2012 International Workshop on Information and Electronics Engineering (IWIEE)

A Lightweight IP Traceback Scheme Depending on TTL

Yan Fen*a*, Zhu Hui*a, Chen Shuang-shuang*a, Yin Xin-chun*a.

*aInformation and Engineering College, Yangzhou University, Yangzhou 225009, China

Abstract

This article proposes a new type of IP traceback scheme, through recording TTL of the IP header to get the attack path, when the algorithm is in processing, you do not need to generate additional packages, so attackers could not detect the traceback happened, the scheme can be applied in at any strength of DDoS attack, solves the problem of traditional PPM algorithm failed in high strength DDoS attack, also improves the efficiency of traceback, compare with the other scheme, this scheme is simple, efficiency and robustness. So this scheme can effectively resolve the reconstruction problem.

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of Harbin University of Science and Technology Open access under CC BY-NC-ND license.

Keywords: DDoS attack; reconstruction; IP traceback.

1. Introduction

With the development of the Internet, many attack design bugs at Protocol or operating system. A "Distribute Denial-of-Service" (DDoS) attack is attackers attempt to prevent legitimate users from using that service. Attackers may use one or more of these ways to achieve their aims: (1) bring down the machine’s providing service ability; (2) flood the link between the user and the service provider; (3) utilize resources available at server using bogus requests. DDoS attackers usually hide their actual address by IP spoofing, especially IP source address spoofing. How to tracking and locating the attacker’s positions using received packets, it is IP traceback technology [2]. Though the source address is false, but

* Yan Fen. Tel.: +86-13912126107; E-mail address: yanfen@yzu.edu.cn
each IP packet from the attackers has to go through the routers between the attacker and victim. Using routers to mark packets or record packets, reconstruction attack path, this is IP traceback’s basic idea.

2. RELATIVE WORKS

Early schemes against DDoS attack and IP traceback technology include link way test, inject debugging, log recording, and so on [3]. In recent years, IP tracking research basically divided into "generates extra tracking information packets" and "IP packets marking" two categories. The former increases network bandwidth, hard to upgrade, but it work when DDoS attack happening. The latter has evolved from the "node appending" and "node sampling" to "edge sampling" development. After 2000, Savage, who found "edge sampling", main scheme of IP traceback technology divide into "tracking technology based on log records" and "probability packet marking technology" two camps [4].

2.1 Node appending and node sampling marks

Nodes appending, each router let its own IP address append to the tag packets. The disadvantage is the mark packet length could not be controlled, attackers can add forge path information [3].

2.2 Traceback scheme based on logging

Traceback scheme based on logging require each transfer routers do some Hash operations, record packet in charge of information summary. Under the high speed link, the packet receive speed and recording speed must be match, so the algorithm suitable field is limited [4].

2.3 Based on probability sampling techniques

Traditional probability sampling techniques is suitable for DDoS attack’s traceback. Routers mark packets with a certain probability, mark including two adjacent router’s addresses, it is an "edge" at attack path. Victim machine recovery attack path according to the received tag information. PPM (Probability Packet Marking) has become IP address traceback study’s main direction. Many researchers contributed a lot in PPM.

2.4 An IP Traceback Scheme Integrating DPM and PPM [1]

DPM: Deterministic Packet Marking, each router in the network writes its own IP into IP packets, target host simply queries check this optional field host can known all of the routers. There are two Shortcomings: a. the rapid growth of packet length; b. packet Length is not known.

PPM: PPM is to solve the DPM’s disadvantage, can be further separated into two types: a. Edge sampling: stored two static fields, respectively stored begin and end IP. b. Node sampling: stored a static "node" field in the header, the size is large enough to store the IP address of a router, IPv4 is 32 -bit. When received a packet, each router with a certain probability to write their own IP. Node sampling gets a sample of node, Edge sampling get a sample record adjacent routers.

Article 1 uses a combination of DPM and PPM algorithm, when the attacks occurred, using PPM algorithm, when attack don’t occur using DPM algorithm, reveal it with Fig 1:

![Fig1 combination of DPM and PPM algorithm Icon](image)

When attack occurred, victims host V sent destination host administratively prohibited packets (T1) to the router R1 near source host, through record router’s IP in the attack path to option field to restore attack path. This scheme has some limitations, each router needs to send a T1 package, but also need T1 package success send back. This scheme is difficult to success in the strong DDoS attack, and if it success
transmission packet back to T1, the T1 packet round-trip time is difficult to determine, this bring an unknown time complexity, in practice, the drawbacks are obvious.

3 A lightweight IP traceback scheme depending on TTL

3.1 Assumptions

A represents an attacker. R represents a node corresponding to a IP address, V represents the victim host, TTL_{HI} is the number of TTL in IP option field, TTL_L is the TTL value when the packet reaches V, IPS represents starting field of IP, IP_E represents the end domain IP. These are premise-based assumptions:
(1) Allow attackers to work together or use zombie machines;
(2) Link is stable over a short period of time;
(3) The routers in the attack path not all have marker function;
(4) The attacker may be aware victim running the IP traceback scheme.

3.2 marking field select

This scheme stores tagging information in the IP head. Only the IPv4 format header analysis in this paper. We use the literature 3 point of view, the tag information field is available. This article uses IP header ID field of the first 8bit, IP optional field the first 64bit, include IPS, IP_E, As Fig 2:

\[
H = \text{TTL}_{HI} - \text{TTL}_L \tag{1}
\]

From References 1 we known, H is the value of the hops R from V, victim host stored such a tree:

3.3 package marking and traceback scheme

If the attack link is Fig3, In order to find the attackers, run Algorithm 1:
(1) When any node receive a package, and the node is not the destination address of the packet, with a certain probability p writes its IP into IPS field, also copy the node's IP to IP_E, in this algorithm, the IPS is the sample edge’s beginning IP, in the same time, copy this package’s TTL to the ID field of the first 8bit recorded as TTL_{HI}, transmit the packet.
(2) Any node found that a packet’s IPS and IP_E is the same, put its own IP to the packet IP_E field, this address is sample edge’s end IP, named IP_E, and transmit the packet.

When V’s packets receiving speed is larger than a given number, attack happens, run Algorithm 2.
Algorithm 2:
(1) In △t time, collect marking packages from each node then store the packages.
(2) In algorithm 1, record the start address of each packet IPS, end address IP_E, and record the packet arrived TTL, record as TTL_L, with the ID domain TTL_{HI} minus TTL_L, set the value H, there are:

\[
H = \text{TTL}_{HI} - \text{TTL}_L \tag{1}
\]

From References 1 we known, H is the value of the hops R from V, victim host stored such a tree:
4. Scheme Explanation

a. V is the root to the node; b. For any node is concerned, H is the node’s floor number, there can be Fig 3; c. For any received mark packet, there are two ordered IP as a path which is the sampling edge IP$_S$, IP$_E$, and then use these paths connect the storage nodes, record the marking number sending by this node; d. After the above treatment there can be Fig 4.

(3) Execute the algorithm 3, analysis the resulting attack tree, and then run the fourth step.

(4) According to the conclusions of the analysis, the victim host sends destination host administratively prohibited packets (ICMP) to the nodes near A, these nodes will abandon the packets which destination address is V.

Algorithm 3:
Each number on node in Fig 4 represents the marked packet number sent from the node to V, in the $\Delta t$ after victim host judge it suffer attack, run the attack tree analysis algorithm, which is algorithm 3:

(1) Let all the node label number in ascending order: mp is the marked points, C is the number of its son, MC$_i$ is the number of children marked number, then: $MC_1 \geq MC_2 \geq \ldots \geq MC_C$

\[ \text{if } mp < \sum_{i=1}^{C} mc_i \text{ decrease } mc_1, mc_2, \ldots, mc_i \]
\[ \text{until } mp = \sum_{i=1}^{C} mc_i \]

(2) In a floor, if a node’s marked number is smaller than its brother number’s $1 / \alpha$, delete the subtree which this node is root, called it the safety factor $\alpha$, $\alpha$ between 5 to 10 the number is reasonable, R$_9$ can be deleted.

(3) If a node is already a leaf, and its marking number is less than $1 / \beta$ of average marking number of the same level, delete it, we call $\beta$ safety speed factor $\beta$. Reasonable range of $\beta$ is 2 to 3.

(4) The remaining leaves are adjacent with A. In Fig 4 they are R$_{10}$ and R$_{11}$;

(5) According to node level with decreasing order: if a node has marking number greater than double of its children marking number sum, this node is also adjacent with A. In Fig 5 it is R$_5$. 

4.1. Security Analysis

Fig 6: no marking function router eg.
Reference 2 suggested that there may have some no marking function router on network, in this case, how to skip no marking function routers to mark the next hop has become a hard problem, this paper checks whether two marker are the same to find out whether the current received this package is close to the router "the next hop", so this is a good solution to the link case of no marking function router exist, as shown in Fig 6.

If R₆ is no marking function router, R₁₁ and R₂ have marking function, when the R₁₁ mark a packet, R₁₁ put its own IP into IP₅ and IP₆, when the packet is transmit through R₆, when arrives R₂, R₂'s IP to be written in IP₆, so you get the sample edge, the modified topology shown in Fig7.

This approach removes no marking function router in the new topology, in order to achieve all the routers are marking router results.

5. Performance analysis

Tₐ represents the time from an attack is detected until find the router adjacent A, it has three parts:

(1) \( \Delta t \): V collects enough packets to reconstruct the path of the time required;
(2) The time required to reconstruct attack graph;
(3) The time required to run the algorithm 3.

The following are analysis if three steps time required for.

(1) For step 1: \( A_i \) as a particular attacker, \( R_n \) represent the packages need to analyze, \( R_A \) is the packets sent by \( A_i \):

\[
R_n = R_A \cdot p \cdot (1 - p)^{d-1} \tag{2}
\]

Assume \( A_i \) generate packages with speed \( p \), there are

\[
\Delta t = \frac{R_A}{P} = \frac{R_n}{P \cdot p} \cdot (1 - p)^{d-1} \tag{3}
\]

\( R_n \) is determined by \( V \), the router can minimum \( p \) by set the \( \Delta t \):

\[
f(p) = p(1 - p)^{d-1} \tag{4}
\]

Its derivative:

\[
f'(p) = (1 - p - p^d + p)(1 - p)^{d-2} = (1 - p^d)(1 - p)^{d-2} \tag{5}
\]

When \( p = 1/d \), \( f(p) \) reached the maximum value, \( t \) reached the minimum value.

In the Internet environment, \( d = 20 \) is widely used to set the formulas 6:

\[
\Delta t = \frac{2650}{p} \tag{6}
\]

In this case, you should set \( p = 0.05 \), you can get \( R_n = 50 \) by former equations.

(2) For step 2, in most cases is less than one second, and, this step could be process with step 1 in the same time, considering the time complexity can be ignored.

(3) For step3, algorithm 3 only involves a single cycle execution time, it should less than one second.

In general, \( T_L \) is mainly determined by the \( \Delta t \), that is, mainly to collect enough marked packets time required. \( \Delta t \) depends on \( p \). When \( A \) is relatively small, in order to make the attack more significant, \( p \) is relatively large, so \( T_L \) is could be less than 10 seconds. When a very large number of an attackers, \( T_L \) might be tens of second or more. In other words, this in any strength of DDoS attacks can be successfully achieved the desired results, especially to have a high efficiency in the attacker’s distribution is not well.

6. Conclusion

We researches the problem of path reconstruction in DDoS attack, and through running routing filter algorithm to defence DDoS. The scheme absorbed the advantages of traditional IP tracking program, not only reach the effect (rapid response, when suffer DDoS attack scheme will not fail, could work in include-no marking function router network, etc.) of traditional IP tracking program, makes router load smaller, improving the efficiency, also has good robustness, and can effectively defence DDoS attacks.
References


