Asian Journal of Information Technology 7 (10): 454-458, 2008
ISSN: 1682-3915

A Comparitive Study of Various IP Traceback Strategies and Simulation of IP Traceback

[1]S. Karthik, [1]V.P. Arunachalam and [2]T. Ravichandran
[1]Department of Computer Science Engineering,
SNS College of Technology, Sathy Main Road, Coimbatore, 641035, India
[3]Hindustan Institute of Technology, Pollachi Main Road, Coimbatore, 641032, India

Abstract: Distributed Denial of Service (DDoS) is a rapidly growing problem. The multitude and variety of both the attacks and the defense approaches is overwhelming. IP traceback-the ability to trace IP packets from source to destination-is a significant step toward identifying and thus, stopping, attackers. The IP traceback is an important mechanism in defending against Distributed Denial of Service (DDoS) attacks. This study constructs a simulation environment via extending ns2, setting attacking topology and traffic, which can be used to evaluate and compare the effectiveness of different traceback schemes. A comparison among some of the Packet Marking schemes is presented with several metrics, including the received packet number required for reconstructing the attacking path, computation complexity and false positive etc. The simulation approach also can be used to test the performing effects of different marking schemes in large-scale DDoS attacks. Based on the simulation and evaluation results, more efficient and effective algorithms, techniques and procedures to combat these attacks may be developed.

Key words: DDoS, IP traceback, simulation, strategies, comparison

INTRODUCTION

The Internet provides a wealth of information and value to its users, but this accessibility makes it extremely vulnerable to motivated and well-equipped users intent on disrupting the flow of information or using it for personal gain. The tools for disruption are readily available to these Internet attackers, ranging from published operating-system weaknesses to executable software ready to exploit such vulnerabilities (CERT Coordination Center, 1999).

A common form of attack is Denial of Service (DoS). DoS attacks consume a remote host or network's resources, thereby denying or degrading service to legitimate users. Typically, adversaries conduct DoS attacks by flooding the target network and its computers with a large amount of traffic from one or (as in the case of distributed DoS, called DDoS) more computers under the attacker's control (CERT Coordination Center, 1998). Such attacks are among the toughest to address because they are simple to implement, hard to prevent and difficult to trace. IP traceback methods provide the victim's network administrators with the ability to identify the address of the true source of the packets causing a DoS (CERT Coordination Center, 1997). IP traceback is vital for restoring normal network functionality as quickly as possible, preventing reoccurrences and ultimately, holding the attackers accountable. Merely identifying the machines and networks that generate attack traffic might seem like a limited goal, but the essential clues it provides can help distinguish the actual attacker. Several efforts are under way to develop attacker-identification technologies on the Internet (CERT Coordination Center, 2001). This study looks at existing DDoS IP traceback methodologies and future trends.

THE ROLE OF IP ADDRESSES

Ideally, the network traffic used in an attack should include information identifying its source. The Internet Protocol (IP) specifies a header field in all packets that contains the source IP address, which would seem to allow for identifying every packet's origin. However, the lack of security features in TCP/IP specifications facilitates IP spoofing CERT Coordination Center (1996), (Ferguson and Senie, 2000) the manipulation and falsification of the source address in the header. The Internet's current routing infrastructure is stateless and

Corresponding Author: S. Karthik, Department of Computer Science Engineering, SNS College of Technology, Sathy Main Road, Coimbatore, 641035, India

largely based on destination addresses, but no entity is responsible for ensuring that source addresses are correct. Thus, an attacker could generate offending IP packets that appear to have originated from almost anywhere. Although, some network-based DoS attacks use IP spoofing by default, only a small percentage of DDoS attacks use forged source addresses; most attack their targets indirectly through other, previously compromised zombie systems (Jelena and Peter, 2002).

IP TRACEBACK

IP traceback is the process of identifying the source machines/nodes that generate the attack traffic and detecting the path traversed by the malicious DDoS traffic. Traceback primarily depends on packet marking (augmenting packets with partial path information) and packet logging techniques (storing packet digest/ signature at intermediate routers) (Savage *et al*., 2000). IP traceback is complicated by various factors which include the distributed anonymous nature of DDoS attacks, stateless nature of the Internet, destination oriented IP routing and the millions of hosts connected to the Internet. The traceback mechanisms fall into four main categories:

- Link testing-hop-by-hop tracing
- Messaging (ICMP based traceback)
- Logging
- Packet marking

Link testing or hop-by-hop tracing: Method tests the network link between routers to determine the origin of the attacker's traffic. Method starts from router closest to the victim and tests the incoming links to determine which link caries the malicious packets. Process is repeated on upstream routers until source node is identified.

Drawback: Attack should remain active until trace is completed (Fig. 1).

Link testing approaches:
- Input debugging
- Controlled flooding

Messaging (ICMP based traceback): ICMP messages are generated by the router and sent along with the network traffic to the victim/destination machine. These messages contain partial path information, including information about where the packet came from, where it was sent and its authentication. This information is used by the victim to trace the path of a packet to its originating source node (Fig. 2).

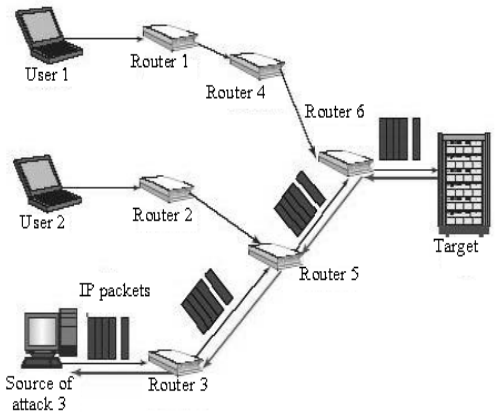

Fig. 1: Link testing or hop-by-hop testing

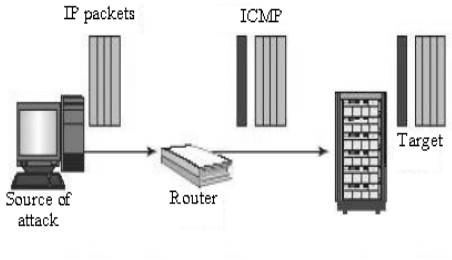

Fig. 2: ICMP based traceback messaging

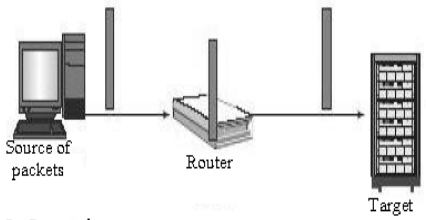

Fig. 3: Logging

Logging: As packets traverse the network towards the destination victim, they are logged at the key routers. This information is analyzed using data mining techniques to extract information about the traffic sources (Fig. 3).

Packet marking: In packet marking method traceback data is inserted into the IP packet by the routers on the path to the destination node. Packet marking information stored in identification field of IP header. Types of packet marking are Probabilistic Packet Marking (PPM) and Deterministic Packet Marking (DPM) (Fig. 4).

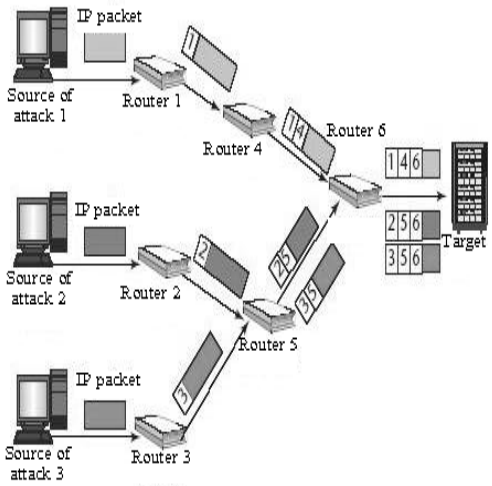

Fig. 4: Packet marking

Probabilistic Packet Marking (PPM): Focuses on reconstructing the entire attack path the malicious packets have traversed. Routers put stamps into packets with a fixed probability and victim reconstructs attack path from these stamps.

Packets are marked with partial path information as they arrive at the routers. This approach exploits the observation that attacks generally comprise large number of p. while each marked packet represents only partial path it has traversed, by combining a modest number of packets a victim can reconstruct the entire path.

Advantages:

- Victim can locate the approximate source of attack traffic without the assistance of outside network operators
- This determination can also be done even after the attack has stopped

Deterministic Packet Marking (DPM): Focuses only on the sources of the malicious packets, no matter which path the malicious packets take to attack the victim.

Each packet is marked when it enters the network. This mark remains unchanged as long as the packet traverses the network. Incoming packet is marked by the interface closest to the source of the packet on an edge ingress router. Marking is done deterministically.

BASIC MARKING ALGORITHMS

Node append: Simplest marking algorithm. Append each nodes address to the end of the packet as it travels

Marking procedure at router R:
for each packet w, append R to w

Path reconstruction procedure at victim v:
for any packet w from attacker
extract path (R_1..R_n) from the suffix of w

Fig. 5: Node append algorithm

through the network from the attacker to the victim. Every packet received by the victim will have the complete path traversed. Algorithm is robust and extremely quick to converge (Fig. 5).

Limitations:

- Router overload-infeasible
- Length of path cannot be predetermined-space constraints in packets
- Length of packet increases-fragmentation

Node sampling: The attack path is sampled one node at a time and avoids recording the entire path. When a router receives a packet it chooses to write its information (address) in the node field based on some probability p. when the victim receives the marked packets it will have atleast one sample for every router in the attack path and the path can be converged (Fig. 6).

Advantages:

- Requires only the addition of a write and checksum update
- Currently high speed routers are available which can handle the marking efficiently

Limitations:

- Inferring the total router order is a slow process
- Routers far away from the victim contribute lower number of samples and cam lead to misordering. (requires more samples to avoid this ie probability of marking in these routers must be higher)
- If multiple attackers are present, then multiple routers may be present at the same distance and hence will be sampled at same probability. Hence technique not robust against multiple attackers

Edge sampling: Instead of encoding individual node information in the packet encode the edge information. This includes the start and end nodes of the link and a distance field. When a router wants to mark the packet it enters its own address as the start information and sets the distance field to zero. If the distance field is already zero indicates that the packet was marked by previous

```
Marking procedure at router R:
        for each packet w
let x be a random number from [0..1]
if x < p then,
      write R into w.node

Path reconstruction procedure at victim v:
       Let NodeTbl be a table of tuples (node,count)
         for each packet w from attacker
           z:= lookup w.node in NodeTbl

           if zz!= NIL then
              increment z.count
           else
             insert tuple (w.node,1) in NodeTbl

sort NodeTbl l by count
extract path (Rf..Rl) from ordered node fields in NodeTbl
```

Fig. 6: Node sampling algorithm

```
Marking procedure at router R:
       for each packet w
         let x be a random number from [0..1]

         if x < p then,
          write R into w.start and 0 into w.distance
         else
          if w.distance = 0 then
              write R into w.end
          increment w.distance

Path reconstruction procedure at victim v:
    Let G be a tree with root v
    Let edges in G be tuples (start, end, distance)
    For each packet w from attacker

         if w.distance = 0 then
           insert edge (w.start, v, 0) into G
         else
           insert edge (w.start, w,end, w.distance) into G

remove any edge (x, y, d) with d ≠ distance from x to v in G
extract path (Rf..Rl) from ordered node fields in NodeTbl
```

Fig. 7: Edge sampling algorithm

router. In this case, the router adds its information to the end field and increments the distance by 1. Even if the router does not mark a packet it has to increment the distance field by 1 (Fig. 7).

SIMULATION OF IP TRACEBACK METHODS

Ns2 was used as our simulative tool. The network topology was constructed as a three layers tree with victim to be the root. The basic assumptions made are that:

- The attacker may generate any number of packets and the packets may be lost or reordered during transit

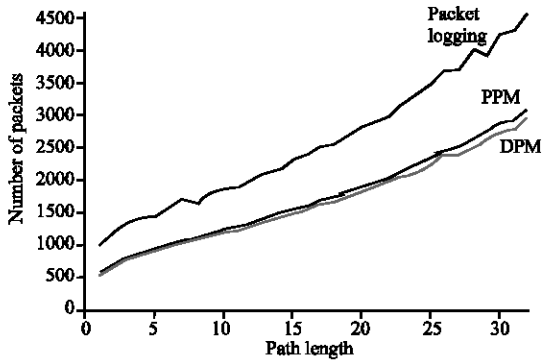

Fig. 8: Experimental results for number of packets required for path reconstruction with marking probability set at 1/25

- Multiple attackers may be involved and attackers may or may not be aware that they are being traced
- The path between attacker and victim is fairly stable
- Routers have limited CPU and memory constraints and are not widely compromised

Assuming a marking probability p, set to 1/25, the experimental results for number of packets needed to reconstruct paths of varying lengths is as shown in Fig. 8.

While IP-level traceback algorithm could be an important part of the solution for stopping denial-of-service attacks, it is by no means a complete solution. These algorithms attempt to determine the approximate origin of attack traffic-in particular, the earliest traceback-capable router involved in forwarding attack traffic from the source that directly generated it. Finally, traceback is only effective at finding the source of attack traffic, not necessarily the attacker themselves. Stopping an attack may be sufficient to eliminate an immediate problem, but long term disincentives may require a legal remedy and therefore the forensic means to determine an attacker's identity (http://www.cisco.com/warp/public/707/newsflash.html). Even with perfect traceback support, unambiguously identifying a sufficiently skilled and paranoid attacker is likely to require cooperation from law enforcement and telecommunications organizations.

CONCLUSION

The Internet has transformed from an information repository to a vital channel for conducting business. Unfortunately, with this positive change has come an increased frequency in malicious attacks (http://ciac.llnl.gov/cstc/nid/intro.html). All the proposed traceback schemes have their own specific advantages and

disadvantages. Currently, no single solution could fulfill all the requirements outlined for an effective trace-back method (Meadows, 1999). For any of these IP traceback solutions to be effective, they would need to be deployed across corporate and administrative boundaries in a substantial portion of the Internet infrastructure. This in itself seems to be one of the biggest obstacles to a unified approach to IP traceback. Also, some measures are ineffective against DDoS attacks, are resource intensive, cause network overhead and cannot be used for post-attack analysis. One conclusion we can draw from this is that unless IP traceback measures are deployed all over the Internet, they are only effective for controlled networks than for the Internet.

REFERENCES

CERT Coordination Center, 1999. Denial of Service Tools. http://www.cert.org/advisories/CA-1999-17.html.

CERT Coordination Center, 1997. IP Denial-of-Service Attacks. http://www.cert.org/ advisories/CA-1997-28.html.

CERT Coordination Center, 2001. Trends in Denial of Service Attack Technology. http://www.cert.org/ archive/pdf/DoS trends.pdf.

CERT Coordination Center, 1998. Smurf attack. http://www.cert.org/advisories/CA-1998-01.html.

CERT Coordination Center, 1996. TCP SYN flooding and IP spoofing attacks. http://www.cert.org/advisories/ CA-1996-21.html.

Cisco, Strategies to Protect Against Distributed Denial of Service Attacks. http://www.cisco.com/warp/public/ 707/newsflash.html.

Computer Incident Advisory Capability. Network Intrusion Detector Overview. http://ciac.llnl.gov/ cstc/nid/intro.html.

Ferguson, P. and D. Senie, 2000. Network Ingress Filtering: Defeating Denial of Service Attacks which Employ IP Source Address Spoofing. RFC 2827.

Jelena, M. and R. Peter, 2002. A Taxonomy of DDoS Attack and DDoS Defense mechanisms. In: Proceedings of the 2nd ACM SIGCOMM Internet Measurement Workshop.

Meadows, C., 1999. A formal framework and evaluation method for network denial of service. In: Proceedings of the 12th IEEE Computer Security Foundations Workshop.

Savage, S., D. Wetherall, A. Karlin and T. Anderson, 2000. Practical Network Support for IP Traceback. In: Proceedings of ACM SIGCOMM.