

Topology Based Packet Marking for IP Traceback

Harendra A. Alwis, Robin C. Doss, Praveen S. Hewage, Morshed U. Chowdhury
School of Engineering and Information Technology
Deakin University
Melbourne, Australia
(haa, rchell, pshew, muc)@deakin.edu.au

Abstract— IP source address spoofing exploits a fundamental weakness in the Internet Protocol. It is exploited in many types of network-based attacks such as session hijacking and Denial of Service (DoS). Ingress and egress filtering is aimed at preventing IP spoofing. Techniques such as History based filtering are being used during DoS attacks to filter out attack packets. Packet marking techniques are being used to trace IP packets to a point that is close as possible to their actual source. Present IP spoofing countermeasures are hindered by compatibility issues between IPv4 and IPv6, implementation issues and their effectiveness under different types of attacks. We propose a topology based packet marking method that builds on the flexibility of packet marking as an IP trace back method while overcoming most of the shortcomings of present packet marking techniques.

Keywords—IP Spoofing; Topology Based Packet Marking (TBPM); Denial of Service Attack (DoS); Ingress Filtering; Egress Filtering; Packet Marking; IP Trace-back; ICMP

I. INTRODUCTION

The Internet has become the backbone of telecommunication networks and a vital tool for personal communication as well as businesses, government and the military. Higher bandwidths have spurred a growth in the use of multimedia. The need for security on the Internet has always been important. However malicious activity on the Internet is increasing [1]. Law and security enforcement on the internet has made slow progress for many reasons [2]. Cyber-stalking, propagation of malware, DoS attacks and gaining unauthorized access to computer systems are threats to Internet security. Most of these attacks are carried out in such a way that the identity of the attacker is not revealed or under the guise of a false identity. In the Internet where a person's identity is represented by a set of digits in a computer or a network, spoofing one's identity and location is simple. It is done primarily by 'spoofing' the Sender Address field in an Internet Protocol (IP) header of a datagram [3]. This field is meant to identify the creator of the IP datagram. The Internet Protocol does not offer any protection to the 'Sender address' field by design [4]. It is a clear example that the Internet Protocol has not been designed with the core security features needed to withstand the types and volume of attacks that threaten it today [5]. Therefore it is a relatively easy and convenient way for attackers to mask their true identity

The widespread use of IP spoofing in network attacks is a good example of the fact that attackers can easily mask their identity while they engage in malicious activities such as Distributed DoS attacks.

Some security vulnerabilities in computer systems are a result of inherent weaknesses in the design and implementation of end user systems such as hardware and software loopholes [6]. In this work, we focus on security issues in the network and specifically, the ease with which the IP Protocol can be abused through source address spoofing.

We examine the weaknesses of present IP trace back methods. It is desirable for IP trace-back mechanisms to be compatible with both IPv4 and IPv6. They also have to be effective when the network is under attack. Packet marking is a versatile tool in meeting these objectives. Mark spoofing and the need for multiple marked packets for the source address to be reconstructed are some of the key weaknesses we aim to overcome. Our approach focuses on a packet marking strategy which embeds information about the route that the packet traverses through a network. We build on principals of existing packet marking technology and propose a new method: Topology Based Packet marking (TBPM) that reasonably addresses these shortcomings. We argue that TBPM would be more effective than source address marking; particularly during DoS attacks.

In section II we present previous work on IP spoofing countermeasures. Section III introduces TBPM which is our main contribution and discusses its implementation methods as well as performance. Section IV is a brief outline of future work and we conclude in section V.

II. PREVIOUS WORK ON IP SPOOFING COUNTERMEASURES

Many different techniques have been proposed as countermeasures against source address spoofing in IP datagrams [7]. There are two main types of countermeasures against IP source address spoofing in use today. One strategy is to focus on eliminating spoofed packets at network gateways. Another is to trace spoofed datagrams from a recipient to their actual source. An alternative approach is to enforce authentication of users at connection time; for example through encryption.

The IP spoofing problem has been exploited in many attacks ranging from session hijacking, gaining unauthorised access based on the weakness of source IP authentication [5] and DoS attacks on networks and network devices. A debate about the need for security at the network layer was opened up with the introduction of the IPSec protocols drafted by the IETF working group [8]. Those who argue for security at the network layer claim that the advantages of security at this level will outweigh the disadvantages. They claim that the network

layer should provide an acceptable level of security to the communication regardless of the security provided by higher levels and the end-to-end applications. Their claims are backed by the fact that an increased number of applications that use real-time and multicast data employ the connectionless User Datagram Protocol (UDP) at the transport layer which is inherently insecure and therefore necessitates the need for security at the network layer.

Opponents for securing network services at the network layer argue that incorporating security in the IP layer will be a complex task. They argue that the use of encryption based authentication and data security will reduce the throughput of the network while also being a hindrance in high speed networks where the processing speeds will not be able to keep up with the inflow of data through a broadband network [8].

A. Filtering

Packet filtering aims to hinder the movement of spoofed IP packets through network gateways. Packet filtering at network gateways falls into two main categories. One is egress filtering which disallows packets with external source IP addresses to leave the network [9]. This eliminates most spoofed IP packets from leaving a network. The other is ingress filtering which blocks packets containing internal source IP addresses from entering a network. This blocks most spoofed packets from entering a network [10]. If implemented globally, these two types of filtering can completely eliminate the threat of IP spoofing.

There is another type of packet filtering [11] which is effective particularly in defusing DoS attacks. During a DoS attack, it uses access logs of a web service to filter out new IP addresses that haven't been logged at the site before. It relies on statistical data which indicates that approximately 86% of DoS traffic comes from IP addresses that have never been logged on a site before.

B. IP trace-back

Tracing back spoofed data packets to their original source is an alternative measure against IP spoofing. IP trace back techniques play a key role in defusing DoS attacks. They enable the trace back and blocking of attack packets as close to their sources as practically possible. IP trace back methods are widely used and researched [7, 12-15]. These techniques aim to trace-back IP packets to their actual source in spite of the spoofed source address in the packet header. Some trace-back techniques use the Internet Control Message Protocol (ICMP) [16].

1) Packet marking

Packet marking is an alternative trace-back technique which is more flexible and versatile and therefore more widely used. Packet marking techniques in turn have many variations. Deterministic Packet Marking (DPM) and Probabilistic Packet Marking (PPM) are two primary methods with variations of their own.

In conventional packet marking techniques, edge routers mark the source address of a packet in the redundant 16bit identification field in the IPv4 header. An IPv4 address is 32

bits long and thus in conventional packet marking; it takes more than two packets to store an IP address length of 32 bits. A fraction of the source address will be marked on the redundant 'ID field' in the IPv4 header of packets passing through an edge router. The ID field of one packet may contain either the first or second half of the source address. A flag in the 'flags' field in the IP header will indicate whether the mark in the ID field is the first or second half of the actual source IP address.

The destination will use the information in multiple packets to reconstruct the source address of a packet stream. A key drawback here is that at least two packets are needed to construct the actual source address. However, when the attack is carried out by multiple sources, more space in the IP header becomes necessary to indicate the identity of the particular source. This can make packet marking almost in-effective during Distributed DoS attacks.

In DPM, each packet that passes through an edge router is 'marked' [13]. In PPM, the marking is done randomly [14]. Both these techniques have been implemented with different enhancements and variations [7].

While the redundant 16-bit 'Packet ID' field in the IPv4 header is utilised for packet marking there aren't any significantly large redundant fields in IPv6. As a result present packet marking techniques are not compatible with IPv6.

In traditional packet marking techniques, the edge routers are responsible for marking the packets with their source address. While packets marked by an edge router of a host network will indicate the address of its immediate neighbour outside the network, that node is unlikely to be the original source of the IP packet. The only way to know the actual source would be if the marking is done at the edge router of the source network. However the source network may not be trustworthy and the marking itself could be spoofed. Furthermore, a mark made by an external router could easily be overwritten by an edge router in another network. Such 'spoofed' marks may invalidate the trace back effort. If the network uses DPM, all previous marks will be overwritten. Even though it is an effective countermeasure against 'mark spoofing', it also nullifies any valid marking done by neighbouring networks.

It is clear therefore, that packet marking techniques that employ the redundant ID field in IPv4 packets are not as effective or efficient as egress and ingress filtering. Their advantage however lies in the fact that it can still be useful in the battle against DoS attacks as it provides information regarding the location where the packets enter a network. This enables the host to take action to stop the flow of packets towards the victim through a particular edge router.

C. Authentication

Authenticating users at connection time is an effective preventive measure against IP source address spoofing. However the IPv4 specification does not have any in-built provision for authentication. Even though authentication is a facility provided for in IPv6, it is not a mandatory implementation. The User Datagram Protocol (UDP) is a

connectionless transport layer protocol that is used for real-time and multicast applications in the TCP/IP protocol suite [17]. Most applications that use UDP cannot enforce end-to-end connectivity and is therefore unable to enforce user authentication. For this reason, perpetrators of DoS attacks increasingly use UDP packets to carry out their attacks [11]. It is clear therefore that authentication at connection time is not a practical solution that can be applied under all circumstances.

Authentication techniques cannot effectively identify legitimate traffic during a network-based DoS attack because a DoS attack can be initiated by an authenticated user. The additional bandwidth and processing overhead involved with cryptographic techniques can compound a DoS attack. Authentication cannot be used in one way communication such as email. Therefore it is clear that host authentication does not provide a comprehensive countermeasure against IP spoofing.

D. Other IP spoofing countermeasures

If the attack is unsophisticated, there might be a specific signature to the traffic [18]. A careful examination of captured packets may reveal a trait on which some rules can be based in router access control lists or firewalls in order to filter out attack traffic. Additionally, a large amount of traffic may originate from a specific origin point which could be temporarily blocked, allowing a portion of legitimate traffic through. This could block legitimate packets as well, but it is an unavoidable sacrifice.

III. TOPOLOGY BASED PACKET MARKING (TBPM)

Embedded topological information in a data packet has many advantages over traditional packet marking methods. Traditional packet marking methods only mark the identity of the edge router through which a packet enters a network. In a DoS scenario for example, as a result of the overwhelming inflow of traffic, the edge router may be unreachable to the node under attack. Having information about the route that the packet had traversed through the network will enable the node to defuse the attack as close to its source as practically possible even when the edge router is unreachable. That would also enable internal network functions to be restored when edge-routers are under a DoS attack.

It may not be desirable however, to reveal the topology of a network to outside sources. Embedding topological information in a readable form, within data packets that leave the network may expose sensitive internal topological information to outsiders. Embedding topological information in each data packet would also increase overheads in the network in terms of bandwidth as well as processing. Therefore the challenge in topology based packet marking is two-fold. The first objective is to conceal the topology of one network from another. The second is to minimise overheads. We propose steps that will help achieve both these objectives including the use of unique node IDs in networks.

A. TBPM Method

We propose to embed information in a data packet that would trace the route it takes from an edge router to its destination node. This would be done by each node appending

a unique signature to the end of the datagram. The destination node will then be able to decode the route that the packet had taken through the network by reading the appended signatures sequentially.

Appending node signatures to packets is not expected to require any modifications to the structure of the IP header. However, it is desirable to keep network and processing overheads to a minimum required level. A single network is almost never large enough to utilise the entire IP address space. A network of 50 nodes for example, needs only 6-bits to assign a unique ID to each node (and have 14 IDs left over). Appending multiple lengthy IP addresses to datagrams is an avoidable overhead. Therefore the IP address of a node can be replaced by an ID that is significantly smaller than its IP address. Listed below are the fundamental principals of TBPM.

1. Assign unique network IDs for each node/router that packets pass through. The length of this ID (node signature) will be pre-defined for a particular network depending on the network size. The length of a hop-count field will also be determined depending on the topology of the network. The optimal lengths for these fields could be derived according to Table (1);

TABLE I. VARIABLE DEFINITIONS AND OPTIMISATION

| | | |
|-----------------------------|------------|---------------------------|
| Network size | X nodes | |
| Node ID Field length | 2^n bits | $(n \geq \log_2 X > n-1)$ |
| Maximum hops in the network | Y | |
| Hop Count Field length | 2^m bits | $(m \geq \log_2 Y > m-1)$ |

2. It is expected that the bit-length of a node ID would be significantly smaller than the length of an IP address. Therefore appending multiple node IDs is not expected to significantly inflate the size of the datagram. Figure (1) depicts a typical datagram marked by an edge router.

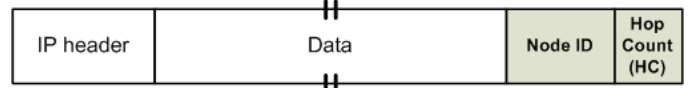


Figure 1. Datagram marked by edge router

3. The edge router will add its unique ID and a hop-count field to the end of the IP datagram. Each router that the packet subsequently passes through will update the Total Length (TL) field in the IP header, append its ID to the IP datagram and increment the hop-count. The destination host will use those two values to decode the network path of the datagram and separate that information from the actual data. Figure (2) illustrates how a datagram changes as it passes through subsequent routers.

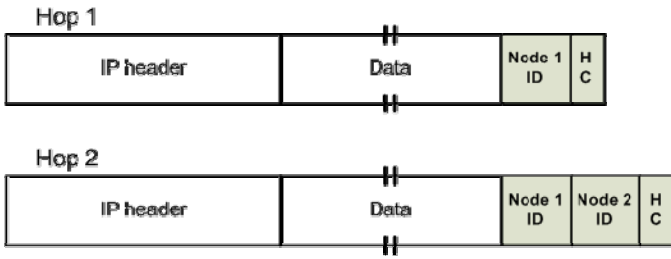


Figure 2. Adding topological data to a packet at each 'hop'

4. If the appending of node signatures increases the size of the datagram beyond either its Maximum Transport Unit, or 64KB which is the maximum possible length of an IP datagram, then in an IPv4 network, the packet will be fragmented as illustrated in Figure (3). IPv6 does not allow routers to fragment datagrams and this can potentially lead to complications when implementing TBPM in IPv6 networks. The simplest way of overcoming this is to set the Maximum Transmission Unit (MTU) [19] at the edge router to a value that accommodates sufficient space for TBPM.

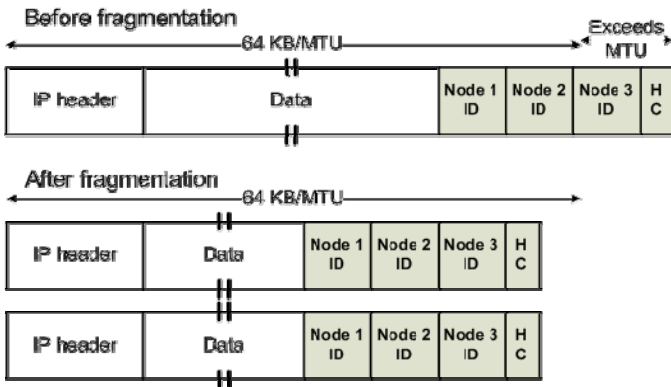


Figure 3. Fragmentation of a datagram in an IPv4 network

Given that each node in the network is aware of the Node ID length and the length of the Hop Count Field, TBPM is therefore compatible with both IPv4 and IPv6 networks.

B. TBPM Performance factors

There are many variable factors that govern how the processing and bandwidth overheads incurred by TBPM may affect performance in a particular network.

TABLE II. VARIABLE DEFINITIONS

| | |
|---|---|
| Number of hops | H |
| Network ID length | L |
| Hop count length | C |
| Number of instructions to increment Hop Count | H |
| Number of instructions to append new ID | A |

Number of instructions to update Total Length (TL) field in IP header

1) Network overheads

Number of additional bits transmitted per datagram;

At node 1:

$$L+C \quad (1)$$

At node 2:

$$L+L+C \quad (2)$$

At Node H:

$$LH+C \quad (3)$$

Total number of additional bits transmitted in the network from edge to destination – per datagram (N);

$$N = \frac{1}{2} HL (H+1) + HC \quad (4)$$

$$N = H (1/2 L (H+1) + C) \quad (5)$$

Given that the average speed of a link in the network is 'S' bits per second (bps) and assuming that data traffic is uniformly distributed in the network, the average additional propagation delay per packet is:

$$N/S \text{ (seconds)} \quad (6)$$

$$H (1/2 L (H+1) + C) / S \text{ (seconds)} \quad (7)$$

2) Processing overheads

Total processing overhead per datagram:

$$H (h + a + t) \text{ instructions} \quad (8)$$

Given that the average processing speed of a node in the network is P (instructions per second) and that data traffic is uniformly distributed in the network, the average processing delay caused by a datagram is:

$$H (h + a + t) / P \text{ (seconds)} \quad (9)$$

C. TBPM performance enhancements

One of the clear advantages of TBPM over traditional packet marking methods is that only one packet is necessary to reconstruct the entire path it has taken through the network including the source of its entry. Therefore it is sufficient that packets are marked randomly at a given frequency. This is expected to dramatically reduce the overheads associated with marking each and every packet that passes through a router and make TBPM comparably more efficient than present PPM methods and more economical than DPM methods without compromising its effectiveness as a IP source identification technique.

Depending on the degree of threat and its severity, the frequency with which packets are marked could be changed at

the edge routers. A new flag has to be set in the IP header for the edge router to indicate whether a given packet has been marked or not.

D. Evaluation of TBPM

It is clear that appending topological information incurs network as well as processing overheads at each node. Traditional packet marking methods only incur processing overheads which are relatively inexpensive compared with network overheads, in terms of cost and delays.

In both IPv4 and IPv6, the total length of the datagram is recorded in the 16 bit Total Length (TL) field. As a result, the total length of an IP datagram has an upper limit of 64 kB. If node signatures are added beyond the 64kB limit, the IP datagram may have to be fragmented – adding to the processing and network overheads.

The fact that it takes only one packet to reconstruct an entire path makes it viable for TBPM to be implemented as a probabilistic marking method even under Distributed DoS where traditional probabilistic packet marking methods fail to be effective [7].

The key strength of TBPM lies in its effectiveness as a tamper-proof method of tracing an IP datagram to its source along the path it has traversed through the network. TBPM also allows not only the source to be identified but also the intermediate nodes between the source and destination. This significantly increases the chances of the destination node to defend against DoS traffic – especially in a situation where the edge router is unreachable.

IV. FUTURE WORK

Our future work will focus on the design and modification of software and protocols that facilitate TBPM. An automated protocol for allocating Node IDs dynamically will allow more flexibility to network administrators. We also plan to measure performance gains that can be achieved by implementing TBPM probabilistically and comparing it with present packet marking methods for efficiency and effectiveness.

V. CONCLUSION

There have been many approaches taken to hinder IP spoofing or alternatively to overcome the challenge of discovering the actual source of spoofed datagrams. We have reviewed the main IP spoofing countermeasures and pointed out their strengths and weaknesses. Our contribution in this paper has been a new approach in anti-IP spoofing techniques that we call Topology Based Packet Marking (TBPM).

TBPM builds on the strengths of the packet marking principal; however it focuses not merely on the source, but also the path traversed by a datagram. We have pointed out how a route discovery method can be more effective, especially during DoS attacks where edge routers that mark packets may themselves be unavailable as a result of the attack. Embedded topological information may enable DoS attacks to be prevented even by intermediate routers. TBPM also enables the source to be identified using a single marked packet; unlike

previous techniques that require multiple packets. We have shown how TBPM techniques are compatible with both IPv4 and IPv6; unlike present packet marking techniques that cannot be effectively implemented in IPv6 networks.

TBPM will give new direction to solving the IP trace back problem and provide more options and flexibility; especially in defending networks from DoS attacks.

REFERENCES

- [1] John Douglas, H., An analysis of security incidents on the Internet 1989-1995. 1998, Carnegie Mellon University.
- [2] Edwards, L. and C. Waelde, Regulating cyberspace : Is there a role for law ? Computers and law (Comput. law), 1997. 8(5): p. 19-23.
- [3] Harris, B. and R. Hunt, TCP/IP security threats and attack methods. Computer Communications, 1999. 22(10): p. 885-897.
- [4] Mosher, D.A., A Study of An Internet Protocol Implementation. 1985: EECS Department, University of California, Berkeley.
- [5] Bellovin, S.M., Security problems in the TCP/IP protocol suite. SIGCOMM Comput. Commun. Rev., 1989. 19(2): p. 32-48.
- [6] Steven, J.T. and L. Karl, A requires/provides model for computer attacks. Proceedings of the 2000 workshop on New security paradigms. 2000, Ballycotton, County Cork, Ireland: ACM Press. 31-38.
- [7] Vadim, K., S. Helena, and S. Andrei, An Evaluation of Different IP Traceback Approaches. Proceedings of the 4th International Conference on Information and Communications Security. 2002: Springer-Verlag. 37-48.
- [8] Oppliger, R., Security at the Internet layer. Computer, 1998. 31(9): p. 43-47.
- [9] Xiang, Y. and W. Zhou. IP Spoofing Attack and Its Countermeasures. in From Information Warfare to Information Operations: Proceedings of the 5th Australian Information Warfare and Security Conference. 2004. Edith Cowan University, Australia.
- [10] Ferguson, P. and D. Senie, Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing. 1998: RFC Editor.
- [11] Peng, T., C. Leckie, and K. Ramamohanarao. Protection from distributed denial of service attacks using history-based IP filtering. in Communications, 2003. ICC '03. IEEE International Conference on. 2003.
- [12] Al-Duwairi, B. and T.E. Daniels. Topology based packet marking. in Computer Communications and Networks, 2004. ICCCN 2004. Proceedings. 13th International Conference on. 2004.
- [13] Belenky, A. and N. Ansari, IP traceback with deterministic packet marking. Communications Letters, IEEE, 2003. 7(4): p. 162- 164.
- [14] Park, K. and H. Lee. On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack. in INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE. 2001. Anchorage, AK, USA.
- [15] Xiang, Y. and W. Zhou. Trace IP packets by flexible deterministic packet marking (FDPM) in IP Operations and Management, 2004. Proceedings IEEE Workshop on. 2004. Beijing.
- [16] Thing, V., et al. Enhanced ICMP Traceback with Cumulative Path. in Vehicular Technology Conference. 2005.
- [17] Marco de, V., O.d.V. Gabriela, and I. Germinal, Internet security attacks at the basic levels. SIGOPS Oper. Syst. Rev., 1998. 32(2): p. 4-15.
- [18] Baba, T. and S. Matsuda, Tracing network attacks to their sources. Internet Computing, IEEE, 2002. 6(2): p. 20-26.
- [19] White, P.P., RSVP and integrated services in the Internet: a tutorial. Communications Magazine, IEEE, 1997. 35(5): p. 100-106.