IPv6 Packet Traceback: A Survey

Rajesh Kumar Singh
M.tech Scholar
Department of Computer Science and Engineering
Graphic Era University Dehradun, India

Sumit Pundir
Assistant Professor
Department of Computer Science and Engineering
Graphic Era University Dehradun, India

Emmanuel S. Pilli,Ph.D
Professor
Department of Computer Science and Engineering
Graphic Era University Dehradun, India

ABSTRACT
The Internet is expanding year by year and providing services of convenience and commercial value. It is also becoming prone for many attacks. Every day, new vulnerabilities are found, new threats are detected and attacks are launched. We need countermeasures for these attacks and IDS and firewalls are not able to defend all the attacks. In this situation we need to traceback the attacker and get to the source of the attacker so that there is deterrence to the cyber criminals, thereby reducing attack rate. In this paper we survey various traceback techniques for IPv6 after introducing the same for IPv4. We also analyze the differences between packet header fields of IPv4 and v6 and list the challenges for IPv6 traceback.

General Terms
IP Traceback.

Keywords
IPv4, IPv6, DoS, DDos, traceback.

1. INTRODUCTION
The Internet expanded very rapidly from last one decade and for every communication it has become a major backbone. To provide connectivity to each and every device we need a huge amount of addresses. With IPv4 around 4.3 million of people or devices can be connected uniquely but as it become one of the important medium of communication soon all addresses will vanishes. To remove this problem IPv6 comes up with 128 bit addresses. As internet becomes pervasive and we are using much functionality it can be consider as one of most integral part of our lives. From last one decade Internet has expand very much and for every transaction starting from communication to e-commerce it has become the prominent choice of anyone. But as the Internet is expanding in various sphere of our life and become a medium for a broad range of transaction, the impact of attacks is getting more and more significant.

Every day new threatening element were coming for this broad medium among them, Denial of Service (DoS) and Distributed Denial of Service (DDoS) are the prominent one. DoS & DDoS attacks consume resources of a remote host or network so that it cannot offer its services to the legitimate users. Such attacks are among the toughest to address because they are simple to implement, hard to prevent and difficult to trace [1].

To prevent these attacks techniques like Intrusion detection system (IDS), Intrusion prevention system (IPS) and firewall are good one but preventing all kind of attacks is nearly impossible. The situation become more panic with the use of spoof IP address means an attacker can hide its entity if he wants. The stateless nature of Internet protocol add advantage to spoofing as the source host itself files source host id in IP packet and in TCP/IP there is no provision for discovery the true origin of the packet [2]. So when prevention fails a mechanism to identify the source of attack needed to at least ensure accountability for these attack and here we need the traceback techniques.

IP traceback is the technology that can traceback the source of spoofed attack packet and recognize attack graph by tracing attack paths and the packet sender/receiver. IP traceback techniques neither prevent nor stop the attack they are only used for identification of offending packets during and after the attack. IP traceback may be limited to identifying the point where the packets constituting the attacks enter the internet [3]. The traceback mechanism is shown in Fig. 1.

IP traceback methods are either reactive or proactive. Reactive traceback technique initiates the traceback in response of an attack and must complete their operation while attack is active means for reactive the attack must be in live. While proactive approach record the trace records as packets traverse through the internet and a victim used recorded data for traceback. IP traceback schemes can be categorized into link testing, messaging, logging and packet marking [4, 5].

Link testing also known as hop-by-hop tracing test network lines between routers to determine the origin of attacker's traffic. In this, testing start from the router closest to the victim and interactively test the upstream links to determine which one carries the attack traffic. This is a reactive method and requires attack to remain active until trace is completed.

Logging is maintaining database for all traffic at every router within the domain and use data-mining technique to extract information about attack traffic source.

In messaging routers send ICMP messages from participating routers to destination. Victims reconstruct the attack path from received ICMP messages.

Packet marking method inserts traceback data into IP packet header. In this router through which the data packets traverse insert partially or complete information of itself as trace data.

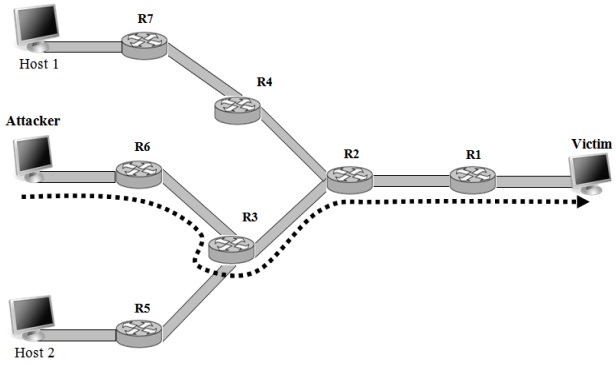

Fig 1: IP Traceback

The victim used this mark and reconstructs the source from where the packet was introduced into the network.

IPv6, like Ipv4 is an Internet layer protocol for packet-switched internetworking and provides end-to-end datagram transmission across multiple IP network. In the early 1990s, researchers argued that the current Internet was insufficient for new applications such as voice and videos. They further argued that growth of Internet would quickly exhaust the set of available addresses which is around 4.29 million. To cope up with tremendous growth of Internet, requirement of new address space is arise so that each and every thing can be connected through Internet with unique IP address. +It took several years for Internet Engineering Task Force (IETF) to formulate the new version of IP [6, 7].

The rest of paper is structured as follows: section2 describes IPv4 techniques and metrics. Section 3 compares IPv6 and IPv4. Section 4 describes the survey of IPv6. Section 5 describes the challenges and section 6 concludes the paper.

2. IPv4 TECHNIQUES AND METRICS

The survey papers [1-5] broadly classify the current IP traceback schemes as:

In link testing, network links between the routers are tested to determine the origin of the attacker's traffic. Most of the techniques start from the router closest to the victim.

Input debugging is one of the approaches, which work by using signature of the attack packets to trace its path backward from the victim to its source. It's a feature available on many routers that helps in determining the incoming link along which the attack packet must have traversed. This is repeated hop-by-hop at every upstream router in network till source or another ISP is reached.

Controlled Flooding, it is conducted by the network administrator by applying burst of traffic systematically to each link, from the victim to its upstream segment and observing how this intentionally generated flood affects the attack traffic intensity.

Logging is maintaining database for all traffic at every router within the domain and use data-mining technique to extract information about attack traffic source. SPIE is a log-based traceback system that uses efficient auditing techniques at network routers to support the traceback of individual IP packets. Transferring auditing is accomplished by computing and compactly storing packet digests rather than storing the packets themselves. Overlay network also called center track is a centralized scheme in which specialized trace route (TR) monitors all the traffic in the network. All the packets have to be routed through TR. This is accomplished by building a generic route encapsulation (GRE) tunnel from every edge router to TR.

ICMP Traceback, This method is based on an approach called iTrace. In this each router selects one packet per 20,000 packets and then generates an ICMP message. The victim receives these packets in addition to information from regular network traffic. These messages contain partial path information including from where the packet came from, when it was sent, and its authentication. Intension-driven iTrace, the idea is to add some intelligence to the marking path so that the information required for path reconstruction may be quickly gleaned by the victim. In this each router needs to modify its routing table information to accommodate the intension information. iCaddie ICMP, this method based upon the number of packets after which to generate iTrace message. Each router is equipped with a timer that indicates how long it hasn't received a traceback message. If this is greater than a certain threshold, the router randomly chooses a ball packet and prepares for it an iCaddie packet, which collects path information of all routers from this point through destination. The attack path can be easily reconstructed by the victim by simply looking at the marking inside a caddie message.

Packet marking methods are characterized by inserting traceback data into the IP packet header to be traced through various routers from the attack source to destination. The traceback data inserted into the packet header either probabilistically or deterministically, the scheme can be Probabilistic Packet Marking (PPM) and Deterministic Packet Marking (DPM). These scheme are based on the idea that router mark the packets that passes through them with their addresses or part of their addresses. To deploy this scheme the routers need to implement two functions: marking and reconstruction. PPM marks the packets with path information in a probabilistic manner and enabled the victim to traceback the path by using the marked packets. DPM, the basic idea is that only the edge routers mark the received packets this mark remains unchanged for as long as the packet traverse the network. IPv4 use 16-bit IP identification field and reserved 1-bitflag in the IP header to store marking.

A number of metrics may be used to evaluate the performance of disparate traceback schemes. Following are the metrics essential in comparing IP traceback approaches [3].

- *ISP involvement*: tracing an anonymous attack is not possible without the involvement of ISP.
- *Number of attacking packet needed for traceback*: scheme must be capable to find source of attack based on few packets once the attack has been identified.
- *Processing overhead*: There are two considerations for processing overhead, where it is incurred and when it is incurred. A scheme must incur minimal processing overhead during traceback only.
- *Bandwidth overhead*: Additional traffic that network has to carry for traceback is considered bandwidth overhead. Schemes should not assume availability of infinite bandwidth.
- *Memory requirements*: Additional memory may be required on routers for traceback, schemes must not use additional memory required on network equipment.
- *Scalability*: Scalability measures how easily the scheme can expand. An ideal scheme should be scalable, and configuration of devices involved should be totally independent of each other.
- *Number of function needed to implement*: Reflects how many functions a vendor of equipment need to implement for a given scheme.
- *Ability to handle major DDoS attack*: An ideal scheme must be able to traceback all attacks.
- *Ability to trace transformed packets*: Packet transformation is a modification of packet during the forwarding process. An ideal scheme would correctly traceback attack consisting of packets that undergo any number of transformations of any type.

3. COMPARING IPV4 AND IPV6

With the rapid development and utilization of addresses in the Internet, IPv4 will be gradually substituted by IPv6. The IPv4 has been in use for over many decades and many of the related devices have been connected to the internet. But with high demand of Internet usage and due to paucity of address space and allocation mechanism some part of the world are beginning to run out of addresses.

Table 1.Comparison between IPv4 and IPv6

IPv4	IPv6
Size of header is of 20 octets.	Size of header is of 80 octets.
IHL or HLEN- Header length, gives datagram header length measures in 32-bit.	Not included.
TOS (Type of Service) - specifies how datagram is handled. It is a hint to forwarding algorithm which helps them to choose among various paths to a destination.	Replaced by Traffic Class and Flow Label. Tags the packet with a value representing a class of traffic that can be used in differentiated services.
Total length- specifies entire length of IP datagram i.e. Header + Data.	Replaced by Payload length- indicates the length of data or payload only.
ID, Flag, Fragment offset fields are used for fragmentation and reassembling of data packet at source and destination.	Not included since fragmentation is done at source node, no intermediate router is doing fragmentation so not included.
Time to Live- indicate maximum time datagram allowed to remain in network.	Replaced by Hop Limit which specifies maximum number of routers a packet traverse before it is considered invalid.
Protocol- specifies format of data area.	Replaced by Next Header.
Header Checksum- an error detecting code applied to header only not on the data. Means packet is examined at every router hop.	Removed, link layer technologies and upper layer protocol handle checksum and error control.
Source address and Destination address of 32 bits.	Source and Destination address of 128 bits.

To solve this new version of Internet protocol has been defined so as to solve the address and other problems in the currently used internet protocol [6, 8-10].

Many of the current IP traceback techniques are designed according to IPv4 so they cannot be directly used in IPv6 network. IPv4 and IPv6 networks differ greatly from each other, which urge the need of techniques specifically for IPv6 Prepare [11]. The need of new traceback techniques arise due to the difference of packet format of the both IP version. IPv4 have packet format of size 20 octets while IPv6 have format of size 40 octets. There is also difference in the fields of the packets. The difference between IPv4 and IPv6 are given in Table 1.

4. SURVEY OF IPV6 TRACEBACK TECHNIQUES

Timothy et al. [12] described the source path isolation engine (SPIE) for IPv6. SPIE is a log based traceback system that uses the efficient auditing techniques at network routers. If a packet is determined to be offensive by some IDS a trace request is sent to the SPIE system which in turn queries routers for packet digests of the relevant time periods. SPIE consists of Data Generation Agent (DGA), SPIE Collection and Reduction agents (SCARs) and SPIE Traceback Manager (STM). DGA produces packet digest of each packet and store the digest in time-stamped digest table. SCRAs are responsible for a particular region of the network, serving as data concentration points for several routers. When a trace request is requested each SCRAs produces an attack graph for its particular region. STM control the whole SPIE system. STM is the interface to the intrusion detection system (IDS) requesting a packet trace. STM checks the authenticity of request, dispatch the request to the particular, gather the resulting attack graphs and assembles them into a complete attack graph. STM replies to the IDS with final attack graph on completion of traceback process.

Lee et al. [13] proposed an authenticated IP traceback mechanism against DDoS attack using a hashed Message Authentication Code (MAC) function on IPv6 header. In this a mechanism is required to authenticate the packet marking. This technique uses one cryptographic MAC computation per marking. In this two parties share a secret key. When one party wants to send the message to other it appends the message with MAC using shared secret key. The other party on receiving the message checks the validity of the MAC. Similarly each router shares a unique secret key with the victim. Routers apply MAC function on its IP address and some packet-specific information with shared secret key to generate the encoding of routers IP address. When an abnormal traffic occur the router starts marking the packet with generated IP address.

Obaid et al. [14] implemented the deterministic packet marking scheme on IPv6 networks. To store the mark a new option in Hop-By-Hop option header is considered. The interface of router closest to the source will mark the new option header. As every packet is mark with complete IP address a single packet would be enough to get the complete traceback. Discuss Managed Security Services (MSS). Different component of MSS might work together to accomplish a traceback. Once the victim aware of an attack with the help of IDS, he could initiate the traceback to get the source of an attack. It also suggested there should be a small data structure to maintain a list of last incoming packets so even after an attack the victim can guess the origin of an attack by viewing the density of a certain interface addresses in that data structure.

Obaid et al. [15] proposed an efficient Lightweight IPv6 traceback algorithm for tracing the actual source of an IP datagram. It uses the probabilistic packet marking scheme for marking the packets. It uses Hop-by-Hop option header to store a mark because it is processed by every router and it provide the larger space to store the mark. On victim side it proposed a data structure called Reverse Lookup Table (RLT) to store the marked packets. To traceback the source victim will sort the RLT by distance field, observe the discontinuity in distance field and apply error correction algorithm to find the missing nodes. Finally victim will resolve the last hop field to complete the RLT.

Obaid et al. [16] extend the previous work by using Policy Based IP Traceback (PBIT) mechanism. Motivation is that, thousand of packets traverse through the router in a second and marking of every packet may affect routing performance. One way to accomplish this is to deploy IDS on victim side and once IDS detect an attack it sends message to intermediate routers to initiate marking. But we do not have path information so we cannot send message to desired router to start marking. Another way is to multicast the message to all backbone routers but it will increase the traffic and for sending these messages we use ICMP which is mainly filtered by many internet service providers. To mitigate the above problems they utilized the Policy Based Management system. Policy-based is an administrative approach that is used to simplify the management of a given endeavor by establishing policies to deal with situations that are likely to occur. Two basic building blocks are Policy Decision Point (PDP) and Policy

Enforcement P- (PEP). PDP is a resource manager handling events and making decision based on those events. PEP exists in network nodes and enforce the policies based on the "if condition then action" rules are set by PDP.

Obaid et al. [11] proposed a traceback architecture using Common Open–Policy Service (COPS) and a novel packet marking scheme. COPS are a messaging protocol for policy based management system. IDS of the victim on detecting attack traffic send request to the PDP to enforce policy, the local PDP send request to all PES to check any packet coming from specified sources and going towards the victims. PDPs next to that PES which produces positive responses are of packet started. Use global unicast address of marked by edge routers. Mark is store in Hop-by-Hop option header.

Shi et al. [17] proposed a deterministic link signature based algorithm for IPv6 traceback. They assign a unique signature to each link and mark the attack packet according to the signature of out link, that to which the packet is sent. The marking involve XORing a packet's signature area with it's out or in links signature. Here packet marking faces two problems: how to assign link signature and how to choose the signature area from IPv6 header. For link signature assign 16-bit long link signatures randomly for each link. For recording the path information they used flow label area of IPv6. When victim receives the marked attack packets to reconstruct the path we only find the path with same signature and hope values as received packet. This provides the IP addresses of all routers in the path of the attack packet.

Kim et al. [18] proposed traceback mechanism for next generation network (NGN) based on IPv6 protocol. NGN is an enhanced intelligent network that can cover voice, data and multimedia with only one integral network and effectively support various value added services. As online based networks are expected to completely transform to NGN therefore, traceback technologies for NGN need to researched and developed. The routers probabilistically mark the path information in packets. Routers store the audit logs of the forwarded packets and victim consults the upstream routers to reconstruct the attack paths. All routers in network that received packet check and process Hop-by-Hop options header.

Ting Ma [19] proposed a link signature based DDoS attacker tracing algorithm. Each link in the network assigned a unique signature. Each router use this signature for marking the packet when forwarding the packets. Marking involve XORing packet signature area with out-link. Use flow-label field for storing mark in IPv6. Since the attack packets contain the marks in the signature area the victim can reconstruct the attack path statistically. Static reconstruction needs entire network topology and link signature of each network link.

You-ye et al. [9] suggested a modified deterministic packet marking for DDoS attack traceback in IPv6 network. The two main advantages of modified deterministic packet is: first, it only needs a small amount of marked packet to reconstruct the DDoS attacking paths; second, it can trace a huge number of simultaneous DDoS attackers. Among several extension header defined in IPv6 Destination Options Header (DOH) is used to carry the information that is needed to be disposed by destination node. The new option data field will increase the length of the IPv6 packet which exceeds the Maximum Transfer Unit (MTU) of then path. To solve this modified path MTU is necessary. To avoid that router did not mark the packet when there is no attack two threshold introduces, L_min and L_max. If load of router is below L_min, it will not mark any packet; if the load is between L_min and L_max, each packet is marked; if the load is above L_max packet will not be marked again. For reconstruction victim hosts decide whether an IPv6 packet has been marked by searching the DOH. If it exists, the victim host will extract the ingress address out of it and put the address into the ingress address table. The sources of DDoS attack will be deduced from this table.

Tripathy et al. [10] focuses on a secure packet marking mechanism that would help in obtaining the marking which are authentic which a desired criteria for an improved marking scheme. Uses modified DPM for marking the packets. In this modified scheme the ingress address is divided into *k* segments. The mark consists of a-bit address bit field, d-bit digest field, and s-bit segment number field. Some padding may be required so that the address is split into segments with equal length. Reconstruction procedure consists of two separate processes: Mark Recording and Ingress Address Recovery. Reason for separating the task is the fact that attack packet may arrive to destination faster than they can be analyzed. Mark recording process will set appropriate bits in RecTbl indicating which marks arrived to the destination. Address recovery will check those bits and will compose address segment permutation and determine which ones are valid ingress addresses.

5. CHALLENGES

The two major challenges of IPv6 traceback techniques if to handle the Mark Spoofing, Dual Stack Implementations and Gradual Deployment. They are described as follows:

A. Ease of Evasion:

An attacker can inject a packet marked with erroneous information into a stream of packets. If the attackers are aware of the marking technique being used, they can place misleading information in the fields being used to store the encoded marks. This is called mark spoofing. The router can totally avoid mark spoofing by ensuring that all packets arriving from an interface connected to the internal network will be definitely marked. Any information placed by the attacker to mislead will definitely be overwritten.

B. Dual Stack Implementations

The network world is moving towards IPv6 but it will still take some time before all the hosts on the internet will be IPv6 enabled. There is a need for both the IPv4 and v6 traffic to be able to talk to each other. This will need dual stack implementations on the hosts to facilitate the communication of two protocols at the same layer. Handling traceback in such situations will also be challenging.

C. Gradual Deployment

Gradual deployment of any traceback technique is possible with only partial routers along the path enforcing the marking mechanism. Any marking mechanism will fail if the network does not enforce the marking mechanism. The challenge is to design a marking mechanism which will ensure traceback even if some networks do not fully support the mechanism.

6. CONCLUSION AND FUTURE WORK

IP traceback techniques are considered to trace back the source of attacker and provide some evidences so that a legal action can be initiated against the attacker who intensely attack the network or the victim. In this paper we have survey different traceback techniques for IPv4 and IPv6. We have also given the comparison between the packet format of IPv4 and IPv6. We would propose a suitable technique for traceback of IPv6 packets overcoming the above limitations. We will simulate the technique using network simulator ns-2 or Omnet++.

7. REFERENCES

[1] H. Aljifri, "IP Traceback: A New Denial-of -Service Deterrent," *IEEE Security & Privacy*, pp. 24-31, 2003.

[2] S. Savage, D. Wetherall, A. Karlin, and T. Anderson, "Network Support for IP traceback," *IEEE/ACM Trans. Networking*, vol. 9, pp. 226-237, 2001.

[3] A. Belenky, "On IP Traceback," *IEEE Communication Magazine*, vol. 41, no. 7, pp. 142-153, 2003.

[4] L. Santhanam, A. Kumar, and D. P. Agrawal, "Taxonomy of IP Traceback," *Journal of Information Assurance and security*, vol. 2006, no. 1, pp. 79-94, 2006.

[5] Z. Gao and N. Ansari, "Tracing Cyber Attacks from the Practical Perspective," *IEEE Communication Magazine*, pp. 123-131, 2005.

[6] D. E. Comer, *Internetworking With TCP/IP*, 5 ed. vol. I: PHI Learning Private Limited.

[7] "What is IPv6," Available: http://technet.microsoft.com/en-us/library/cc738582(v=ws.10).aspx, [May 15, 2013]

[8] Jivika Govil, Jivesh Govil, Navkeerat Kaur, and H. Kaur, "An Examination of IPv4 and IPv6 Networks: Constraints and Various Transition Mechanisms.," pp. 178-185, 2008.

[9] S. You-ye, Z. Cui, M. Shao-qing, and L. Kai-ning, "Modified Deterministic Packet Marking for DDoS Attack Traceback in IPv6 Network," in *Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on*, 2011, pp. 245-248.

[10] A.Tripathy, "A Secure Packet Marking Scheme for IP Traceback in IPv6," *(ICACCI '12)*, 2012.

[11] S. Obaid, "A novel IPv6 traceback architecture using COPS protocol," *Ann. Telecommun*, vol. 63, pp. 207-221, 2008.

[12] Strayer WT, Jones CE, Tchakountio F, and H. RR, "SPIE-IPv6: single IPv6 packet traceback," *IEEE International Conference on Local Computer Networks*, pp. 118-125, 2004.

[13] Hyung-Woo Lee and S.-H. Yun, "Authenticated IPv6 Packet Traceback Against Reflector Based Packet Flooding Attack," pp. 1118-1124, 2005.

[14] S. Obaid, "On IPv6 Traceback," *(ICACT '06)*, pp. 2139-2143, 2006.

[15] S. Obaid, "A Lightweight IP Traceback Mechanism on IPv6," *(IFIP '06)*, pp. 671-680, 2006.

[16] S. Obaid, "IPv6 Traceback using Policy Based Management System," *(KNOM '06)*, vol. 9, no. 2, 2006.

[17] Y. Shi, "Deterministic link signature based IP traceback under IPv6," *(ICACT '08)*, 2008.

[18] RH Kim, JH Jang, and H. Youm, "An Efficient IP Traceback Mechanism for the NGN based on IPv6 Protocol," *(IITA '09)*, 2009.

[19] T. Ma, "A Link signature based DDoS attacker tracing algorithm under IPv6," *(IJSA '09)*, vol. 3, no. 2, April 2009 2009.