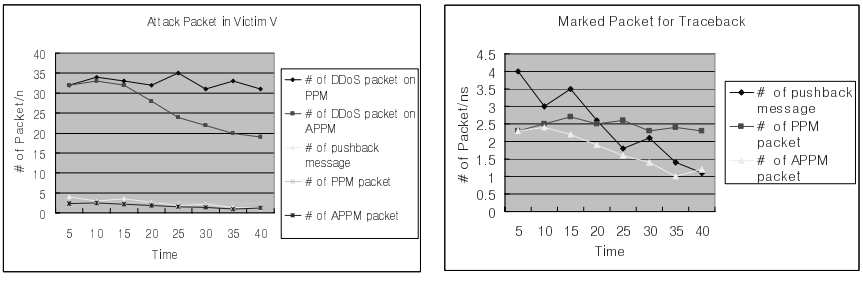

Fig. 7. Traffic Simulation Results by the Proposed Method .

We can control the DDoS traffic by issuing Pushback message to upper router and marking router's own address in IP packet. So, proposed mechanism can identify/control DDoS traffic by using existing Pushback module and trace back its spoofed real origin address with fewer marking packet compared with previous PPM mechanism.

6.2 Analysis and Discussions

[Table 1] shows the comparison of the performance of the proposed method with that of existing IP traceback-related technologies. Filtering, which provides an access control function at routers, does not reduce the load of the entire system

and the victim system but inspect packets at routers like SYN flooding. Thus the method does not need additional memory, but it cannot provide a traceback function nor security and cannot cope with DDoS attacks. The method that manages the log of packet information at routers requires a large volume of memory at routers. Although it provide a partial traceback function it is poor in security and vulnerable to DDoS attacks.

Table 1. Comparison of performance with existing IP traceback methods.

	Net. load	Sys. load	Memory	Traceback	Security	DDoS	# of packet
Filter	×	×	×	×	×	×	×
SYN fld.	×	↓	×	×	×	×	×
Logging	×	×	↑	▽	◇	▽	1
PPM	↓	↑	↑	△	◇	▽	↑
iTrace	↓	↑	↑	△	◇	▽	↑
APPM	↓	↓	↑	△	◇	△	n

×:N/A ↓:low ↑:high △:good ◇:moderate ▽:bad

Existing packet marking methods and iTrace methods based on node and edge sampling cause low load on the management system and the network but create heavy load when a victim system restructures the traceback path. These methods are considered suitable in terms of traceback function and scalability. However, they are vulnerable to DDoS attacks. As a whole, most of IP traceback techniques proposed up to now modify existing established routers and cause additional load on the network and the system.

The method proposed in this study runs in a way similar to existing PPM, so its management load is low. But the overall computational overhead will be increased compared with common PPM algorithm as the proposed mechanism is combined with ACC-based pushback module and TTL based traceback module. But, we can ignore these overhead because it provide better traceback functionality. Furthermore, because it applies identification/control functions to packets at routers it reduces load on the entire network when hacking such as DDoS attacks occurs. What is more, while existing PPM methods mark packets by sampling them at an arbitrary probability of p, the method proposed in this study uses an ACC-based congestion control function and marks path information using the value of TTL field, which reduces the number of packets necessary for restructuring a traceback path to the victim system.

Thus the method improves the bandwidth of the entire network and can restructure the path to the source of DDoS attacks with a small number of marking packets. In the method, the path to the source of attack can be restructured only with n traceback messages if the packet has been transmitted via n routers on the network. As a disadvantage, the method requires additional memory at routers for the DDoS-related identification function performed by the ACC-based push-

back module. Proposed method requires about 18.8% of additional memory for pushback function compared with common traceback mechanism.

7 Conclusions

As a technology to cope with rapidly increasing hacking and viruses on the Internet, this study proposed a method for a victim system to trace back the actual IP address of the attacker for spoofed traffic when a DDoS attack happens. Reviewing the structure, current state and problems of existing traceback technologies, the author proposed a new marking technique that provides the functions of identifying/controlling DDoS hacking attacks on the network and at the same time enables victim systems to trace back the spoofed source of hacking attacks.

The proposed method is superior to existing ones in load, performance, stability and traceback function. Recently mobile networks and ad-hoc-based networks are found to have vulnerable points to DDoS attacks. Thus it is necessary to study how to provide filtering for packets and trace back the source of attacks in mobile environment. Furthermore, traceback functions should be considered in IPSec-based environment, which provides security protocol at the IP layer, and in ordinary IP layers. Lastly, it is necessary to inquire into methods that provide safety to the entire network including routers, which has been performed by firewall and IDS, and at the same time provide an improved traceback function for packets.

References

1. Computer Emergency Response Team, "TCP SYN flooding and IP Spoofing attacks", CERT Advisory CA-1996-21, Sept, 1996.
2. L. Garber, "Denial-of-Service attacks trip the Internet", Computer, pages 12, Apr. 2000.
3. S. Floyd, S. Bellovin, J. Ioannidis, K. Kompella, R. Mahajan, V. Paxson, "Pushback Message for Controlling Aggregates in the Network," Internet Draft, 2001.
4. P. Ferguson and D. Senie, "Network ingress Filtering: Defeating denial of service attacks which employ IP source address spoofing", May 2000. RFC 2827.
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 '01, pages 338-347, 2001.
6. D. X. Song, A. Perrig, "Advanced and Authenticated Marking Scheme for IP Traceback", Proc, Infocom, vol. 2, pp. 878-886, 2001.
7. Steve Bellovin, Tom Taylor, "ICMP Traceback Messages", RFC 2026, Internet Engineering Task Force, February 2003.
8. Tatsuya Baba, Shigeyuki Matsuda, "Tracing Network Attacks to Their Sources," IEEE Internet Computing, pp. 20-26, March, 2002.
9. Deering, S. and R. Hinden, "Internet Protocol, Version 6, (IPv6) Specification", RFC 2460, December 1998.
10. Andrey Belenky, Nirwan Ansari, "On IP Traceback," IEEE Communication Magazine, pp.142-153, July, 2003.