Ain Shams Engineering Journal (2015) 6, 715–722

Ain Shams University

Ain Shams Engineering Journal

www.elsevier.com/locate/asej
www.sciencedirect.com

ENGINEERING PHYSICS AND MATHEMATICS

IP Traceback through modified Probabilistic Packet Marking algorithm using Chinese Remainder Theorem

Y. Bhavani [a,*], V. Janaki [b], R. Sridevi [c]

[a] Dept of Information Technology, Kakatiya Institute of Technology and Science, Warangal, India
[b] Dept of Computer Science, Vaagdevi College of Engineering, Warangal, India
[c] Dept of Computer Science, Jawaharlal Nehru Technological University, Hyderabad, India

Received 14 September 2014; revised 9 November 2014; accepted 2 December 2014
Available online 20 January 2015

KEYWORDS

DOS attack;
IP Traceback;
Chinese Remainder Theorem;
Modified Probabilistic Packet Marking algorithm

Abstract Probabilistic Packet Marking algorithm suggests a methodology to identify all the participated routers of the attack path by probabilistically marking the packets. In this approach, these marked packets contain partial information regarding the routers of the attack path. At receiver, to get the complete information of every router, it requires more number of marked packets and hence more combinations and more false positives. To overcome this drawback we have presented a novel idea in finding the exact IP address of the routers in the attack path by applying Chinese Remainder Theorem. The result of our implementation reveals that our idea requires less number of marked packets and takes no time in constructing the attack path. The same idea is true even in the case of multiple attackers.

1. Introduction

Information transfer became very easy due to the invention of Internet. The speed of transmission has been tremendously increased and along with this, the attack rate has also grown exponentially. An "attack" is defined as a method of creating obstruction during the transmission of information. Due to the attacks all authorized persons are unable to retrieve the information while unauthorized people are successful in getting the information.

These attacks are broadly categorized as passive and active attacks. Generally passive attacks are difficult to detect but to some extent easy to prevent. Active attacks are difficult to prevent and simple to detect. In the active attacks, one of the most upsetting and very difficult task is to trace the adversary, called DOS attack, in which the legitimate people are unable to access the information. This is due to the intense logging of redundant packets sent by the attacker. This problem can be solved by finding the IP address of the attacker, but the IP

* Corresponding author.
E-mail addresses: yerram.bh@gmail.com (Y. Bhavani), janakicse@yahoo.com (V. Janaki), sridevirangu@yahoo.com (R. Sridevi).

Peer review under responsibility of Ain Shams University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.asej.2014.12.004

address can be spoofed. Hence the best solution in dealing with DOS attacks is to find the attack path from the victim to the adversary [10–13]. The process of constructing the attack path using the information from the received packets is called IP Traceback.

Different techniques have been proposed for IP Traceback [2–6] but they all have short comings, such as limit on their usability in practice. Probabilistic Packet Marking algorithm (PPM) was originally suggested by Burch and Cheswick [1] and later it was designed and implemented by Savage et al. [2] to solve the IP Traceback problem. This PPM algorithm has two procedures one packet marking procedure and second graph reconstruction procedure. In the packet marking procedure, when a router receives a packet, a random number is generated and the packets will be marked by comparing it with the threshold value P_m which is a predefined value. At the victim, graph reconstruction procedure uses these marked packets to construct the graph [9]. This approach called as fragment marking scheme (FMS), and it has a very high computation overhead at the victim to construct the attack path and when an attack originates from multiple sources then the false combination rate is high(false positives) which is a major drawback of this algorithm.

Song and Perrig [3] elucidated an Advanced Marking Scheme (AMS) to rectify the problem of IP Traceback. This technique uses the hash value of IP address to encode the packets rather than the IP address itself. Similar to FMS this has low network and router overhead. This AMS also has lower false positive rate and lower computation overhead at the victim. The major drawback of this algorithm is the victim can construct the attack path only when he has the map of the upstream routers.

Dean et al. [4] proposed an Algebraic Traceback Approach (ATA) which encodes router's IP address as a polynomial in Identification field of IPv4 packet. However this algorithm does not scale in large number of DDoS attacks. Source Path Isolation Engine (SPIE) was explicated by Snoren et al. [5]. This could perform Traceback using just a single packet, however it requires large amount of storage space and hardware changes for packet logging due to which it is not practically deployable.

Kiremire et al. [14] made a comparative study of different PPM algorithms [2–5] in various network topologies. Kiremire et al. [15] explained that three network-dependent factors affect different PPM-based schemes uniquely giving rise to a variation and discrepancy between scheme performances from one network to another.

Lih-Chyau et al. [7] explained an IP Traceback process based on Chinese Remainder Theorem. In this paper the IP address of the routers and remainder values, calculated using Chinese Remainder Theorem, is sent through the marked packets. These packets are used to construct the attack path. When compared to the previous procedures false positive rate has been reduced to some extent but still have more combinations and hence false positives.

In our present paper, we have proposed a novel idea of finding the IP addresses of the attack path, by drastically reducing the comparisons and also false positives, by applying Chinese Remainder Theorem on every IP address. The success rate of finding the IP addresses of attack path is 96.484375 and rate of false positives has been significantly reduced to 3.515625. Our paper is organized as follows. We have explained our proposed procedure for IP Header encoding using CRT in Section 2 and the corresponding Probabilistic Packet Marking algorithm using CRT is presented in Section 3. Method for constructing attack path is elucidated in Section 4. Results are shown in Section 5 and the paper is concluded in Section 6.

2. IP Header encoding using CRT

The DOS attacks can be solved by constructing the attack path there by finding the source router of the attacker. IP Traceback is a technique to find the IP addresses of the routers in the attack path. In this section we explain the technique of IP Traceback by applying Chinese Remainder Theorem.

Chinese Remainder Theorem (CRT) states that

$$\begin{aligned} X &\equiv a_1 \bmod m_1 \\ X &\equiv a_2 \bmod m_2 \\ &\text{---------} \\ &\text{---------} \\ X &\equiv a_k \bmod m_k \end{aligned} \tag{1}$$

where $m_1, m_2, m_3, \ldots, m_k$ are pair wise relatively prime. Then the system of congruences $X \equiv a_i \bmod m_i$ where $1 \leqslant i \leqslant k$ has a unique solution modulo M.

According to our proposed methodology at each router, a unique X value is calculated using CRT for its IP address. The corresponding X values of every router are sent to the receiver by placing the X value in the Identification field of IPv4 packet header. At the receiver, by applying modular inverse on X, the IP address of every router can be deduced. As the value of X is unique, the occurrences of fallacy IP addresses are drastically reduced.

The main aim of the DOS attack is to find the IP addresses of the routers through which the packets are traversed. The address of each router cannot be encoded as a whole due to the limitation of packet header size. The IP packet header format is shown in Fig 1. The 16 bit Identification field of IPv4 is used to store the IP address. As the address of each router is of 32 bits, it cannot be fitted into the identification field as a whole. This forced us to fragment the IP address into 4 equal parts each of 8 bits.

The IP address is split at each dot, into four parts denoted as IP_1, IP_2, IP_3, IP_4 (for example the IP address 192.168.0.1 is split as $IP_1 = 192$, $IP_2 = 168$, $IP_3 = 0$, $IP_4 = 1$). If these fragments are sent as it is (without applying CRT) it will be very difficult to combine the corresponding fragments of each IP

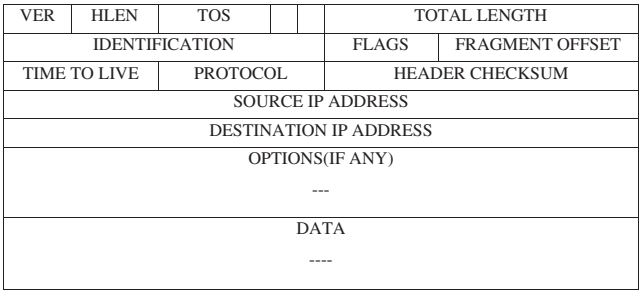

Figure 1 IPv4 packet format.