通过自适应随机数据包标记实现实时IP回溯*

梁 丰1+, 赵新建1, David Yau2

1(浙江工业大学 浙江省光纤通信技术重点实验室,浙江 杭州 310014)
2(普度大学 计算机科学系,印第安那州 IN47907,美国)

Real Time IP Traceback with Adaptive Probabilistic Packet Marking

LIANG Feng1+, ZHAO Xin-Jian1, David Yau2

1(Zhejiang Provincial Key Laboratory of Fiber Optical Communication Technology,Zhejiang University of Technology,Hangzhou 310014,China)

2(Department of Computer Science,Purdue University,West Lafayette IN 47907,USA)

+ Corresponding author: Phn: 86-571-88320562, E-mail:liangf@zjut.edu.cn

http://www.zjut.edu.cn

Received 2002-01-07; Accepted 2002-08-27

Liang F,Zhao XJ,Yau D.Real time IP traceback with adaptive probabilistic packet marking *Journal of Software*,**2003,14(5);1005～1010**
http://www.jos.org.cn/1000-9825/14/1005.htm

Abstract: Probabilistic packet marking (PPM) is a practical and effective method for IP traceback of denial-of-service(DoS) attack.In this paper,an adaptive PPM algorithm is presented:a router marks a passing packet with a probability which is adaptive to the distance that the packet has traversed,so that a minimum convergence time for an attacking path can be achieved in the victim.With a new IP header overloading scheme,the labeled fragment encoding scheme,a real-time reconstruction is provided,so that thousands of paths can be traced simultaneously.Compared with previous PPM schemes,a 50% decrease in convergence time is achieved,while the computation overhead and false positives in reconstruction are greatly reduced

Key words: network security;DDoS;router;IP traceback;PPM

摘 要: 随机数据包标记(PPM)是对拒绝服务攻击进行 IP 回溯的一种实用而有效的方法.提供了一种自适应的 PPM 算法:一个路由器按一个与路过的数据包已传输距离自适应的概率标记该数据包,从而被攻击者可以以

* Supported by the Natural Science Foundation of Zhejiang Province of China under Grant No.602112(浙江省自然科学基金)；the Center for Education and Research in Information Assurance and Security (CERIAS);the National Science foundation of US under Grant No.CCR-9875742(CAREER)

LIANG Feng was born in 1967.He is a professor and master supervisor at the Department of Telecommunication,Zhejiang University of Technology.His research interests are network security and multimedia communication.ZHAO Xin-Jian was born in 1955.He is an associate professor and master supervisor at the Department of Telecommunication,Zhejiang University of Technology.His research interests are network security and mobile communication.David Yau was born in 1967.He is an assistant professor and doctoral supervisor at the department of Computer Science,Purdue University.His current research areas are network,operating system architectures,algorithms for Qos provisioning,multimedia communication,software-programmable router technologies,database and information system.

最短的收敛时间重构一个攻击路径.通过一个新的称为标注片段编码的 IP 重载方案,实现了实时的重构,从而能同时回溯数千条路径.与以前的 PPM 方案相比,收敛时间减少了 50%,同时大大减少了重构计算量和伪证性.

关键词: 网络安全;分布拒绝服务;路由器;IP 回溯;随机数据包标记

中图法分类号: TP393 **文献标识码**: A

Denial-of-Service (DoS) attack has become one of the most serious problems in Internet security. With the advance of distributed DoS (DDoS)[1], the victim list even includes large sites such as yahoo, eBay and Amazon. In a typical DoS, an attacker sends a great amount of packets to a victim (e.g. a server), so that the victim has to deny service to legitimate users because of resource shortage. However, the harm made by DoS attack is limited by the capability of the single attacker, and the attacker is somehow easy to be found as it must send an extremely larger traffic than normal users. On the other hand, a DDoS attack, which origins from hundreds or thousands of attackers simultaneously, can overload any victim with the aggregation of many relatively small traffic flows.

While many efforts on defending DoS attack focus on improving the victim 's resistance[2~4], IP traceback[5], which tries to locate the origin of the attack by probabilistic packet marking (PPM), gains more and more interests. PPM is simple and effective when dealing with single attacker DoS, but when the number of attackers increases, the number of false positives and the computation overhead of the reconstruction increase extremely fast. So it is absolutely ineffective and impractical in case of DDoS. To settle these problems, an advanced scheme is proposed by Song[6], where a router is represented by a set of hashes of its IP Address. This scheme results almost no false positives for less than 1500 attackers, and takes less than 100 seconds to reconstruct the attack paths. However, an extremely big and dynamic map of Internet must be maintained for the reconstruction, and the computation overhead of the reconstruction is still too big for an online performance.

Notice all previous works on PPM are based on the assumption that only the attacking packets are traced. However, although most available DDoS tools have observable signatures in packet contents nowadays, it's not difficult to make the attacking packets no different from normal ones. So all legal and illegal packets have to be checked, which means thousands of paths have to be traced simultaneously. This makes the false positive and computation overhead problem of PPM be very serious even for only several attackers.

In this paper, we improve PPM for real-time traceback of DDoS attack: An adaptive PPM algorithm is introduced to minimize the convergence time; a labeled fragment encoding scheme is provided to keep the rate of flows from thousands of paths under surveillance and find DDoS attackers with less false positives and computation overhead.

1 System Model

Assume during a certain short time period, there is a set of attackers, $V=\{a_1,a_2,\ldots,a_n\}$, sending attacking flows to a victim server v. The transmission path from a_i to v, P_i, is called the attacking path of a_i. Notice one attacker may have several attacking paths due to Internet's connectionless nature, however, we can assume that only one path is used in a short time-scale (seconds). All attacking paths $\{P_i\}$ form a directed tree $T=(V,E)$ (see Fig.1), where E is the set of edges. The leaf nodes of T are a_i. The root of T is v. Other nodes of T, r_i, are routers. P_i can be represented by an ordered list of edges, e.g., $P_2=(e_7,e_5,e_2,e_1)$. The length of P_i, l_i, is the number of the edges in it. An edge e_i is represented by an ordered list such as $e_1=(r_1,v)$, $e_2=(r_2,r_1)$, and $e_4=(a_1,r_2)$. A packet traverses an edge from its initial node (the first one in the

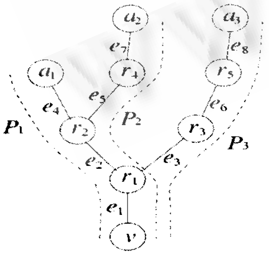

Fig.1 The attacking paths form a tree

list) to its terminal node (the second one). The distance from e_i to v, d_i, is denoted as the number of edges from e_i to v, and we say e_i is d_i hops away from v. Notice that e_i may belong to more than one attacking paths.

The task of IP traceback is to get T or $\{P_i\}$. As all nodes of T are connected, T can also be simply represented by E[7]. So the task of IP traceback becomes getting E.

2 Adaptive PPM: Minimization of the Convergence Time

The convergence time is defined as the minimum number of packets from a path that the victim needs to observe before messages of all edges of the path arrive[5]. It's a very important metric of PPM for two reasons: First, if the number of packets from a host is less than the convergence time of the transmission path, the path can't be whole reconstructed; Secondly, as the reconstruction process of PPM can't start until all edges' messages are available, so the convergence time also represents delay of the reconstruction process. For a DDoS attack, the traffic rate of one attacker is much smaller than that of single attacker DoS, and there can be dynamics of attacker distribution, e.g. groups of attackers send attacking traffic on shift. So minimization of the convergence time is significant for finding the attackers in real time.

2.1 Estimate of the convergence time of PPM

The expectation of the convergence time N of PPM for a single path is bounded by $E(N) < \dfrac{k\log(kl)}{p(1-p)^{l-1}}$, where l is length of the path, k is either the number of fragments for one edge in the fragment scheme[5] or the threshold value in the hash scheme[6]. N has its minimum when $p=1/l$. As length of a transmission path for Internet traffic is rarely longer than 25, $p=l=25$ is usually selected in PPM. For any path with length little than 25, N is only a little bigger than the minimum.

2.2 APPM: Minimization of convergence time

Assume $E=\{e_i|i=1,2,\ldots,m\}$. An edge e_i is represented by k messages, and every message is taken by a packet to v with equal probabilities. Computation of the convergence time of E falls in the *COUPON COLLECTOR PROBLEM WITH UNEQUAL PROBABILITY*. The expectation of the convergence time of E is[8]:

$$E(N)=\int_0^\infty [1-\prod_{i=1}^{m}(1-\exp(-q_i t))]\mathrm{d}t\ , \tag{1}$$

where q_i is the probability of a packet taking a message of e_i to v. When $q_i=1/km$ for all e_i, $E(N)$ has its minimum:

$$\min(E(N))=\sum_{i=0}^{km-1}\frac{km}{km-i}=km\log(km)+O(km)\,. \tag{2}$$

Assume e_i is d_i hops away from v. Let p_i denote the marking probability of e_i on a packet passing it. We have

$$q_i=\frac{N_i p_i}{kN}\prod_{j}^{e_j\in P_i, d_j<d_i}(1-p_j)\,, \tag{3}$$

where N_i is the number of packets passing e_i, N is the total packets from all paths, P_i is the path from e_i to v. So $E(N)$ reaches its minimum when

$$\frac{N_i p_i}{N}\prod_{j}^{e_j\in P_i, d_j<d_i}(1-p_j)=\frac{1}{m},\forall e_i\in E\,. \tag{4}$$

In PPM, $p_i=p$ for all e_i. As N_i, N and P_i are different for different edges, Eq.(8) can not be satisfied, and $E(N)$ cannot reach it is minimum.

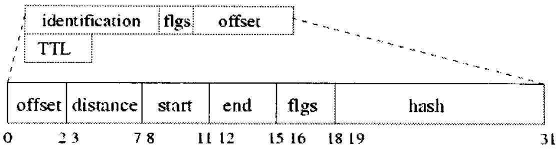

Fig.2 IP header overloading scheme of APPM

To satisfy Eq.(8), p_i must be adaptive to different N_i, N, P_i and other edges' marking probabilities. However, as N is related to the number of packets from every leaf, m is related to the topology of the tree, and N_i to either, all of them are unknown to marking routers, there is no way for an edge to pre-modulate its marking probability satisfying Eq.(8). However, we can achieve the minimum convergence time for a single path of E independently, which is still significant as what we are really interested is not topology but the attacking paths.

Assume a path P has l edges, the expectation of convergence time of P has its minimum

$$\min(E(N))=\sum_{i=0}^{lk-1}\frac{lk}{lk-i}=lk\log(lk)+O(lk) \qquad (5)$$

```
Algorithm 1. Marking procedure at router $r_i$
    for each packet $w$
        let $x$ be a random number from [0,1]
        let $o$ be a random integer from [0,7]
        let $f$ be a fragment of $r_i$ at offset $o$
        let $h$ be a fragment of $r_i$ at offset $w.offset$
        if $w.TTL > 64$
            write 64 to $w.TTL$; $p$=1;
        else
            $p = 1/(65-w.TTL)$
        fi
        if $x < p$
            write $f$ into $w.start$
            write 0 into $w.distance$
            write o into $w.offset$
            write Hash($r_i$ ) into $w.hash$
        else
            if $w.distance$=0
                write $h$ into $w.end$
            fi
            increment $w.distance$
        fi
    end for
Algorithm 2. Reconstruction procedure at $v$
    let $T$ be a tree with root $v$
    let an edge of $T$ be a tuple ($start$, $end$, $count$)
    let $E$ be a two dimension array of the tuples
    for each packet $w$
        replace the fragment at offset $w.offset$ of
          $E[w.distance][w.hash].start$ with $w.start$
        replace the fragment at offset $w.offset$ of
          $E[w.distance][w.hash].end$ with $w.end$
        increment $E[w.distance][w.hash].count$
    end for
```

Fig.3 Marking and reconstruction procedure of

$$\text{when } q_i=\frac{p_i}{k}\prod_{j}^{e_j\in P,d_j<d_l}(1-p_j)=\frac{1}{lk},\forall e_i\in P. \qquad (6)$$

Theorem 1. Equ.(11) is satisfied when p_i=1/(l–d_i+1) for all e_i∈P.

Proof. Be omitted.

According to Theorem 1, to achieve the minimum of a path's convergence time, we introduce an adaptive PPM (APPM) method: every router marks a passing packet with a probability p_i=1/(l_k–d_i+1).

3 IP Header Overloading: The Labeled Fragment Encoding Scheme

As in PPM[5,6], APPM overloads the 16bit identification field and the 13bit offset field of IP header by 5 subfields, which are shown in Fig.2. This will not work if the field is already used for packet fragmentation. Discussions on this topic and a complementary scheme can be found in Ref.[5].

Distance a 5bit field to record distance of an edge from v;

Start a 4bit field to record a fragment of the IP address of the initial node of an edge;

End a 4bit field to record a fragment of the IP address of the terminal node of an edge;

Offset a 3bit field to record the offset of the two fragments above.

Hash a 13bit hash of the IP address of the initial node of an edge.

Figure 3 shows the marking procedure. The IP address of a router splits into 8 4bit fragments. A packet's *TTL* field is used for computing the marking probability, *p*. As most packets are sent with a default initial value of *TTL* as 64, we make $p=1/(65-w.TTL)$. Here we assume the decrement of *w.TTL* is occurred after the marking procedure. The case of a packet's *TTL* with an initial value less than 64 will be discussed in next section.

Figure 3 also shows the reconstruction procedure too, which is actually a storing procedure. A two dimension array *E* is used for storing edges. Every element of *E* is an edge, which is represented by a tuple (*start*, *end*, *count*). *start* stores IP address of the edge's initial node, *end* stores IP address of the terminal node, and *count* records the arrival rate of packets which bring messages of this edge. Both *start* and *end* are composed of 8 fragments, and the fragmentations of *start* and *end* are just the same as in routers for marking process. One dimension of *E* is the distance from an edge to *v*, another dimension of *E* is 13bit hash of IP address of the initial node. As *w.start*, *w.end* and *w.offset* are used together as a label to locate which fragment of which edge in *E* the message from the packet should be written, our scheme is called labeled fragment scheme. Whenever a packet is recorded, the *count* of corresponding edge in *E* is incremented. After a fixed time interval, *count* can be checked as traffic rate from that edge, then be reset to 0.

4 Evaluation and Discussions

Convergence time: For a path of *l* edges, there are *kl* kinds of edge messages, where *k* is the number of fragments for one IP address. If the packets are sent along the path with an initial *TTL* value t_0=64, then by APPM, a packet takes every message to *v* with same probability $1/kl$, and the convergence time of the path reaches its minimum, $kl\log(kl)+O(kl)$. However, if $t_0 < 64$, the expectation of the convergence time of the path becomes

$$E(N)=\frac{\frac{kl\log(kl)}{l}}{64-t_0+l}+O(kl)=(64-t_0+l)k\log(kl)+O(kl)\,. \tag{7}$$

Figure 4 shows *E*(*N*) of APPM with different t_0 and of PPM with $p = 1/25$ over various path lengths. Notice for t_0=8 and 16, the curves are shorter than others, because t_0 can't be smaller than the path length, otherwise the packet will be dropped before arriving the destination. It shows that when the packets are sent by default t_0=64, APPM takes less than 50% of the convergence time of PPM. Notice that if t_0<64, the convergence time of APPM increases. So the attackers can use a small t_0 for their attacking packets to eliminate our effort on minimizing the convergence time. However, we argue here that the attackers will not do this, as the philosophy of the attackers is to make the attacking packets undistinguishable from normal ones, and using a small t_0 is just against this philosophy.

Reconstruction time and false positives: All matching work and hash computation of previous schemes is avoided in APPM, so the reconstruction time is proportion to the convergence time, and there's no false positives at all if the labels are unique for one edge.

Reconstruction error from hash collisions: If more than one routers have the same distance-hash, messages of these edges will be recorded in the same position of *E* in the reconstruction procedure, then collision occurs. Assume there are *m* paths in *T*, then for any particular distance *d*, the maximum number of distinct routers is *m*. For a random hash of length *h*, the expectation numbers of routers has the same hash is $m=2^h$. So if the hash are perfect random functions, our scheme can trace 2^{13} paths with less error.

Forged marking: For every kinds of packet marking scheme, there is a potential weakness: An attacker can write the marking field to confuse *v* if the marking routers are comprised. One possible way of against forged marking is authentication. The hash of IP address is replaced by a Message Authentication Codes (MAC) of the IP address. The MAC is encrypted with a secret key. Detail discussion on this can be found on Ref.[6].

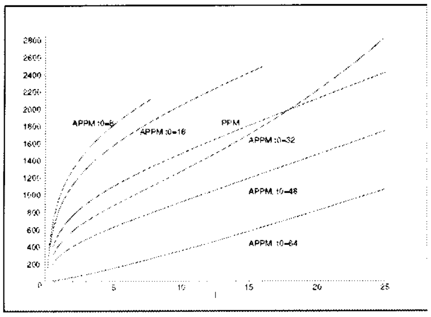

Fig.4 The convergence time comparison between APPM and PPM

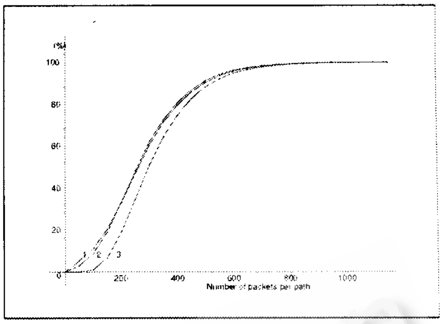

Fig.5 Simulation results for speed of convergence and reconstruction

5 Simulation Result

The total convergence and reconstructed time for all paths is a function of the topology of the tree and the amount of packets from each path as discussed before. Figure 5 shows simulation results of the percentage of edges of a tree being reconstructed over number of packets received per path. Three trees are tested, each of them has 2000 leaves with same depth *l*=15. The packets are coming from all leaves with same rate. The topology of the first and the second trees is abstracted from paths selected from real trace-route paths from a single source[9]. The third tree is one that all 2000 paths are separated from each other. Here we suppose a perfect hash function is used.

References:

[1] Anderson D. Distributed denial of service Attacks (DDOS). 2000. http://wind.lcs.mit.edu/ dga/ddos.txt.

[2] Banga G, Drusched P, Mogul J. Resource containers: A new facility for resource management in server systems. In:OSDI ed. Proceedings of the 1999 USENIX/ACM Symposium on Operating System Design and Implementation(OSDI'99). New Orleans, LA: OSDI, 1999. 45~58.

[3] Spatscheck O, Peterson L. Defending against denial of service attacks in scout In: OSDI, ed. Proceedings of the 1999 USENIX/ACM Symposium on Operating System Design and Implementation(OSDI'99). New Orleans, LA: OSDI, 1999. 59~72.

[4] Meadows C. A formal framework and evaluation method for network denial of service In: PCSFW, ed. Proceedings of the 1999 IEEE Computer Security Foundations Workshop. Mordano IEEE Computer Society Press, 1999. 4~13.

[5] Savage S, Wetherall D, Karlin A, Anderson T. Practical network support for IP traceback. In: ACM, ed. Proceedings of the ACM SIGCOMM 2000. Sweden: ACM, 2000. 295~300.

[6] Song D, Perrig A. Advanced and authenticated techniques for IP traceback In: INFOCOM, ed. Proceedings of the IEEE INFOCOM 2001. Anchorage: INFOCOM, 2001.

[7] Mayeda W. Graph Theory. NewYork: Wiley-Interscience, 1972.

[8] Klamkin M, Newman D. Extensions of the birthday surprise. Journal of Combinatorial Theory, 1967. 279~282.

[9] The Internet mapping project at AT&T. http://cm.bell-labs.com/who/ches/map/dbs/index.html.