Flexible Deterministic Packet Market: An IP Traceback Scheme

Pooja G. Kukreja¹, D.N. Rewadkar²
¹, ²Computer Department, R.M.D. Sinhgad Technical Institute, Warje, Pune, India

Abstract: Internet Protocol (IP) traceback is the enabling technology to control Internet crime. IP traceback system is called as Flexible Deterministic Packet Marking (FDPM) which provides a defense system with the ability to find out the real sources of attacking packets that traverse through the network. While a number of other traceback schemes exist, FDPM provides innovative features to trace the source of IP attack and can obtain better tracing capability than others. In this paper Any File is first divided into packets. Then all Packets are marked on marker side according to marking Scheme algorithm. Then there is a hacker who receives packet then he send it either normally or delete/alter some data in some packets then send further. Then there is one “Reconstructor” he only trace the source of IP attack and can obtain better tracing capability than others. In this paper Any File is first divided into packets. The implementation and evaluation demonstrates that FDPM needs moderately a small number of packets to complete the traceback process and requires little computation work; therefore this scheme is powerful to trace.

Keywords: DDoS attacks, IP traceback, security, Flexible Deterministic Packet Marking.

1. Introduction

A great amount of effort in recent years has been directed to the network security issues. In this paper, we address the problem of identifying the source of the attack. We define the source of the attack to be a device from which the flow of packets, constituting the attack, was initiated. This device can be a zombie, reflector, or a final link in a stepping stone chain. While identifying the device, from which the attack was initiated, as well as the person(s), behind the attack is an ultimate challenge, we limit the problem of identifying the source of the offending packets, whose addresses can be spoofed. This problem is called the IP traceback problem.

Several solutions to this problem have been proposed. They can be divided in two groups. One group of the solutions relies on the routers in the network to send their identities to the destinations of certain packets, either encoding this information directly in rarely used bits of the IP header, or by generating a new packet to the same destination. The biggest limitation of this type of solutions is that they are focused only on flood-based (Distributed) Denial of Service (DDoS) attacks, and cannot handle attacks comprised of a small number of packets. The second type of solutions involves centralized management, and logging of packet information on the network. Solutions of this type introduce a large overhead, and are complex and not scalable.

IP traceback is a name given to any method for reliably determining the origin of a packet on the Internet. Due to the trusting nature of the ip protocol, the source IP address of a packet is not authenticated. As a result, the source address in an IP packet can be falsified (IP address spoofing) allowing for Denial Of Service attacks (DoS) or one-way attacks (where the response from the victim host is so well known that return packets need not be received to continue the attack). The problem of finding the source of a packet is called the IP traceback problem. IP traceback is the enabling technology to control Internet crime. Most existing approaches to this problem have been tailored toward DoS attack detection. Such solutions require high numbers of packets to converge on the attack path(s). In this paper we present a novel and practical IP traceback system called Flexible Deterministic Packet Marking (FDPM) which provides a defense system with the ability to find out the real sources of attacking packets that traverse through the network. While a number of other traceback schemes exist, FDPM provides innovative features to trace the source of IP packets and can obtain better tracing capability than others.

Figure 1: A Scenario of DOS Attack

Volume 4 Issue 1, January 2015

www.ijsr.net
Licensed Under Creative Commons Attribution CC BY
the IP packets. It can be applied in many security systems, such as DDoS defense systems, Intrusion Detection Systems (IDS), forensic systems, and so on.

2. Current IP Traceback Schemes

There are some tradeoffs of different IP traceback schemes. Current IP traceback schemes can be classified into five categories: link testing, messaging, logging, packet marking, and hybrid schemes.

A. Link testing
The main idea of the link testing scheme is to start from the victim to trace the attack to upstream links, and then determine which one carries the attack traffic.

**Disadvantage**
It consumes huge amount of resources, introduces additional traffic, and possibly causes denial of service when the number of sources needed to be traced increases.

B. Messaging
Messaging schemes use routers to send ICMP messages from the participating routers to destinations. For a high volume flow, the victim will eventually receive ICMP packets from all the routers along the path back to the source, revealing its location.

**Disadvantage**
The disadvantages of messaging schemes are that the additional ICMP traffic would possibly be filtered by some routers, and huge number of packets is required by the victim to identify the sources.

C. Logging
Logging schemes include probabilistic sampling and storing transformed information. Logging schemes maintain a database for all the traffic at every router within the domain and to query the database to identify the sources of an IP packet. Hash function or Bloom filter is used to reduce the data stored.

**Disadvantage**
The main disadvantage of logging schemes is that they heavily overload the participating routers by requiring them to log information about every packet passing by, although it is claimed that it needs only a single packet to find its origin.

D. Hybrid Traceback scheme
A hybrid traceback scheme combining logging and packet marking is presented to achieve the small number of packets needed to trace a single source and the small amount of resources to be allocated to the participating routers. Although the hybrid schemes try to overcome the disadvantages of each traceback scheme, the complexity of such combination and the practicability of their implementation still need more research.

E. ICMP Traceback
Internet Control Message Protocol (ICMP) in need of trace out full path of the attacks. This approach was originally introduced by Bellovin. The principle idea in these schemes is for every router to generate an ICMP traceback message or iTrace directed to the same destination as the selected packet. The iTrace message itself consists of the next and previous hop information and a time stamp.

**Figure 2:** Packet marking

As packets travel through the network, they gather and store information about the routers they traverse.

A router creates an ICMP traceback message, which contains part of a traversing IP packet, and sends the message to the packet's destination. We can identify the traversed router by looking for the corresponding ICMP traceback message and checking its source IP address. Because creating an ICMP traceback message for every packet increases network traffic, however, each router creates ICMP traceback messages for the packets it forwards. If an attacker sends many packets the target network can collect enough ICMP traceback messages to identify its attack path.

F. Packet Marking Algorithm
In Packet Marking Algorithm schemes, each router in addition to forwarding a packet also inserts a mark in the packet. This mark is a unique identifier corresponding to this particular router.

As a result the victim can determine all the intermediate hops for each packet by observing the inserted marks. There are two variants to this marking scheme. First is the Deterministic Packet Marking (DPM) scheme in which each router marks all the packets passing through it with its unique identifier. This scheme is thus similar to the IP record-route option. This makes the reconstruction of the attack path at the victim trivial. But the downside to this scheme is that routers are slowed down as they have to perform additional functionality. An attacker who controls a trusted router can forge any path up to that router unless some further authentication scheme is used. A router that trusts data from an attacker effectively allows that attacker to act like a compromised router. Authentication methods could be used, but these add significant cost in the form of processing time and space in the marked packets. A downside of this scheme is that some packets will not be overwritten by any of the routers. The attacker can therefore write bogus information in all the packets knowing that some of these packets will get through and confuse the victim. This method also does not work well for DoS attacks that can work without a lot of packets as it requires a large number of packets to converge. The second instances is probabilistic packet marking (PPM), DoS attacks can be prevented if the spoofed source IP address is traced back to its origin which allows assigning penalties to the offending party or isolating the compromised hosts and domains from the rest of the network.
Recently IP traceback mechanisms based on probabilistic packet marking have been proposed for achieving traceback of DoS attacks. In this paper, we show that probabilistic packet marking of interest due to its efficiency and implementability vis-à-vis deterministic packet marking and logging or messaging based schemes suffers under spoofing of the marking field in the IP header by the attacker which can impede traceback by the victim. Attacks on PPM: Attacks involving spoofed traceback data are described in. In general the two major problems in PPM reliability are the probabilistic nature of the algorithm causes some packets not to be marked by cooperating routers and these retain whatever marks are given them by the senders. Attackers can simply mark their original packets to intentionally mislead the traceback mechanism. In DPM routers mark all forwarded packets with link identifying data. With PPM, multiple routers on the paths overwrite the same data, and each packet identifies at most one link. With DPM, each cooperating router adds link identifying data to the packet and each packet ends up with data that identifies all of the links (under universal co-operation) that it traversed.

Disadvantages of Packet Marking
a) Mark Length: It cannot adjust the length of marking field according to the network protocols deployed.
b) Marking Rate is not flexible according to the load of the participating router.
c) Number of Packets required is comparatively more.
d) False Positive rate is large.
e) Tracing Capability is less.
f) The path reconstruction process requires high computational work, especially when there are many sources. For example, a 25-source path reconstruction will take days, and thousands of false positives could happen.
g) When there are a large number of attack sources, the possible rebuilt path branches are actually useless to the victim because of the high false positives.

G. FDPM Traceback:
Flexible Deterministic Packet Marking (FDPM) is the optimized version of DPM. This scheme provide more flexible features to trace the IP packets and can obtain better tracing capabilities over other previous IP traceback mechanisms, such as Link testing, logging, ICMP traceback, probability packet marking (PPM) and Deterministic packet marking (DPM). In FDPM schemes, the Types of Services (ToS) fields will be used to store the mark under some circumstances. The two fields in the IP header are exploited, one is fragment ID and other is Reversed flag. An identifying value is assigned to the ID field by the sender to aid in assembling the fragments of a datagram. Given that less than 0.25% of all internet traffic is fragments, this field can be safely overloaded without causing serious compatibility problems. FDPM reconstruction process includes two steps: mark recognition and address recovery. Compared to DPM, the reconstruction process is simpler and more flexible. When each packet that is used to reconstruct the source IP address arrives at the victim, it is put into a cache, because in some cases the processing speed is lower than the arrival speed of the incoming packets.

The FDPM scheme utilizes various bits (called marks) in the IP header. The mark has flexible lengths depending on the network protocols used, which is called flexible mark length strategy. When an IP packet enters the protected network, it is marked by the interface close to the source of the packet on an edge ingress router. The source IP addresses are stored in the marking fields. The mark will not be overwritten by intermediate routers when the packet traverses the network. At any point within the network, e.g., the victim host, the source IP addresses can be reconstructed when required.

The flow-based marking scheme is proposed to solve the overload problem. When the load of a router exceeds a threshold, the router will discern the most possible attacking packets from other packets then selectively mark these packets. The aim is to alleviate the load of the router while still maintaining the marking function.

Advantages
a) Easy to find out packet loss and Duplicate packets.
b) Reduces the network traffic.
c) Bandwidth consumption is less.
d) Flexible mark length: The length of marking field can be adjusted according to the network protocols deployed.
e) Flexible mark rate: The marking rate can be changed adaptively according to the load of the participating router.
f) Low false Positive rate.
g) Number of packets required is comparatively less.
h) Better Tracing Capability.
i) It has Different probabilities that a router marks the attack packets.

H. TBPM Method:
Topology aware single packet IP traceback system is namely TOPO. It is based on the bloom filter which utilizes router’s local topology information, i.e., its immediate predecessor information, to traceback. TOPO can significantly reduce the number and scope of unnecessary queries and thus, significantly decrease the false attributions to innocent nodes. The main goals of TOPO as follows:

1. to design a single packet IP traceback system, this has fewer unnecessary query messages and fewer false attributions to innocent nodes.
2. to design a single packet IP traceback system this needs not to be fully deployed in the entire network.
3. to design a mechanism which helps achieve the best performance of Bloom filters by adaptively adjust using parameter.

Topology Based Packet Marking (TBPM) has been a new approach in Anti-IP spoofing techniques.
Still maintaining the marking function. The flexibility of threshold, the router will discern the most possible attacking packets. The aim is to alleviate the load of the router while overload problem. When the load of a router exceeds a packets from other packets then selectively mark these flow-based marking scheme is proposed to solve the of arrival packets waiting for FDPM to mark them. The participation of a router to be overloaded when there are a large number of CPU time of a participating router. Therefore, it is possible to take days, and thousands of false positives could happen.

Disadvantages of Current IP Traceback Schemes
a) Mark Length: It cannot adjust the length of marking field according to the network protocols deployed.
b) Marking Rate is not flexible according to the load of the participating router.
c) Number of Packets required is comparatively more.
d) False Positive rate is large.
e) Tracing Capability is less.
f) The path reconstruction process requires high computational work, especially when there are many sources. For example, a 25-source path reconstruction will take days, and thousands of false positives could happen.
g) When there are a large number of attack sources, the possible rebuilt path branches are actually useless to the victim because of the high false positives.

Flexible Deterministic Packet Marking

1. System Overview
The FDPM scheme utilizes various bits (called marks) in the IP header. The mark has flexible lengths depending on the network protocols used, which is called flexible mark length strategy. When an IP packet enters the protected network, it is marked by the interface close to the source of the packet on an edge ingress router. The source IP addresses are stored in the marking fields. The mark will not be overwritten by intermediate routers when the packet traverses the network. At any point within the network, e.g., the victim host, the source IP addresses can be reconstructed when required. Processing packets consume resources such as memory and CPU time of a participating router. Therefore, it is possible for a router to be overloaded when there are a large number of arrival packets waiting for FDPM to mark them. The flow-based marking scheme is proposed to solve the overload problem. When the load of a router exceeds a threshold, the router will discern the most possible attacking packets from other packets then selectively mark these packets. The aim is to alleviate the load of the router while still maintaining the marking function. The flexibility of FDPM is twofold. First, it can use flexible mark length according to the network protocols that are used in the network. This characteristic of FDPM gives it much adaptability to current heterogeneous networks. Second, FDPM can adaptively adjust its marking process to obtain a flexible marking rate. This characteristic prevents a trace back router from the overload problems.

2 Advantages
a. Easy to find out packet loss and Duplicate packets.
b. Reduces the network traffic.
c. Bandwidth consumption is less.
d. Flexible mark length: The length of marking field can be adjusted according to the network protocols deployed.
e. Flexible mark rate: The marking rate can be changed adaptively according to the load of the participating router.
f. Low false Positive rate.
g. Number of packets required is comparatively less.
h. Better Tracing Capability.
i. It has Different probabilities that a router marks the attack packets.

3 FDPM Schemes
a) Encoding
Before the FDPM mark can be generated, the length of the mark must be determined based on the network protocols deployed within the network to be protected. According to different situations, the mark length could be 24 bits long at most, 19 bits at middle, and 16 bits at least. Therefore, the flexible length of the marks results in three variations of the encoding scheme, which are named as FDPM-24, FDPM-19, and FDPM-16. FDPM encoding scheme is shown in Fig. 1. The ingress IP address is divided into k segments and stored into k IP packets. The padding is used to divide the source IP address evenly into k parts. For example, if k = 6, the source address is added with 4 bits of 0, making it 36 bits long, then each segment will be 6 bits long. The segment number is used to arrange the address bits into a correct order. The address digest enables the reconstruction process to recognize that the packets being analyzed are from the same source. Without this part, the reconstruction process cannot identify packets coming from different sources, thus will not be able to trace multiple IP packets.

Diagram:

In FDPM, before the encoding process begins, the length of the mark must be calculated. If the TOS field in the IP
packet is not used by the protected network, the 1-bit Reserved Flag in the header is set to 0, and the length of mark is set to 24. Under other situations, the length of mark will be 19 or 16, with relevant bit(s) in TOS marked. If the network supports TOS Precedence but not TOS Priority, fourth to sixth bits of TOS are utilized for marking; and if the network supports TOS Priority but not TOS Precedence, first to third bits of TOS are utilized for marking.

b) Reconstruction
The reconstruction process includes two steps: mark recognition and address recovery. When each packet arrives at the point that requires reconstruction, it is first put into a cache because, in some cases, the reconstruction processing speed is slower than the arrival speed of the incoming packets. The cache can also output the packets to another processing unit, by this design the reconstruction methods can be applied in a parallel mode (e.g., if the router has multi core architecture).

The mark recognition step is the reverse process of the encoding process. By reading the control fields in the mark, the length of the mark and which fields in the IP header store the mark can be recognized. If the RF is 0, the mark length is 24 (both TOS and ID are deployed). If the RF is 1, according to different protocols of TOS used, the mark length is 16 or 19. The second step, address recovery, analyzes the mark and stores it in a recovery table. It is a linked-list table; the number of rows is a variable, and the number of columns in the table is k, representing the number of segments used to carry the source address in the packets. Here, the segment number is used to correlate the data into the correct order. The row of the table means the entry; usually each digest owns one entry (source IP address). However, different source IP addresses may have the same digest because the digest is a hash of the source IP address, and is shorter than an IP address. In this case, hash collision is unavoidable. When the hash collision occurs, more than one entry may be created in order to keep as much information as possible. The advantage of this design is that it can reconstruct all possible sources but the disadvantage is it also brings possible irrelevant information. Compared with DPM, our reconstruction process is compatible with different protocols and will not lose any sources even when hash collision occurs. Fig.2 shows the reconstruction scheme. When all fields in one entry are filled according to the segment number, this source IP address is reconstructed and the entry in the recovery table is then deleted.

3. Flow Based Marking
The possibility of the overload problem always exists because the resources of a router are always limited. If the router is overloaded, the marking scheme can be totally ineffective. All packet marking traceback schemes consume the computing power and storage capacity of routers as they need to overwrite many bits in the IP header. Therefore, overload prevention is important to all packet marking traceback schemes. There are many methods to lighten the burden of a router. One is to increase the computing capability of the router, for example, by using multi core-based architecture. Another is to apply an adaptive algorithm to reduce the load of processing of packets when the load of the router exceeds a threshold, which is our novel approach. The idea of flow-based marking is to selectively mark the packets according to the flow information when the router is under a high load. Therefore, it can reduce the packet marking load of a router but still maintain the marking and traceback function. Because the main application of FDPM in our research is DDoS defense, the flow-based marking mainly deals with the packets in DDoS attack scenarios. For other applications, this overload prevention mechanism can be modified accordingly to target most possible attacking packets.
4. Modules Description

A. User Login

User can send or receive messages, for this they have to get login. Message Transfer window is the window in which one can type or browse the messages which have to send to other nodes. Receiving window is used when a user receives the message the receiver window automatically opened on the receiver side and the DOS attacker packet are shown in alert box, if there are any. Graph construction window can see the graph in the way which it is traversed.

B. Packet Encoding

The FDPM algorithm is designed to automatically determine when the algorithm should terminate. We aim at achieving the following properties:
1. The algorithm does not require any prior knowledge about the network topology.
2. The algorithm determines the certainty that the constructed graph is the attack graph when the algorithm terminates.

C. Used Methods

1. Routing Traceback System: Packet marking schemes insert trace back data into an IP packet header to mark the packet on its way through the various routers from the attack source to the destination, then the marks in the packets can be used to deduce the sources of packets or the paths of the traffic. This property makes it a promising trace back scheme to be part of DDoS defense systems.

2. FDPM Scheme:
The FDPM scheme utilizes various bits (called marks) in the IP header. The mark has flexible lengths depending on the network protocols used, which is called flexible mark length strategy. When an IP packet enters the protected network, it is marked by the interface close to the source of the packet on an edge ingress router. The source IP addresses are stored in the marking fields. The mark will not be overwritten by intermediate routers when the packet traverses the network. At any point within the network, e.g., the victim host, the source IP addresses can be reconstructed when required.

3. Flow-Based Marking Scheme:
The idea of flow-based marking is to selectively mark the packets according to the flow information when the router is under a high load. Therefore, it can reduce the packet marking load of a router but still maintain the marking and trace back function. Because the main application of FDPM in our research is DDoS defense, the flow-based marking mainly deals with the packets in DDoS attack scenarios. For other applications, this overload prevention mechanism can be modified accordingly to target most possible attacking packets.

Flexible Deterministic Packet Marking (FDPM) is the optimized version of DPM. This scheme provides more flexible features to trace the IP packets and can obtain better tracing capabilities over other previous IP traceback mechanisms, such as Link testing, logging, ICMP traceback, probability packet marking (PPM) and Deterministic packet marking (DPM). In FDPM schemes, the Types of Services (ToS) fields will be used to store the mark under some circumstances. The two fields in the IP header are exploited, one is fragment ID and other is Reversed flag. An identifying value is assigned to the ID field by the sender to aid in assembling the fragments of a datagram. Given that less than 0.25% of all internet traffic is fragments, this field can be safely overloaded without causing serious compatibility problems. FDPM reconstruction process includes two steps: mark recognition and address recovery. Compared to DPM, the reconstruction process is simpler and more flexible. When each packet that is used to reconstruct the source IP address arrives at the victim, it is put into a cache, because in some cases the processing speed is lower than the arrival speed of the incoming packets.

The FDPM scheme utilizes various bits (called marks) in the IP header. The mark has flexible lengths depending on the network protocols used, which is called flexible mark length strategy. When an IP packet enters the protected network, it is marked by the interface close to the source of the packet on an edge ingress router. The source IP addresses are stored in the marking fields. The mark will not be overwritten by intermediate routers when the packet traverses the network. At any point within the network, e.g., the victim host, the source IP addresses can be reconstructed when required. Processing packets consume resources such as memory and CPU time of a participating router. Therefore, it is possible for a router to be overloaded when there are a large number of arrival packets waiting for FDPM to mark them.

The flow-based marking scheme is proposed to solve the overload problem. When the load of a router exceeds a threshold, the router will discern the most possible attacking packets from other packets then selectively mark these packets. The aim is to alleviate the load of the router while still maintaining the marking function.

Advantages

a) Easy to find out packet loss and Duplicate packets.
b) Reduces the network traffic.
c) Bandwidth consumption is less.
d) Flexible mark length: The length of marking field can be adjusted according to the network protocols deployed.
e) Flexible mark rate: The marking rate can be changed adaptively according to the load of the participating router.
f) Low false positive rate.
g) Number of packets required is comparatively less.
h) Better Tracing Capability.
i) It has different probabilities that a router marks the attack packets.

5. Challenges and Future Work

Future difficult and challenging issues IP traceback should address include:

- Identifying the indirect sources of reflector based DDoS attacks.

Volume 4 Issue 1, January 2015

www.ijsr.net

Licensed Under Creative Commons Attribution CC BY
• Identifying the attacker who conceals himself/herself with stepping stones.
• Integrating IDS or defensive measures with traceback so that one mechanism may perform tracing as well as detection and/or defense.
• Automatic traceback to speed up tracing and reduce human intervention.

At present, all the above are still open problems. A scheme contrived to address reflector based DDoS attacks has to address one important issue: some kind of trust relationship must exist between the victim and the reflectors so that the reflectors may authenticate the querying requests from the victim, and the victim may obtain from the reflectors their tracing results. The trust relationship must be deliberately established and efficiently maintained. Otherwise, an attacker may exploit it to mount a DDoS attack by frequently sending bogus querying requests. Here, scalability is still a big challenge.

In addition to spoofed source IP addresses, a sophisticated attacker may use a series of stepping stones to further conceal its trail. A stepping stone is a host that is remotely logged in by a user whose physical location may be pretty far away. Employing many stepping stones can effectively hamper efforts to identify the attacker. No sound scheme has been presented to tracing through stepping stones yet.

Integrating IP traceback with other functionalities such as detection and defense is another topic of interest. Currently, a common assumption is that there exists IDS at the victim or at the TR in Center Track. IP traceback may identify attack sources. However, IP traceback itself is not a detection or defense scheme. A scheme that may effectively and efficiently combine detection, defense, and traceback may significantly enhance performance and mitigate false positives. Instead of the current practice of human manipulation, automatic tracing is very useful, especially in a large network made up of a huge number of hosts. Automatic traceback requires more intimate coordination between IDS and traceback. To decrease the false alarm rate, the accuracy of detection needs to be significantly improved. However, improving the accuracy of DDoS detection is a daunting task given the fact that a DDoS attack may be a hybrid of different types of attacks using different protocols, ports, and attack rates.

6. Conclusion

One conclusion we can draw from this is that unless IP traceback measures are deployed all over the Internet, they are only effective for controlled networks than for the Internet.

Today we can find many tools for doing DoS attacks. DoS attacks have become very popular. Hence we need to design proper mechanisms to protect systems from such attacks. Mechanisms has been developed and deployed to prevent such attacks. But DDoS is still a problem as it is difficult to trace DDoS attackers and its effect is too bad. We need to start development towards defending DDoS. Some schemes are present which very well defends such attacks, but without the cooperation of ISPs it will be difficult to deploy any scheme. Though RFC asks to deploy ingress filtering, still very less number of ISPs have deployed that. Mechanisms like hash based traceback leads to many management issues, which in current scenario doesn’t seem to be working. Mechanisms are there which talks about single packet traceback, but there are lots of overheads for such methods.

References