I.J. Wireless and Microwave Technologies, **2012, 6, 1-9**
Published Online December 2012 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijwmt.2012.06.01

Available online at http://www.mecs-press.net/ijwmt

Design and Simulation Implementation of an Improved PPM Approach

Guo Fan[a,*], Feng Bo[a], Yu Min[a]

[a]*College of Computer Information Engineering, Jiangxi Normal University, Nanchang, China*

Abstract

Different from recent probabilistic packet marking (PPM) methods, Dynamic PPM may solve many problems of traditional methods, such as loss of marking information, hard to reconstruct attack path, low accuracy, and so on. A novel DPPM approach is proposed and the network simulation software (NS2) is used to verify the performance and efficiency of the approach by constructing simulation DOS environments. In comparison with PPM methods, simulation results show that DPPM is much better.

Index Terms: IP Traceback; DDOS; PPM; Network Security

1. Introduction

In recent years, distributed denial of service (DDOS) has become the most serious threat and has made great damages to Internet. Hackers may make use of system vulnerabilities to control thousands of remote victim hosts which are further triggered to attack Internet services, so that it is very hard to find out the origin of DDOS attack using current intrusion detection techniques.

IP trace back is a new research direction to obtain the really original address of the attack packets so as to find out the real attacker. Probabilistic packet marking(PPM), having many merits such as low management cost, no extra network traffic, low additional workload for routes, is one of the most popular methods for tracing IP addresses[1].

In 2000, PPM was firstly brought up by Savage [2] marking the addresses of routers in the packets to improve the accuracy rate of path reconstruction and to reduce network transmission overhead as well. A route is considered as an independent "node" if it is passed through by the packet. The "node" writes its own IP address into the fixed positions of the packet in the specified probability P. As the number of packets the victim receives increases, it will receive at leas one modified packets for each "node" at some moment so that the attack paths can be reconstructed.

The key problem of PPM is that the probability (P) is fixed and it is hard to select an accurate P. If P is too big, the mark of a distant "node" in the packets is probably overwritten by that of a closer "node". Past researches

* Corresponding author:
E-mail address: guofan771210@yahoo.com.cn

show that the optimal value of P should be 0.04.

The traditional packet marking procedure in PPM is depicted in the following:

Packet Marking Procedure (Packet w)

```
probability P=0.04
marking procedure at router A:
  write address of A to m in probability P
path reconstruction procedure at victim n:
  for each packet m from attacker
      extract path(A[i] ....A[j])from the suffix from m
  end for
end
```

Optimized PPM methods may decrease the occurring frequency of mark overwritten but they could not fundamentally solve the problem which further raises the problem of the weakest node and weak convergence.

2. Dynamic PPM

An improved PPM approach, dynamic PPM is proposed in order to solve the problems of traditional PPM. The approach is composed of three parts: reconstructing marking fields, using dynamic probability to mark packets and improving the marking algorithm.

2.1. Reconstructing Marking Fields

In PPM methods, although the whole IP address is 32bits long, only 17bits are used as marking fields in each packet head. The new marking field is constructed by dividing the IP address into four segments, that is to say, a packet only carries one forth of an IP address these 4 segments. If four consecutive marked packets are received, a whole IP address is obtained.

The format each segment is shown as below:

2 bit	5 bit	8 bit	1 bit
C	Distance	Node[C]	flag

Fig. 1. Marking fields

C: Value 0~3, denotes the current segment index. Distance: Value 1~31, denotes the distance between the current node and the victim. Node[C]: Segment C of the four segments (8bits). Flag: Value 0~1, if flag=1, the current packet has been marked in the past. Node[c]~{node[0],node[1],node[2],node[3]},for example {192.168.1.3}.

The distance field is used to compute the hops from the attacker to the victim. The network topology is normally about 20 hops and not more than 32 hops, so five bits is enough for the field.

2.2. Dynamic Probabilistic Packet Marking (DPPM)

The core idea is using two marking probabilities in replace of a fixed one so as to flexibly and intelligently mark the packets and to solve mark missed and overwritten problems. The two probabilities are 0.04 and 1. The former one is usually selected in PPM methods and has been demonstrated to be optimal. The latter means the packet is to be definitely marked so that the packet will not be missed. The approach uses the marking field "flag" to make decision which probability is chosen.

The "flag" field of each packet is checked by the router. Zero value of "flag" denotes the packet is never marked before so that the packet is marked by PPM methods with probability 0.04. After the packet is marked, "flag" is set to 1. The value of the "distance" field is increased by 1 whether the packet is marked or not since it only represents the hops from attackers to victims. If the "flag" value is 1, it means the packet has been marked before so that the router does not mark the packet again and uses the probability 1 for the subsequent packet. If the following packet is not marked before, it is marked 100 percent and the probability is restored to 0.04 for the next packets. In this way, marking missed and overwritten problems in PPM is effectively solved.

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

Packet Marking Procedure(Packet w)

```
1)  let x be a random number in [0 . . . 1)
2)  ip[c]={ip[0],ip[1],ip[2],ip[3]}
3)  if x < p and w.flag=0 then
4)     w.node = ip[c]
5)     w.flag = 1
6)     p = 0.04
7)     w.c = (c+1) % 4
8)  else
9)  if w.flag == 1 then
10)    p = 1
11) end if
12) w.distance ++
13) end if
```

Path reconstruction procedure at victim v:

```
1)  let Node Tb1 be a table of pairs(node, distance)
2)  for each packet w from attacker
3)     z:=lookup w.node in Node Tb1
4)     if z!=null then
5)        z.distance ++
6)     else
7)        insert pair (w.node,1) into Node Tb1
8)  sort Node Tb1 by distance
9)  extracting path(Ri…Rj) from ordered nodes  in Node Tb1.
```

Fig. 2. The improved PPM aprroach

2.3. In Comparison with PPM Methods

The new approach has many advantages in comparison with PPM methods.

1) More spaces for marking fields. 17 bits allocated for marking are hard to cover 32 bits space of a whole IP address in PPM methods, while the new approach marking four packets to represent an IP address. The IP address is partitioned into four fragments (ip[0]-ip[3]) each of which is marked, transmitted and reassembled respectively. Experiments show it is a feasible way for the real router environments.

2) More flexible. The path reconstruction results are significantly affected by the fixed probability value the PPM method selects and the network topology. The fixed probability may not change during the marking procedure. The new approach replaces the fixed one by two exchangeable ones and in addition, a new "flag"

field is used to determine which one is being used. In this way, the approach may be flexibly adapted to different network topologies and environments.

3) No marking missed routers. One of the consequences of the fixed probability is that some router may always have no opportunity of inserting the mark of themselves into the packets so that the router is marking missed. The new approach solves the problem by using the two probabilities alternately and dynamically as the lines 3-12 in figure 2 show. If one of the four fragments for the router is marked, the next fragment unmarked before is sure to be marked by the router.

4) No marking overwritten problems. A mark in a packet may be overwritten when the packet pass through many routers. The mark for the distant router may be covered by the closer one since both routers are probably marking the same packet because of the fixed marking probability. The "flag" field in the new approach avoids marking overwritten problems. If the value of "flag" is 1, the router will not mark the packet anymore.

5) Better convergence time. Assuming t is the distance between the farthest router and the victim, d is the router hops, the probabilities of receiving the packet marking for the farthest one R_k are λ_k in a PPM method, and λ'_k in the new approach. And the following equations are not hard to infer.

$$\lambda_k = p(1-p)^{d-1} / 2^{d-1} \ (0 < n \le d-1)$$

$$\lambda'_k = p / 2^{d-1}$$

$$E(N) = \int_0^{\infty} [1 - \prod_{k=1}^{2^{d=-1}} (1 - e - \lambda^j t)] dt$$

$$E(N') = \int_0^{\infty} [1 - \prod_{k=1}^{2^{d-1}} (1 - e - \lambda'_j t)] dt$$

$$\lambda_k < \lambda'_k, \ \mathrm{E}(N) > \mathrm{E}(N')$$

It is obvious that the new approach needs less time for convergence.

3. Using the Template

Network Simulator version 2(NS2), a popular open-souce network simulation software developed by UC Berkley, is chosen to construct the simulation environment for the evaluation of the new approach

3.1. DOS Simulation

Fig. 3(a) shows the simulation environment for a DOS attack. There are four nodes each of which represents a route in real world. Fig. 3(b) shows the core part of corresponding NS2 implementation.

1) Node 0: the normal router for regular transmission with speed 0.5 MB/ms.

2) Node 1: the original source of DOS. Since a huge number of packets are to be passed through to exhaust the network resource of the victim, the transmission speed is 20MB/ms.

3) Node 2: the intermediate router as the inevitable node for the attack to succeed. It is responsible for marking packets. The speed is configured as 10 MB/ms so that if DOS occurs, the path between it and Node 3 are obviously congested.

4) Node 3: the victim. It reassembles, filters and analyze the packets received to reconstruct the correct attack path of (1, 2, 3).

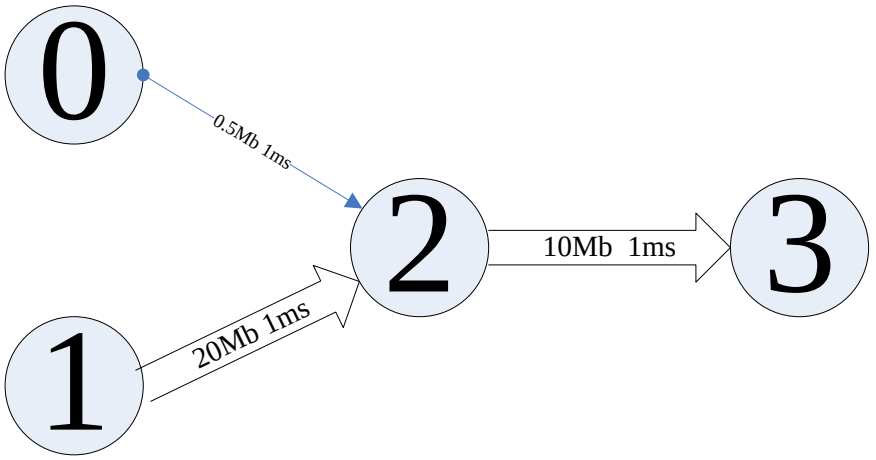

Fig. 3(a) the network topology of a DOS enviroment

```
#Create links between the nodes
    $ns duplex-link $n0 $n2 0.5Mb 1ms DropTail
    $ns duplex-link $n1 $n2 20Mb 1ms DropTail
    $ns duplex-link $n3 $n2 10Mb 1ms SFQ
    $ns duplex-link-op $n0 $n2 orient right-down
    $ns duplex-link-op $n1 $n2 orient right-up
    $ns duplex-link-op $n2 $n3 orient right
#Monitor the queue for the link between node 2 and node 3
    $ns duplex-link-op $n2 $n3 queuePos 0.5
#Create a UDP agent and attach it to node n0
    set udp0 [new Agent/UDP]
    $udp0 set class_ 1
    $ns attach-agent $n0 $udp0
    set cbr0 [new Application/Traffic/CBR]
    $cbr0 set packetSize_ 500
    $cbr0 set interval_ 0.005
    $cbr0 attach-agent $udp0
```

Fig. 3 (b) the corresponding NS2 implementation

Fig. 3. DOS simluation environment and NS2 implementation

The simulation program lasts for 10 milliseconds. Node 0 transmits totally 5 Mb data in speed of 0.5 Mb/ms, and node 1 transmits 200 Mb data in speed of 20Mb/mss, so that the total number of data transmitted is 205 Mb. Since the bandwith between node 2 and node 3 is configured as 10Mb/ms, packets may be discarded when they all reaches node 2 at the same time, and it means the sign of a DOS attack. As DPPM is used in the environment, the number of marked packets received at node 3 is much less than the number of totally sent packets. The purpose of the experiment is to evaluate whether the approach is capable of accurately reconstructing the attack path with limited marking resources.

The configuration of the experiment includes: IP address of Node 0 is 10.10.10.10, Node 1 is 1.1.1.1, Node 2 is 2.2.2.2 and Node 3 is 3.3.3.3. Node 0 sends out 5Mb data. Each packet is 500 bit long, and the number of sent packets is 10485. Node 1 sends out 200Mb data and the number of sent packets is 419430.

1) Marking results with the PPM method

Table 1. Marking information from node 3

Edge	distance	flag	attack(y/n)	packets
1.1.1.1	1	1	Y	49382
10.10.10.10	1	1	N	2592

Table 2. Marking information from node 2

Edge	distance	flag	attack(y/n)	packets
1.1.1.1	1	1	Y	49382
10.10.10.10	1	1	N	2592

2) Marking results with the DPPM approach

Table 3. Marking information from node 3

Edge	distance	flag	attack(y/n)	packets
2.2.2.2	2	1	Y	133056
2.2.2.2	2	1	N	3438
1.1.1.1	1	1	Y	74073
10.10.10.10	1	1	N	1965

Table 4. Marking information from node 2

Edge	distance	flag	attack(y/n)	packets
1.1.1.1	1	1	Y	98765
10.10.10.10	1	1	N	2592

Table 1 and Table 2 show the marking results obtained with the PPM method, which succeeds in reconstructing the attack path according to the signs of a DOS attack including the big flow rate, network congestion. Table 3 and Table 4 show the marking results obtained with the new approach. It is obvious that the new approach has many advantages.

1) Much more marked packets. The received number of packets in Table 3 and 4 is much more than that received in Table 1 and 2. The reason is the PPM method may miss marking packets or overwrite marked packets while the new approach solves the problem.

2) More marking information. Since there are only 17 bits available for marked fields in the PPM method, the marked results only indicates the number of received packets. The marked results of the new approach include not only the distance, the source, the received number of packets, but also the signs whether the packet is part of the attack and whether the packet is marked or not.

3) More simple for reconstruction. Since the new approach knows whether a packet is part of the attack with the corresponding sign marked in the packet, the algorithm is able to reconstruct the correct attack path much faster.

3.2. DDOS Simulation

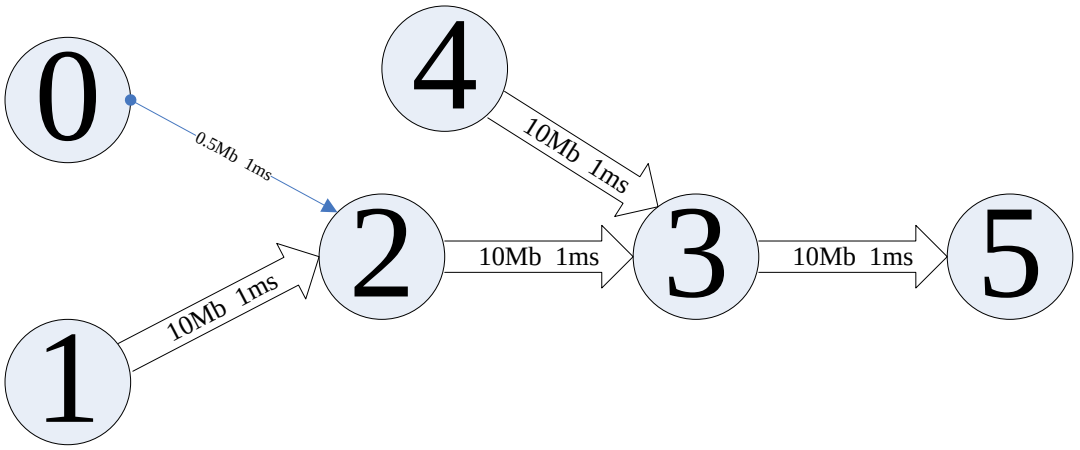

Fig. 4. DDOS simulation environments

1) Marking results with the PPM method

Table 5. Marking information from victim node 5

Edge	distance	packets
3.3.3.3	1	26321
2.2.2.2	2	16254
4.4.4.4	2	21543
1.1.1.1	3	8756
10.10.10.10	3	127

Table 6. Marking information from node 3

Edge	distance	packets
2.2.2.2	1	12474
4.4.4.4	1	15124
1.1.1.1	2	15889
10.10.10.10	2	489

Table 7. Marking information from node 2

Edge	distance	packets
1.1.1.1	1	48564
10.10.10.10	1	980

As Fig. 4 shows, DDOS simulation adds several attack sources to the DOS simulation environment as before. Multiple sources will attack the victim simultaneously in the experiment. Node 4, cooperating with node 1, is going to attack node 5. The flow rate of the two attack nodes (node 1 and node 4) is 10Mb/ms. The flow rate of normal node 0 is 0.5 Mb/ms. The attack lasts for 10 milliseconds. the IP address of node 4 is 4.4.4.4 and node 5 is 5.5.5.5. The other configuration is the same as that in 4.1. Node 1 sends 209715 packets and the total size is 100 Mb. Node 3 does the same.

In DDOS simulation, as the number of attack sources is more and the length of attack path is larger than that in DOS simulation, the marking results is more significantly impacted by missed marking and mark overwritten problems in the PPM method. In comparison with the number of packets received in the new approach, the number received in Table 5, 6,7 is much less. From table 5, 6 and 7, it is hard to reconstruct the correct path since many useful information are lost during marking procedures in the PPM method. In the mean while, from the marking results in Table 8, 9 and 10, the new DPPM approach easily reconstructs the two attack paths which are 1.1.1.1→2.2.2.2→3.3.3.3→5.5.5.5 and 4.4.4.4→3.3.3.3→5.5.5.5.

2) Marking results with the new approach

Table 8. Marking information from node 2

Edge	distance	flag	attack(y/n)	packets
3.3.3.3	1	1	Y	30719
3.3.3.3	1	1	N	368
2.2.2.2	2	1	Y	14745
2.2.2.2	2	1	N	737
4.4.4.4	2	1	Y	26214
1.1.1.1	3	1	Y	19660
10.10.10.10	3	1	N	982

Table 9. Marking information from node 3

Edge	distance	flag	attack(y/n)	number of packets
2.2.2.2	2	1	Y	22118
2.2.2.2	2	1	N	1327
4.4.4.4	1	1	Y	42564
1.1.1.1	2	1	Y	33422
10.10.10.10	2	1	N	1

Table 10. Marking information from node 2

Edge	distance	flag	attack(y/n)	packets
1.1.1.1	1	1	Y	49382
10.10.10.10	1	1	N	2592

4. Using the Template

The paper presents a novel DPPM approach solving the key problems of traditional PPM methods that include marking missed and overwritten, limited marking fields and low accuracy, by using dynamic probability and fragments. Experiments show the new approach is much better than the PPM method in detecting DOS attack. Future work includes:

a) How to add authentication information in marking fields;
b) How to improve Edge-Marking with dynamic probability.

References

[1] K. Park, H. Lee. ON the effectiveness of probabilistic packet marking for IP traceback. In : Proc. Of the IEEE Int'l Conf.on Communication 2004.2204.1008-1013
[2] T. Baba, S. Matsuda TracingNetwork Attacks to Their Sources[J].IEEE Internet Computing,2002,16(2):20-26
[3] R. Stone. Center Track: An IP Overlay Network for Tracking DoS Floods[A]. In :Proceedings of USENIX Security Symposium, 2000
[4] P. Ferguson, D. Senie. Network Ingress Filtering:Defeating Denial of Service Attacks which EmployIP Source Address Spoofing(RFC2827) The Internet Society, 2000
[5] R. Bajaj, P. Dharma. Improving Scheduling of Tasks in a HeterogeneousEnvironment[J]. IEEE Transactions,on Parallel and Distributed Systems, 2004, 15(2):107~118
[6] S. Savage. Network Support for IP Traceback[J]. .IEEE ACM tractions on networking, 2001, 9:226-237
[7] D.X. Song, A.Perrig. Advanced and authenticated marking schemes for IP traceback[C]. In: Proc of IEEE INFOCOM 2001, pp 878-886
[8] H. Burch, B. Cheswick. Tracing Anonymous Pacekets to Their Approximate Source[C]. Proc of Usenix LISA,New Orleans.200-12:313-3