ISSN 2047-3338

A Resolved IP Traceback through Probabilistic Packet Marking Algorithm

Swathy Vodithala[1], S. Nagaraju[2] and V. Chandra Shekhar Rao[3]

[1,2]Department of CSE, KITS, Warangal

[3]Department of MCA, KITS, Warangal

***Abstract*– The major problem of network security in present years is DoS (Denial of Service) attacks, in order to protect the network from these attacks a research is implemented in the key streams of network security. Packet marking is always required to track few details of packet like its source and the status toward reaching the destination. In most of the cases, packets transmitted by a source are lost or data in it is corrupted and may lose the packets permanently. A perfect packet marking algorithm is always required to mark the packet by the IP address of source and the current routers traversed by it. We suggest not marking each and every packet with equivalent probability; instead the marking probability is computed for the purpose of every packet by all the routers depending on field value of TTL (Time to Live).**

***Index Terms*– Attack Graph, Denial of Service, Probabilistic Packet Marking Algorithm and Resolved Probabilistic Packet Marking Algorithm**

I. INTRODUCTION

THE critical problem of network security in present years is DoS (Denial of Service) attacks. Many techniques are offered by the dynamic field of research, some of these techniques are the packet filtering techniques, pushback message and ICMP [8], [5] trace back. Savage [2] has designed an algorithm called Probabilistic Packet Marking Algorithm which is also called as PPM algorithm. This algorithm is designed for contributing the IP trace back idea. The IP trace back method permits the routers to encode particular data on the attack packets. The routers encode the data based on some programming probability. The victims can build a set of paths which are navigated or traversed by the attack packets when they receive marked packets in sufficient number. As a result of this, the location of attacker can be recognized by the help of victim.

Obtaining a built graph is the most significant objective of PPM algorithm [7], a constructed graph which is obtained by the PPM algorithm is similar to the attack graph. The set of paths traversed by the attacker are represented by the attack graph and the constructed graph represents the graph that is obtained from the algorithm of PPM. Savage et al., [2] have been suggested a technique to fulfill this aim. In this technique the data of the attack graph edges are encoded into the packets of DoS attacks with the help of routers that are present in the attack graph from the victim's location. The algorithm of PPM is implemented very specially using two different procedures:

- Graph reconstruction procedure
- Packet marking procedure.

The graph reconstruction procedure is implemented on the victim location and the procedure of packet marking is implemented on the router's location.

II. RELATED WORK

A. Probabilistic Packet Marking

It is defined to be the most famous packet identification techniques. In this methods, the packets are marked with the router's IP address from which they traversed or the path edges from which the packet is being transmitted.

Marking the packets with the router's address is the best approach when compared to the two alternatives provided here, where if a packet is lost of affected with any attack, the source router address can be fetched and send back to the actual router. Now the router checks the packets and re-transmits the packet to the actual destination. With this implementation, an accuracy of 95% can be achieved to identify the actual attack path.

Second approach considered in probabilistic packet marking is edge marking and here the IP address of two nodes are required to mark the packets. This approach is much complicated when compared to marking the IP address of the router, where much state information of the packet is required in the former case.

There are few techniques to reduce the state information required in this case and they are discussed here. A simple XOR operation can be performed between the two nodes that form the edge. Address of a node made XOR with another node and this result in an edge ID and thus the actual state information required is reduced by half. Even this approach can be made easy by splitting the Edge ID in to number of small fragments. To pass the state information, any fragment can be considered and encoded to hold the state information. There are few limitations with this technique and important among them is that, this method assumes that the affected packets when compared to normal packets are more. Path reconstruction of the lost packets requires many computation cycles in this method and this is not practically possible for the systems with low resources.

Burch and Cheswick [5] proposed this algorithm and was designed carefully later it was implemented by Savage [2] for solving the trace back problem present with IP address. This algorithm is used to trace out the attacker with the help of an internet map or by the usage of an attack graph when the attack related to distributed denial-of service is going on.

The attacked graph edges are encoded by the packets in random. The encoded information is used to construct the new graph. The graph obtained newly should be as that of the older attacked graph.

The PPM algorithm constructs the new graph where as the graph attacked consists of the set of paths where the packets are traversed.

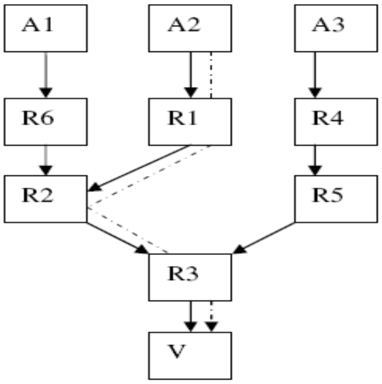

Fig. 1. The attacked graph containing the path attacked

The view of network can be defined as a directed graph having G = (V,E), here E represents the edges set, and V represents the nodes set. The single host that is under attack can be considered as V or device that one present at the border of a network which may be firewall or a system that is intrusion detected representing a number of paths. The origin of all the potential attacks is at A_i which is represented as a leaf in tree that is being embedded at V, and there are routers in the path namely R_i that are present among A_i as well as V. The routers ordered list that is between A_i and V having the packets traversed is considered as the "attack path" which is represented in the figure1 with a dotted line in the above figure1 it is (R_1, R_2, R_3). The number of routers that are present in between the R_i and V in a path is considered as the "distance" which is represented in the figure for the path $R_{3,}$ R_1, R_2. Those packets that are utilized in the attacks of DDoS are considered as the 'attack packets'.

B. Packet Marking Procedure

For the implementation of IP trace back service, the packet header is allocated with enough space making one to use this space for recording the path traversed by a packet. Consider an example where every router performs the forwarding of packets and also performs routing functions besides appends or records the ID of its own into the space that is pre-allocated in the header of a packet. In this kind of situation a marked packet is received by the victim who has the permission to examine the header of packet and then the entire traverse information that is related to marked packet is being collected.

There is a problem with this type of technique, as there is no fixed length for a packet to traverse and making the pre-allocation of space in a preceding packet header is impossible. There is also other practical problem to be considered regarding the complete information of path of every packet being capable of recording victim as an attacker can influence the information of path if needed, and then the identity of router can be modified as false in the header of packet which misleads to the victim site who is ready to attack.

The PPM algorithm records the traversed edges details in probabilistic manner from the attacker site to the correct site instead of recording the entire information that is related to the packet that is traversed. The information of an attack packet is being encoded in three marking fields namely start, end and distance, by the help of routers. The start and end points are used to accumulate the router's IP addresses which are being present at the marked edge end points. The space ground is utilized to trace several hops that are taken by the victim in order to pass through marked edge and site.

Marking procedure to be followed at router 'R':

```
for each packet w
assume that x be a random number from [0..1]a
if x < $p_m$ then
write 0 into w.distance and R into w.start
else
if w.distance = 0 then
write R into w.end
increment w.distance
```

Fig. 2. Packet marking procedure

C. Graph Reconstruct Procedure

The victim 'V' after acquiring the packets is required to filter the unmarked packets and then the algorithm of graph construction is required to be implemented by the victim to all the collected packets, and is noticed to reconstruct the attack graph.

Construction procedure of attack graph at the victim 'V':

```
Assume G be a tree with root being victim V;
Suppose edges in G be tuples (for e.g. start, end,
distance);
for (each obtained marked packet w)
{
if (w.distance==0) then
insert edge (w.start,V ,0) into G ;
else
insert edge (w.start, w.end, w.distance) into G ;
}
extract any edge (x,y,d) from x to V in G with $d \neq$ distance;
remove path (Ri…Rj) by means of enumerating acyclic paths in G ;
```

Fig. 3. Graph reconstruction procedure

III. PROPOSED ALGORITHM

A. Resolved Probabilistic Packet Marking

There are different packet marking procedures [6] in literature to mark the packets as node append, node sampling, edge sampling and TTL based packet marking [1]. The combination of both the edge sampling and TTL based algorithms give a better performance .In this algorithm we are going to add a field for the TTL in the marking field of a packet, along with the start, end and distance fields if the value of $r \leq 1/h$ (from Fig. 5) then encode the values of the corresponding edge.

The proposed algorithm is:

Marking procedure to be followed at router 'R':

```
For each packet Pkt
$t \leftarrow t - 1$
if $t_p > t$
$h \leftarrow t_p - t$
else
$h \leftarrow 1; t \leftarrow t_p$
Let r be a random number in [0, 1)
if $r \leq 1/h$
write 0 into pkt.distance and R into pkt.start
else
if pkt.distance = 0 then
write R into pkt.end
increment pkt.distance
```

Fig. 4. Resolved packet marking procedure

Here, t is the TTL value being marked, tp is the maximum path length and h is the maximum remaining distance that packet would traverse.

The design of the packet marking procedure can be explained as follows:

MARKING FIELD	DATA

TTL	START	END	DISTANCE	DATA

Fig. 5. Design of packet marking

The TTL is a counter to limit the packet lifetime. It is supposed to count the time in seconds, allowing a maximum lifetime of 225 seconds. It must be decremented on each hop. When it hits zero, the packet is discarded and a warning packet is sent back to the source host.

In the IP address format i.e., IP header we have a field by the name Time to live (TTL).We make use of this field ,while marking the packet along with the router address.

We consider a source, destination and three routers R1, R2, R3 [3].

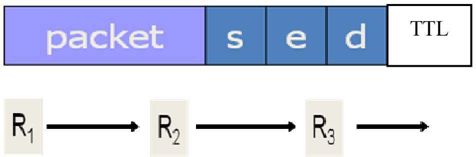

Fig. 6. Design pattern of packet pkt

The packet contains the data and the marking fields are start, end, distance and Time to live. This is shown in Fig. 6.

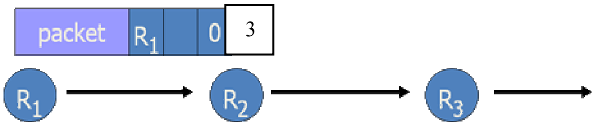

Fig. 7. Router R1 encoding the edge

Let R1 starts encoding the edge, its router address will be placed in the data packet. Then its starts checking the value of TTL, let it be 3.This is shown in Fig. 7.

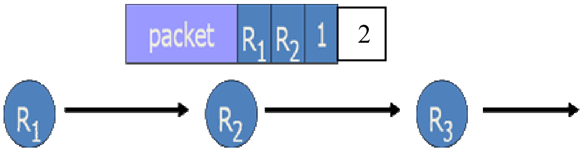

Fig. 8. Router R2 encoding the edge

The value of TTL decrements by 1 and the distance gets incremented by 1 as the packet passes from R1 to R2. This is shown in Fig. 8.

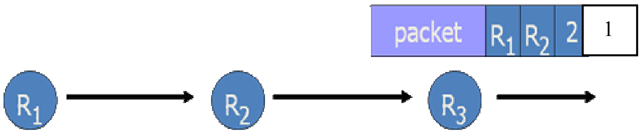

Fig. 9. Router R3 encoding the edge

The value of TTL decrements by 1 and the distance gets incremented by 1 as the packet moves from R2 to R3.This is shown in Fig. 9. The packet gets passed to the destination or another router (for example may be R4) and the value of TTL then becomes zero, as on when the TTL=0 packet gets discarded and a warning packet of loss is sent to the source.

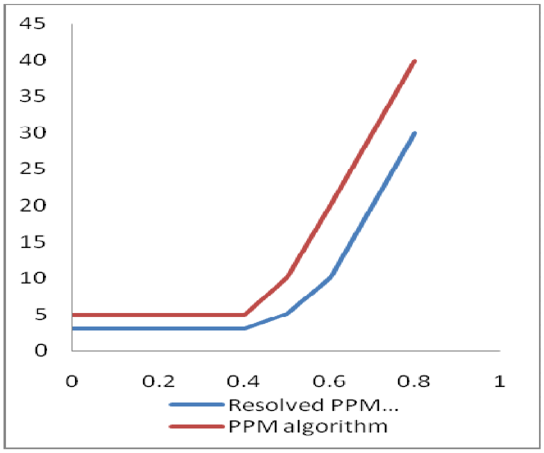

Fig. 10. Marking probability vs. no. of average packets

IV. EXPERIMENTAL RESULTS

In the proposed algorithm the marking probability depends on TTL [1] (Time to Live). The average number of marked packets required for a correct graph reconstruction against different values of marking probability (Fig. 10) has been decreased by the Resolved PPM when compared with PPM algorithm [2], [4]. The time required to reconstruct the graph also decreases compared to PPM algorithm.

V. CONCLUSION AND FUTURE SCOPE

In this paper, we proposed a new scheme for tracing back DDoS packets, where each router marks a packet in a manner such that packets traversing longer distances are marked with high probability, while packets with short distances to traverse are marked with lower probabilities. Such a type of marking ensures that each attack is marked with much higher probability by intermediate routers, which greatly reduces the impacts of spoofed marks [1]. We have suggested an algorithm as Resolved PPM algorithm. The algorithm of Resolved PPM does not need any preceding skills regarding the graph of network. In the Resolved PPM algorithm we focused on the packet marking procedure. The algorithm of Resolved PPM is used as an efficient means of enhancing the actual PPM algorithm's reliability. The future scope can be focused on Graph reconstruction procedure for the multiple attacks.

REFERENCES

[1] Vamsi Paruchuri, Arjan Durresi and Sriram Chellappanernational, "TTL based Packet Marking for IP Traceback", 2001.

[2] S. Savage, D. Wetherall, A. Karlin and T. Anderson, "Practical Network Support for IP Traceback", Proc. ACM SIGCOMM, 2000.

[3] Niranjan reddy and Bhavani, "An efficient ip traceback through probabilistic packet marking algorithm", 2003.

[4] Tsz-Yeung Wong, "A precise termination condition of probabilistic packet marking algorithm", 2007.

[5] Burch, Hal; Bill Cheswick, "Tracing Anonymous Packets to their Approximate Source", LISA, pp. 319–327, 2000.

[6] D.X. Song and A. Perrig, "Advanced and Authenticated Marking Schemes for IP Traceback", Proc. IEEE INFOCOM, 1st April 2001.

[7] M. Adler, "Trade-Offs in Probabilistic Packet Marking for IP Traceback," Journal of ACM, March 2005.

[8] M. T. Goodrich, "Efficient Packet Marking for Large-Scale IP Traceback", Proc. Of ACM CCS 2002, November 2002.

Swathy Vodithala received her B.Tech degree in computer science and engineering in 2005 from KITS, huzurabad (Jawaharlal Nehru Technological University). And M.Tech degree in software engineering from KITS, Warangal (Kakatiya University) in 2011.she worked with computer science department for 2 years in KITS, huzurabad as an Assistant Professor. At present, she is working as an Assistant Professor in computer science department at KITS, Warangal, Email: vinna_m@yahoo.com

S. Nagaraju, Associate Professor at Department of CSE, KITS, Warangal, India
Email: nag_sangepu@yahoo.com

V. Chandra Shekhar Rao, Associate Professor at Department of MCA KITS, Warangal, India
Email: vcsrao.kitswgl@gmail.com