International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Forensics Tracking for IP Spoofers Using Path Backscatter Messages

Mithun Dev P D[1], Anju Augustine[2]

[1, 2]Department of Computer Science and Engineering, KMP College of Engineering,
Asamannoor P.O Poomala, Odakkali, Kerala, India

Abstract: ***Attackers may use spoofed IP addresses to conceal their real locations. A number of different mechanisms are suggested to track the spoofers. However, due to the difficulties of implementation, there has been no commonly adopted IP traceback mechanism, at least at the Internet-level. Consequently, the mist on the places of spoofers has never been dissipated until now. This paper suggests a novel passive IP spoofer tracking mechanism that bypasses the implementation difficulties of IP traceback methods. This mechanism uses Internet Control Message Protocol error messages (named path backscatter) triggered by spoofing packets, and traces the spoofers depending on publicly available information (e.g., topology). In this way, the mechanism can find the spoofers without any further deployment requirements. This work discusses the causes and collection of path backscatter messages. Furthermore, by employing the TTL field in IP packets, the geographical location details of routing device near to IP spoofers are found. Though the proposed scheme cannot work in all the spoofing attacks, it may be the most useful mechanism to trace spoofers before an Internet-level traceback system has been deployed in real.***

Keywords: Network Security, IP traceback, Time to live, Denial of Service, IP spoofing.

1. Introduction

IP Spoofing, which is technique used by attackers for initiating attacks using forged source IP addresses, is considered as a serious security issue on the internet. Attackers use addresses that are allocated to others or unassigned addresses, to prevent revealing their actual locations, or improve the impact of attack, or to launch reflection based attacks. Some well-known attacks that depend on IP spoofing are SYN flooding, SMURF, DNS amplification etc. A DNS amplification attack which seriously deteriorated the functioning of a Top Level Domain (TLD) name server is reported in [1]. A report of ARBOR on NANOG 50th conference reveals spoofing is still important in observed DoS strikes [2].

Identifying the origins of IP spoofing traffic is of great importance. As long as their locations are not revealed, they cannot be discouraged from launching further attacks. Even just nearing the spoofers, for example, determining the ASes (Autonomous Systems) or networks they live in, attackers can be located in a compact sized place, and filtration mechanisms can be placed closer to the attacker, before the spoofing traffic gets bundled. Furthermore, this can help develop a reputation system for ASes, which would be beneficial to force the corresponding ISPs to verify IP addresses.

2. Related Work

The related work can be categorized into two parts. The first one describes the existing IP traceback mechanisms, and the second one introduces IP spoofing observation activities.

2.1 IP Traceback

IP traceback methods are developed to reveal the real origin of IP traffic or track the path. The existing IP traceback approaches can be classified into the following: packet marking, ICMP traceback, logging on router, link testing, overlay, and hybrid tracing.

In packet marking techniques, the routers are required to modify the header of packets to contain information of the router and the forwarding decision. The received packets can then be utilized, by the receiver to reconstruct the path of the packets. There are two types of packet marking schemes: probabilistic packet tagging [4], [8]–[11] and deterministic packet tagging [12]–[15]. As packet marking is not widely supported by routers, it is challenging to enable packet marking in the network.

In ICMP traceback [5], [16], additional ICMP messages are generated to a collector or the destination. It can be used to rebuild the attack path. The drawback of ICMP traceback is that it utilizes more bandwidth by generating considerable additional traffic. Additionally, if the attack is against the bandwidth of the victim, the additional traffic will favour the attack.

Logging on router [6] involves routers keeping a history of all the packets it has forwarded. Attack path can be rebuilt from log on the router. In link testing scheme, the upstream of hop-by-hop attacking traffic is determined, while the attack is in progress.

Overlay scheme [17] involves employing special tracking routers where suspect traffic is offloaded from edge router to

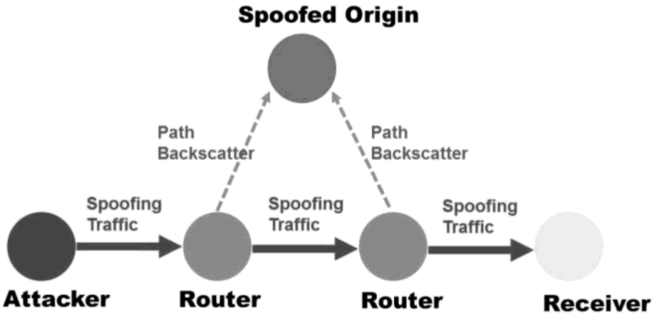

Figure 1: Path backscatter generation and collection

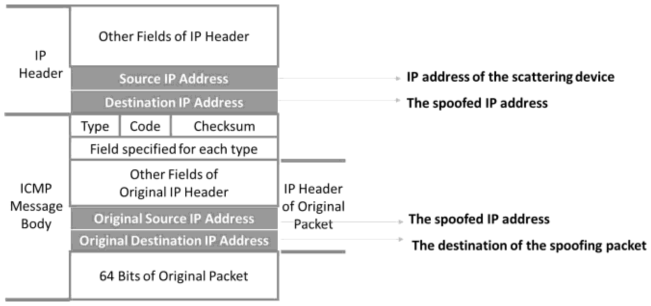

Figure 2: Path backscatter message format

them through an overlay network. Hybrid schemes employs combination of the above mentioned techniques to achieve better traceability and to reduce the cost.

Though there are a huge variety of appealing traceback mechanisms, these techniques are not accepted and implemented widely, especially at the Internet level.

2.2 IP Spoofing Observation

A fundamental technique for passive observation of spoofing activities is the use of Network Telescopes [3].Network telescope catches non-solicited messages, which are mainly generated by victim systems struck by traffic with source prefix set in the range of the telescope. At present, the biggest range telescope is the CAIDA UCSD telescope, which holds 1/256 of all the IP addresses and is mainly used to monitor DDoS activities. Moore el at. [7] provided a technique known as backscatter analysis which infers features of DoS strikes based on records gathered by the network telescope.

The MIT Spoofer Project [18] tries to reveal which networks are able to release spoofing centred strikes. Volunteer participants set up a client that assess the spoofing capability of their hosts and networks. The results reveals 6700 ASes out of 30205 do not filter spoofing.

3. Path Backscatter Messages and TTL

3.1 Overview of Path Backscatter

Many packets may not reach their intended destination. A router may fail to forward a packet due to various factors. It may produce an ICMP error message, i.e., path backscatter message, under some circumstances. The source IP address indicated in the original packet will receive the path backscatter messages. If the source address is spoofed, then the messages will be sent to the node who actually own the address. This means that the victims of reflection based attacks, and hosts whose addresses are used by spoofers, may collect such information. This situation is shown in Figure 1.

The structure of the path backscatter message, is shown in Figure 2. as specified by RFC792 [19]. Each message contains mainly two parts: IP header and ICMP message body. The IP header part contains 1) the IP address of the scattering device i.e. router, which is on the path from the attacker to the destination of the spoofing packet; 2) the spoofed IP address i.e. the victim. The ICMP message body part contains 1) the spoofed IP address; 2) the original destination of the spoofing packet. The original IP header also contains the remaining TTL of the spoofing packet.

3.2 Classes and Causes of Path Backscatter

The path backscatter messages may be generated due to various reasons. There are totally 5 kinds of path backscatter messages. There are a variety of codes associated with each type. The combination of code and type determines the cause that the router decides to send the ICMP message. The combination of code and type may be named as a class. The different types and the associated classes are explained below.

3.2.1 Time Exceeded

Packets with zero TTL value triggers TIMXCEED_INTRANS messages. These are the most common path backscatter messages.

3.2.2 Destination Unreachable

The filtering mechanisms such as ACLs deployed between the spoofing origin and the victim may trigger UNREACH_FILTER_PROHIB,UNREACH_NET_PROHIB and UNREACH_HOST_PROHIB. If there is no route to destination, then UNREACH_HOST and UNREACH_NET messages are generated. If the size of attacking packets are larger than MTU of a hop on the path, and if the DF (Don't Fragment) flag is set, UNREACH_NEEDFRAG messages are generated.

3.2.3 Source Quench

When the router has no buffer to queue the original packet, SOURCEQUENCH messages are generated. It may be resulted from heavy aggregated attacking traffic which the router cannot forward.

3.2.4 Redirect

If the spoofing origin has two or more gateways, and one of the gateway finds that the packet should be sent through another gateway as it is the shortest route, REDIRECT_HOST and REDIRECT_NET messages are generated.

3.2.5 Parameter Problem

If the router finds a problem with the header parameters in the original packet, PARAMPROB messages are generated.

3.3 Collection of Path Backscatter

In the following section, we classify spoofing based attacks into four categories, and discuss whether path backscatter message can be gathered in each category of attacks.

3.3.1 Multiple Sources, Single Destination

In this type of attack, the source address of spoofed packets are selected from a set of candidate addresses. This type of attack, named as random spoofing, is used to deplete the resource of a target. In random spoofing, path backscatter messages can be collected using Network Telescopes.

3.3.2 Single Source, Multiple Destinations

In this type of attack, all the spoofed packets have the same source address. This type of attack, named as reflection attacks, sends spoofed packets to different destinations. In reflection attacks, the victim of the attack can collect the path backscatter messages.

3.3.3 Multiple Sources, Multiple Destinations

These type of attacks are usually regarded as a combination of the above two category of attacks.

3.3.4 Single Source, Single Destination

In this type of attack, spoofed packets will be send from a single source address to a single destination address. It is usually done to steal or break a session between two communicating parties. The spoofed origin will receive all the path backscatter messages.

3.4 TTL

Place Time-to-live (TTL) is a value in an Internet Protocol (IP) packet that informs a router whether or not the packet has been in the network for too long and should be discarded. A packet may not reach its destination, in a reasonable time frame due to several reasons. Such packets may loop endlessly on the network.

To avoid this situation, the packet may be discarded after a certain time and send a message to the originator, who can decide whether to resend the packet. The initial TTL value is set, usually by a system default, in an 8 bit field of the packet header. The original concept of TTL was that it would specify a certain time frame in seconds that, when exhausted, would cause the packet to be discarded. Since each router is needed to deduct at least one count from the TTL field, the count is usually used to mean the number of router hops the packet is permitted before it must be removed. Each router that gets a packet subtracts one from the count in the TTL field. When the count gets to zero, the router discovering it discards the packet and sends an ICMP message back to the originator.

The standard Windows XP/7 TTL value is 128. The ping and the traceroute utilities both make use of the TTL value to attempt to reach a given host computer or to trace a route to that host. Using the multicast IP protocol, the TTL value indicates the scope or range in which a packet may be forwarded. By convention:
0 is restricted to the same host
1 is restricted to the same subnet
32 is restricted to the same site
64 is restricted to the same region
128 is restricted to the same continent
255 is unrestricted

4. Proposed Tracking Mechanism

By employing the concepts of path backscatter messages and TTL, we can converge the location of a spoofer to a small geographical region. i.e we can track upto the Autonomous System (AS), or the edge router nearest to the spoofer or his network.

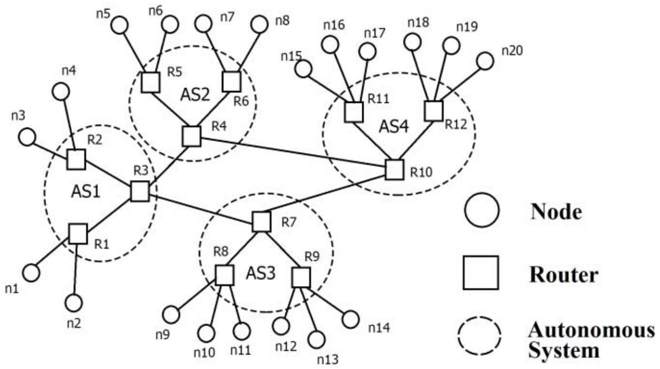

Figure 3: Sample system layout

Fig. 3. shows a sample layout of the proposed scheme. The attacker node, say node n1, spoofs IP address of victim node, say node n5 and sends packets to a destination node, say n15. The spoofed packets flows through the network and reaches the destination node. In our work, the shortest path to destination is considered for packet transmission. Meanwhile, as these are spoofed packets, each router on the path sends path backscatter messages to victim node. The victim node gets the IP addresses of each individual routers, from each path backscatter messages. Upon receiving all the path backscatter messages, the victim knows about all the routers that forwarded the spoofed packet. These set of routers constitute a suspect set. The original spoofer may be attached or nearby any one of those routers.

For finding the router nearest to the spoofer, we employ the concept of TTL. The ICMP message body contains the original IP header information. Moreover, the original IP header contains the remaining TTL value of the spoofing packet. The TTL value is set by the sender of a packet, and is reduced (by 1) at every router on path to the destination. Hence we can conclude that the router with the largest TTL value will be the one closest to the sender. Using this scheme the router nearest to the spoofer is obtained. With the help of different IP location trackers and databases available, the geographical location of these routers can be obtained.

5. Performance Evaluation

The proposed tracking scheme has been implemented at the IP (network layer) level in network simulator 2. The sample system layout in Figure 3. was simulated in normal single attacker scenario. The simulation is done as follows. To begin with, three nodes are selected as the attacker (spoofer), victim and the destination. The attacker node sends spoofing packets with the victim node's IP address to the destination node. The default initial TTL value is considered as 255, which will be decremented by each router on the path to the destination. Since the packets are spoofed ones, each of the routers sends path backscatter messages to the victim node. Victim node determines the suspect set using information from the path backscatter messages. The router nearest to the attacker is determined by comparing the remaining TTL value of the original spoofed packet from each of the routers.

The following performance metrics are recorded: Accuracy, Cumulative Fraction, and Number of Bytes Received. The simulation results are shown below.

5.1 Accuracy

The proposed scheme is compared with a packet marking scheme. The accuracy of proposed scheme is found to be more when considering accuracy against the number of packets available. Existing mechanisms such as packet marking schemes needs a large amount of packets to accurately locate the spoofer.

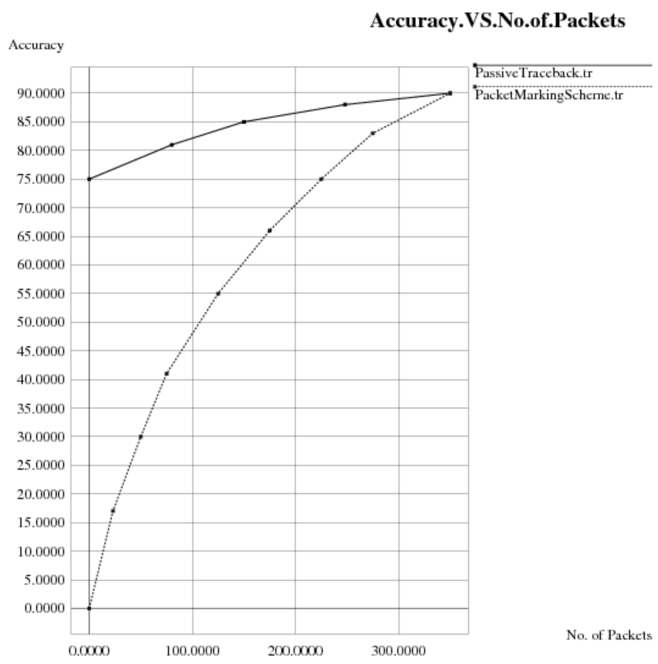

Figure 4: Accuracy

5.2 Cumulative Fraction

The CDF (Cumulative Distribution Fraction) of involved original destinations IP numbers of top reflectors are plotted. The results show that a small number of reflectors forwarded spoofing traffic to a large number of original destinations.

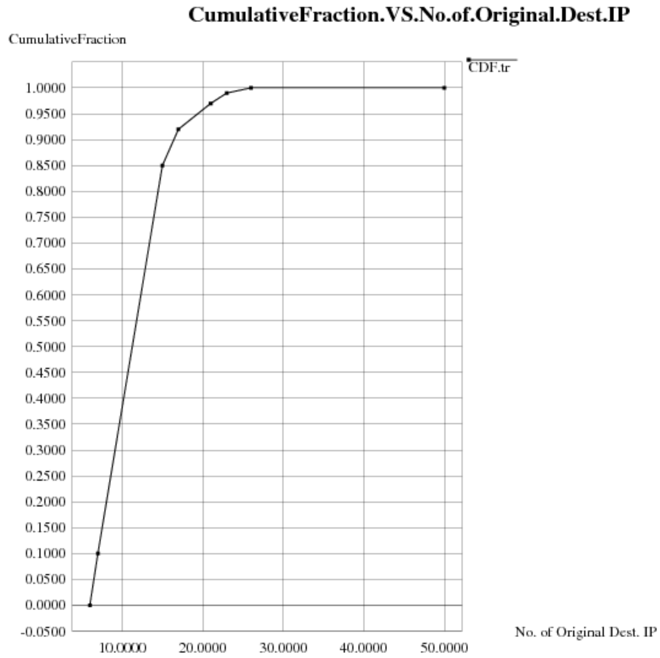

Figure 5: Cumulative Fraction

5.3 Number of Bytes Received

The number of bytes received is plotted against time. The results shows that the number increases without any remarkable abrupt changes with time.

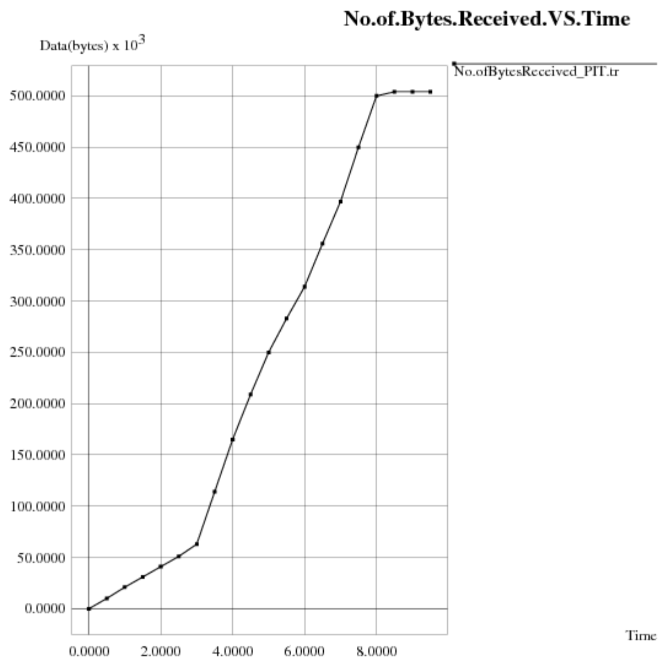

Figure 6: Number of bytes received

6. Conclusion

In this work, a new IP Traceback procedure which tracks spoofers depending on path backscatter messages and Time-to-live (TTL) is introduced. The different causes, classes, and collection of path backscatter messages are illustrated. The effectiveness of the proposed scheme is demonstrated based on deduction and simulation. The proposed scheme may be applied to available path backscatter dataset to find location of spoofers on a large scale. The results may further be used to reveal IP spoofing, and prevent attacks such as Denial-of-service (DoS) attacks.

References

[1] ICANN Security and Stability Advisory Committee, "Distributed denial of service (DDOS) attacks," SSAC, Tech. Rep. SSAC Advisory SAC008, Mar. 2006.

[2] C. Labovitz, "Bots, DDoS and ground truth," presented at the 50th NANOG, Oct. 2010.

[3] The UCSD Network Telescope. [Online]. Available: http://www.caida.org/projects/network_telescope/

[4] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Practical network support for IP traceback," in Proc. Conf. Appl., Technol., Archit., Protocols Comput. Commun. (SIGCOMM), 2000, pp. 295–306.

[5] S. Bellovin. ICMP Traceback Messages. [Online]. Available: http://tools.ietf.org/html/draft-ietf-itrace-04, accessed Feb. 2003.

[6] A. C. Snoeren et al., "Hash-based IP traceback," SIGCOMM Comput. Commun. Rev., vol. 31, no. 4, pp. 3–14, Aug. 2001.

[7] D. Moore, C. Shannon, D. J. Brown, G. M. Voelker, and S. Savage, "Inferring internet denial-of-service activity," ACM Trans. Comput. Syst., vol. 24, no. 2, pp. 115–139, May 2006. [Online]. Available: http://doi.acm.org/10.1145/1132026.1132027

[8] M. T. Goodrich, "Efficient packet marking for large-scale IP traceback," in Proc. 9th ACM Conf. Comput. Commun. Secur. (CCS), 2002, pp. 117–126.

[9] A. Yaar, A. Perrig, and D. Song, "FIT: Fast internet traceback," in Proc. IEEE 24th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM), vol. 2. Mar. 2005, pp. 1395–1406.

[10] J. Liu, Z.-J. Lee, and Y.-C. Chung, "Dynamic probabilistic packet marking for efficient IP traceback," Comput. Netw., vol. 51, no. 3, pp. 866–882, 2007.

[11] M. Adler, "Trade-offs in probabilistic packet marking for IP traceback," J. ACM, vol. 52, no. 2, pp. 217–244, Mar. 2005.

[12] A. Belenky and N. Ansari, "IP traceback with deterministic packet marking," IEEE Commun. Lett., vol. 7, no. 4, pp. 162–164, Apr. 2003.

[13] Y. Xiang, W. Zhou, and M. Guo, "Flexible deterministic packet marking: An IP traceback system to find the real source of attacks," IEEE Trans. Parallel Distrib. Syst., vol. 20, no. 4, pp. 567–580, Apr. 2009.

[14] R. P. Laufer et al., "Towards stateless single-packet IP traceback," in Proc. 32nd IEEE Conf. Local Comput. Netw. (LCN), Oct. 2007, pp. 548–555. [Online]. Available: http://dx.doi.org/10.1109/LCN.2007.160

[15] M. D. D. Moreira, R. P. Laufer, N. C. Fernandes, and O. C. M. B. Duarte, "A stateless traceback technique for identifying the origin of attacks from a single packet," in Proc. IEEE Int. Conf. Commun. (ICC), Jun. 2011, pp. 1–6.

[16] H. C. J. Lee, V. L. L. Thing, Y. Xu, and M. Ma, "ICMP traceback with cumulative path, an efficient solution for IP traceback," in Information and Communications Security. Berlin, Germany: Springer-Verlag, 2003, pp. 124–135.

[17] R. Stone, "CenterTrack: An IP overlay network for tracking DoS floods," in Proc. 9th USENIX Secur. Symp., vol. 9. 2000, pp. 199–212.

[18] R. Beverly, A. Berger, Y. Hyun, and K. Claffy, "Understanding the efficacy of deployed internet source address validation filtering," in Proc. 9th ACM SIGCOMM Conf. Internet Meas. Conf. (IMC), 2009, pp. 356–369.

[19] J. Postel. Internet Control Message Protocol, RFC792. [Online]. Available: https://tools.ietf.org/html/rfc792, accessed Sep. 1981.

Author Profile

Mithun Dev P.D received the B.Tech Degree in Information Technology from Cochin University of Science and Technology, Kerala, India, in 2012. He is currently pursuing M.Tech Degree in Computer Science and Engineering with Specialization in Cyber Security from Mahatma Gandhi University, Kerala, India. His research interests include information/network security.

Anju Augustine received the B.Tech Degree in Information Technology, and M.Tech Degree in Computer Science and Engineering with Specialization in Information Systems from Mahatma Gandhi University, Kerala, India, in 2010 and 2014 respectively. She is currently working as Assistant Professor at K.M.P College of Engineering, Kerala, India.