Efficient Dynamic Probabilistic Packet Marking for IP Traceback

Jenshiuh Liu, Zhi-Jian Lee
Dept. of Information Eng. and Computer Science
Feng-Chia University
Taichung, Taiwan 407
ROC
{liuj,lee}@iecs.fcu.edu.tw

Yeh-Ching Chung
Dept. of Computer Science
National Tsing Hua University
HsingChu, Taiwan 300
ROC
ychung@cs.nthu.edu.tw

***Abstract*— Recently, Denial-of-service (DoS) attack has become a pressing problem due to lack of efficient method to locate the real attackers and easy to execute with readily available source codes on the Internet. Traceback is a subtle scheme to tackle the DoS attacks. The probabilistic packet marking (PPM) is a new way for practical IP traceback. Although the PPM enables a victim to pinpoint the attacker's origin to within 2-5 equally possible sites, it have been shown that PPM suffers from uncertainty under attack with spoofed packets. In this work, we present a new approach, called dynamic probabilistic packet marking(DPPM), to further improve effectiveness of PPM. Instead of using a fixed marking probability, we propose to deduce how far a packet has traveled and then choose the marking probability as an inverse function of hop count traveled. The DPPM may remove uncertainty completely and enable victims to precisely pinpoint attacking origin under DoS attacks. Our proposed DPPM can be applied to DDoS attacks with a very limited uncertainty.**

***Index Terms*— Denial of service, IP, Network security, Probabilistic packet marking, Traceback.**

I. INTRODUCTION

In recent years, Denial-of-service (DoS) attack has become a pressing problem on the Internet [1]. As opposed to other types of attacks, DoS attacks do not alter, delete or steal information stored on victims' computer systems, but prevent legitimate access to services normally provided by victims. In February 2000, Yahoo, was held back by DoS attacks over twelve hours. Many DoS attacks (e.g., eBay, Amazon, and other .com sites) occurred before and after Yahoo's event. DoS attacks have become more prevalent recently due to lack of efficient method to locate the real attackers and easy to execute with readily available source codes on the Internet. Research work observes that there are 12,805 attacks on over 5,000 distinct hosts belonging to more than 2,000 distinct organizations during a three-week period[2]. Even worse, recent reports indicate that hackers have developed tools to coordinate attacks from many separate sources simultaneously. This is so called distributed denial-of-service(DDoS) attack [3], [1]. As the Internet attracts more and more applications, coping with DoS becomes an important issue.

Most work on solution to DoS attack has been along the following two directions. One is to tolerate attacks by mitigating their effect on the victims[4], [5], [6]. The other is to attempt to locate the origin of attacks and hopefully to stop the attacks at the source. The process to identify the machines that directly generate attack packets and the network paths these packets follow is called the *traceback* problem [7]. Traceback is a subtle scheme to tackle the DoS attacks. Because, if it could provide us with precise attacking origin, then we may apply some proper action to stop attacks completely; even incomplete or approximate information is valuable, since applying packet filtering the closer to the attacking source the more we are able to control and contain attacks.

It is surprisingly difficult to determine the origin of attacks in the Internet due to its characteristics of IP routing: each packet is routed to its destination independently, and moreover, attackers routinely disguise their origin using incorrect or "spoofed" address in the IP source address field. Much research work has been done on the traceback problem [8], [9], [7], [10], [11]. Recently, Savage *et. al.* [7] have proposed the *probabilistic packet marking* (PPM) as a network support for practical IP traceback. In their work, each router probabilistically marks packets with path information as they pass by. By collecting certain number of packets, a victim is able to identify the network path(s) traversed by attack traffic without requiring interactive assistance from outside network operators. The traceback is a game between the victim and the attacker. Under the PPM, the victim may raise the marking probability in order to collect path information with least number of packets. On the other hand, the attacker may choose spoofed marking value and spoofed source address to lessen the effectiveness of PPM. Park and Lee[12] have shown that PPM suffers from uncertainty under attack with spoofed packets, which may impede traceback by the victim. Their interesting findings are as follows. With the PPM, the victim can pinpoint the attacker's address to within 2-5 equally possible sites under single source DoS attack. However, under DDoS attack, the uncertainty introduced by the attacker will be amplified significantly, which may diminish the effectiveness of PPM.

In this work, we present a new approach, called *dynamic probabilistic packet marking* (DPPM), to further improve effectiveness of PPM. The DPPM may remove uncertainty completely. With support from hosts and routers in the Internet community, the DPPM enables any victim to precisely

This work is sponsored in part by a grant from National Science Council of ROC under contract no. NSC91-2213E-035-031.

0-7803-7788-5/03/$17.00 ©2003 IEEE.

Authorized licensed use limited to: The Chinese University of Hong Kong CUHK(Shenzhen). Downloaded on November 11,2024 at 07:26:37 UTC from IEEE Xplore. Restrictions apply.

pinpoint attacking origin under the DoS attack.

The rest of this paper is organized as follows: SectionII contains an introduction to PPM and some issues with it. We present our DPPM and its implementation issues in SectionIII. Performance analysis is given in SectionIV. Finally, summary and remarks are given in SectionV.

II. PRELIMINARIES

One feature of the Internet Protocol(IP) is that the source host itself fills in the IP source address field before it sends the packet. This would permit anonymous attacks, which has been long understood. Packet marking[9], [7]is one way to enable IP traceback, where each router puts some path information as packets are forwarded to their destinations. By collecting certain number of packets, a victim is able to identify the network path(s) traversed by attacking packets. Probabilistic packet marking(PPM) is one of the most prominent method for traceback in DoS attack. In this section, we will briefly review the PPM.

A. Probabilistic Packet Marking

A traceback can be divided into *marking* and *reconstruction* phases. During the marking phase, each router marks, with some probability p, packets with path information as they pass by. The victim performs the reconstruction, where it uses the path information recorded in packets to create a network graph leading back to source or sources of the attack. Node append, node sampling and edge sampling are three different schemes for recording path information. Savage *et. al.*[7] proposed to use edge sampling and distance for path information. Interested readers should refer to [7] for more details.

B. Issues in Choosing Probability

Traceback is a game between attackers and victims. Attackers may use spoofing and may limit the number of attacking packets to hide their identity. On the other hand, victims may choose proper marking scheme to pinpoint the attacker(s). In the following, we will argue that it is very difficult for victims to determine a proper marking probability for efficient PPM, since many issues are involved.

Consider an attack path $A = (a, r_1, r_2, \cdots, r_D, v)$, where a and v denote the attacker and victim of a DoS incident, $D+1$ is the distance between them, and $r_i (i = 1, 2, \cdots, D)$ denote D routers in the attack path.

1) At least one marking per router: Let p_i represent the marking probability of router r_i. Define *leftover* probability for router r_i, denoted by α_i, to be the probability that an attacking packet is lastly marked at router r_i and nowhere further down the path. Thus,

$$\alpha_i = p_i \times \Pi_{j=i+1}^{D}(1 - p_j). \quad (1)$$

All routers have a fixed probability p for marking in the PPM. By Eq. 1, we have $\alpha_i = p(1-p)^{D-i}$. Therefore, the leftover probability is geometrically smaller the closer it is to the attacker.

The victim must collect at least one marking from each router along the attack path in order to construct a path to the source. Let N denote the total number of attacking packets (attack volume) from an attacker to a victim. One constraint for successful traceback by PPM is

$$N\alpha_1 = Np(1-p)^{D-1} \geq 1. \quad (2)$$

Figure 1 shows the values of α_1 (leftover probability for r_1) with respect to p and D. It can be seen that α_1 is a bell

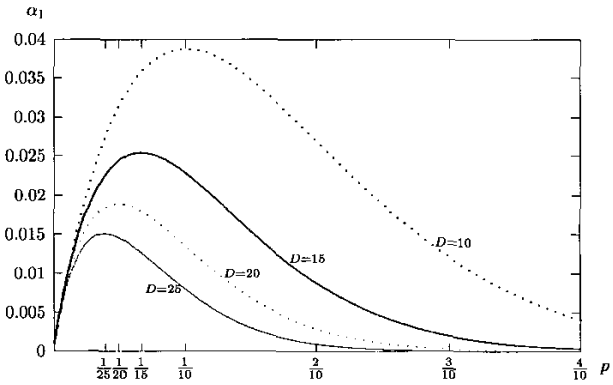

Fig. 1. Leftover probability(α_1) for r_1.

shape function of p. Moreover, one can show that its peak value occurs at $p = 1/D$. Since D is usually not known to victims, it is difficult for users to determine the optimal marking probability in advance.

2) Spoofed packets: The probability that a packet reaching the victim without any marking is $\alpha_0 = (1-p)^D$. Attackers may spoof the marking field with false value in order to hide themselves or the attack path. If a packet is not marked by any router along the path, the spoofed packet may result in false information during the path reconstruction. Figure 2 shows the unmarked probability (α_0) for a packet with respect to p and D. It clear that α_0 is a decreasing function of p.

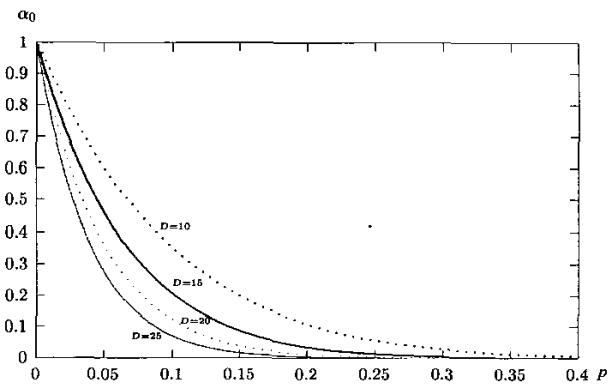

Fig. 2. Unmarked probability(α_0).

3) Uncertainty: Spoofed packets also introduce *uncertainty* in traceback, which was first studied by Park *et. al*[12]. Figure 3 helps us to grasp the idea of uncertainty. Consider

Authorized licensed use limited to: The Chinese University of Hong Kong CUHK(Shenzhen). Downloaded on November 11,2024 at 07:26:37 UTC from IEEE Xplore. Restrictions apply.

an attack path $\mathcal{A} = (a, r_1, r_2, \cdots, r_D, v)$. The attacker a may spoof its marking field with the edge (u_1, r_1). If this spoofed packet is not marked by any router along the attack path, during the traceback the victim may conclude that u_1 is a source of attack. Similar conclusion may be drawn for

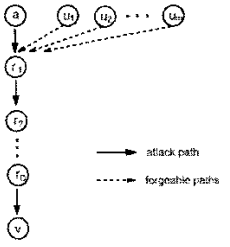

Fig. 3. Forgeable paths.

$u_2, u_3, \cdots$ and u_m. Hence, the traceback may give m false sources of attack in addition to the real one. Park *et. al.*[12] have defined m as the *uncertainty* factor, which is a function of p. To maximize entropy, an attacker sends spoofed packets with equal probability as from m different sources. For a given marking probability p, the maximal uncertainty factor of a DoS attack is shown[12] to be $m = \frac{1}{p} - 1$.

We can see that uncertainty decreases as p increases. Therefore, users tend to choose large p in order to diminish the impact of spoofed packets. In fact, the victim may choose $p = 1$ to try to completely eliminate the uncertainty. However, the traceback would not work under this condition, since in such a condition all packets arriving at the victim would bear the marking of the last node on the attack path.

The uncertainty is a key issue in the traceback problem. Two determinant factors of uncertainty are marking probability (p) and the total number of attacking packets(N). The victim tries to minimize uncertainty with a large p, while the attacker tries to maximize it with various spoofed packets. Moreover, as we mentioned, the at least one marking per router constraint(Eq. 2) also affects the victim's choice of p. The larger p the victim chooses, the larger N is required for successful traceback. However, the attacker controls the value of N. A min-max model has been used to study the uncertainty problem in [12].

Analyzing PPM's uncertainty in DDoS attack is more complicated. Given a desired attack volume N, the attacker may mount M separate attacks each with N/M packets in a DDoS attack. Even without spoofed packets, the victim needs to process M attack paths. With spoofed packets, uncertainty may be amplified by DDoS attack. It is shown in [12] that the larger M, the higher the amplification. Thus, PPM has very limited application in the case of DDoS attack due to amplification of uncertainty.

III. Dynamic Probabilistic Packet Marking

The PPM uses a fixed probability for marking. As we have seen in Section II-B that smaller p would enable traceback with smaller attack volume(N). However, smaller p would lead to larger uncertainty. The major cause of this conflict is due to uneven leftover probability for routers along the attack path. In the following, we will present a new packet marking scheme based on dynamic probability, called dynamic probabilistic packet marking(DPPM), where the marking probability of a packet is dynamically determined as a function of how far the packet has traveled.

One approach to minimize the number of packets required for successful traceback is to have an uniform leftover probability for all routers. Additionally, the uncertainty introduced by spoofed packets can be removed completely if every packet got a legitimate marking along the path. Our proposed DPPM meets both of these conditions.

To achieve an uniform leftover probability, routers should decrease the marking probability as a packet traveling along the path. Instead of a fixed p, in the DPPM, each router uses different marking probability to mark packets. A router chooses a high marking probability if the packet is just sent out from its source. On the other hand, a route chooses a low marking probability if the packet is far away from its source. More precisely, the DPPM works in the following way. For a given attack path, let $i (1 \leq i \leq D)$ be the distance of a packet w from its source. Router r_i chooses a marking probability $p_i = 1/i$ to mark packet w.

One question needs to be answered is: for each arriving packet how can the route determine its distance from the source? We will answer this in the following section.

A. Determination of distance

Our dynamic marking scheme is based on the distance of a packet from its source(origin). One challenge we have to face is how to determine the distance for each packet. Our solution lies in the Time-to-live(TTL) value in the IP header. The TTL serves two purposes. It limits the lifetime of an IP datagram, and it also terminates an internet routing loops. The source node of a packet sets the TTL to a default initial value, which is system and protocol dependent. As the packet travels through routers on the network, each router decrements TTL by one[13]. Routers drop any packet with a zero in its TTL field. If a router knows the initial TTL value of a packet, then the distance of that packet from its source could be calculated accordingly. One question remains: how can a route find out the initial TTL value for every packet passing by? A formal solution to this is to require all hosts use the same initial TTL value as suggested by assigned numbers in [14]. For all systems to comply with this may take time. Although systems may use different TTL values on different protocols, reports[15], [16] show that most initial TTL values fall in the set of $\mathcal{S} = \{32, 64, 126, 255\}$. Recent studies [17], [18], [19] show that there are very few packets will pass through 25 routers. Hence, the most likely initial TTL value for a packet with a TTL value of 47 is 64, additionally, this packet is most probably at a distance of 17(64-47) from its origin. More precisely, a router can determine the initial TTL value of any packet in the following way. Let t be the TTL value of a packet. Its initial TTL value should be the least value in set $\mathcal{S}$ that is equal to or greater than t.

Authorized licensed use limited to: The Chinese University of Hong Kong CUHK(Shenzhen). Downloaded on November 11,2024 at 07:26:37 UTC from IEEE Xplore. Restrictions apply.

B. Leftover probability and Uncertainty

Leftover probability is a good index of effectiveness. We now proceed to study it for DPPM. The leftover probability is computed as in Eq. 1:

$$\begin{aligned}\alpha_i &= p_i \cdot \Pi_{j=i+1}^{D}(1-p_j) \\ &= p_i \cdot (1-p_{i+1}) \cdot (1-p_{i+2}) \cdots (1-p_D) \\ &= \frac{1}{i} \cdot (1-\frac{1}{i+1}) \cdot (1-\frac{1}{i+2}) \cdots (1-\frac{1}{D}) \\ &= \frac{1}{D}. \end{aligned} \quad (3)$$

Eq. 3 shows that each router along the attack path has the same probability to leave its information in the marking field. In other words, the victim has an equal probability to obtain each router's information along the path despite their distance from the victim. This is a subtle feature of our DPPM. We have seen in Section II that spoofed packets may introduce uncertainty in PPM. There are D routers in the attack path. Each one has a leftover probability of $1/D$. Therefore, the unmarked probability for any packet under the DPPM is zero, *i.e.*, $\alpha_0 = 0$. This implies that there is no uncertainty in DPPM, since each packet got a legitimate marking. Therefore, the effectiveness of our DPPM will not be affected by spoofed marking.

C. Implementation

The only difference between PPM and DPPM is the determination of marking probability. Since implementation issues of PPM has been studied extensively. We will move our focus to the IPv6 environment and point out some issues in implementing the DPPM.

The distance of a packet is determined by its TTL value. The TTL in IPv6 has a different name: *hop limit*, which serves the same purpose and works the same way as its counterpart. Savage *et. al.*[7] proposed to overload the *identification* field in IP header for the marking. There is no identification or similar field available in IPv6 basic header to carry marking. However, IPv6 supports extension headers for additional functionality[20]. Hop-by-Hop options header is one choice to carry the marking information. Options in Hop-by-Hop extension header are represented in type-length-value(TLV) format[20]. Once we define a new option type for DPPM. All marking information can be encoded into the TLV format.

IV. Performance Analysis

In this section, we will compare our DPPM to the PPM. All our findings indicate that the DPPM is superior to the PPM.

A. Minimal number of packets required for traceback

To satisfy the requirement of at least one marking per router, a victim needs to collect certain amount of packets. The expected minimal number of packets required for successful traceback, denoted by N_{min}, depends on the leftover probability. As we saw in Eq. 2 that $N_{min}p(1-p)^{D-1} \geq 1$ for PPM. Thus, for a fixed p and D, PPM needs

$$N_{min} \geq \frac{1}{p(1-p)^{D-1}}. \quad (4)$$

On the other hand, we learned from Eq. 3 that $\alpha_i = 1/D$ for all routers on the attack path. Therefore, we have $N_{min} = D$ for DPPM. Table I displays some numerical values of N_{min} for different setups of PPM and DPPM. It is clear that DPPM always needs less number of packets to get its job done. The difference gets bigger if we prefer a low uncertainty(high p) from PPM.

TABLE I
MINIMAL NUMBER OF PACKETS REQUIRED BY PPM AND DPPM

p	D			
(PPM)	10	15	20	25
0.01	109	115	121	127
0.02	59	66	73	81
0.04	36	44	54	66
0.06	29	39	54	73
0.08	26	40	60	92
0.1	25	43	74	125
0.2	37	113	346	1,058
0.3	82	491	2,924	17,399
0.35	137	1,188	10,246	88,310
DPPM	10	15	20	25

B. Uncertainty

It is shown in[12] that the maximal uncertainty for PPM is $m = 1/p - 1$. We saw in Section III-B that the unmarked probability is zero and hence there is no uncertainty in the DPPM. This suggests that the DPPM would enable us to pinpoint the exact attacker under DoS attack. On the other hand, the PPM may give few sites for possible attacker if spoofed packets present.

C. Overhead of routers

Each marking poses some cost to a router. We now compute the overhead(with respect to no marking) of PPM and DPPM. It seems that each marking by DPPM costs more than by PPM, since the DPPM needs to find the distance of a packet in order to determine its marking probability. However, this is not necessarily true, because routers have to examine and decrease TTL by one for each arriving packet. The distance of a packet(hence, marking probability) can be obtained by a table lookup when routers examine the TTL. Therefore, there is very little cost difference between a marking by DPPM and by PPM. Both can be achieved in about the same time. For simplicity, we choose to use number of markings performed as our overhead measurement.

Consider a DoS attack with D routers between the attacker and the victim. We examine two types of overhead. The *individual* overhead is the cost experienced by each route along the attack path. The *total* overhead is the cost summed over all D routers.

In PPM, each router uses a fixed probability p to mark packets. If there are N packets in a DoS attack, the individual overhead for each router is Np. On the other hand, in DPPM, router r_i uses a probability of $1/i$ to mark packets, where $i = 1, 2, \cdots, D$. Therefore, the individual overhead for router r_i is N/i. Figure 4 compares the individual overhead of PPM and DPPM, where $N = 100,000$, $D = 25$ and $p = 0.35$.

Authorized licensed use limited to: The Chinese University of Hong Kong CUHK(Shenzhen). Downloaded on November 11,2024 at 07:26:37 UTC from IEEE Xplore. Restrictions apply.

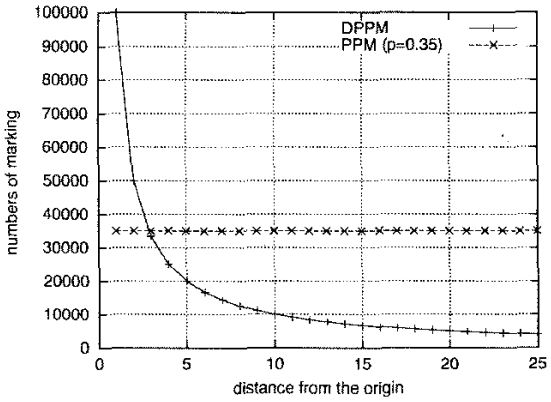

Fig. 4. A comparison of individual overhead.

It can be seen that all routers in PPM have the same individual overhead. On the other hand, the first two routers in the DPPM suffers a very high overhead. However, it drops very rapidly. The high overhead for the first two routers will be an advantage of the DPPM, since any router experiences a sudden surge of workload may indicate some sort of DoS attack. Hence, proper action can be taken at the first place.

The total overhead is the price paid by all routers along the attack path. In general, the total overhead is a cost index for the Internet community to perform marking. Let $O_{ppm}(O_{dppm}$, respectively) denote the total overhead for PPM(DPPM, respectively). There are D routers in an attack path. Therefore,

$$O_{ppm} = NpD.$$

For the DPPM, the total overhead is obtained by summing D terms:

$$\begin{aligned} O_{dppm} &= N(\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{D}) \\ &= NH_D, \end{aligned} \quad (5)$$

where H_D is the D^{th} harmonic number. Table II compares the total overhead of PPM and DPPM, where the overhead has been normalized to the number of times (frequency) each packet will be marked in the network. An asterisk(*) in the table entry indicates that the at least one marking per router constraint can not be satisfied at that condition. The uncertainty factor of PPM also depends on p. The last column of Table II shows the uncertainty for our reference.

Some interesting points deserve our attention. For a fixed D, O_{ppm} increases as p increases, whereas O_{dppm} remains fixed. Small p could lower O_{ppm}. However, it also introduces large uncertainty. This would present a challenge to PPM applications. In general, O_{ppm} is large than O_{dppm} if we want the uncertainty to be less than 2.33(*i.e.*, $p \geq 0.3$).

D. Distributed DoS attack

Given a desired attack volume N, an attacker may mount a DDoS attack from M different sites each with a volume of N/M. Although this may be considered as M separate DoS attacks, two issues deserve our further attention. In general, users of PPM will choose a large p to ease the uncertainty

TABLE II

NORMALIZED TOTAL OVERHEAD OF PPM AND DPPM.

p (PPM)	D				m
	10	15	20	25	
0.04	0.4	0.6	0.8	1	24
0.05	0.5	0.75	1	1.25	19
0.067	0.6	1	1.2	1.5	14
0.1	1	1.5	2	2.5	9
0.2	2	3	4	5	4
0.3	3	4.5	6	7.5	2.33
0.35	3.5	5.25	7	8.75	1.85
0.42	4.2	6.3	8.4	*	1.38
0.56	5.6	8.4	*	*	0.78
0.721	7.21	*	*	*	0.38
DPPM	2.92	3.31	3.59	3.81	0

problem. For a crafty attacker, he/she would choose to mount a DDoS attack from a large number of sites in order to hide him/herself. Under the PPM, with a fixed probability p and attack volume N, the attack volume from each site (N/M) may fall below the constraint of at least one marking per router (N_{min}). Therefore, the PPM may not be able to get its job done if its marking probability is not adjusted accordingly and properly. Adjusting p properly with network condition is a challenge to PPM's application. On the other hand, the DPPM performs much better, since its N_{min} is less than that of PPM(cf. Table I) and no attention is needed to adjust marking probability.

As we just saw that PPM needs to lower p in order to combat with DDoS attack. This also introduces the amplification[12], which amplifies the uncertainty factor with respect to the DoS attack. It is shown in [12] that the amplification could be up to 20 if the attack path length is sufficiently large. Since there is no uncertainty in DPPM, the DPPM suffers no amplification in the DDoS attack.

E. Challenge on spoofed TTL value

We have learned that the performance of DPPM will not be affected by spoofed marking field. A subtle attacker may observe one possible weakness of the DPPM. The marking probability of any packet is completely determined by its distance from its origin, which is equivalent to the TTL value of that packet. By spoofing the initial TTL value of a packet, any attacker may beat the DPPM. For example, by sending all packets with TTL values of 129, a crafty attacker would definitely get away without any trace, since the router would deduce that those packets are at a distance of 126(=255-129) from their origin and marks these packets with a probability of $1/126$. In the following we will discuss how to prevent this and some other issues on spoofed TTL attack.

Routers examine and decrease TTL value by one when they forward packets[13]. To defeat spoofed TTL attack, we propose an *unified* initial TTL value, denoted by T_{ini}. Any route sees a packet with a TTL value greater than T_{ini} should rewrite it as T_{ini} and mark this packet as it is at a distance of 1 from its origin. If all systems comply with the unified initial TTL value, only attackers or compromised routers could put a TTL value greater than T_{int}. The best way to handle such a packet is to view it as one hop away from some attacker. A

Authorized licensed use limited to: The Chinese University of Hong Kong CUHK(Shenzhen). Downloaded on November 11,2024 at 07:26:37 UTC from IEEE Xplore. Restrictions apply.

good choice for T_{ini} could be 32 or 64, since studies show that most applications on the Internet are within this limit[17], [18], [19]. In the following, we will show that DPPM still provides a uniform leftover probability and suffers very little uncertainty under spoofed TTL attack.

1) Leftover probability and Overhead: With the unified initial TTL value, attackers still may try to beat the DPPM by sending packets with spoofed TTL values. However, they should choose TTL values less than T_{ini}, otherwise it will be corrected as we mentioned before. Assume that an attacker sends a packet with a spoofed TTL value that is z ($0 \leq z < T_{ini}$) less than the T_{ini}, *i.e.*, TTL$= T_{ini} - z$, through attack path $\mathcal{A}$. All routers can not distinguish such a packet between normal ones. Router r_1 views this packet as originated at $1+z$ hops away and marks it accordingly with a probability of $p_1' = 1/(1+z)$. Similarly, router r_i ($1 \leq i \leq D$) will mark this packet with a probability of

$$p_i' = \frac{1}{i+z}. \tag{6}$$

The leftover probability for routers can be calculated similarly as in Eq. 3. Hence, we have

$$\alpha_i' = \frac{1}{D+z}. \tag{7}$$

Eq. 7 indicates that we still have uniform leftover probability under spoofed TTL attack. Moreover, the expected minimal number of packets required for successful traceback, denoted by N_{min}', can be obtained immediately from Eq. 7. Thus, we have

$$N_{min}' = D + z,$$

which is a small increase.

By Equations 6 and 7, both marking and leftover probability decrease under the spoofed TTL attack. This is a gain for DPPM, since both individual and total overhead decrease. However, we have to pay this with a presence of uncertainty. It can be shown that DPPM suffers an uncertainty of z under the spoofed TTL attack.

2) Uncertainty: The unmarked probability is no longer zero under DoS with spoofed TTL values. By Eq. 7, the sum of all leftover probability is $D/(D+z)$. Hence, the unmarked probability is $z/(D+z)$. This suggests that our DPPM will suffer from spoofed marking. Recall that the leftover probability for each route is $1/(D+z)$. To maximize entropy, the attacker may spoof markings as from z different sites(cf. Figure 3). Therefore, DPPM may suffer a uncertainty of z under the spoofed TTL attack, where z is the difference between T_{ini} and the spoofed packet's TTL value. It is clear that an upper bound for uncertainty is T_{ini}. However, a more tight bound for uncertainty is $T_{ini} - D$, since no attacking packet can reach the victim if its initial TTL value is set below D.

V. Concluding Remarks

Traceback is a subtle scheme to tackle the DoS attacks. The PPM opens a new avenue for practical IP traceback. Although the PPM enables a victim to pinpoint the attacker's origin to within 2-5 equally possible sites, it have been shown that PPM suffers from uncertainty under attack with spoofed packets. The uncertainty can be amplified significantly under distributed DoS attack, which may diminish the effectiveness of PPM.

In this work, we present a new approach, called DPPM, to further imporve effectiveness of PPM. The DPPM may remove uncertainty completely and enable victims to precisely pinpoint attacking origin under DoS attacks. Our proposed DPPM can be applied to DDoS attacks with a very limited uncertainty. A subtle feature of the DPPM is that it allows incremental deployment. Implemntation issues on IPv6 is dicussed. Formal analysis indicates that the DPPM outperforms the PPM in all respects.

References

[1] Computer Emergency Response Team, "CERT Advisory CA-2000-01 Denial of Service Developments," January 2000, http://www.cert.org/advisories/CA-2000-01.html.

[2] D. Moore, G. M.Voelker, and S. Savage, "Inferring Internet Denial-of-Service Activity," in *In Proceedings of the 10th USENIX Security Symposium*, August 2001, pp. 9–22.

[3] Computer Emergency Response Team, "CERT Incident Note IN-99-07 Distributed Denial of Service Tools," July 1999, http://www.cert.org/incident_notes/IN-99-07.html.

[4] G. Banga, P. Druschel, and J. Mogul, "A New Facility for Resource Management in Server Systems," in *USENIX/ACM Symposium on Operation System Design and Implementation*, February 1999, pp. 45–58.

[5] O. Spatscheck and L. Peterson, "Defending Against Denial of Service Attacks in Scout," in *1999 USENIX/ACM Symposium on Operating System Design and Implementation*, February 1999, pp. 59–72.

[6] Cisco Systems, "Configuring TCP Intercept (Prevent Denial-of-Service Attacks)," *Cisco IOS Documentation*, 1997.

[7] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Network support for IP traceback ," *IEEE/ACM Transactions on Networking*, vol. 20(2), pp. 226–237, June 2001.

[8] H. Burch and B. Cheswick, "Tracing Anonymous Packets to Their Approximate Source," in *2000 USENIX LISA Conference*, December 2000, pp. 319–327.

[9] T. Doeppner, P. Klein, and A. Koyfman, "Using router stamping to identify the source of IP packets," in *Proceedings of the 7th ACM Conference on Computer and Communications Security*, November 2000, pp. 184–189.

[10] S. M. Bellovin, "ICMP Traceback Messages," March 2000, http://www.research.att.com/~smb/papers/draft-bellovin-itrace-00.txt.

[11] A. C. Snoeren, C. Partridge, L. A. Sanchez, C. E. Jones, F. Tchakountio, B. Schwartz, S. T. Kent, and W. T. Stayer, "Single-Packet IP Traceback," *IEEE/ACM Transactions on Networking*, vol. 10, pp. 721–734, December 2002.

[12] K. Park and H. Lee, "On the Effectiveness of Probabilistic Packet Marking for IP Traceback under Denial of Service Attack," in *2001 IEEE INFOCOM Conference*, June 2001.

[13] J. Postel, "RFC 791: Internet Protocol," Sept. 1981.

[14] J. Reynolds and J. Postel, "RFC 1700: ASSIGNED NUMBERS," Oct. 1994.

[15] T. S. Education and R. Network, "Default TTL Values in TCP/IP," 1999, http://secfr.nerim.net/docs/fingerprint/en/ttl_default.html.

[16] O. Arkin and S.-S. Group, "ICMP Usage in Scanning," 2001, http://www.sys-security.com/html/projects/icmp.html.

[17] R. Carter and M. Crovella, "Server Selection Using Dynamic Path Characterization in Wide-Area Networks," in *IEEE INFOCOM Conference*, April 1997.

[18] W. Theilmann and K. Rothermel, "Dynamic Distance Maps of the Internet," in *Proceedings of the 2000 IEEE INFOCOM Conference*, March 2000.

[19] Cooperative Association for Internet Data Analysis, "Skitter analysis," 2000, http://www.caida.org/tools/measurement/skitter/.

[20] S. Deering and R. Hinden, "Internet Protocol, Version 6 (IPv6) specification," December 1998, RFC 2460.

Authorized licensed use limited to: The Chinese University of Hong Kong CUHK(Shenzhen). Downloaded on November 11,2024 at 07:26:37 UTC from IEEE Xplore. Restrictions apply.