On the Issues of IP Traceback for IPv6 and Mobile IPv6

Henry C.J. Lee, Miao Ma, Vrizlynn L.L. Thing, Yi Xu
Institute for Infocomm Research (I^{2}R)
{hlee, miaom, vriz, yxu}@i2r.a-star.edu.sg

Abstract

As the Internet becomes pervasive, the vulnerability of some fundamental design aspects of the Internet has also become significant. Among which, Denial-of-Service (DoS) and Distributed DoS (DDoS) pose significant problems, as they are disruptive to the useful traffics and are hard to prevent. One solution consists in instituting accountability, which hold the attackers accountable for the attack. The key issue is to identify the "real" sources of the attacks as attackers use spoofed IP address to hide their actual network location. However, the Internet architecture does not provide intrinsic support for identifying the real sources of IP packets. Numerous mechanisms have been proposed to "traceback" the real sources. Most of such works have been addressing the IP version 4. In this paper, we address the issues of IP traceback in the context of IPv6 and Mobile IPv6. This paper provides a detailed analysis of these issues and problems. The main problem lies with the transformations that are introduced by IPv6 and Mobile IPv6 protocols, namely tunneling and addresses manipulation. We then propose a solution, including new ICMPv6 messages for traceback co-ordination, to facilitate the traceback mechanism.

1 Introduction

Without any doubt, the Internet is becoming the pervasive means of communications for data in particular. However, its pervasiveness has also generated many security problems, such as authentication, data confidentiality, data integrity, intrusion etc. In this paper, we look at the issue of determining the actual source of IP packets, in the context of the problems of Denial of Service (DoS) [1] and Distributed DoS (DDoS) attacks. In a DoS attack, typically huge quantity of packets are generated and directed towards one or many victims, causing congestion at the intermediate routers as well as the end systems. DDoS is a variation of DoS in that the attacker launches an attack not from one single source, but from several sources that the attacker has already penetrated. As a result, the legitimate data traffics are disrupted and services are denied to the legitimate users. In such attacks scenario, attackers usually send packets with spoofed IP addresses so as to hide its true network location from the victims and the network infrastructure. Besides the scenario of DoS/DDoS, other form of attacks such as network intrusion is also typically carried out with spoofed IP address. In order to institute accountability, the actual origin of IP packets has to be determined. However, the nature of the DoS attacks are different from network intrusion attack in that the amount of attacker traffic in the former case is usually very significant while the traffic may not be significant in the latter attack scenario. In this paper, we focus on the traceback issue for the DoS/DDoS scenario, although some of the analysis can be generalized to other cases.

The IP [2] packet contains two addresses: source and destination. The destination address is used by the routing architecture to deliver the packet. The IP network routing infrastructure does not verify the authenticity of the source address carried in IP packets. The source address is used by the destination host to determine the source for message reply. In general, no entity is responsible for the correctness of source address. The scenario is the same as sending letters using postal service; the postal service does not care about the correctness or authenticity of the source address, it merely makes sure that the letters are delivered to the correct destination.

The design of the IP protocol and forwarding mechanism makes it difficult to identify the originator of a packet. This characteristic of the Internet is exploited by some malicious users of the Internet to hide their source and identify. Some mechanisms such as "Ingress filtering" [3] are used to enforce the validity of source IP address originating from a stub network. However, such mechanisms are quite limited as they can only be used in edge networks and the enforcement of which is difficult.

This paper is organized as follows: section 1 gives the introduction. Section 2 gives some background information on IP traceback, IPv6 and Mobile IPv6. Section 3 provides an analysis of current traceback solutions applied to the IPv6 and Mobile IPv6 protocols, the research issues as well as the proposed solutions. Section 4 concludes the paper.

2 Background

2.1 IP Traceback

The challenge of IP Traceback is to find an efficient and scalable way to track the sources of an arbitrary IP packet. The source can be an Ingress point to the traceback-enabled network, the actual host or network of origin, or compromised routers within the enabled network. It depends on the extent to which the traceback framework is deployed. In an attack, it is possible that the routers may be subverted, hence there is a need to construct the attack path, which comprises the routers traversed by packets from the "source" to the victim. In the case of a DDoS attack, packets come from potentially many secondary sources, hence many attack paths. The attack graph is defined as the set of attack paths.

The objective of the IP traceback mechanism is to construct the attack graph with the constraint that it should minimize the time that routers spend on tracking and minimize the storage used to keep the tracking information. Lastly, the solution should not adversely impact the privacy of legitimate users.

As an IP packet travels from the source to the destination, it may be modified by various mechanisms such as: TTL (Time to Live), checksum recomputation, fragmentation due to MTU (Maximum Transfer Unit), IP option processing, NAT, Encapsulation (IP in IP) or Mobile IP. As a result, the packet received by the victim may have taken various forms en route from the attacker to the victim. Consequently, the traceback mechanism has to be able to trace the route taken by the packet, even though the packets that it receives is not identical to what the routers in the network has processed.

2.1.1 Approaches

There are two main approaches to perform traceback: infrastructure scheme and end host scheme.

In the first approach, infrastructure scheme, the network is responsible for maintaining the traceback state information, which is the information necessary for the victim and the network to reconstruct the attack graph. This approach employs the traceback logging (or IP logging) technique. In other words, the network routers log the passage of IP packets. The key challenge here lies in the potential huge amount of information storage. For example, if the router were to log all the packets in its entirety, each OC-192 link at 1.25GB/s at the router requires 75 GB of storage for a 1-minute query buffer. The storage requirement quickly becomes prohibitive as the number of router links increases.

In the second approach, end host scheme, the traceback state information is stored at the end hosts. The approach is usually achieved by one of the following techniques: traceback marking (or IP marking) or ICMP messaging [10]. In the traceback marking technique, the path information is encoded within the IP header, typically in rarely-used fields. In the ICMP messaging technique, the path information is sent separately in dedicated ICMP messages from the routers to the victim.

2.1.2 IP logging

One mechanism for IP logging, SPIE (Source Path Isolation Engine) [8], has been proposed for IP version 4. The mechanism is designed to identify the true source of a particular IP packet given a copy of the packet to be traced and an approximate time of receipt. The proposed scheme requires that the intermediate routers log the passage of all IP packets. In order to take care of packets transformation, the mechanism consists in identifying the invariant portions of the 20-byte IPv4 header. The fields that are susceptible to changes include: TOS (Type of Service), TTL (Time to Live), Checksum and Options field. The logging is based on the invariant portion of the IP header and the first 8 bytes of payload. Based on statistics collected, the 28-byte prefix described above results in a rate of collision approximately 0.00092% in a WAN environment and 0.139% in a LAN environment. To further reduce the storage requirement, instead of storing the entire 28-byte prefix, hashing is performed on it, followed by a Bloom filter processing. The scheme reduces memory storage requirement in the router to 0.5% of link bandwidth per unit time. It also maintains privacy, prevents eavesdropping of legitimate traffic stream.

2.1.3 IP marking

This mechanism marks the IP packets with additional information so that the victim can use them to determine the attack path. Approaches proposed include node append, node sampling and edge sampling [9]. The node append mechanism is similar to the IP Record Route Option [2], in that the addresses of successive routers traversed by an IP packets are appended to the packets. The victim can thus easily traceback the source of such attack packets. However, this method introduces very high overhead in terms of router processing and packet space. The node sampling approach reduces such overhead by the probabilistic marking of IP packets. The edge sampling approach, as its name imply, marks an edge of the network topology, traversed by the IP packets, instead of just the node. Most proposed algorithms put the marking information in the Identification field of the IP header. This type of mechanisms has an inherent disadvantage in that it affects the format of IP packets. The necessary changes in the IP packet format depends on the algorithm. The standardization of format for IP marking becomes an issue.

2.2 IPv6

The 40-byte IPv6 protocol [4] header is shown in Figure 1.

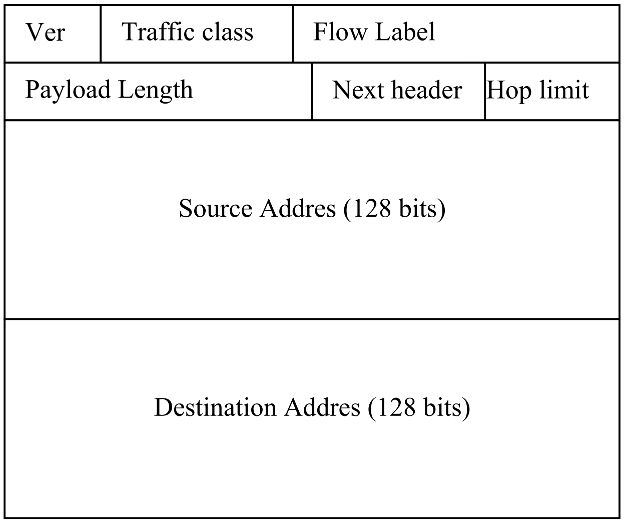

Figure 1. IPv6 header format

In the IPv6 header, the invariant fields are version (4 bits), payload length, next header and source address (128 bits). The fields of the header affected by transformation are hop limit, traffic class, destination address and flow label.

The 8-bit Traffic Class field in the IPv6 header can be used by originating nodes and/or forwarding routers to identify and distinguish between different classes or priorities of IPv6 packets. Nodes that support a specific use of some or all of the Traffic Class bits are allowed to make changes to them when they originate, forward, or receive, as required for that specific use. For nodes which do not support the specific usage of the Traffic class fields, they should ignore the bits and leave them unchanged. An upper-layer protocol must not assume that the value of the Traffic Class bits in a received packet are the same as the value sent by the packet's source.

The 128-bit Destination address field is not necessarily the ultimate destination. It is set as the next hop address when source routing is used. The source route is kept inside the Type 0 routing header. The 8-bit hop limit field is decremented as the packet is forwarded by the routers.

The 20-bit Flow Label field in the IPv6 header may be used by a source to label sequences of packets for which it requests special handling by the IPv6 routers, such as non-default quality of service or "real-time" service. Currently, this aspect of IPv6 is still not finalized as the requirements for flow support in the Internet remains unclear. Nodes that do not support the functions of the Flow Label field are required to set the field to zero when originating a packet, pass the field on unchanged when forwarding a packet, and ignore the field when receiving a packet. Consequently, we assume that this field is not an invariant one.

2.3 Mobile IPv6

In the current routing framework in the Internet, the IP address is used to identify the topological location of an interface of a router or a host. The connections between applications are identified by their IP addresses. With the advent of wireless communications, host computers or PDA can roam geographically and topologically, which results in the change of IP address, thus affecting ongoing connections and higher layer applications. The Mobile IP [7] protocol for IPv6 has been designed to enable mobile nodes to maintain connections using a fix home address while roaming to foreign domains and networks.

Each mobile node (MN) is given a home address (HoA) by the home network. A special node “Home Agent” (HA) is designated to act as a proxy for the MN when it moves away from the home network. When the MN roams into a foreign network, it obtains a Care-of-Address (CoA) and registers it with the HA so that the HA knows the current location of the MN. When correspondent nodes (CN), which are not aware of MN's movement, sends packets to the MN’s home address, they are intercepted in the MN’s home network by the HA and tunneled [6] to the MN. Subsequently, the MN updates the CN of its new CoA so that subsequent data traffic can be sent directly between the CN and MN, without going through the home network.

When the MN sends packets to the CN, it uses its CoA as the source address and puts its actual home address in the Destination Option (Home address option). Upon reception by the CN, the source address is replaced by the MN's home address so that the application perceives that it is still communicating with the MN at its HoA. On the other hand, when the CN sends packets to the MN, it puts the MN's CoA in the destination field and the home address in Type 2 Routing header. Figure 2 depicts the addressing mechanism.

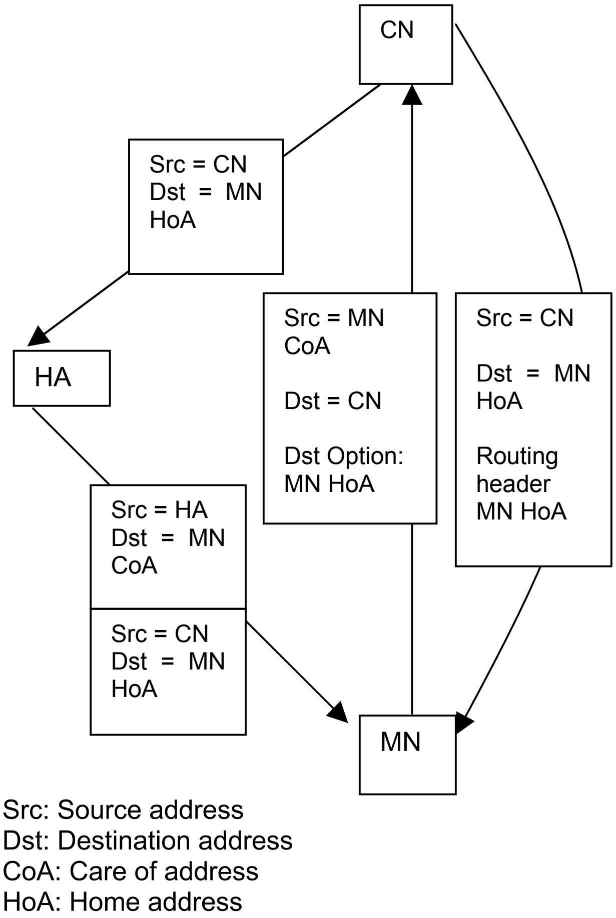

Figure 2: Mobile IPv6 addressing mechanism

3 IP Logging for IPv6 and Mobile IPv6

3.1 Improved IP Logging for IPv4

The SPIE solution proposed in [8] states that the IPv4 destination address field is invariant. However, we note that this is valid only when the Losse/Strict Source routing mechanism in IPv4, which replaces the destination field with the successive hops of the source route, is not used. Otherwise, the destination address field contains successively the addresses of the routers on the source route path. The SPIE mechanism is thus vulnerable to attackers who exploit the source routing option. In fact, when the attacker puts a few intermediate routes into the packet, at each intermediate router, the packet is "transformed" from the viewpoint of the SPIE mechanism. When the victim tries to traceback the attack path, it is only able to traceback to the last router in the source path route. We propose two solutions to address this problem.

Solution 1: Final destination address logging

This solution is to log the IP packets by always using the final destination instead of the address found in the destination field. This is achieved by extracting the final destination address from the Loose/Strict Source routing option. Since the destination (victim) node receives the IP packet with the final destination, the traceback will be straightforward. However, this method will incur additional computation in the routers. In fact, all the router traversed by the packets have to process the source routing option to include the appropriate destination address for IP logging, regardless of whether the current router is part of the source route.

Solution 2: Destination address transformation logging

In this solution, the routers log the IP packet transformation due to source routing, in addition to the logging of the IP packet itself. As compared to the previous solution, this method is likely to incur less computation, as this processing will only be done when the current router is on the source route. Packets without source routing options, or packets with source routing option but not including the current router as part of the source route are not affected. In this scenario, when the victim initiates a traceback with the packet containing the final destination address, the routers on the source path will replace it successively with the addresses stored in the transformation log.

3.2 IP Logging for IPv6

We have seen in the previous section that the invariant fields in IPv6 headers are version, payload length, next header and source address. The destination address can be modified by the source routing mechanism. When source routing is used, the successive addresses are stored in the Type 0 routing header. As the packet reaches each intermediate destination router, the destination address is replaced by the next hop address.

As in IPv4, it is difficult to perform the traceback if the destination address is not used for logging. If the destination address is used, the traceback can only be performed to the last hop router. Similarly to IPv4, two mechanisms can be used to address this issue arising from the change of destination address field. We propose that the router log the IP packet transformation due to the destination address, i.e. the Destination address transformation logging (DTTL). In such a way, the source can be successively traceback from the victim, even if source routing has been used. For the following discussion, we assume that the routers log the transformation of the destination IPv6 address.

3.3 IP Traceback for Mobile IPv6

As discussed earlier, the key transformations introduced by the Mobile IPv6 protocols are the IPv6 tunneling and the swapping of home address with the CoA. We analyze successively the various possible scenarios of attacks, where the HA, CN or MN is under DoS or DDoS attacks.

3.3.1 HA under attack

This case is not specific to Mobile IPv6. It can be treated as a normal attack for IPv6.

3.3.2 CN under attack

When CN is the victim of an attack and that the attack packets contains a "destination home address option", the CN should reconstruct the actual packet that has traversed the network by substituting the source address field with the MN's CoA and the destination option containing the MN's home address.

3.3.3 MN under attack

The MN can be attacked in two different ways: directly and indirectly. When it is attacked directly, packets are destined to its CoA. In the other case, packets are sent to its home address. From the viewpoint of the types of packets, the MN can receives on its interface the three following formats:

1. IP packet without routing header. These are packets sent to the MN's CoA, without any reference to its HoA.
2. IP packet with routing header containing the MN's home address, for example, packets sent to MN by CN who has acquired the MN's CoA binding.
3. Tunneled IP packet, for example, packets sent by CN to the MN's HoA and then tunneled to the MN.

The first case is actually the generic scenario that can be processed normally. In the second case when the routing header is used, the MN has to reconstruct the actual packet for traceback by setting the appropriate destination address before triggering the traceback procedure.

In the third case, the tunneled packet can come in two forms:

1. The attacker sends packets with spoofed source address to the MN's home address. Such packets are routed to HA, which are then tunneled to the MN
2. The attacker constructs fake tunneled packets with HA as the source address in the outer packets and spoofed source address in the inner packets. Such packets are routed directly to the MN without passing through the HA.

In the first scenario, when the HA receives the attack packets destined for the MN, it is unable to identify any attack. Only when the packets reach the destination node, i.e. the MN, that an attack can be identified. However, given the nature of traceback, the MN cannot initiate the traceback directly due to the tunneling. Consequently, MN has to notify HA that it is under attack so that the HA can trigger the traceback mechanism to determine the attack path. A new signaling mechanism and protocol is required for MN to signal to HA of the attack.

The second scenario is ambiguous. If the MN trigger the same process above and inform the HA to begin traceback for the inner packet, the HA will not be able to obtain any result, simply because the packet has never been routed by HA in the first place. This loophole can be easily exploited by attackers to bypass currently proposed traceback mechanism. However, this situation can be prevented if the IPv6 AH (Authentication Header) is used for all tunneled traffic between the HA and MN. This would enable the MN to verify the authenticity of the incoming tunneled packets so as to detect illegitimate tunneled packets and initiate the right traceback procedure, i.e. either the MN initiate the traceback directly or request the HA to perform the traceback.

3.4 Proposed ICMPv6 protocol for attack notification

As we have seen previously, a notification mechanism is required for the MN to signal the HA that it is the victim of a DoS/DDoS attack. We propose two ICMPv6 messages [5] for the attack notification: ICMPv6 Traceback request (Figure 3) and ICMPv6 Traceback reply (Figure 4).

Type	Code	Checksum
IPv6 packet that HA tunnels to MN		

Figure 3: ICMPv6 Traceback request

Type	Code	Checksum
Source address of attacker		

Figure 4: ICMPv6 Traceback reply

As the MN detects an attack in the form of authenticated tunneled messages from the HA, it will send the ICMPv6 Traceback request message to the HA. The HA will commence the traceback using the tunneled packet content. Once the traceback is completed, the HA will notify the MN of the attacker's source address using the ICMPv6 Traceback reply message.

4 Conclusion

In this paper, we studied the issues and problems relating to the IP logging mechanism for IPv6 and Mobile IPv6. We analyzed the applicability of the IP logging mechanism in the context of IPv6 and Mobile IPv6. Our analysis shows that the problem arises from the packet transformations performed by IPv6 and Mobile IPv6 mechanisms. We studied the IPv6 protocol formats and processing mechanism and identified the fields in IPv6 that can be used for IP logging. We identified the problem posed by the use of source routing which has not been identified in previous work on IP logging. We then proposed a computationally efficient solution for the routers to log the transformation of the destination address so that the traceback mechanism can effectively trace back to the true source. For Mobile IPv6, we studied the problems when the MN is the victim. We identified a vulnerability which can be easily exploited by attackers to attack the MN and rendering traceback mechanism ineffective. We concluded that the tunneled traffic between the HA and MN has to be authenticated. In addition, we studied the effect of tunneling and concluded that the DoS/DDoS attack in the presence of Mobile IPv6 necessitate notification mechanism in order to co-ordinate traceback. Finally, we propose a new ICMPv6 protocol message for the traceback coordination.

5 Future Work

In this paper, we have focused our analysis based on the IP logging traceback mechanism. Other traceback methods like IP marking and ICMP traceback messaging give rise to new set of problems. For example, in the case of IP marking. assuming that the attack is aimed at the MN's home address, the HA will need to obtain more attack packets from MN to reconstruct the attach path and to effectively traceback the source.

6 Reference

[1] K.J. Houle, G.M. Weaver, "Trends in Denial of Service Attack Technology", CERT Coordination Center, Oct 2001. http://www.cert.org/archive/pdf/DoS_trends.pdf

[2] J. Postel, “Internet Protocol”, Request for Comments 0791, Internet Engineering Task Force, 1981.

[3] P. Ferguson, D. Senie, “Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing”, Request for Comments 2827, Internet Engineering Task Force, May 2000.

[4] S. Deering, R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification”, Request for Comments 2460, Internet Engineering Task Force, December 1998.

[5] Conta, S. Deering, “Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification”, Request for Comments 2463, Internet Engineering Task Force, December 1998.

[6] Conta and S. Deering, “Generic Packet Tunneling in IPv6 Specification”, Request for Comments (Proposed Standard) 2473, Internet Engineering Task Force, December 1998.

[7] David B. Johnson, Charles E. Perkins, “Mobility Support in IPv6”, Internet Draft <draft-ietf-mobileip-ipv6-18.txt>, Internet Engineering Task Force, June 2002.

[8] Alex C. Snoeren et al, “Hash-Based IP Traceback”, ACM SIGCOMM 2001, August 2001.

[9] Stefan Savage et al, “Practical network support for IP traceback”, ACM SIGCOMM 2000.

[10] Steve Bellovin et al, "ICMP Traceback messages", IETF Internet Draft draft-ietf-itrace-03.txt, Jan 2003.