Efficient Flow Marking IP Traceback System

Sonali H. Mane[1], S. Pratap Singh[2]

[1]ME Student Dept. computer Engineering, IOK College of Engineering, Pune, India

[2]Professor, Project Guide Dept. computer Engineering, IOK College of Engineering, Pune, India

Abstract: ***For controlling the offense on internet the Internet Protocol (IP) traceback technology is used. In this paper we represents a novel and practical IP trace back system, known as Efficient Flow Marking (EFM) for getting IP address of the attacker when IP spoofing technique is used by attacker. The services given by the resources to the justifiable clients Denial of service attack deny that services. The DOS Attacker uses IP spoofing technology for hiding their own uniqueness, for defending against the denial of service attack the first step is to find out IP address of the attacker to take further action. EFM is belongs from the packet marking family of IP trace back scheme. The new characteristics of EFM is that: at first it can fix the length of marking field according to the network protocols set up (flexible mark length strategy); second, by observing the load of the participating router by a flexible flow based marking method the proposed method adaptively change its marking rate. This paper focuses on implementation of EFM on network processor.***

Keywords: EFM; network processor; IP traceback; packet marking; DDoS, PPM, DPM.

1. Introduction

The general definition of IP traceback is that the capability to trace the IP packets to their origins. Existing IP traceback system represents four major groups which are link testing, web management message protocol (ICMP)-based traceback, logging, and packet marking. The Packet marking approaches are describe by adding the traceback data into the IP packets. The mark in the packets can be accustomed surmise the trail of the malicious traffic. Efficient Flow Marking (EFM) applies flexible length mark for holding the IP address segment. When a packet reaches its destination host, EFM reconstruction method reconstruct an IP address that is ingress routers IP address. The benefit of IP address is that a small variety of packets are required to accomplish the process of traceback. The bandwidth outburst has compressed every part our lives and this growth will continue for many years. Most of the networks are demand equipment with high output. Additionally they want the resilience for supporting the new protocol and applications. Network Processor (NP) is useful for its architecture is meant and imposed to satisfy the requirement. Many networks are demanding equipment with high output. The Intel IXP2400 network processor is a member of Intel's second-generation network processor family. The processor is a fully programmable network processor which achieves high-performance process architecture on one chip. Processor contains 9 programmable processors from which one is Intel XScale core and 8 are micro-engines on an equivalent die. The intel Xscale core is an reduced instruction set computing (RISC), it is a computer design architecture machine which compliant with ARM framework. The micro engines are reduced instruction set computing processors optimized for processing the fast path packet. Due to intelligence and adaptability of IXP 2400 network processors permit their customers competently, handle their network resource and bandwidth. EFM encoding process has been enforced on Intel(R) IXP 2400 network processors and that we shall introduce the performance of the EFM on IXP2400.

2. Related Work

Savage suggested probabilistically marking packets as they traverse routers through the internet. They proposed that the router blot the packets with either the router's IP address or the edges of the path from which the packet traverse to reach the router. Accordingly, Song and Perrig propose the following trace-back scheme: instead of encoding the IP address interlaced with a hash value, they recommended that encoding the IP address into an 11 bit hash and maintaining the 5 bit hop count, both store in the 16-bit fragment ID field. With router based way, the router has the responsibility to maintain the information regarding packets which pass through it. Let's take on example, Ram request to log packets and after that data mine them. This is the advantage of being out of band and thus not delaying the fast path. However we proposed Efficient Flow Marking (EFM) I, one of the IP trace-back approaches. It only needs moderately a small number of packets to complete the IP trace-back process. The implementation of EFM in network processor demonstrates that EFM is a good trace-back method and the network processor can defend effectively against DDOS attacks.

Internet Protocol (IP) trace-back is the enabling technology to control Internet crime. In this paper, we propose a novel and practical IP trace-back system called Efficient Flow Marking (EFM) which provides a defense system with the ability to find out the real sources of attacking packets that traverse from networks. There are number of traceback methods are exist, EFM provides innovative features to trace the source of IP packets and can obtain better tracing capability than others. In particular, EFM adopts a flexible mark length strategy to make it compatible to different network environments; it also adaptively changes its marking rate according to the load of the participating router by a flexible flow-based staining method. From the results of both the simulation and real system it concludes that the proposed method require the small number of packet to accomplish.

Probabilistic Packet Marking Schemes

Probabilistic packet marking (PPM) method [6] is one of the parts of the packet marking scheme. The assumption of PPM is that the attacking packets are much more frequent than the usual packet. It scripts the packets with path information in a probabilistic manner and permits the sufferer for reconstructing the attack path by using the marked packet. PPM encodes the information in rarely used 16-bit Fragment ID field in the IP header. To reduce the data that is to be stored in 16 bits, the compressed edge fragment sampling algorithm is used. Although PPM is simple and can support incremental deployment, it has many shortcomings that can seriously prevent it from being widely used. First, for the very high computational work the path reconstruction process is required; mainly when more than one source are exist. With the help of example, a 25 source path modernization will take days, and thousands of false positives could happen [7]. Second, 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 positive. Since, the routers which are long distance from the sufferer have very low chances for pass their recognition to the sufferer for the reason that the information has been vanished because of overwriting by the mediator routers. Many approaches were proposed to overcome the above deficiencies. For example, Song and Perrig proposed an advanced and authenticated PPM based on the assumption that the victim knows the mapping of the upstream routers. It not only reinforces the capability to trace more sources at one time but also solves the problem of spoofed marking. Another method to reduce the overhead of reconstruction was proposed in. It uses counters to complement the loss of marking information from upstream router, for storing the computation time and reducing the false positives. Adler examines the tradeoff among mark bits essential in the IP header and the number of packets essential to restructure the paths.

Deterministic Packet Marking Schemes

An additional stream of packet marking methods, which does not use the above probabilistic assumption and stores the source address in the marking field, it exist in the category termed as the deterministic approaches, such as Deterministic Packet Marking (DPM) [8], [9], our FDPM (the first version of FDPM was published in [10]), and Deterministic Bit Marking. Recently, in [11], the DPM scheme was modified to reduce false positive rates by adding redundant information into the marking fields. Unlike PPM, deterministic approaches only keep the first ingress edge router’s information in the marks (but not the whole path). Moreover, they record marks in a deterministic manner (but not a probabilistic manner as in PPM). This category of schemes has many advantages over others, including simple accomplishment, which does not required any additional bandwidth, and less computation overhead. However, enough packets must be collected to reconstruct the attack path. Importantly, all existing works neither execute well in terms of, nor have addressed the problems of, the maximum number of sources that the trace back system can trace in a single trace back process, the number of packets needed to trace one source, and the overload prevention on participating routers.

3. System Architecture

Flexible Deterministic Packet Marking scheme is novel packet marking IP traceback scheme. It contains two main parts one is encoding scheme another is reconstruction scheme. System architecture of proposed scheme is shown in figure no. It contains sender, Network, Destination machines. In network there is no. of routers. Ingress router is the closest router from sender. Each packet send by sender is passes through ingress router. Each Ingress router is deployed with encoding scheme and each router is deployed with reconstruction scheme. As shown in figure 1, each packet send from sender is marked with marking information at ingress router.

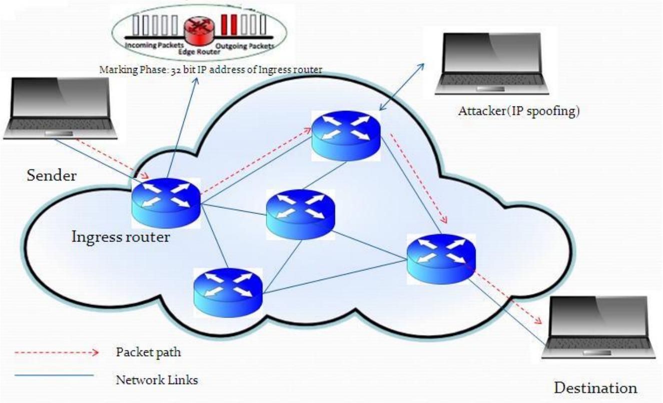
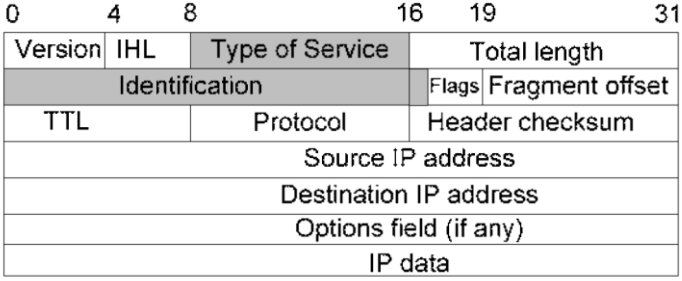


Figure 1: EFM Architecture

(A) Header Utilization for Marking Purpose



0	4	8	16	19	31
Version	IHL	Type of Service	Total length		
Identification			Flags	Fragment offset	
TTL		Protocol	Header checksum		
Source IP address					
Destination IP address					
Options field (if any)					
IP data					

Figure 2: Header Utilization

In proposed method we have to mark an IP address of the source machine from which packets are originate. System needs space to store mark (IP address) in packet header. Proposed scheme will use rarely used fields in the packet header by current network framework. Refer above figure no 2.Type of service is the 8 bit field which denotes what quality of service should be given to the packet. Details of Type of service are discussed in [13]. Support for Type of service is still under work, so we can use Type of service field for marking purpose. Less than 0.25 percent of all Internet traffic is fragments [14], Fragment ID can be safely overwrite without causing severe compatibility troubles. Dealing with the fragmentation problems has been discussed in [15].System can get space of 25 bits (8 +16+ 1) for marking purpose. Reserved bit will be used as flag to show weather system is using Type of Service field or not.

(B) Mark

This scheme is deployed on the ingress router in network. In this scheme, IP address of the source is marked in the packet header of packets. We get maximum space for marking for single packet is 25 bits, so minimum two packets are required to mark 32 bits of IP address. When 32 bit IP address is marked on the two different packets there is need to sequence them for reconstruction, so system will use sequence ID for that purpose. At the time of reconstruction on any router in the network, reconstruction router needs to know which packets are from which router, so each packet must contain such a field which identifies that on which router marking is done. Our system will use, digest for such purpose. Digest is calculated by applying hash function on IP address of the marking router. Our mark for single packet contains sequence number +Digest+ part of IP address of the source.

(C) Encoding Scheme

As per name of the scheme marking information encoded and marked on packet header of each packet at ingress router in this scheme. First system decides the mark length, if network is not using TOS field then system can use TOS field for marking then total marking length will be 24 bits and 1 bit to flag that system is using TOS field for marking. If System is not using TOS field i.e. network is using TOS then mark length would be 16 bits. If system is using TOS then flag would be marked as 0 otherwise marked as 1.If network is using TOS field partially (precedence field using but priority fields not using and vice versa) then mark length would be 19 bits. 2 Bits of the TOS field would be marked as 10 or 01 when TOS is used partially by system and 11 when complete TOS field is used by system. When length decision is executing par alley, digest is calculated using hash function where input is IP address of the marking router. Each packet is marked with sequence number, digest of marking router and part of IP address of the source as shown in figure no 3 and then send randomly using randomly selector.

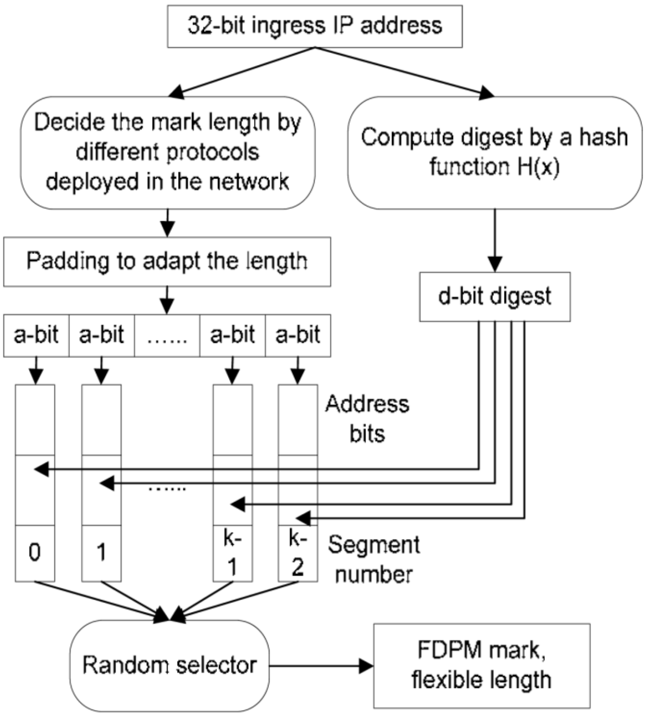

32-bit ingress IP address
Decide the mark length by different protocols deployed in the network
Compute digest by a hash function H(x)
Padding to adapt the length
d-bit digest
a-bit
a-bit
……
a-bit
a-bit
Address bits
……
0
1
k-1
k-2
Segment number
Random selector
FDPM mark, flexible length

Figure 3: Encoding Scheme

(D) Reconstruction scheme

Reconstruction scheme is exact opposite of the encoding scheme, where IP address of the source reconstructed using marks in the packet header. Refer figure no 4.Incoming packets are stored in cache because rate of incoming packets is more than reconstruction speed. First step is recognizing length of the mark. Reconstruction scheme first see RF bit in the header if it is 1 then mark is of length 24 bits. If it is zero then, then system checks 7th and 8th bits of the TOS field, if they are 01 or 10 then mark length is 19 bits and if they are 11 then mark length is 16 bits. Packets of same digest number would be taken in single data structure and after that all packets with same digest number are sorted according to sequence number. Finally IP address of the source is extracted from packets. IF there is double segment number for same digest then they are put in new data structure.

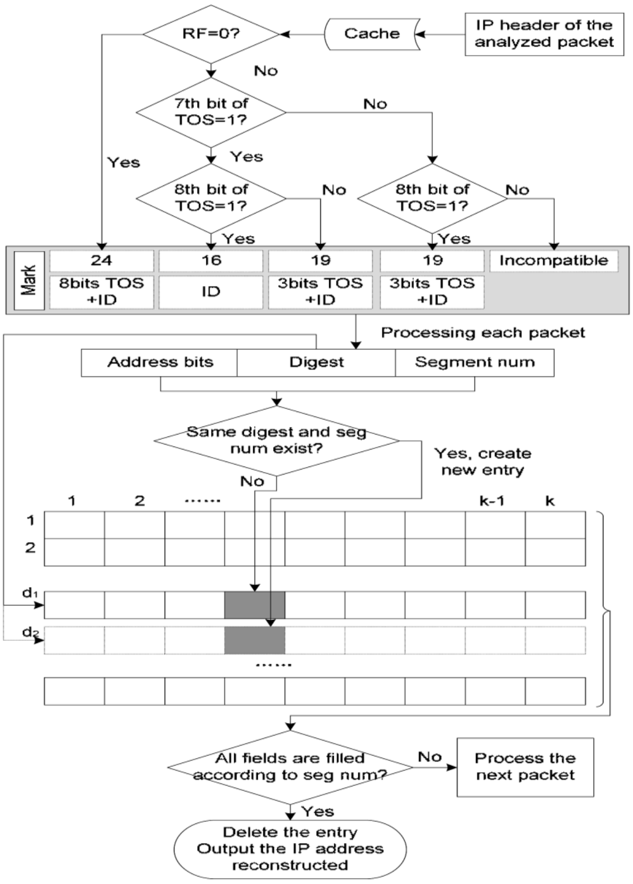


Figure 4: Reconstruction Scheme

4. Modules

(A) File Transfer module:

In this module, we will design basic experiment set up. Our basic experiment set up contains one sender machine, destination machine, 2-3 machines acting as router. In this module we will implement scenario in which, sender sends file to destination by using socket programming. In this module sender selects file to send, and clicks send button on the panel then, packet formation of file takes place, and we can access fields in the packet header. Packets generated are sent along the socket to destination machine. Destination Machine gets file by receiving all packets sent by sender and can get sender machine's IP address by accessing IP header fields of the received packets.

(B) Encoding-Decoding Module:
In this module, Packet marking and decoding IP from marked packets will be covered. IP address of the sender machine will be marked to the packets at the ingress router by using Encoding Scheme. IP address of the sender machine can be retrieved by using Decoding scheme. Marking of the packets will be either 16-bit or32 bit or 19 bit depends on the network. In this module, there is no hacker in to the picture. Normal sender sends file but behind the screen when packets are form marking of IP address of the sender will be mark into the marking fields of the packets. When packets are received at the destination machine IP address of the sender is retrieved by using marks in marking fields in the packet header.

(C) Hacker module:
In this module we will implement attack; how attacker will capture the packets on the path, after capturing the packets how attacker will manipulate those packets. Attacker will capture packets and form file from it then he may modify data in the file, delete data in the file, or only reads file and forwards to the destination and spoofs own IP address with Senders IP address.

(D) Final EFM:
In this module, we will integrate all previous modules and develop final GUI for demo purpose. This module does all necessary remaining work. When we will develop this module, attacker maybe active or maybe not, EFM system will find out IP address of the real sender of the packets without depending on the source IP address fields in the IP header of the packet.

5. Result

	APPM	DPM	EFM
Computational Overhead	Moderate	Low	Very Low
Adaptability According to Network	No	No	Yes
Flow Marking	No	No	Yes
Minimum Packets required to Trace IP	More than DPM	More than EFM	Minimum 4

6. Application

(A) Our work is motivated with for enhancing security of current network system by utilizing rarely used field in the packet header, so our proposed system can be applied to current network system on large scale. This application requires more research and more experimental work before deploy it to the real system

(B) To install proposed method to small private network for any private bank network.

7. Conclusion

In our work, we studied DDoS attack, IP spoofing technique and different available countermeasures for the same. We studied packet marking IP traceback scheme, like Probabilistic Packet Marking, Deterministic Packet Marking scheme. Then we proposed and designed new Flexible Deterministic packet marking scheme which has flexibility of changing mark length as per network protocol deployed.

8. Future Work

In the proposed work, we trace IP address of the attacker which uses IP spoofing for Dos attack; in future our system can get enhanced with block IP address functionality. We proposed our system on IPV4, in future system can enhanced with support for IPv6. In our project work, we will implement system on small scale, future work will be to implement on large scale.

References

[1] en.wikipedia.org/wiki/Denial-of-service attack
[2] TCP/IP spoofing fundamentals, Hastings, N.E. Dept. of Electr. Eng. & Comput. Eng., Iowa State Univ., Ames, IA, USA McLean, P.A.1995
[3] H. Farhat, "Protecting TCP Services from Denial of Service Attacks," Proc. ACM SIGCOMM Workshop Large-Scale Attack Defense (LSAD '06), pp. 155-160, 2006
[4] H. Wang, C. Jin, and K.G. Shin, "Defense against Spoofed IP Traffic Using Hop-Count Filtering," IEEE/ACM Trans. Networking, vol. 15, no. 1, pp. 40-53, 2007
[5] H. Aljifri, "IP Traceback: A New Denial-of-Service Deterrent," IEEE Security and Privacy, vol. 1, no. 3, pp. 24-31, 2003
[6] A. Belenky and N. Ansari, "On IP Traceback," IEEE Comm., vol. 41, no. 7, pp. 142-153, 2003.
[7] S.M. Bellovin, ICMP Traceback Messages—Internet Draft, Network Working Group, 2000
[8] A.C. Snoeren, C. Partridge, L.A. Sanchez et al., "Single-Packet IP Traceback," IEEE/ACM Trans. Networking, vol. 10, no. 6, pp. 721-734, 2002.
[9] C. Gong and K. Sarac, "IP Traceback Based on Packet Marking and Logging," Proc. IEEE Int'l Conf. Comm. (ICC), 2005
[10] S. Savage, D. Wetherall, A. Karlin et al., "Network Support for IP Traceback," ACM/IEEE Trans. Networking, vol. 9, no. 3, pp. 226-237, 2006
[11] A. Belenky and N. Ansari, "P Traceback with Deterministic Packet Marking," IEEE Comm. Letters, vol. 7, no. 4, pp. 162-164, 2003
[12] A. Belenky and N. Ansari, "On Deterministic Packet Marking," Computer Networks, vol. 51, no. 10, pp. 2677-2700, 2007
[13] Type of Service in the Internet Protocol Suite, RFC1349, Network Working Group, 1992.
[14] I. Stoica and H. Zhang, "Providing Guaranteed Services without Per Flow Management," Proc. ACM SIGCOMM '99, pp. 81-94, 1999
[15] A. Belenky and N. Ansari, "On Deterministic Packet Marking," Computer Networks, vol. 51, no. 10, pp. 2677-2700,2007