BASE: An Incrementally Deployable Mechanism for Viable IP Spoofing Prevention

Heejo Lee, Minjin Kwon
Dept. of Computer Science and Engineering
Korea University
Seoul, South KOREA
{heejo,mjkwon}@korea.ac.kr

Geoffrey Hasker, Adrian Perrig
CyLab / CMU
Pittsburgh, USA
ghasker@andrew.cmu.edu
adrian@ece.cmu.edu

ABSTRACT

DoS attacks use IP spoofing to forge the source IP address of packets, and thereby hide the identity of the source. This makes it hard to defend against DoS attacks, so IP spoofing will still be used as an aggressive attack mechanism even under distributed attack environment. While many IP spoofing prevention techniques have been proposed, none have achieved widespread real-world use. One main reason is the lack of properties favoring incremental deployment, an essential component for the adoption of new technologies. A viable solution needs to be not only technically sound but also economically acceptable. An incrementally deployable protocol should have three properties: initial benefits for early adopters, incremental benefits for subsequent adopters, and effectiveness under partial deployment. Since no previous anti-spoofing solution satisfies all three of these properties, we propose a new mechanism called "BGP Anti-Spoofing Extension" (BASE). The BASE mechanism is an anti-spoofing protocol designed to fulfill the incremental deployment properties necessary for adoption in current Internet environments. Based on simulations we ran using a model of Internet AS connectivity, BASE shows desirable IP spoofing prevention capabilities under partial deployment. We find that just 30% deployment can drop about 97% of attack packets. Therefore, BASE not only provides adopters' benefit but also outperforms previous anti-spoofing mechanisms.

Keywords

DDoS attack, IP spoofing, packet marking and filtering, BGP anti-spoofing extension

1. INTRODUCTION

The prevention of IP spoofing continues to be an important challenge. In *IP spoofing*, an attacker forges the source IP address to deceive the receiver of the true packet origin [5]. IP spoofing is used by attackers to defeat source-based filtering and resource allocation mechanisms to mount a variety of attacks, mainly denial-of-service (DoS) attacks. For example, IP spoofing is used in TCP SYN flooding, an attack that ties up server resources [5].

Currently, most large-scale Internet attacks do not use IP spoofing attacks, as evidenced by the distributed denial-of-service (DDoS) attack on the domain name system (DNS) [9], and the various network worms. Recent attacks do not use IP spoofing simply because they *do not need to spoof*—the massive distribution of attack agents is already enough to paralyze a victim. However, it would be short-sighted to disregard IP spoofing as a threat, because as DDoS and worm defense mechanisms start to get deployed, IP spoofing will become again an attractive mechanism to circumvent the deployed defense mechanisms.

In "Crossing the Chasm", Moore points out that the customers for a product range from *early adopters*, to *early majority*, to *late majority*, and finally to *laggards* [19]. A central difference in the deployment of networking protocols is the availability of hardware and software that implements the protocols. In the absence of active networks [28], network operators need to wait for the availability of hardware or software upgrades. Although the market for Internet technology differs from a mainstream product, we can still draw some analogies. Initially, some specialized hardware may be available that implements the new protocols; as the demand increases, the mainstream router manufacturers may implement the new protocols as well.

We conjecture that the deployment of networking protocols follows a similar trend: early adopters with a critical need for some new technology purchase specialized equipment. As the larger router manufacturers recognize the business case and observe customer demand, they also implement the feature, causing the early mainstream customers to deploy it. Finally, some customers may not update their routers frequently and thus they may require a longer time period until they implement the functionality.

Assuming that our network protocol provides good incremental deployment properties, we are only concerned with bootstrapping the process from the early adopters to the early majority. Thus, in this paper we study a viable protocol which needs to have the following three properties:

- **Initial benefit** : The protocol needs to provide initial benefits for early adopters. The first two deployments should already result in a benefit.

ASIACCS'07, March 20-22, 2007, Singapore.

- **Incremental benefit** : The protocol needs to provide incremental benefits for the early majority. Such benefits should increase substantially as deployment proceeds.
- **Partial deployment** : The protocol needs to provide properties such that a proportionally small deployment becomes sufficiently effective. Broad deployment requires a prolonged period; thus, a practical protocol approaches full strength as approximately 30–50% of routers deploy the mechanism, which requires about 10% of the larger ASs.

An anti-spoofing protocol needs to be not only technically sound but also economically acceptable. Unfortunately, currently proposed IP spoofing prevention mechanisms are inadequate, especially in the dimension of providing incentives for incremental deployment.

An alternative mechanism to drive deployment is to provide *disincentives for non-deployers*, for example through legal pressure. In some countries, for example Italy, the computer crime law holds people liable for damages caused by their computers, even if the computer was externally compromised [4]. However, positive incentives are preferable, as they do not rely on globally drafted laws for global deployment.

The typical filtering scheme that shows incrementally deployable effect is DPF [20]. However, it does not give direct benefits to the adopters. Also, DPF should store the routing information of all nodes,[1] and moreover it should update all nodes' routing information for each node whenever even one route is updated. Therefore, DPF requires tremendous space and produces high messaging costs for each node. Compared with DPF, the proposed scheme, BASE(BGP Anti-Spoofing Extension) gives direct incentives to the adopters. Also, BASE works on the existing routing protocol, BGP, to distribute marking values, so there is not much messaging overhead. Therefore, DPF is not suitable for implementing in the real world, but BASE is a viable solution in current Internet environment.

This paper's main contribution lies in advancing an incrementally deployable mechanism for viable IP spoofing prevention. We propose an anti-spoofing solution that satisfies the three incremental deployment properties. This scheme works with legacy BGP routers transparently. We take a marking and filtering approach using BGP update messages, so the marking value inscribed in a packet should exist as the source's entry in each node's filtering table, or otherwise drop the packets. The power for detecting spoofed packets rises substantially as the number of deployers increases. With only about 30% of ASs deploying the mechanism, we can almost perfectly filter attack packets. Thus, BASE is the only viable solution suggested so far which satisfies the three deployment properties. Furthermore, we show that BASE outperforms previous schemes in terms of detecting spoofed packets under the same deployment circumstances.

2. RELATED WORK

Researchers have followed two main directions in the investigation of techniques to mitigate spoofed source IP addresses: IP traceback and detection of spoofed packets. The goal of IP traceback is to find the true origin(s) of attack packets; probabilistic packet marking is one such IP traceback mechanism [15, 24, 25]. However, IP traceback has several drawbacks; for example, spoofed packets can destroy a victim network before being reactively curtailed, and the uncertainty of IP traceback amplifies under distributed attacks, which eventually makes IP traceback useless under massive DDoS attacks.

In this paper we discuss the second direction: how to detect spoofed packets. Once given an ability to discriminate between attack packets and legitimate packets, it is a simple task to filter attack packets before they reach a victim. In the remainder of this section, we discuss these approaches.

Ingress filtering [2, 8] has been proposed for dropping packets with invalid source addresses before the packets leave their local networks. However, the usefulness of ingress filtering depends on the deployment, providing little incentive to early adopters. Moreover, the incentives for deployment of ingress filtering are structured in an awkward fashion. Consider an ISP that deploys ingress filtering—this does not benefit the customers of the ISP directly, but it protects other ISPs customers *from* its own customers (because they cannot send spoofed IP packets). Thus, ingress filtering does not provide significant benefits for early adopters; except in the case where laws make the sender of malicious packets liable for the damage caused (as is the case in Italy), because the customers won't be able to use IP spoofing to attack a victim.

Reverse path forwarding (RPF) is an extension of ingress filtering, which uses IP routing tables for dropping spoofed packets [2]. RPF has become an optional function of mainstream routers in order to mitigate the problems caused by IP spoofing [6]. RPF-enabled routers forward only packets that have valid source addresses consistent with the IP routing table. However, there is one topological restriction; RPF can only be used for symmetric routing environments. Moreover, RPF does not provide sufficient benefits to adopters, which is the same as ingress filtering.

Route-based distributed packet filtering (DPF) has been proposed for filtering spoofed packets using routing information [20]. DPF determines whether a packet travels an unexpected route from its specified source and destination addresses, and if so, discards the packet. DPF can be viewed as a generalized address-based filtering scheme which eliminates the limitations of ingress filtering and RPF. The DPF filter can be located in transit ASs; thus, only a part of the Internet needs to be used for filtering a significant fraction of spoofed packets. But DPF does not provide direct incentives to deployers—everyone shares the benefits.

Path identification (Pi) is a reactive filtering scheme based on packet marking [29]. In Pi, each packet in the same path has the same identifier, which can be used for filtering attack packets. Thus, Pi is beneficial to adopters who can use the Pi-filter for protecting their network. However, Pi gives little benefit for early adopters because it becomes especially effective after substantial deployment.

Hop-count filtering (HCF) is another filtering technique for spoofing attacks [12]. The idea behind HCF stems from the fact that packets coming from the same location travel the same path to the destination. Thus, time-to-live (TTL) values in IP headers can be used for classifying the attack packets. However, the TTL is only an estimation of hop count, so HCF provides higher false-positive results than

[1] A node can be a BGP router or an AS which holds a group of routers. Hereafter, a node represents a BGP router since router-level description gives us clear understanding.

Pi [7], and attacking with fake TTL values further reduces the effectiveness of HCF. Thus, even if HCF does not require any deployment, it is not a sufficient countermeasure against IP spoofing.

3. THE BASE MECHANISM

This section proposes a new mechanism called "BGP Anti-Spoofing Extension" (BASE) which combines Pi [29] with DPF [20], and further enables the three deployment properties for a viable protocol addressed in Section 1. As discussed in Section 2, DPF is incrementally deployable, but deployers have no more incentive than non-deployers. Pi provides a benefit to deployers, but only at relatively high levels of deployment. Thus, Pi does not give enough benefit to early adopters. In BASE, path-based marking enables in-network filtering and provides significantly more benefit to early adopters than non-deployers. BASE enables for deployers to utilize the distributed filters for collectively dropping attack packets. As well, in the occurrence of an attack on other nodes, the deployer can provide the proof of innocence by showing that the marking value propagated to the victim does not originate from the deployed node.

We first assume an attacker sends spoofed packets to the target node to hide the identity of the attacker. Second, a victim has the ability to recognize a spoofing attack. Once the attack is recognized, the victim can utilize BASE for protecting itself from the attack. Third, we assume BGP-enabled routers can utilize the BASE filter. Each BGP router has a marking and filtering policy, so BASE has its roots in network-based filtering. Forth, each BGP-enabled router within an AS can be updated for performing the BASE mechanism with the following assumptions:

- **Per-AS key:** Each AS has a secret key which it uses for computing marking values; the key is shared by the routers within the AS.
- **Enough marking space:** We assume that the IP header has sufficient space to store a marking value.
- **Router marking and filtering:** The BASE router(s) on the border of an AS mark every outgoing packet and filter every incoming packet without a correct mark.

The BASE mechanism distributes valid marking values through BGP update messages [23]. Thus, every BASE-enabled BGP router learns the valid marks for every source network. We takes advantage of prevalent BGP routers and further deploys BASE-enabled routers in an incremental way. When a BASE-enabled victim is under an attack, the victim can invoke packet marking and filtering for checking the validity of every packet destined to the victim network. Thereafter, only legitimate packets with a correct mark can go through the BASE network; spoofed packets are dropped before entering the BASE-enabled secure network. After proper handling of the attack, the victim can revoke the marking and filtering function for its network.

The marking in BASE is "path-based" instead of "IP-based", which means the use of network addresses (prefixes of an IP address) instead of individual IP addresses. This reduces the required storage for marking values, however, collective filters can detect spoofed packets effectively. Also, the marking value in the filtering table of each router is mapped based on source's network address, similar to the use of network addresses in the routing table of each router.

A BASE router communicates with other BASE routers by the use of BGP update messages. One BGP update message invokes distributed filters, propagating through all legacy BGP routers, reaching all BASE-enabled routers. This mechanism works transparently with the legacy BGP speakers by using *optional transitive attributes* [23], in which the information stored in transitive attributes is forwarded even through non-BASE routers. However, it is possible that an AS's routing policy may prevent an update from propagating to a neighboring AS, even though it sends packets to that AS. The effect of BGP routing policy decisions requires further study.

The next subsection describes the four phases of BASE. First, marking values are computed by a one-way hash chain and they are distributed using BGP update messages. Second, once the victim notifies that an attack is happening, the victim propagates invocation messages. Third, BGP-enabled routers activate the BASE mechanism, so BGP-enabled routers inscribe marking values and filter packets with the false markings. Finally, the victim can propagate revocation messages after the attack subsides. Then, we show how the proposed mechanism works in environments of full deployment, partial deployment, and asymmetric routing paths.

3.1 Four-Phase BASE Mechanism

The framework of BASE extends the concept of distributed packet filtering (DPF) with cryptographic packet marking, which enables non-adjacent BASE-enabled ASs to verify path correctness. As well, the BASE mechanism runs on-demand filtering for specific destination addresses. Thus, only when undergoing DoS attacks, the victim can initiate collective filtering of attacking packets crowding into the victim's network.

For the purpose of distributing filtering information, one approach is the use of BGP update messages to coordinate between routers [20]. Alternatively, we can design our own distribution protocol using piggybacking on regular packets or generating information packets. The SAVE protocol [16] is an example of designing a new protocol for verifying the correctness of the source address of each incoming packet. ICMP traceback messages [3] is an example of generating extra ICMP messages to send traceback information to the victim. There are advantages and disadvantages for using legacy protocols or designing our own protocols as an information distribution scheme. When using BGP messages, BASE easily distributes marking values and maintaining up-to-date marking information. Even though there is the issue of false-positive according to AS's routing policy in asymmetry environments by using BGP messages, that is not a big problem because we have false-positive only during an attack. In this paper, it is assumed that BGP update messages are used for distributing filtering information, while preserving the primary properties of deployment issues.

BASE works according to following four phases: *distribution of marking values, filter invocation, packet marking and filtering,* and *filter revocation.* We let s denote the source AS, and t the destined AS. And, v is the current filtering AS. Then, the situation is for considering a packet of (s, t) traveling the filter v.

Phase 1: Distribution of marking values..

The *distribution* phase is for distributing marking values

among BASE filters. The marking value is computed by a one-way hash chain, i.e., $m_i = MAC(k_i, m_{i-1})$, where k_i is the secret key and m_0 is the prefix of the IP address of the victim. The marking values are distributed using BGP updates. This is a once-only operation unless the BGP path has been changed.

Phase 2: Filter invocation..
The *invocation* phase invokes packet marking and filtering for packets destined to the victim network. Invocation is propagated through BGP update messages. Upon receiving an invocation message, a BASE node starts packet marking and filtering for the corresponding addresses.

Phase 3: Packet marking and filtering..
The *filtering* phase marks outgoing packets and filters incoming packets without a correct mark. Every packet with the same source address will have the same mark when it leaves a BASE node, even though it may have arrived with different marks through different interfaces. This replacement scheme allows the BASE mechanism to work in asymmetric routing paths.

Phase 4: Filter revocation..
The *revocation* phase terminates marking and filtering of packets destined to the victim. Revocation is also propagated via BGP update messages.

Internet connectivity can be represented by a graph $G = (V, E)$ where V is the set of nodes and E is the set of edges. The graph G represents the AS-level connectivity such that a node is an AS and an edge is a link between two nodes. A path $\mathcal{P}(s,t)$ is an ordered set of consecutive edges from a source s to a destination t such that $\mathcal{P}(s,t) = \{v_1, v_2, .., v_n\}$ where $v_1 = s$ and $v_n = t$. The marking values are computed as shown in Fig. 1. The marking value of v_i for (s,t) is defined by $m_i = MAC(k_i, m_{i-1})$ where k_i is the key of v_i and m_0 is the prefix of s. The computed marking value for each node is distributed to next nodes described as Fig. 2

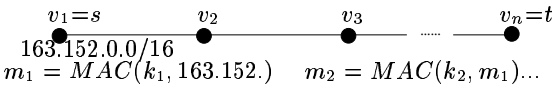


Figure 1: Marking value computation for each node

Each BGP node has additional features for packet marking and filtering. We call this node a BGP filter and the function of a BGP filter is formally described as Fig. 3 shows. Each BGP filter v_i has a filtering table F. We can store all the legitimate marking values in source's entry of the filtering table. If we can store only one marking value in each

Distribution of marking values(m_i)
2 FOR each BGP filter v_i from $i = 1$ to $i =$ all filter nodes
3 $m_i = MAC(k_i, m_{i-1})$
4 forwards m_i to next BGP filter nodes
by using the BGP optional transitive attributes
5 ENDFOR
END of Algorithm

Figure 2: Distribution algorithm of marking values

Packet marking and filtering
2 FOR each BGP filter v_i from i=1 to i=all filter nodes
3 IF $m_{i-1} \in F(s)$ // F is a filtering table for a filter
4 forwards a packet (s,t) to $\mathcal{R}(t)$ with a new mark m_i
5 ELSE
6 drops (s,t)
7 ENDIF
8 ENDFOR
END of Algorithm

Figure 3: Packet marking and filtering algorithm in a node

record of a filtering table, we call this "one mark" and if we can store multiple marking values, we call this "multiple marks". This will be explained in detail in Section 3.4. In distribution phase, when a filter receives a marking value, the filter stores the marking value in its filtering table. In marking and filtering phase, when a filter receives a packet (s,t), the filter forwards the packet to $\mathcal{R}(t)$ with a new mark m_i only if $m_{i-1} \in F(s)$, otherwise it drops (s,t). $\mathcal{R}(t)$ denotes t's entry in v_i's routing table.

Cryptographic approaches improve the strength of packet marking under the attacker's forgery of marking values as well as source addresses. Since the marking field spoofing diminishes the effectiveness of packet marking [15], we use a cryptographic approach such as message authentication codes (MAC) to check the validity of marking values. Pseudo-random functions (PRF) [10] can be used as a hash function for MAC. A PRF takes two arguments, a key and an input, then produces an output which is indistinguishable from a random value as long as the key is secret.

One-way hash chains—cryptographic primitives frequently used in the design of secure protocols—are used for computing marking values. Computation of a chain of marking values has advantages of reducing forgeability of marking values and enhancing routability of legitimate packets, without prior knowledge of traveling paths. Marking values are pre-computed and distributed between neighbor BGP filters, then they are used for marking and filtering. No additional computation of marking values is required during packet processing except for comparing and changing a marking value with table lookups.

Several fields have been proposed for storing a marking value in a packet. They include the record route option in the Internet Protocol (IP) [22], the IP identification (ID) field [24], the IP header available by compression [1], and in conjunction with the IP ToS field or the IP TTL field. There are pros and cons on their usages, and the ID field has received most attention based on overloading the 16-bit IP identification field used for fragmentation in previous works [24, 25, 29]. For further discussion, it is assumed that the BASE uses the 16-bit IP ID field for storing marking values.

3.2 BASE Protocol Design

A BASE router communicates with other BASE routers by using BGP update messages. One BGP update message invokes distributed filters, propagating through all legacy BGP routers, reaching all BASE-enabled routers. This mechanism works transparently with the legacy BGP speakers by using *optional transitive attributes* [23], in which the infor-

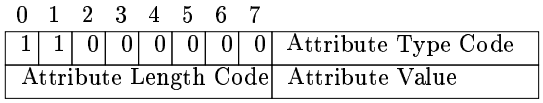


Figure 4: The Path Attributes field using optional transitive attributes of the BGP update message

mation stored in transitive attributes is passed on to other BGP speakers, even if it is not understood by non-BASE routers.

Each path attribute is a triple of Attribute Type, Attribute Length and Attribute Value [23]. The high-order bit(bit 0) of Attribute Type is the optional bit and the second high-order bit(bit 1) of Attribute Type is the transitive bit. Accordingly, we should set these two bits to 1 as you can see in Fig. 4. The Attribute Type Code octet should contain the attribute type code which is not currently defined. We can create new BGP attribute type code and should send it to the IANA [17]. For example, we can use value 32 for attribute type code of BASE marking mechanism. Also, as Fig. 5 shows, we can construct Attribute Value field.

Fig. 5 (a) is a format of an attribute value for distribution of marking values. Fig. 5 (b) is a format of an attribute value for BASE filter invocation and revocation. The Type field defines whether the BGP update message is for distribution of marking values(if set to 1) or invocation(if set to 2) or revocation(if set to 3). The Source field is for the source's AS number of the BGP message. The Marking Value field gets a 16-bit marking value for distributing it to next BGP filter nodes.

3.3 Fully-Deployed BASE on Symmetric Paths

A routing path is called symmetric if the path has the same forward and backward path between two nodes. A symmetric routing path of (s,t) implies that the forwarding path of (s,t) is a subgraph of the BGP tree such that $\mathcal{P}(s,t) \subseteq \mathcal{B}(s)$, where the BGP tree $\mathcal{B}(s)$ is a tree expanded by BGP updates for s. This BGP tree is an s-rooted spanning tree constructed by the best routes destined to s. Thus, the propagation of marking values for (s,t) follows the path $\{v_1, .., v_n\}$ in the spanning tree, and that becomes the routing path for (s,t) in symmetric routing. Since the BGP updates flow to the opposite directions of the chosen best routes to s, the BGP flooding paths are not always equivalent to the routing path starting from s. We will discuss asymmetry of routing paths in the next subsection.

Fig. 6 shows how BASE works for two nodes, s and t, in a network. Marking values for s are distributed by the use of BGP updates, as shown in Fig. 6 (a). Using a secret key k_i of v_i, an unpredictable marking value m_i is computed by $MAC(m_{i-1}, k_i)$ as follows.

$$\begin{aligned} m_1 &= MAC(k_1, Pref(s)) \\ m_2 &= MAC(k_2, m_1) \\ m_3 &= MAC(k_3, m_2) \end{aligned}$$

Fig. 6 (b) shows the *invocation* messages being propagated from t, which is under a spoofing attack. After that, each node drops packets destined to t without a correct mark, which are attacking packets with spoofed source addresses and/or fake marks. Fig. 6 (c) shows that each node checks the marking value of a packet (s,t) and inscribes a new marking value before forwarding to the next node.

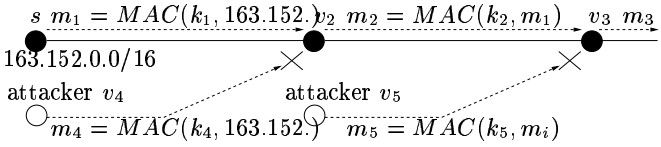


Figure 7: Spoofed packets dropping case

We utilize a MAC that uses an algorithm for generating authenticated output, which is used to ensure the authenticity and integrity of the marking values. As shown in Fig. 7, an attacker v_4 cannot produce a correct marking value m_1 without the secret key k_1 of s because m_1 is computed through $m_1 = MAC(k_1, 163.152.)$. Therefore, the attacker v_4 cannot send packets with a correct marking value m_1 because v_4 cannot know the secret key of s. The attacker just computes an incorrect marking value $m_4 = MAC(k_4, 163.152.)$. Hence even though an attacker tries to send packets with spoofed IP address, the attacker cannot produce a correct marking value as long as an attacker cannot predict the secret key of the node. In node v_2, the packets with a correct marking value m_1 can go through the path, but the packets with an incorrect marking value m_4 are dropped. Additionally, an attacker v_5 cannot send packets with a correct marking value m_2. The attacker v_5 doesn't know not only the secret key k_2 of v_2 but also the marking value m_1 which is propagated from v_2's upstream node. Therefore, the attacker v_5 has an incorrect secret key k_5 and an incorrect marking value m_i. As a consequence, the attacker v_5 computes an incorrect marking value m_5, so those packets with the incorrect marking value are dropped in node v_3.

3.4 Asymmetric Paths in Full Deployment

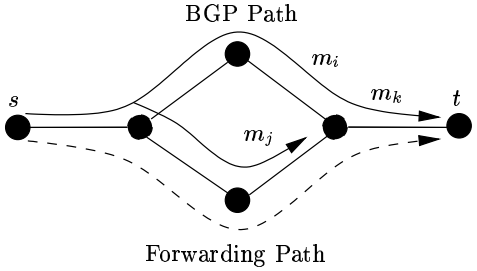


Figure 8: Asymmetric routing paths when a routing path does not go through BGP paths

There are non-negligible portions of routing asymmetry in the current Internet [21]. A recent study measured US academic networks display about 14% routing asymmetry while commercial networks show about 65% routing asymmetry on the AS-level [11]. Routing asymmetry is often caused by routing policies or traffic engineering such as hot-potato routing [26], load distribution [27], or delayed BGP convergence [14] that may result in random selection among multiple shortest paths.

Fig. 8 shows the case of asymmetric routing paths, where the routing path from s differs from the BGP paths from s. This causes packets to travel a different path from the path transferring marking values. But, since BASE replaces the received mark with the new one, the node in a merged

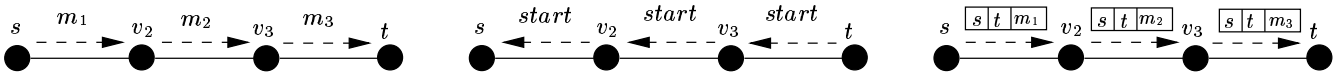
Type (1 byte)	Source (2 byte)	Marking Value (2 byte)

(a) The Attribute Value for distribution of marking values

Type (1 byte)	Source (2 byte)

(b) The Attribute Value for filter invocation and revocation

Figure 5: The Attribute Value of the Path Attributes field of BGP update message: (a) for distribution of marking values, (b) for BASE filter invocation and revocation.



(a) Distribution of marking values for s

(b) Filter invocation by t

(c) Packet marking and filtering for t

Figure 6: Working example of fully-deployed BASE. (a) Distribution of marking values along with the BGP update messages from s. (b) Filter invocation for packets destined to t with the BGP update messages from AS t. (c) Packet marking and filtering of spoofed packets destined to t, without correct marking values for the source address s.

point can update a packet in asymmetric routes with a correct mark. For instance, if the filter receives any of the valid marks m_i and m_j through the corresponding interfaces, it changes the mark to m_k before forwarding to the next node, as shown in Fig. 8. This replacement resolves the risk of dropping legitimate packets so that BASE works under asymmetric forwarding paths.

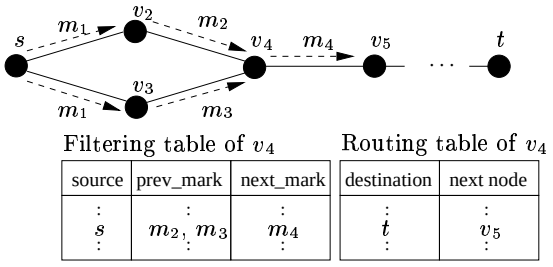


Figure 9: Storing multiple marks in asymmetric routing paths in order to admit every possible legitimate packet

Fig. 9 shows the filtering table and routing table at v_4. If a node v_4 receives the packets with m_2 or m_3, v_4 checks its filtering table to check whether it has the m_2 or m_3 in source s's entry. Once it has, it checks its routing table to send the packets to the next node. Then, it sends the packets to next node v_5 with the replaced marking value m_4.

We allow packets to travel through every possible path with a valid marking value from source to destination in order to prevent the dropping of legitimate packets, whether the path is asymmetric or not. Therefore, legitimate packets will not be dropped because packets with one of the possible marking values are passed through the routing path to the next nodes.

3.5 Partial Deployment with Asymmetric Paths

BASE has a salient feature in that it works in partial deployment, and substantially increases in power for larger rates of adoption. Also, any individual deployer receives an additional reward from the more powerful BASE network, simply by deploying to their networks. This unique characteristic comes from the incremental deployability of BASE. This is excellent motivation to adopt this mechanism, and thus BASE can be deployed to the current Internet.

We now explain how BASE works for partially-deployed environments. Assume k filters among n nodes such that $w_1, w_2, .., w_k \in \mathcal{P}(s,t)$ for $0 \leq k \leq n$. Then, any filter in the path, i.e., $w_i \in \mathcal{P}(s,t)$, can communicate with the next filter w_{i+1}, for $0 < i < k-1$, the same way as in the fully-deployed case. Non-BASE BGP speakers just relay the marking information because it is stored as optional transitive attributes.

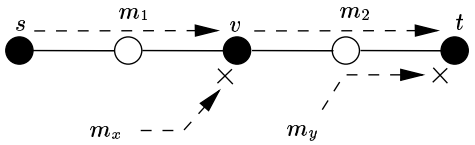


Figure 10: Partial deployment tunneling effect

Fig. 10 shows how BASE works in a partially deployed network. Any filter node can communicate with other filters across non-filter nodes. Also, each filter node in the partial-deployment can mark and filter spoofed packets. Furthermore, even though an attacker may succeed in injecting spoofed packets into the normal flow through a non-filter node, these packets will be distinguishable at the next filter. An example is filtering packets with m_y at the node t, as shown in Fig. 10. We call this the tunneling effect in which surrounding filter nodes protect non-filter nodes.

The proposed BASE scheme is simple but powerful for protecting against spoofing attacks. Nevertheless, there are certain cases that BASE cannot deal with when using one mark for each row of the filtering table. Fig. 11 shows one case of partial deployment in asymmetric routing paths. Since the BGP path is different from the routing path, the

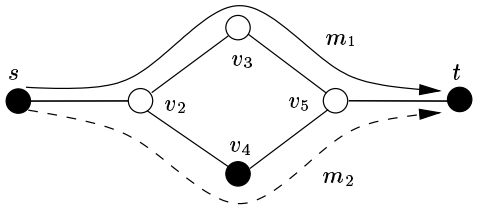


Figure 11: Legitimate packet dropping case on asymmetric routing paths

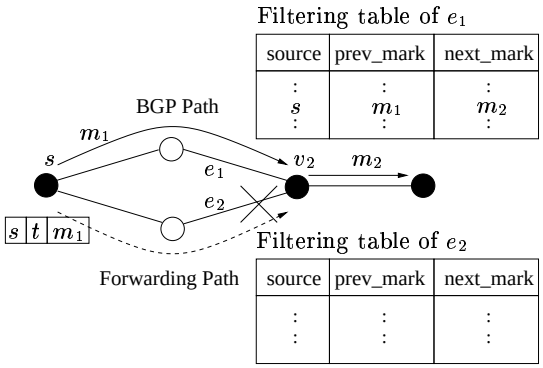


Figure 12: Asymmetric routing paths when using one mark in each record of filtering table

recorded mark, m_1, is different from the arriving mark m_2, which is an unregistered but legitimate mark. This will create a blackhole effect for a portion of legitimate traffic toward node t. To eliminate the black hole effect, we can store legitimate multiple marks in each record of a node's filtering table.

In order to mark and filter spoofed packets, we use a filtering table for each BASE filter node. If we store one mark in each record of a filtering table, each BASE filter has a filtering table F_e for an interface e. However, there is the issue of false-positive under asymmetric environments in the case of using one mark in each record of the filtering table. To eliminate the issue of false-positive under asymmetric environments, we can store multiple marking values in each record of a filtering table for each filter.

Fig. 12 and Fig. 13 are respectively the cases that one marking value and multiple marking values in each record of a filtering table. In case of one marking value as you can see in Fig. 12, there is a filtering table for each interface. When the BGP path and forwarding path are different, an interface e_1 in BGP path has a filtering table and also has a marking value m_1 mapped into source's network address s. However, once packets are forwarded through interface e_2, these packets would be dropped in spite of their legitimacy, because a filtering table of e_2 doesn't have the source's marking value.

However, in case of multiple marking values as you can see in Fig. 13, there is a filtering table for each node. Even though the BGP path and the forwarding path are different, packets would not be dropped. By storing multiple marking values, we can stop dropping legitimate packets. Also, we can get smaller space complexity and time complexity. Each BASE filter node has fewer tables to store the same data when compared with the use of one marking value. The number of marking values in each record is at most the maximum of the node's degree. Also, filtering table lookup time is similar to routing table lookup time. The fewer the entries in each BASE filtering table, the smaller the filtering table lookup time is. Therefore, there is no more overhead in terms of space and time complexity when using multiple marks.

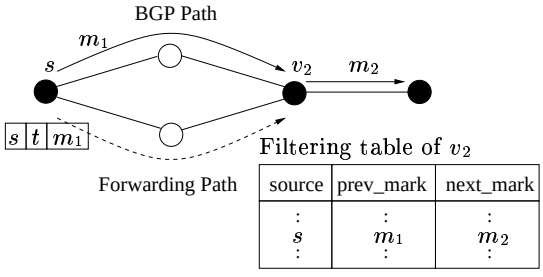


Figure 13: Asymmetric routing paths when using multiple marks in each record of the filtering table

4. SECURITY ANALYSIS: WHICH ATTACKS DID WE STOP?

4.1 IP Spoofing Attacks

The first measure of the effectiveness of an anti-spoofing mechanism is the proportion of spoofed packets that are dropped before arriving at a victim's location. To compare the filtering ability, let us consider a network with partial deployment which is shown in Fig. 14. In the case of ingress filtering [8], an attacker can mount an attack with spoofed packet (s, t) at 5 locations $(v_1, .., v_5)$. In the case of DPF [15], only 3 locations $(v_1, .., v_3)$ are available to an attacker for mounting a spoofing attack. In the case of BASE, no location can spoof a packet from s to t; only s can send a packet (s, t) to a host t. Even at v_2 and v_3, an attacker cannot send spoofed packets with the source address of s since the valid mark coming from s can be verified at t.

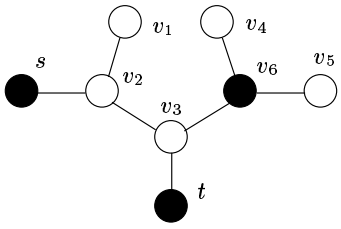


Figure 14: A partially-deployed network

Enhanced protection power of BASE comes from the use of packet marking to allow non-adjacent BASE-enabled ASs to verify the validity of the source address of traveling packets. This property enables BASE to be more effective than other schemes when partially deployed.

4.2 Marking Field Spoofing

Fake marks inscribed in a packet before being sent by an attacker greatly reduce the effectiveness of packet marking [15]. Pi [29] suffers from marking field spoofing since an

attacker can forge the marking field with the knowledge of the addresses of routers in an attacking path.

However, the BASE mechanism prevents an attacker from predicting the marking value, since a one way hash chain is used for computing marking values. Such a cryptographic marking mechanism renders a filter's marking values unpredictable by an attacker.

Also, valid marking values are not visible to an attacker, since the marking values, even in the process of marking and filtering by BASE, flow towards the victim. Thus, an attacker cannot gather valid marks from the attacking locations, making BASE resilient to replay attacks.

Ingress filtering and DPF do not use packet marking, so they have no threat of marking field spoofing. Instead of packet marking, they are location-aware filtering schemes. BASE takes advantages of location-awareness by means of tamper-resistant packet marking. Thus, BASE enhances anti-spoofing capability in addition to being insusceptible to marking field spoofing.

4.3 Compromising BASE Routers

We have shown that BASE has a strong ability to protect end-hosts from spoofed packets. Nonetheless, we need to consider other weaknesses in terms of security. First, malicious BASE speakers at compromised routers can be used to pass attack packets and drop legitimate packets. Clearly, a compromised router can control all packets it forwards. Another potential issue is an attacker who computes correct markings without access to the secret key, or who can compute the key independently. However, assuming a secure MAC function with an output of sufficient size, e.g. 16 bits, we do not need to consider these possibilities. UMAC is a fast message authentication code and we can determine the output of UMAC into 32-bit tag [13]. When UMAC produces 32-bit tag, the probability that an attacker could produce a correct tag for any message of its choosing is about $1/2^{30}$. Among the 32-bit tag, we can take 16 bits and those can be our marking value. Then, the strength of marking value would be $1/2^{15}$, the half of probability of UMAC. That means when an attacker choose a marking value randomly, 0.003% of attack packets can be passed to the target node. That would be negligible in attack situation.

5. EVALUATION

In order to measure the effectiveness of different anti-spoofing mechanisms, we first need to develop a means to produce an accurate model of today's Internet connectivity. We used the AS connectivity graph archived by NLANR from the Oregon Route Views project [18]. The AS graph used is the connectivity at April 2006, which consists of 22,000 nodes. We also used various 300-node subgraphs of the AS graph for the diversity of network topology and for faster simulation. To compare the filtering performance, we simulated on 4 mechanisms, which are ingress filtering, RPF, DPF and BASE. We selected filter nodes according to two filter placement policies, random filter placement and priority filter placement. Random filter placement means that filter nodes are chosen randomly and priority filter placement means that filter nodes are chosen according to the priority. A node that has many connections to other nodes gets higher priority than other nodes. Therefore, the highest degree node becomes filter node first. Also, we simulated on large and small asymmetric environments. In the Internet, many asymmetric environments exist, so we wanted to simulate on various situations. The subgraph, sub_large, has 46.7% asymmetry ratio and the subgraph, sub_small, has 12.2% asymmetry ratio.

The output of simulation is attack packet dropping ratio and legitimate packet dropping ratio as the number of filter nodes increases. The horizontal axis of graphs means % of deployed routers. The vertical axis of graphs means % of dropped attack packets or dropped legitimate packets. To acquire more exact results, we repeated 10 times for each simulation and took an average as a result.

Fig. 15 shows the dropping ratio of packets in AS_graph and subgraph. These two results show similar pattern in dropping packets. By using subgraph, we can make many map circumstances in selecting 300 subnode such as the graphs that have different asymmetric ratio. Also, we can save time for simulating the mechanisms. Therefore, we used the subgraph in the rest of the simulation.

Fig. 16 shows the filtering performance for dropping attack packets using random filter placement and priority filter placement in a large asymmetry path(46.7%). To compare the filtering performance, we used 4 mechanisms, which are ingress filtering, RPF, DPF, and BASE. DPF and BASE are more powerful than others in dropping attack packets. When using priority filter placement, the filtering performance increases especially in DPF and BASE. Because the node that has high priority could be a transit AS, the filtering performance of priority filter placement is higher than that of random filter placement. As we can see, even though about 30% transit ASs deploy the mechanism, we can filter almost all of attack packets.

The following simulation is to measure the filtering performance in terms of target's benefit. If the target is a filter node, it can guarantee the safety of itself in that it cannot receive attack packets. Fig. 17 shows the difference of filtering performance in DPF and BASE depending on whether the target node is a filter node or not. When the target node is a filter node or a non-filter node under partial deployment in DPF, the difference of attack packet dropping ratio is small. However, the difference is larger in BASE than DPF. That is, The BASE mechanism gives much larger benefit to early adopters than DPF.

Fig. 18 shows the difference of filtering performance in DPF and BASE when the node that a spoofed IP address belongs to is a filter node or not. If the node that a spoofed IP address belongs to is a filter node, it prevents the attack from happening by using IP address of itself. If packets originated from that node, the node would make and transfer its own marking value on the packets. The valid mark inscribed in a packet coming from the node can be verified at the target node. When the attack happened in the other node, the node that a spoofed IP address belongs to can provide the proof of innocence as showing that the packets didn't originate from the node by using the marking value of the node. Also, if the node that a spoofed IP address belongs to becomes filter node, target's attack packet dropping ratio would be increased. Therefore, it provides incremental benefit.

Fig. 19 shows the result of simulation in large and small asymmetric environments. There is not much difference in large asymmetric paths and small asymmetric paths in ingress filtering, DPF and BASE, but RPF shows much difference according to asymmetry ratio. Because RPF uses

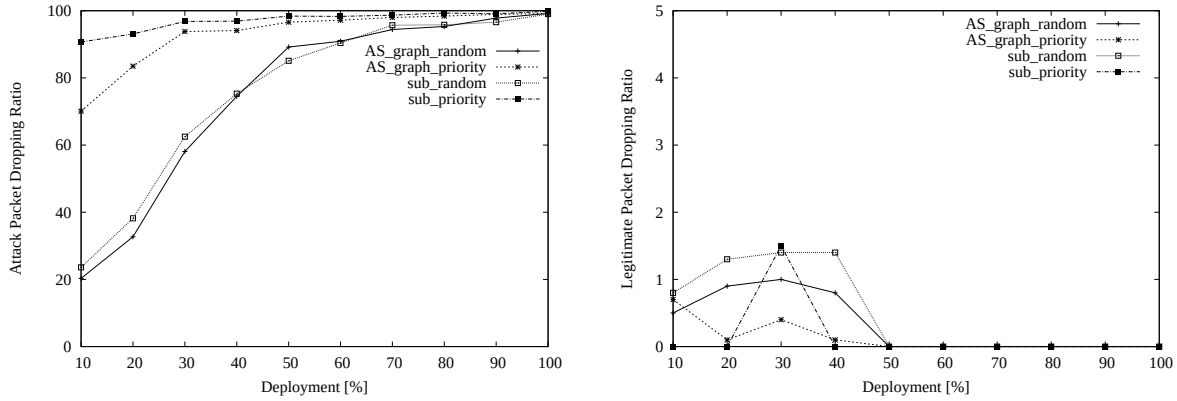


Figure 15: Dropping ratio of packets using BASE filter in random and priority filter placement in a AS_graph and subgraph: (left) dropping ratio of attack packets (right) dropping ratio of legitimate packets

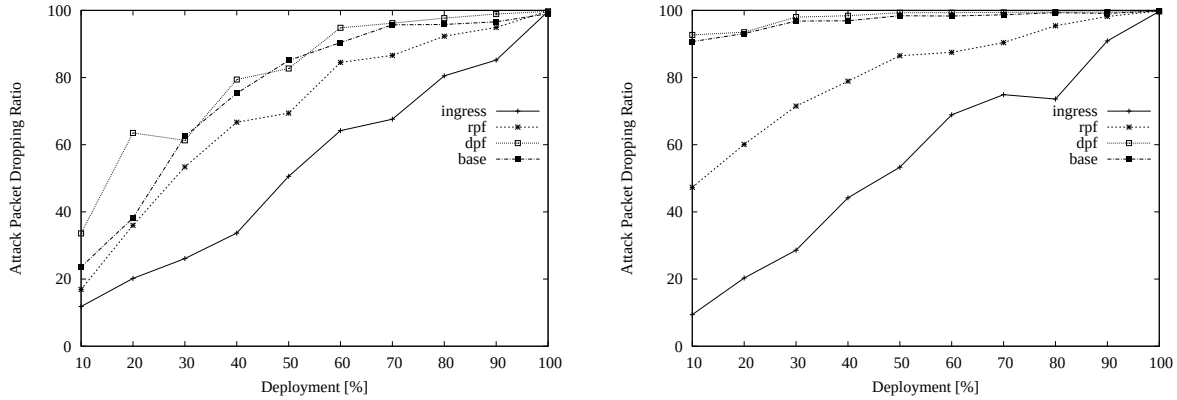


Figure 16: Dropping ratio of attack packets using different anti-spoofing mechanisms in a subgraph: (left) random filter placement (right) priority filter placement

IP routing tables for dropping attack packets, it is effective in symmetric routing environment.

Fig. 20 shows the filtering performance for dropping attack packets and legitimate packets in one and multiple mark environments. When BASE has one marking value in filtering table, filtering performance for dropping attack packets is a little bit better than in multiple mark environment. However, the number of dropped legitimate packets decreases when using multiple marking values as opposed to one.

From the experiments, we confirm that BASE shows good performance in dropping attack packets and not dropping legitimate packets. Especially, when the transit ASs deploy BASE, the deployers get much higher filtering effect than non-deployers as shown in Fig. 16. When the target node is a filter node, attack packet dropping ratio is much higher than when the target node is not a filter node. Therefore, early adopter gets benefit as shown in Fig. 17. As shown in Fig. 18, if the node that a spoofed IP address belongs to and target node are filter nodes, BASE can achieve almost perfect attack packet dropping performance. As shown in Fig. 19, BASE doesn't have much different filtering performance in large and small asymmetric environments because BASE uses multiple marking policies to prevent legitimate packet dropping. When BASE uses multiple marks, the filtering performance is a little bit worse than when BASE uses one mark. However, BASE can considerably prevent dropping legitimate packets by using multiple marks as shown in Fig. 20.

In terms of viable protocol as discussed in Section 1, the BASE mechanism satisfies initial benefit, incremental benefit, and partial deployment. From the target's point of view, it can drop attack packets. Therefore, BASE gives direct benefits to early adopters. As the filter nodes increase, the filtering performance increases. Therefore, it also satisfies incremental benefit to subsequent adopters. Finally, almost all the attack packets would be dropped even though only about 30% of transit ASs are deployed. Therefore, BASE is effective when partially deployed.

6. DISCUSSION

6.1 Adopter's Benefit

Ingress filtering and DPF are more powerful when deployed near the attacking location, but less effective near the victim. Also, they provide no specific benefit to the adopters, which has been a barrier to deployment. Only Pi gives an obvious benefit to a victim for defending against spoofing attacks. Nevertheless, Pi still has significant weaknesses—the full benefit of Pi occurs only after huge deploy-

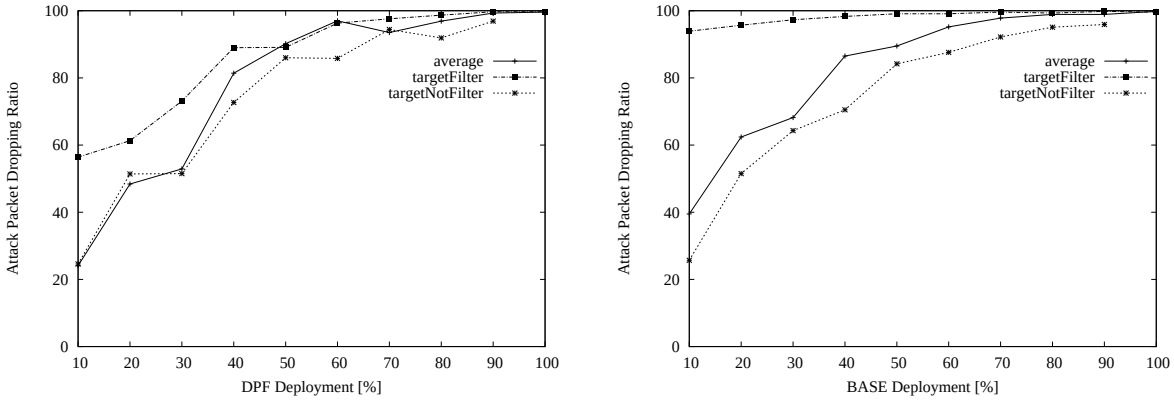


Figure 17: Dropping ratio of attack packets according to target is filter node or not in a subgraph: (left) DPF filter (right) BASE filter

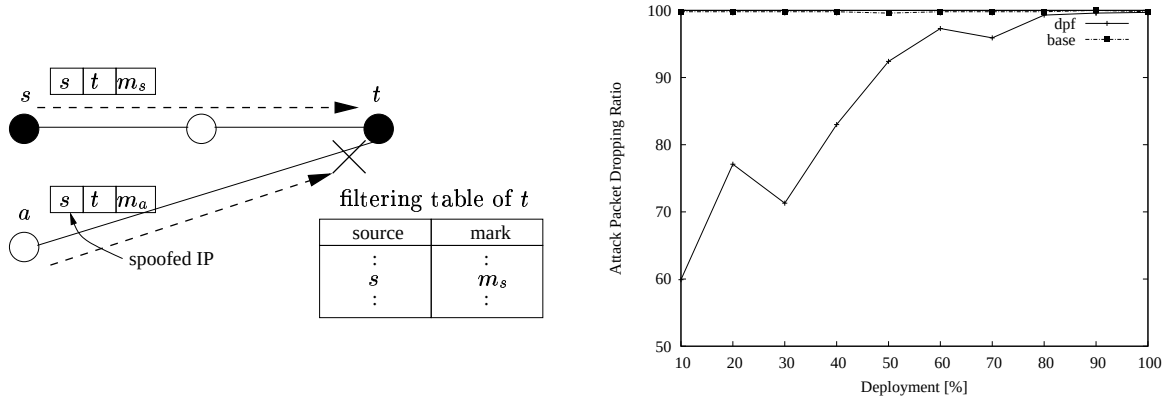


Figure 18: Dropping ratio of attack packets when both the node that a spoofed IP address belongs to and target node are filter nodes

ment. Therefore, Pi cannot be an immediate solution for a current victim of spoofing attacks. BASE remains the only current solution which gives direct benefits to an adopter and can be a viable solution for defending against spoofing attacks.

6.2 Cost Effectiveness

We need to analyze anti-spoofing schemes in terms of processing cost, messaging cost and filtering cost.

In the distribution phase, BASE requires small processing cost to create marking values. The marking values should be computed once before they are distributed using BGP update messages. This process does not happen frequently unless the BGP path has been changed. Also, BASE does not incur messaging costs because BASE uses the existing protocol, BGP, to distribute the marking values.

In invocation and revocation phase, there is not much overhead for messaging the marking values because only a single BGP update for each start or stop signal is required. This is the minimum cost for saving a victim from overwhelming garbage traffic. Also, BASE works only when the victim wants to invoke it. Therefore, it will not cause much overhead.

In the packet marking and filtering phase, filtering table lookup should be done for each node to filter spoofed packets. Filtering table lookup time would be similar to routing table lookup time, so that is not much overhead.

6.3 Implementation Issues

To apply any IP spoofing mechanism to current Internet environments, it should satisfy the three deployment properties addressed in Section 1. Also, it should be a practical mechanism which does not need much overhead and cost.

When compared with BASE, DPF is difficult to apply to current Internet because it doesn't give direct benefits to the adopters. Also, each node should have routing information of all nodes, so it needs much space for each node. Once the BGP information is updated, the routing information in each node should be updated promptly. Once some of the routing information is updated, it should inform its updated routing information to all nodes. It requires in general $O(n)$ messaging overhead for each node. On the other hand, BASE uses BGP update messages only once to distribute marking values. Therefore, it can be denoted by $O(1)$.

It is difficult not only to store the routing information of all nodes in each node and but also to change the updated routing information in all node whenever even one of the routing paths is changed. Therefore, DPF is not good for implementing because of space cost and messaging overhead.

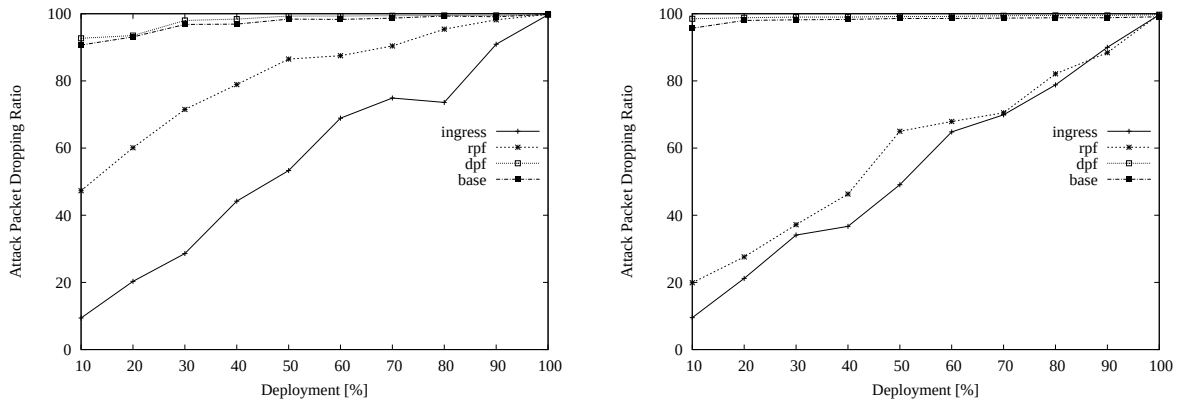


Figure 19: Dropping ratio of attack packets using different anti-spoofing mechanisms in sub_large and sub_small: (left) dropping ratio of attack packets in sub_large (right) dropping ratio of attack packets in sub_small

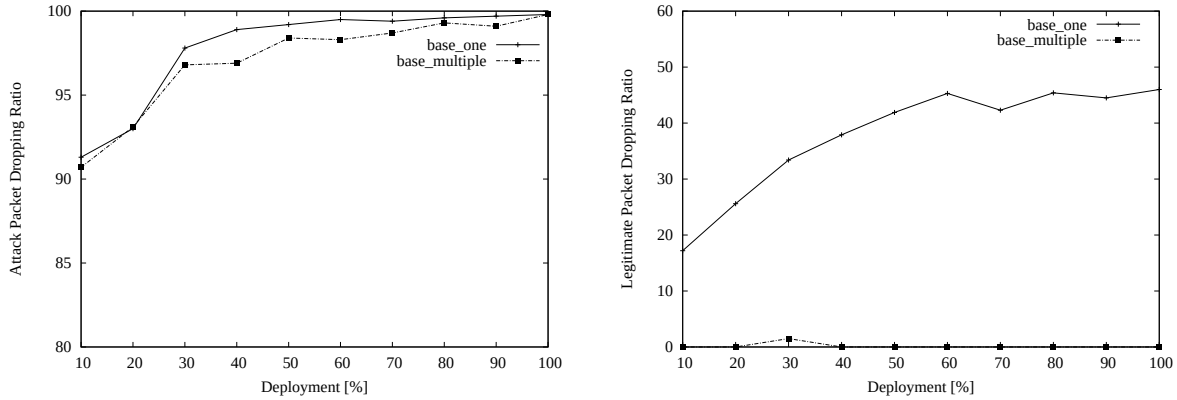


Figure 20: Dropping ratio of packets according to the number of marks: (left) dropping ratio of attack packets (right) dropping ratio of legitimate packets

However, BASE has no more additional cost because it only distributes marking values on the existing routing protocol and doesn't need any further messaging overhead.

6.4 BASE Limitations

While BGP distributes routing information between ASs, BASE uses BGP to distribute marking values. There are a few limitations coming from the use of BGP as a broadcasting mechanism. Routing asymmetry in a certain BASE configuration and routing policies that do not forward update messages provide opportunities to drop legitimate traffic. However, legitimate packet dropping according to AS's routing policy and routing asymmetry occurs only during an attack because BASE is turned on only during an attack when the victim wants to filter attack packets. Without any defense mechanism, almost all the legitimate packets destined to the victim would be dropped during an attack. Therefore, legitimate packet dropping during an attack is not a big problem.

Policy issues remain more problematic with smaller ISPs or stub ASs as opposed to transit ASs of major tier-1 or tier-2 ISPs. AS-level asymmetry will not happen between neighboring ASs, but between ASs in a long path. Thus, asymmetry depends on the distance between ASs, implying that BASE can protect attacks from near ASs but may provide misleading information from remote ASs. Nonetheless, attacks from distant ASs have more opportunities to encounter BASE filters than attacks from near ones.

We could design our own protocol instead of utilizing BGP to spread BASE information. A dedicated protocol would work in a manner similar to the BGP-enabled BASE scheme, while retaining favorable properties for incremental deployment. This would resolve asymmetry issues because we can store every possible mark, eliminating the issue of false-positive. However, distributing marking values remains a significant problem without direct integration with the routing protocol because of the difficulty in maintaining up-to-date marking information.

7. CONCLUSIONS

The BASE mechanism is suggested for fulfilling the incremental deployment properties which are essential for adoption in current Internet environments. Along with distributed filtering, cryptographic packet marking, and on-demand filtering for the destination addresses of the victim's network, the protective power is enhanced as BASE filters are distributed gradually. The BASE mechanism is superior to existing solutions because a) it is more effective than others

when partially deployed, and b) it offers a greater immediate benefit to deployers.

BASE is a likely candidate for overcoming the barriers to wide-spread adoption that have prevented other mechanisms from taking hold. This is because of its ability to prevent spoofing of a large percentage of the IP address space when it has only been deployed to a comparatively small percentage of that space. However, some challenges still must be surmounted. AS's routing policies may prevent the BGP update messages from propagating to a neighboring AS, and also malicious BASE speakers at compromised routers can pass attack packets and drop legitimate packets. Additionally, the effect of real-world routing policies on distribution of BASE control data needs further examination. Despite this, BASE is a promising new direction in IP spoofing prevention.

8. ACKNOWLEDGMENTS

We are grateful to Damon Smith and Jihoon Son for conducting the simulations. This work was supported in part by the ITRC program of the Korea Ministry of Information & Communications, and the Basic Research Program of the Korea Science & Engineering Foundation. This research was supported in part by CyLab at Carnegie Mellon under grant DAAD19-02-1-0389 from the Army Research Office.

9. REFERENCES

[1] H. Aljifri, M. Smets, and A. P. Pons. IP traceback using header compression. *Computers & Security*, Feb. 2003.

[2] F. Baker and P. Savola. Ingress filtering for multihomed networks. RFC 3704, Mar. 2004.

[3] S. Bellovin, M. Leech, and T. Taylor. The ICMP traceback message. Internet-Draft, draft-ietf-itrace-01.txt, Oct. 2001. Work in progress, available at `ftp://ftp.ietf.org/internet-drafts/draft-ietf-itrace-01.txt`.

[4] D. Bruschi, L. Cavallaro, and E. Rosti. Less harm, less worry or how to to improve network security by bounding system offensiveness. In *Annual Computer Security Applications Conference*, Dec. 2000.

[5] CERT. TCP SYN flooding and IP spoofing attacks. Advisory CA-96.21, September 1996.

[6] Cisco. Strategies to protect against distributed denial of service (DDoS) attacks. Updated News Flash, Apr. 2003.

[7] M. Collins and M. K. Reiter. An empirical analysis of target-resident DoS filters. In *Proceedings of IEEE Symposium on Security and Privacy*, May 2004.

[8] P. Ferguson and D. Senie. Network ingress filtering: Defeating denial of service attacks which employ IP source address spoofing. RFC 2827, May 2000.

[9] D. Fisher. Internet survives massive ddos attack. `http://www.eweek.com/article2/0,1759,1498701,00.asp`, Oct. 2002.

[10] O. Goldreich, S. Goldwasser, and S. Micali. How to construct random functions. *Journal of ACM*, Oct. 1986.

[11] Y. He, M. Faloutsos, S. Krishnamurthy, and B. Huffaker. On Routing Asymmetry in the Internet. In *Proceedings of IEEE Globecom*, 2005.

[12] C. Jin, H. Wang, and K. G. Shin. Hop-count filtering: An effective defense against spoofed DDoS traffic. In *Proceedings of ACM Conference on Computer and Communications Security*, Oct. 2003.

[13] T. Krovetz. Umac: Message Authentication Code using Universal Hashing. RFC 4418, Mar. 2006.

[14] C. Labovitz, A. Ahuja, A. Abose, and F. Jahanian. Delayed internet routing convergence. In *Proceedings of ACM SIGCOMM*, Aug. 2000.

[15] H. Lee and K. Park. On the effectiveness of probabilistic packet marking for IP traceback under denial of service attack. In *Proceedings of IEEE Infocomm*, Apr. 2001.

[16] J. Li, J. Mirkovic, M. Wang, P. Reiher, and L. Zhang. Save: Source address validity enforcement protocol. In *Proceedings of IEEE INFOCOM*, June 2002.

[17] B. Manning. Registering New BGP Attribute Types. RFC 2042, Jan. 1997.

[18] D. Meyer. University of Oregon Route Views archive project. `http://archive.routeviews.org`, 2005.

[19] G. A. Moore. *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers.* HarperCollins Publishers, 1995.

[20] K. Park and H. Lee. On the effectiveness of route-based packet filtering for distributed DoS attack prevention in power-law internets. In *Proceedings of ACM SIGCOMM*, Aug. 2001.

[21] V. Paxson. End-to-end routing behavior in the internet. In *Proceedings of ACM SIGCOMM*, 1996.

[22] J. Postel. Internet protocol. RFC 791, Sept. 1981.

[23] Y. Rekhter, T. Li, and S. Hares. A Border Gateway Protocol 4 (BGP-4). RFC 4271, Jan. 2006.

[24] S. Savage, D. Wetherall, A. Karlin, and T. Anderson. Practical network support for IP traceback. In *Proceedings of ACM SIGCOMM*, August 2000.

[25] D. Song and A. Perrig. Advanced and authenticated marking schemes for IP traceback. In *Proceedings of IEEE Infocomm*, April 2001.

[26] N. Spring, R. Mahajan, and T. Anderson. Quantifying the causes of path inflation. In *Proceedings of ACM SIGCOMM*, Aug. 2003.

[27] R. Teixeira, K. Marzullo, S. Savage, and G. M. Voelker. Characterizing and measuring path diversity of internet topologies. In *Proceedings of ACM SIGMETRICS*, June 2003.

[28] D. Wetherall, J. Guttag, and D. L. Tennenhouse. ANTS: A toolkit for building and dynamically deploying network protocols. In *IEEE OPENARCH'98*, Apr. 1998.

[29] A. Yaar, A. Perrig, and D. Song. Pi: A Path Identification mechanism to defend against DDoS attacks. In *Proceedings of IEEE Symposium on Security and Privacy*, May 2003.