Detection of Spoofing Attacks Using Intrusive Filters For DDoS

V.Shyamaladevi
Asst.Prof.Dept of IT
KSRCT

Dr.R.S.D. WahidaBanu
Prof/Head. Dept of ECE
GCE, SALEM

Abstract
Internet hosts are threatened by large-scale Distributed Denial-of-Service (DDoS) attacks. The Path Identification DDoS defense scheme has recently been proposed as a deterministic packet marking scheme that allows a DDoS victim to filter out attack packets on a per packet basis with high accuracy after only a few attack packets are received. This paper proposes the Stack Path identification marking, a new packet marking scheme based on path identification, and new filtering mechanisms. The Stack Path Identification marking scheme consists of two new marking methods that substantially improve Path identifier's incremental deployment performance i.e., Stack-based marking and Write-ahead marking. The proposed scheme almost completely eliminates the effect of a few legacy routers on a path, and performs better than the original Path identification scheme in a sparse deployment of path identifier enabled routers. For the filtering mechanism, derive an optimal threshold strategy for filtering with the Path identification marking. The system develops the path identification IP filter, which can be used to detect IP spoofing attacks with just a single attack packet. Finally, evaluate the Stack path identification's compatibility with IP Fragmentation, applicability in an IPv6 environment, and several other important issues relating to potential deployment of Stack path identification.

Keywords:
DDoS, IP spoofing, ISP security, Network, Stack-based marking, Write-ahead marking.

1. Introduction

Internet security is of critical importance to our society, as the government and economy increasingly rely on the Internet to conduct their business, and people use the Internet as a convenient vehicle for simplifying a wide range of tasks, from banking to shopping. Unfortunately, the current Internet infrastructure is vulnerable to a Distributed Denial of Service (DDoS) attack. Because DDoS attacks typically rely on compromising a large number of hosts to generate traffic to a single destination, the severity of DDoS attacks will likely increase as greater numbers of poorly secured hosts are connected to high bandwidth Internet connections.

An attacker can intentionally modify, or spoof, the source address of the packets it sends from a compromised host. One of the DDoS attacks that rely on IP address spoofing is TCP SYN Flooding, in which an attacker sends TCP SYN packets as if to initiate a TCP connection with its victim. These SYN packets contain spoofed source IP addresses, which cause the victim to waste resources that are allocated to half-open TCP connections which will never be completed by the attacker.

Another one is Reflector Attack in which the attacker attempts to overwhelm the victim with traffic, by using intermediate servers to amplify the attacker's bandwidth and/or hide the attacker's origin. The attacker simply sends requests to the intermediate server with a spoofed source IP address matching the victim's IP address. The intermediate server only sees that a number of requests are supposedly coming from the victim, and so sends its responses to the victim. When properly coordinated, a group of attackers can cause a flood of packets to hit the victim, without sending any packets directly to the victim itself. To amplify the traffic, the attacker selects intermediate servers whose responses to the spoofed requests are larger than the requests themselves.

These types of DDoS attacks, which use large amounts of traffic to disable a victim server, are the focus of this article. However, source IP address spoofing is also used in many other attacks. An attacker who wants to evade source IP address based packet filtering will use source IP spoofing. Finally, some DDoS attacks do not rely on source IP address spoofing, because the attacker simply does not care whether or not the machine that it has compromised is implicated in the attack, so long as the attacker itself remains unknown. However, as source IP address filtering mechanisms become widely deployed, it is likely that attackers will have to resort to source IP address spoofing to increase the effectiveness of their attacks.

Traits of Defense Mechanisms
Because the current Internet infrastructure has few capabilities to defend against DDoS attacks, it is needed to design an adaptable network level defense mechanism against these attacks. Most predominant defense traits are

Fast response
The solution should be able to rapidly respond to and defend against attacks. Every second of Internet service disruption causes economic damage. We would like to immediately enable blocking of attack traffic.

Manuscript received October 5, 2008
Manuscript revised October 20, 2008

Scalable

Some attacks, such as TCP SYN flooding, involve a relatively small number of packets. However, many DDoS attacks are large scale and involve thousands of distributed attackers. A good defense mechanism must be effective against low packet count attacks, but also scale up to handle large-scale attacks.

Victim filtering

Some DDoS defense schemes in the literature assume that once the attack path is revealed, upstream routers will install filters in the network to drop attack traffic. This is a weak assumption because such a procedure may be slow, since the upstream ISPs have no incentive to offer this service to non-customer networks and hosts. A defense mechanism should enable sites to perform local filtering, which is especially effective if the attack does not cause network congestion.

Efficient

The solution should have very low processing and state overhead for routers and, to a lesser degree, victim servers.

1.2 Proposal

This paper proposes the Stack Path identification marking, a packet marking scheme based on Pi, and new filtering mechanisms. The Stack Path identification marking scheme consists of two marking methods that substantially improve path identification's incremental deployment performance i.e., Stack-based marking and Write-ahead marking. This scheme eliminates the effect of legacy routers when they constitute less than 20% of the topology, and performs 2–4 times better than the original Path identification scheme. For the filtering mechanism, derive an optimal threshold strategy for end hosts and edge servers for filtering based on the Path identification marking. The proposed system develops the path identification IP filter, which an be used to detect IP spoofing attacks with a single attack packet. It also examines the conflicts between IPv4 fragmentation and path identification marking, and path identification deployment in an IPv6 environment.

2. Literature Review

Many approaches have been proposed for securing against DoS and DDoS attacks. Ferguson and Senie propose to deploy network ingress filtering to limit spoofing of the source IP address [5]. However, unless every ISP implements this scheme, there will still be entry points in the Internet where spoofing can occur. Also, the additional router configuration and processing overhead to perform this filtering is another reason why it may not be widely deployed. Stone suggests a mechanism whereby ISPs use routers capable of input debugging connected through IP tunnels to an ASes border routers to enable AS-level tracing [14].

Park and Lee propose a distributed packet filtering (DPF) mechanism against IP address spoofing [13]. DPF relies on BGP routing information to detect spoofed IP addresses. Their approach is interesting, but requires high levels of router participation.

Bellovin et al. suggest adding a new type of ICMP message for traceback [2], [7], and Mankin et al. present an improvement to this scheme [12]. Several researchers propose to embed traceback information within the IP packet [1], [3]. Most of these schemes use the 16-bit IP Identification field to hold traceback information. Routers along the packet's path probabilistically mark certain bits in the IP Identification field in certain ways. While the traceback schemes could be used to find the origins of the attacks, they often require a large number of packets and thus cannot be used to filter out packets on a per-packet basis.

Ioannidis and Bellovin, and Mahajan et al. propose Pushback, a packet filtering infrastructure leveraging router support to filter out DDoS streams [6], [11]. Jin, Wang and Shin propose the use of packet TTL as an effective means of identifying spoofed traffic. The mechanism proposed in this article can be used to greatly increase the effectiveness of Pushback and Hop-count filtering, as the filters can take the packet markings into account and thus distinguish packets from various origins (increasing the accuracy of filtering).

Sung and Xu propose an altered IP traceback approach, where the victim tries to reconstruct the attack path but also attempts to estimate if a new packet lies on the attack path or not [15]. Their scheme is probabilistic and each router either inserts an edge marking for the IP traceback scheme or a router marking identifying the router. Unfortunately, their approach requires the victim to collect on the order of 105 attack packets to reconstruct a path, and once the path is reconstructed, this scheme will likely have a high false positive rate as the routers close to the victim will all lie on some attack path and frequently mark legitimate packets which will then get rejected.

The original Path identification marking is based on the use of the packet's TTL field as an index into the IP Identification field where a router should add its marks. This method is not as lightweight as the Stack Path identification method. Legacy routers have a harmful affect on the original Path identification scheme because they decrement the TTL of a packet but do not add any markings. The Stack Path identification scheme is robust to legacy routers and even includes the write-ahead scheme to incorporate markings for single legacy routers in the path.

Collins and Reiter use a novel approach of combining Cisco NetFlow data from a large network with Skitter map data, to compare DDoS defense mechanisms [4]. They measure the effectiveness of path aware defense systems (Path identification and Hop-Count Filtering), as well as Static and Network-aware clustering.Recently, network capability-based systems have been proposed for DDoS defense. Machiraju et al. propose a secure Quality-of-Service (QoS) architecture that is based on network capabilities [10]. Lakshminarayanan et al. leverage the i3 infrastructure to enable a receiver to cut off unwanted senders [8]. Anderson et al. [2] present an infrastructure where the sender uses a capability to set up a path to the receiver. We subsequently proposed SIFF, a capability-based system that allows a receiver to enforce flow-based admission control [16].

Yang et al. propose a capability-based mechanism with fine-grained service levels that attempts to address the denial-of capability attack [17]. They leverage Path identification markings to filter out floods of request packets in their scheme routers attempt to provide fair sharing among capability request packets based on their Path identification markings. Path identification are complementary to capability-based systems, and can be used to mitigate spoofing and flooding in the capability request channel.

3. Path Identification Scheme

The Path identification DDoS defense scheme is composed of a packet marking algorithm that encodes a complete Path Identifier in each Packet, and a packet filtering algorithm, that determines how a DDoS victim will use the markings of the packets it receives to identify and filter attack packets. The uniqueness of path identification lies in the fact that the path identification marking scheme is deterministic at the path level i.e., all packets traversing the same path receive the same marking. Because each packet contains the complete path marking, and the marking for a path is unchanging, then the victim need only identify a single attack packet or flow (through some high level algorithm based on packet contents or flow behavior) in order to block all subsequent packets arriving from the same path, and presumably, from the same attacker.

The Path identification marking scheme defines how the Pi-marks are generated as a packet traverses the routers along its path to its destination. Each path identification enabled router marks n bits into the IP Identification field of every packet it forwards. The IP Identification field is broken into 16/n. sections, and each router marks its n bits into the section indexed by the packet's current TTL modulo 16/n.. Because the IP Identification field is 16 bits in length, each Pi-mark can hold markings from the last 8 (n = 2) or 16 (n = 1) routers away from the packet's destination, a new router marking simply overwrites the marking of a previous router.

Our research on Path identification shows that the markings of the last 8 or 16 routers suffice for filtering out the majority of DDoS traffic, even though many different paths carry the same marking. The average Internet path length is roughly 15, which is almost double the number of hops that the n = 2 bit scheme can hold. Thus, the victim receives the markings from only the last 8 routers in the n = 2 bit scheme. It is found that the filtering power of Path identification improves if prevent the local domain routers from marking, thus preserving the markings from routers further away.

Internet packets would thus carry the markings from routers 4 to 11 hops away (assuming an n = 2 marking scheme). It is critically important that the individual router's markings have as high an entropy as possible, so that the probability of two distinct paths sharing or, colliding at the same marking is as small as possible.

For this reason, the router's marking bits are computed as the last n bits of the MD5 hash of the current router's IP address concatenated with the last hop router's IP address. A Path identification enabled router would cache its marking bits for each interface to avoid recalculating the hash for each forwarded packet.

The original Path identification mark works well in a network where all routers implement Path identification marking. Unfortunately, performance degrades substantially if legacy routers are present, as they decrement the TTL but do not mark the packet. This paper introduces two techniques that greatly enhance the performance of Path identification in the presence of legacy routers i.e., the Stack marking and the Write Ahead improvement.

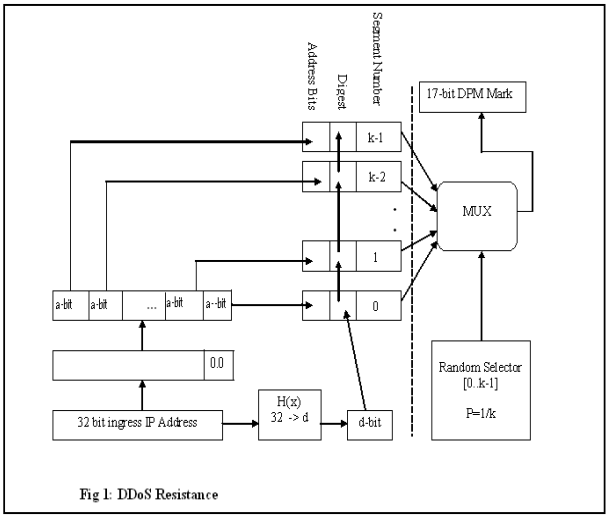

Fig 1 shows the mechanism of DDoS attack resistance.

The incoming 32 bit ingress address is processed for calculating the hash functional value. The ingress address is also parsed for the identification of the segment number and address bit. Then the hash functional value and parsed values are fed into the multiplexer for evaluating the marking process as indicated in fig 2.

4. Path identification Filtering Scheme

The Path identification filtering scheme defines how a DDoS victim uses the Pi-marks of the packets it receives to accept the least amount of attack traffic while accepting the most amount of legitimate traffic. The simplest Path identification filtering scheme is as follows i.e., upon identifying a particular mark as belonging to an attacker, the victim drops all subsequent packets bearing the same mark.

Unfortunately, because there are a constant number of Path identification marks, as the number of attackers increases it is more and more likely that any given Path identification mark will receive some attack packets, hence causing all legitimate user traffic to be dropped as well. This effect is called marking saturation.

To cope with marking saturation, the victim needs to have more flexibility in deciding whether or not to reject all packets with a particular Pi-mark. This flexibility can be defined in terms of a, a value measured as the maximum allowable ratio of attack packets bearing a particular Path identification mark to the total number of packets arriving with that Path identification mark. In a threshold filter, the victim will only drop all packets with a particular mark if the ratio of attack to total traffic on that mark equals or exceeds the threshold value.

Stack Marking

In order to generate a path identifier that is representative of a particular path from a source to a destination in the Internet, each router along the path must contribute some small amount of information whose aggregate among the routers of the path will be the Path identification marking. However, instead of using the packet's TTL to aggregate the markings from different routers, each router instead treats the IP Identification field as though it were a stack.

Upon receipt of a packet, a router shifts the IP Identification field of the packet to the left by n bits and writes its marking bits (calculated in the same way as in TTL marking) into the least significant bits that were cleared by the shifting. In other words, the router simply *pushes* its marking onto the stack. Because of the finite size of the identification field then most significant bits, which represent the oldest mark in the packet, are lost in this process just as in TTL marking.

The differences between TTL and Stack marking become evident when legacy routers are introduced into the topology. Unlike TTL marking, which interacts poorly with legacy routers because of its reliance on the packet's TTL which is modified by legacy routers, Stack marking does not rely on the TTL, and hence, has no interaction with legacy routers at all. There are no longer any marking holes because each marking router places its mark adjacent to the last marking router's mark, in the least significant bits of the IP Identification field. Completely marking the whole field using Stack marking requires only that there be 16/n non-legacy routers anywhere in the path.

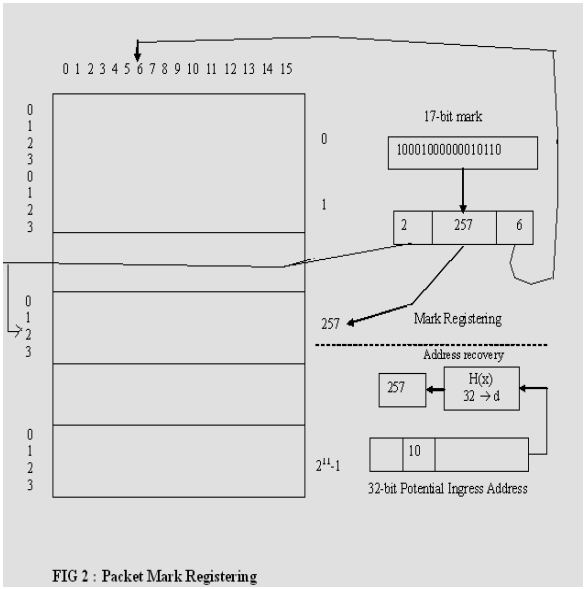

FIG 2 : Packet Mark Registering

The packet mark registration process is shown in the Fig: 2. The 32 bit ingress address is evaluated with the hash mark function to identify the value set on the incoming packet. The value obtained from the incoming packet is marked for register if the abnormal event marks is associated with the value set. The positional indication of the mark values are refered in the 32 bit address.

5. Experimentation and Performance Evaluation

5.1 DDoS Attack Model

In order to model Pi's performance under a DDoS attack, must have some way for the DDoS victim to identify attack packets, so that it can bootstrap the Path identification filter. Unfortunately, this requires the simulation of a higher-level algorithm that is likely to be

dependent on the content of the traffic (HTTP or DNS etc.) to make its classifications.

To compensate for this, model our DDoS attack in two phases i.e., the learning phase and the attack phase. In the learning phase, the victim is considered omniscient, and can determine, for each packet received, whether that packet originated from an attack or a legitimate user. This phase of the attack is used to simulate the effect of a high-level traffic and content analysis algorithm, without specifying the algorithm itself. The knowledge gained in the learning phase is used to bootstrap the Path identification filter with the Path identification markings of known attackers. In the attack phase, the victim can no longer differentiate attack and user packets and is forced to use the Path identification filter to make accept or drop decisions for every packet it receives. All of the results presented are taken during the attack phase. The length of the learning phase is 3 packets per legitimate user and 30 packets per attacker. The length of the attack phase is 20 packets per legitimate user and 200 packets per attacker.

The DDoS simulations is done NS2. A certain number of *paths* are selected, at random, from the topology file and assigned to be either attack or legitimate user paths. All of the DDoS simulations have 50 legitimate users and vary the number of attackers. The system use an n = 2 bit marking scheme and assume, that the last three hops of any path are under the victim's ISP control and thus, do not add their marks to the packet. The results presented are the averages of 6 runs of each attack.

The threshold value of the Path identification filter is used to give a DDoS victim some flexibility in deciding whether or not to drop all packets arriving with a particular mark by setting a minimum acceptable level of user traffic to that Path identification mark. Derive the formula for the optimal threshold value as a function of attack and user traffic, and confirm the optimality of our result using our DDoS simulation.

In order to quantify the performance of the Path identification filter, we first define two metrics, representing the two different types of errors a Path identification filter can make i.e., false positives, where legitimate users' packets are dropped; and false negatives, where attackers' packets are accepted. For the purpose of our evaluation, refer the following two metrics, the user acceptance ratio, which is 1 minus the false positive rate, and the attacker acceptance ratio, which is exactly the false negative rate.

Stack Path identification in IPv6.

Although the Path identification scheme has been specifically designed for deployment in IPv4, its principal ideas are equally applicable in an IPv6 environment. The IPv6 protocol does not support en-route packet fragmentation, and thus does not have an equivalent field to the IP Identification field of IPv4. There are, however, two possibilities for marking space in IPv6, in the flow identification field or in a hop-by-hop option. The advantage of marking in the flow identification field of the header is that because the field is part of the standard header, router markings will not add to the packet's size (which might cause the packet to exceed the MTU of an intermediate network and be dropped). The flow identification field is 20 bits in length, which allows more routers to include their markings in each mark.

The other option is to include the Path identification marking in a hop-by hop option inserted by the first Path identification enabled router in the path. The benefit of this approach is that the length of the option need not be limited to 20 bits, as is the flow identification field. However, inserting such an option into the packet may cause it to exceed the MTU of a link somewhere along the path. In either case, DDoS protection is a critical feature that should be present at the network level, and IPv6's current limited deployment makes it a good candidate for modification to include the Path identification scheme.

Deployment

Previous DDoS defense mechanisms do not provide a good incentive structure to foster adoption. For example, consider the benefits to an ISP deploying ingress filtering. That ISP protects other ISPs' customers from its own customers, as ingress filtering stops its customers from spoofing their source IP address.

Ingress filtering does not directly benefit the customers of the ISP, yet it introduces more complexity, higher router management overhead, lower performance due to filtering, and potential customer problems (when some legitimate customer's packets get filtered out). In contrast, the Path identification scheme offers very good incentives for deployment that encourage adoption. If an ISP deploys Path identification marking on all its routers, a customer can immediately start using the filtering techniques we describe in this article to determine from where the attack traffic enters its ISP's topology.

A victim can already perform filtering if only 17% of the routers implement Path identification marking. Ideally, this creates a market pressure for ISPs to deploy Path identification enabled routers. If ISPs want to deploy Pi, this creates an incentive for router manufacturers to produce path identifier enabled routers.

It is anticipated that the benefits of Path identification will produce these market incentives that drive deployment. The main difference with previous techniques is that Path identification deployment immediately benefits the customers of an ISP, and helps those customers defend against DDoS attacks.

Fig 3 shows the graphical display of the DDoS attack resistance (time delay) against the cause of attack (path length variation). The graph values shown are taken by combining the overall attack path variation in the given simulation conditions of 40 nodes. It is inferred from the graph the proposed system is more effective as shown by the fact that the ddos attack resistive mechanism works in direct ratio for the combined path variations.

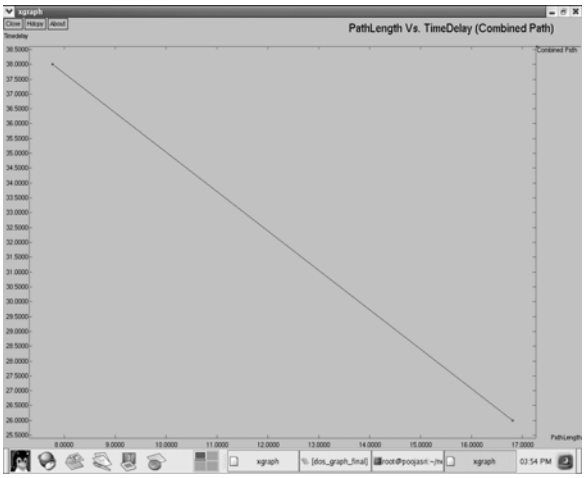

Fig 3: Pathlength Vs Time Delay (Combined path)

The ddos attack resistance time delay is calculated for each and every node attack is shown in fig 4. In this display, time delay is directly proportionate to path length variation in normal traffic shown in the upper curve is less effective (as path length increases time delay also increases). Time delay is inversely proportional to path length attack variation in uneven network traffic shows (as path length increases time delay decreases) that the proposed system is more effective.

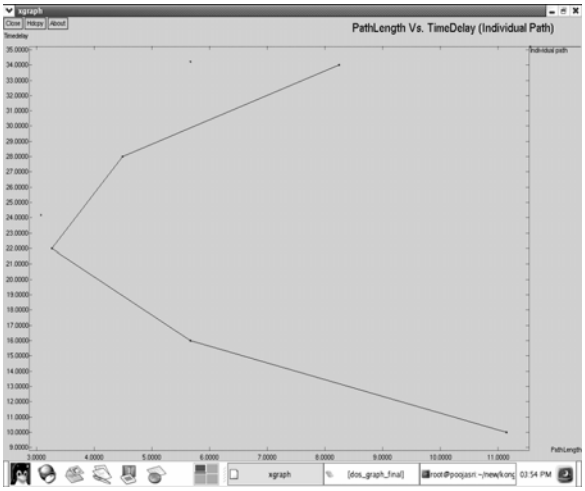

Fig 4: Pathlength Vs Time Delay (Individual path)

Changing Path identification Marks

One of the basic assumptions of the Path identification scheme is that the paths from specific senders remain constant over the timescale of an attack. Attackers can exploit this assumption in a variety of ways. Instead of focusing a DDoS traffic flood on a particular victim, an attacker can try and flood the routers along the path to the victim, potentially causing a disruption in the paths packets take to reach the victim, resulting in new Path identification marks arriving at the victim.

A clever victim may be able to identify the router under attack by comparing the Path identification marks of traffic before and after the attack begins. Colluding attackers may try to poison the PiIP filter by coordinating to complete a TCP connection, while spoofing an address belonging to a single attacker. The end-host may be fooled into including the Path identification marks of all attackers as legitimate Path identification marks of the one attackers IP address. This attack is limited, however, because the attackers would need to spoof the same address, or set of addresses, (if they poison more than one attackers address) during the flooding phase of their attack.

6. Conclusion

This paper presented approaches for packet marking and filtering in the Path identification DDoS defense scheme. The Stack Path identification marking improvements, stack-based and write-ahead marking, eliminate the marking holes generated by legacy routers and include the markings from single legacy routers immediately following Pi-enabled routers in a path. We derive an equation that allows a DDoS victim to select the optimal threshold value for the Path identification filter.

The system introduces a novel filter which relies on the hash path identification, IP path identification tuple of each packet, making it far less likely that an attacker will successfully bypass the filter.

With these improvements, our evaluation shows that Path identifier provides measurable DDoS protection, even when only 20% of routers in the Internet participate in the marking scheme. The Path identification scheme is very general and quite promising in performance. These properties promise to make Path identification a critical deterrent to today’s most common Internet attacks.

References

[1] Micah Adler. Tradeoffs in Probabilistic Packet Marking for IP Traceback. In Proceedings of 34th ACM Symposium on Theory of Computing (STOC), pages 407 418, 2002.

[2] S. Bellovin, M. Leech, and T. Taylor. The ICMP Traceback Message. Internet-Draft, draft-ietf-itrace-01.txt, October 2001. Work in progress, available at ftp://ftp.ietf.org/ internet-drafts/draft-ietf-itrace-01.txt.

[3] Hal Burch and Bill Cheswick. Tracing Anonymous Packets to Their Approximate Source. In Proceedings of Usenix LISA, pages 319–327, December 2000.

[4] Michael Collins and Michael K. Reiter. An Empirical Analysis of Target- Resident DoS Filters. In IEEE Symposium on Security and Privacy, May 2004.

[5] P. Ferguson and D. Senie. Network Ingress Filtering: Defeating Denial of Service Attacks Which Employ IP Source Address Spoofing. RFC 2267, January 1998.

[6] John Ioannidis and Steven M. Bellovin. Implementing Pushback: Router- Based Defense Against DDoS Attacks. In Proceedings of the Symposium on Network and Distributed Systems Security (NDSS 2002), San Diego, CA, February 2002.

[7] ICMP Traceback (itrace). IETF working group,http://www.ietf.

[8] K. Lakshminarayanan, D. Adkins, A. Perrig, and I. Stoica. Taming IP Packet Flooding Attacks. In Proceedings of ACM HotNets-II, pages 45–50, November 2003.

[9] Heejo Lee and Kihong Park. On the Effectiveness of Probabilistic Packet Marking for IP Traceback under Denial of Service Attack. In Proceedings IEEE Infocomm 2001, April 2001.

[10] S. Machiraju, M. Seshadri, and I. Stoica. A Scalable and Robust Solution for Bandwidth Allocation. In International Workshop on QoS, May 2002.

[11] Ratul Mahajan, Steven M. Bellovin, Sally Floyd, John Ioannidis, Vern Paxson, and Scott Shenker. Controlling High Bandwidth Aggregates in the Network. CCR, 32(3):62–73, July 2002.

[12] A. Mankin, D. Massey, C.L. Wu, S.F. Wu, and L. Zhang. On Design and Evaluation of Intention-Driven ICMP Traceback. In Proceedings of the IEEE International Conference on Computer Communications and Networks, October 2001.

[13] Kihong Park and Heejo Lee. On the Effectiveness of Route-Based Packet Filtering for Distributed DoS Attack Prevention in Power-Law Internets. In ACM SIGCOMM ’01, pages 15–26, 2001.

[14] Robert Stone. CenterTrack: An IP Overlay Network for Tracking DoS Floods. In Proceedings of the 9th USENIX Security Symposium, pages 199–212, Denver, Colorado, August 2000.

[15] Minho Sung and Jun Xu. IP Traceback-based Intelligent Packet Filtering: A Novel Technique for Defending Against Internet DDoS Attacks. In Proceedings of IEEE ICNP 2002, November 2002.

[16] Avi Yaar, Adrian Perrig, and Dawn Song. SIFF: A Stateless Internet Flow Filter to Mitigate DDoS Flooding Attacks. In Proceedings of the IEEE Symposium on Security and Privacy, pages 130–143, May 2004.

[17] Xiaowei Yang, David Wetherall, and Tom Anderson. A DoS-limiting Network Architecture. In Proceedings of ACM SIGCOMM, pages 241– 252, August 2005.