International Symposium on Information Theory and its Applications, ISITA2006
Seoul, Korea, October 29–November 1, 2006

An IP Traceback Scheme with Variably Probabilistic Packet Marking

Takeaki Terada†, Masakazu Soshi‡ and Atsuko Miyaji‡

† Fujitsu Laboratories Ltd.
4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki-shi,
Kanagawa 211-8588 Japan
E-mail: tterada@labs.fujitsu.com

‡ School of Information Science
Japan Advanced Institute of Science and Technology
1-1 Asahidai, Nomi, Ishikawa 923-1292 Japan
E-mail: {soshi,miyaji}@jaist.ac.jp

Abstract

One of the effective countermeasures against Denial of Service (DoS for short) attacks is IP traceback, which tries to identify the attacker by determining the attack path. In particular *Probabilistic Packet Marking* (PPM) for IP traceback is so promising that many researches on it have so far been done. However, most of the previous PPM schemes have the disadvantage such that they require a large number of packets to reconstruct the attack path. Therefore we propose an efficient traceback scheme based on PPM in this paper.

Our scheme has a novel feature that routers on the attack path mark packets with *variable probabilities*. In this paper, we show this mechanism can reduce the number of packets for path reconstruction significantly.

1. Introduction

In recent years Denial of Service (DoS for short) attacks have posed serious threats on the Internet. The attacker in an DoS attack sends the victim server a large number of packets and tries to disturb the services the victim provides to legitimate users. Therefore it is urgent to solve the problems caused by DoS attacks.

The countermeasures proposed so far against DoS attacks can be roughly classified into two types. One is *filtering* (for example, [7]), which drops excessive packets to alleviate congestion caused by the DoS attack, and the other is *(IP) traceback*, which tries to identify the attacker by determining the path the packets of the DoS attack followed (*attack path*). In this paper we focus our attention on the latter approach, in particular, probabilistic packet marking schemes.

Probabilistic Packet Marking (PPM) is the most effective technique of the traceback methods. In PPM, each router on the attack path probabilistically marks packets with its own information and then the victim reconstructs the path by using the information on the packets. PPM is so promising that many researches on it have so far been done. However, most of the previous PPM schemes have the disadvantage such that they require a large number of packets to reconstruct attack paths [1, 2, 3, 6]. The main reason for this is that the previous schemes are not sufficiently flexible, for instance, the packet marking probabilities are fixed in them. Therefore in order to counter these problems, we propose an efficient traceback scheme based on PPM in this paper.

Our proposed scheme is different from the previous work in that in our scheme routers on the attack path mark packets with *variable probabilities*. As we shall show later in this paper, this mechanism can reduce the number of packets for path reconstruction significantly.

The paper is organized as follows. In section 2, we discuss the related work to ours and point out the problems in the previous approaches. In order to overcome the drawbacks of them, we propose a new IP traceback scheme in section 3. Next in section 4 we evaluated our proposed scheme and conclude this paper in section 5.

2. Related Work

Before going into details of our approach, this section discusses some related work.

Savage et al. propose two fundamental forms of PPM, that is, node sampling and edge sampling [6]. Furthermore, they also discuss how to accommodate attack path information in the IP header of a packet. Their work is seminal, although, it is disadvantageous because it requires a large number of packets to reconstruct attack paths.

Al-Duwairi et al. reduced the number of the required packets by making each router prepare a buffer for storing information for the path reconstruction [2]. If a packet includes the information for the reconstruction, then each router marks the packet after storing the information in its buffer. Otherwise, it only

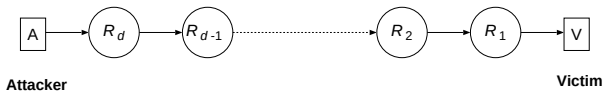

Figure 1: Assumed Attack Path

does marking without saving the information. Via the buffers, the information once stored on packets by a router certainly reaches the victim server without being overwritten by other routers, which reduces the number of the necessary packets for the path reconstruction. However, the obvious disadvantage of the scheme is that the routers suffer from heavy load by the buffer processing in the presence of DoS attacks, which results in severe performance degradation.

In summary, what we need is efficient PPM schemes with requiring no buffers on routers to counter DoS attacks.

3. Our Approach

In this section we propose efficient IP traceback schemes with variably probabilistic packet marking.

3.1. Basic Idea

First of all, in this section we present an basic idea behind our approach.

The main reason for the inefficiency of previous PPM schemes is that in a DoS attack, the more distant from the victim a router is located, the higher is the probability that the information marked on packets by the router is overwritten by other succeeding routers on the attack path. So a large number of packets are required for recovering information put by further routers.

In order to solve this problem, we propose an efficient PPM scheme, where routers on an attack path mark packets with *variable probabilities*. Especially in our scheme further routers from the victim mark packets with higher probabilities than nearer ones, so that a less number of the packets are required for the path reconstruction.

3.2. Proposed Scheme

Let us assume that there exist d routers between the attacker (host) A and the victim server V, which is shown in Figure 1. Let p_i be the probability with which router R_i marks packets ($i = 1, 2, \ldots, d$). Here note that p_i is a function of the distance i from the victim. Furthermore let q_i be the probability that information marked by router R_i reaches the victim without being overwritten by other succeeding routers on the attack path. Then the following recurrence equations hold:

$$\begin{aligned} q_1 &= p_1, \\ q_2 &= p_2(1-p_1), \\ q_3 &= p_3(1-p_2)(1-p_1), \\ &\vdots \\ q_d &= p_d(1-p_{d-1})(1-p_{d-2}) \\ &\quad \cdots (1-p_2)(1-p_1). \end{aligned} \tag{1}$$

Now solving for p_i, we obtain:

$$p_i = \frac{q_i}{1 - \sum_{j=1}^{i-1} q_j}. \tag{2}$$

where p_i needs to satisfy $0 \leq p_i \leq 1$, which in turn implies that

$$0 \leq \sum_{j=1}^{i} q_j \leq 1,$$

for $1 \leq i \leq d$. Note that we can determine p_i from q_i and vice versa.

The path reconstruction procedure at the victim is just the same as in [6] and so refer to it for more details.

Now let us discuss the form of p_i or q_i ($1 \leq i \leq d$). In this paper we shall mainly pay attention on q_i, since the functions for q_i are directly connected with the number of the required packets for path reconstruction, as will be demonstrated later in this paper. Furthermore, by paying attention only on q_i, we can drastically simplify the computation of the number of required packets.

We can consider various candidate functions for q_i, for example, $\frac{c}{\ln(i+1)}$, $\frac{c}{(i^2+1)}$, and so on. In this paper, we present only one of them, that is,

$$q_i = \frac{c}{i} \tag{3}$$

for some c. The reasons for this choice of q_i are given as follows. First, the form $q_i = c/i$ is essential and simple. Second, it results in better performance than the previous work. Third, even if q_i should have another form, we could analyze the behavior of our scheme easily as is just done bellow.

4. Analysis

In this section we estimate the required number of packets for path reconstruction for our proposed scheme.

4.1. General Analysis

First of all, we evaluate the number of required packets for attack path reconstruction in general. For such analysis, the one given by Savage et al. in [6] is well known. That is, an upper bound of the required packets for path reconstruction of PPM (with the same marking probability p for all routers) is given by $\frac{\ln d}{p(1-p)^{d-1}}$. However, the analysis is conducted in a very rough manner and it is inappropriate for our analysis of PPM performance with variable marking probabilities. Fortunately we can overcome the problem of the analysis in a simple way.

With a little thought we can easily see that from Eq. 1 an upper bound of the number of the necessary packets N for path reconstruction for PPM with variable packet marking probabilities is given by:

$$N \leq \sum_{i=1}^{d} \frac{1}{q_i}. \tag{4}$$

For example, we consider the scheme given by Savage et al. [6] In this case, because the marking probability p is the same for all routers, we have $q_i = p(1-p)^{i-1}$ $(1 \leq i \leq d)$. Therefore from Eq. 4 for upper bound N_S we obtain

$$\begin{aligned} N_S &\leq \sum_{i=1}^{d} \frac{1}{q_i} \\ &= \sum_{i=1}^{d} \frac{1}{p(1-p)^{i-1}} \\ &= \frac{1-(1-p)^d}{(1-p)^d} \cdot \frac{1-p}{p^2} \\ &\approx \frac{1}{p^2(1-p)^{d-1}}. \end{aligned} \tag{5}$$

4.2. Analysis of our approach

Now we shall evaluate the number of the required packets N_V for our scheme. At this moment it should be easy to perform the analysis.

As presented in section 3.2, we have $q_i = c/i$ for $1 \leq i \leq d$. Since $\sum_{i=1}^{d} q_i$ must be one, we find that c should be $\frac{1}{\ln d}$ approximately. By using this value of c, we can easily obtain the expected number of packets required for path reconstruction in our scheme.

Next, we evaluate our scheme through the comparison with previous schemes with respect to the number of required packets. However, the accurate comparison is difficult because a certain scheme uses variable marking probability, while another not. For the comparison under the same condition, we set up the probability of each scheme according to the following policy.

- Savage scheme: $p_i = 1/10$, which is the given value in [6].
- Al-Duwairi scheme: We set $p_i \simeq 0.17$, which is the mean of the marking probability of our scheme. In this case, the probability u succeeding in reconstructing a path is $u = 0.999$.
- Our scheme.

In the above-mentioned items, we evaluate each scheme, as depicted in Figure 2. From this graph, we find that our scheme is better than Savage scheme in general and that it reduces the number of the packets within the range $2 \leq d \leq 6$ than Al-Duwairi scheme. In the Internet it is well known that the hop counts (with respect to AS level) are not so large (at most 15 or so) and we can see that our scheme is efficient from the practical point of view.

5. Conclusion

One of the effective countermeasures against Denial of Service (DoS for short) attacks is IP traceback, which tries to identify the attacker by determining the attack path. In particular *Probabilistic Packet Marking* (PPM) for IP traceback is so promising that many researches on it have so far been done. However, most of the previous PPM schemes have the disadvantage such that they require a large number of packets to reconstruct the attack path. Therefore we propose an efficient traceback scheme based on PPM in this paper.

Our scheme has a novel feature that routers on the attack path mark packets with *variable probabilities*. In this paper, we show this mechanism can reduce the number of packets for path reconstruction significantly.

Acknowledgments

This research is conducted as a program for the "Fostering Talent in Emergent Research Fields" in Special Coordination Funds for Promoting Science and Technology by Ministry of Education, Culture, Sports, Science and Technology, and is partially supported by Grant-in-Aid for Scientific Research (C), 17500035, 2005.

References

[1] M. Adler. Trade-offs in probabilistic packet marking for IP traceback. *J. ACM*, 52(2):217–244, Mar. 2005.

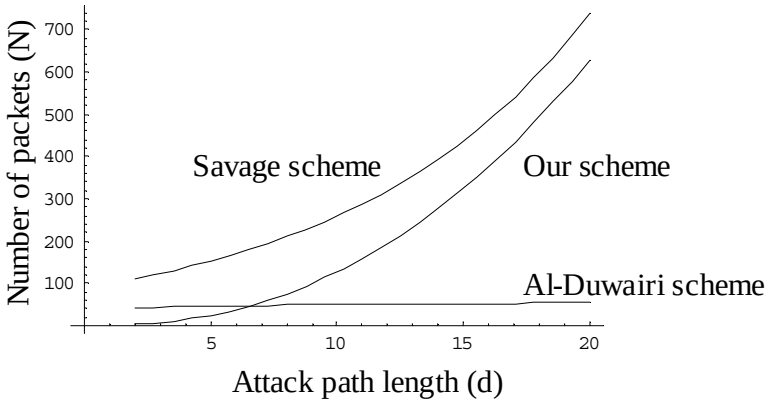

Figure 2: Number of packets required for path construction

[2] B. Al-Duwairi and G. Manimaran. A novel packet marking scheme for IP traceback. In *10th International Conference on Parallel and Distributed Systems (ICPADS'04)*, pages 195–202, July 2004.

[3] D. Dean, M. Franklin, and A. Stubblefield. An algebraic approach to IP traceback. *ACM Transactions on Information and System Security*, 5(2):119–137, May 2002.

[4] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. John Wiley & Sons, 3 edition, 1968.

[5] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge Univ Press, 1995.

[6] S. Savage, D. Wetherall, A. R. Karlin, and T. Anderson. Practical network support for IP traceback. In *Proceedings of the ACM SIGCOMM*, pages 295–306, 2000.

[7] A. Yaar, A. Perrig, and D. Song. Pi: A path identification mechanism to defend against DDoS attacks. In *IEEE Symposium on Security and Privacy*, pages 93–109, May 2003.