Modeling Adoptability of Secure BGP Protocols*

Haowen Chan, Debabrata Dash, Adrian Perrig, and Hui Zhang
Carnegie Mellon University
{haowenchan, ddash, perrig, hzhang}@cmu.edu

ABSTRACT

Despite the existence of several secure BGP routing protocols, there has been little progress to date on actual adoption. Although feasibility for widespread adoption remains the greatest hurdle for BGP security, there has been little quantitative research into what properties contribute the most to the adoptability of a security scheme. In this paper, we provide a model for assessing the adoptability of a secure BGP routing protocol. We perform this evaluation by simulating incentives compatible adoption decisions of ISPs on the Internet under a variety of assumptions. Our results include: (a) the existence of a sharp threshold, where, if the cost of adoption is below the threshold, complete adoption takes place, while almost no adoption takes place above the threshold; (b) under a strong attacker model, adding a single hop of path authentication to origin authentication yields similar adoptability characteristics as a full path security scheme; (c) under a weaker attacker model, adding full path authentication (e.g., via S-BGP [9]) significantly improves the adoptability of BGP security over weaker path security schemes such as soBGP [16]. These results provide insight into the development of more adoptable secure BGP protocols and demonstrate the importance of studying adoptability of protocols.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: General

General Terms

Security, Design

Keywords

Adoptability, Adoption dynamics, incentives-compatibility

*This research was supported in part by grants CNS-0433540 and ANI-0331653 from the National Science Foundation, and by a gift from Cisco. The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of CMU, Cisco, NSF, or the U.S. Government.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
SIGCOMM'06, September 11–15, 2006, Pisa, Italy.
Copyright 2006 ACM 1-59593-308-5/06/0009 ...$5.00.

1. INTRODUCTION

The security problems of BGP are well known [12]. S-BGP was the first proposal to address BGP security [9], and it has been followed by numerous alternative proposals including soBGP [16], IRV [3], SPV [6], Listen and Whisper [14], and psBGP [15]. Despite the availability of this wide range of innovative technologies for BGP security, none of these protocols have been adopted by ISPs. The reasons for this lack of adoption are complex and involve many unmeasurable socio-political and economic factors.

From a more general viewpoint, the lack of adoption of secure BGP protocols is a specific instance of the problem of predicting interdomain protocol adoption, where the different parties considering protocol adoption do not necessarily have the same agenda. There remains to date no quantitative analytical framework that can assist computer-networking researchers in assessing the potential for adoption of new protocols. In this paper, we present a new approach to the problem of analyzing interdomain protocol adoption: instead of focusing on the phenomenological and strategic aspects favored in economics and the social sciences, or on the standard metrics typically favored by protocol researchers such as communications and memory overhead, we propose a new metric for protocol design, *adoptability*. Intuitively, adoptability measures the strength of a protocol's properties in driving the adoption process. Under this definition, a protocol with stronger properties will provide greater benefits to its adopters and thus have greater adoptability. We propose a simulation-based model for quantitatively deriving the adoptability of a protocol in any proposed context by iteratively considering the decision process of each potential adopter. Using this methodology, we studied the problem of protocol adoption for BGP-security protocols under various assumptions.

Typically, attractiveness for adoption is not quantitatively studied in proposed Internet protocols. Most newly developed protocols claim at most *incremental deployability*, which means that the protocol can be gradually adopted over a period of time. During this adoption period, adopters of the new protocol have full compatibility with non-adopters running a legacy protocol, while enjoying some level of benefit even though adoption is not universal.

Although incremental deployability helps in adoption, it is neither a necessary nor sufficient condition. This is because incremental deployability is an inherent property of the *protocol*, while adoptability must necessarily involve the *context* in which the protocol is deployed (e.g., the Internet). Simply observing that a protocol possesses the property of incremental deployability does not give any indication about the likelihood of widespread deployment in any context; in fact, it does not even imply the existence of a set of feasible scenarios in which widespread adoption could take place. Similarly, *not* possessing incremental deployability does not imply that adoption is impossible—a clearly superior technology

with a low transition cost might easily gain a sufficiently large base of early adopters to ensure global adoption regardless of whether it is incrementally deployable.

As mentioned, the adoptability of a protocol must be measured with respect to some assumed *deployment context*—this includes the group of potential adopters and their inter-relationships, and various assumptions about their decision-making strategies. We make the assumption that each potential adopter is rational and selfishly motivated, and model the *greedy incentives-compatible adoption dynamics* of a range of secure BGP protocols under various assumptions. Under the greedy bounded-rationality assumption, an Autonomous System (AS) adopts the new protocol if and only if the *immediate security benefits* of adopting the protocol is greater than some *switching threshold*, which represents the cost of adoption. Typically, the more ASes that currently support a protocol, the greater is the benefit enjoyed by a new adopter—this is the well-known *network effect* in economics, a specific example of which is Metcalfe's Law. Hence, the adoption process across the Internet is dynamic—as more ASes decide to adopt, their decisions will drive new adoptions by other ASes which had formerly found adoption unappealing. An *incentives-compatible adoption scenario* is a scenario in which, starting with a pre-set group of initial adopters, we can iterate over the set of ASes and continually find ASes for which adoption is greedily rational, until either there are no new adopters of the protocol, or all the ASes in the Internet have adopted the scheme. By simulating these adoption scenarios over a range of switching thresholds, we can chart the space of switching thresholds for which the incentives-compatible adoption process will yield widespread adoption of a given protocol. The larger the range of switching thresholds that a protocol can support, the greater its *adoptability*. Such an analysis yields a quantitative evaluation of the practical attractiveness of a given set of security properties in terms of how likely it is that these properties might drive eventual full adoption of the protocol.

Using our model, we collect the adoptability results for each of five classes of known security protocols. We observe that under a standard strong attacker model, any scheme that provides weak partial security, by implementing origin authentication with first hop authentication in the AS_PATH, already has closely comparable adoptability as a scheme with full path security (e.g., S-BGP). This implies the surprising result that, under this attacker model, the incremental gain to adoptability for increasingly strong security properties is very small or nonexistent. In contrast, under a more realistic weak attacker model, the full-path security property has up to ten times the adoptability of a scheme compared with partial path security (e.g., soBGP) or simple origin authentication. This implies that, for its trade-off of weaker security properties to be attractive, soBGP needs to offer switching costs that are at least ten times lower than S-BGP.

Modeling and measurement of adoptability are of great importance both to researchers seeking to create more viable protocols, and to policy-makers seeking to select the best new technologies to promote. In the course of formulating our approach to the problem, we made many simplifying assumptions. Hence our results cannot be used as direct predictions of the likelihood and/or cost of adoption. However, the value of our work lies in the formulation of the problem model, and the methodology for calculating a new metric which can be used to compare the relative strengths of different protocols in driving adoptability. This is an important problem that has thus far not been the subject of intensive research.

2. BACKGROUND AND RELATED WORK

Many protocols for secure BGP have been proposed. The main security problems of BGP are outlined in an IETF draft by Murphy [12]. S-BGP was proposed by Kent et al. [8, 9]. It approached BGP security by securing the complete Update message by use of *attestations*, which are essentially signatures within the context of a public key infrastructure (PKI). Origin ownership is authenticated through a PKI, while AS_PATH attributes are similarly signed by each contributing AS using *route attestations*. When an AS receives a BGP advertisement, it appends the next hop (i.e., the next AS to which it will readvertise this prefix) to the AS_PATH and signs the new AS_PATH along with all previous route attestations. This provides assurance of the integrity and authenticity of the path.

White et al. [16] propose soBGP, where origin authentication is accomplished in an oligarchy PKI similar to that in S-BGP. Unlike S-BGP, soBGP does not use cryptographic mechanisms to secure the authenticity of the entire AS_PATH. Instead, AS_PATHs are verified against a database of AS-to-AS routing relationships. Any path consisting of edges that are not present in the database is considered malformed and is rejected. For example, if a path contains two consecutive local ASes, neither of which claims to have a relationship with the other, then it is detected as malformed by soBGP and is rejected. Kruegel et al. augment this approach with a topological anomaly detection heuristic [10].

Goodell et al. propose IRV [3], which proposes maintaining dedicated verification servers to verify the authenticity of BGP advertisements. Yu et al. propose a reputation-based scheme to evaluate authenticity of BGP advertisements[18]. Aiello et al. also address the problem of origin authentication through the use of Origin Authentication Tags (OATs) [1]. Zhao et al. propose techniques for detecting invalid multiple origin AS (MOAS) conflicts in the Internet [22]. Subramanian et al. propose Listen and Whisper [14], which protects AS_PATH integrity while performing anomaly detection by observing traffic flow. Hu et al. propose SPV [6] which addresses AS_PATH authentication through the use of one-time signatures and symmetric cryptographic primitives, limiting the use of expensive public-key cryptography. SPV possesses the property that secure ASes further down the AS_PATH can act for any insecure ASes earlier in the path by performing signatures on their behalf. Wan et al. propose psBGP [15], which provides equivalent path security benefits to S-BGP along with slightly less secure but more efficient prefix ownership authentication. Zhao et al. propose improved cryptographic primitives to make S-BGP efficient [21].

To our knowledge there has not been any work on studying adoptability as a metric for the usefulness of specific Internet protocols, in order to guide design and policy decisions. However the general process of adoption of new technologies is well studied in social networks and economics [7]. He et al. have proposed a framework for measuring incremental deployment properties of router-assisted services [4], however they did not study the adoptability properties, which are distinct from incremental deployability as explained in Section 1.

3. A TAXONOMY OF PROPERTIES

We classify secure BGP protocols into the following categories. **Origin Authentication (OA)** refers to the ability to authenticate that a given AS is the legal owner of a prefix that it originates. An origin authentication protocol ensures that if some AS *A* (which speaks the protocol) originates a prefix, any other speaker AS can verify that *A* is the legitimate owner of the prefix. While many BGP security schemes contain an OA component, OA is considered a relatively weak property such that no pure OA scheme (e.g., OATs [1] or MOAS detection [22]) is meant as a self-contained solution for BGP security. In practice, the owner of a prefix could authorize a different organization to originate the prefix. Since this implies an explicit trust relationship between the owner and the originator, this does not change our analysis.

First-hop Authentication (OA+1) refers to a hypothetical origin authentication protocol where the originator of the prefix additionally encodes the identity of the first-hop AS on the path from itself (for example, by signing the identity of the next AS into the prefix ownership attestation). This ensures the integrity of first *two* ASes on the AS_PATH. We abbreviate this property as "OA+1" to indicate that it can be enabled by only a small additional step to origin authentication. As with OA, no actual secure BGP protocol implements just OA+1. An example of OA+1 could be a limited-functionality version of S-BGP protocol where only the originating AS signs any attestations (i.e., the address attestation and the first-hop route attestation only).

Routing Topology Path Verification (RTPV) is the path security model employed in soBGP [16]. Under RTPV, any advertised AS_PATH must conform to some authenticated map of the AS-level routing structure of the Internet. IRV [3], the reputation mechanism of Yu et al. [18], and the heuristics of Kruegel et al. [10] are also examples of protocols which achieve this property. In our analysis, we assume that each speaker AS only has a partial view of the Internet routing topology corresponding to the neighborhood of the other speaker ASes, since the topological information of insecure nonspeaker ASes can be spoofed. This excludes the protocols from using "well-known" but unauthenticated information in performing path verification—for example, soBGP cannot use the well-known fact that all Tier-1 ASes have peering agreements with each other unless all the Tier-1 ASes are soBGP speakers. Such an assumption is unrealistic, but it is a necessary simplification. Without this assumption, it would be necessary to label all $n \times n$ potential edges in the AS graph with some assumed function of confidence, which would make our analysis intractable.

Path Authentication (PA) is the path security model employed in S-BGP [16]. psBGP [15] provides a more efficient method of origin authentication but secures its path information in a manner similar to S-BGP, and hence possesses identical PA properties. In S-BGP, every S-BGP speaker AS on the AS_PATH is involved in signing the path and providing assurance to its complete authenticity up to the first non-speaker AS (e.g., AS_i) in the AS_PATH. Since AS_i is not an S-BGP speaker, it will not have the requisite keypairs to perform any S-BGP signatures. In particular, it will not be able to sign the next AS (AS_{i+1}) into the route attestation chain. This yields a gap in the chain of signatures that an attacker can exploit by stripping away all cryptographic information for any ASes after AS_i, giving it the ability to arbitrarily forge the remainder of the AS_PATH after AS_i.

Retroactive Path Integrity (RPI) is the path security model of SPV [6]. RPI addresses the drawbacks of PA by allowing subsequent secure ASes to perform digital signatures on behalf of earlier non-deploying ASes so that the chain of integrity is not broken. This ensures that the integrity of the path is protected up to the latest secure AS on the path (rather than the first non-deploying AS on the path under PA). This model possesses stronger path integrity (the attacker is less free to remove information from the path) but may lose the property of AS authentication (for example, in SPV, an attacker is able to add arbitrary deploying ASes to the path, which it was not able to do in the PA model). Also, under this model, ASes may be able to perform signatures on behalf of other ASes using cryptographic information that is revealed from one AS to another. Hence RPI's resistance to path forgery may depend strongly on the attacker's eavesdropping ability.

We do not consider path-expansion attacks in this paper because we make the simplifying assumption that shorter paths are always preferred regardless of the ASes on the path (we explain why this assumption is necessary and reasonable in Section 4.5). Recall that in RTPV, we only consider a view of the Internet restricted to the neighborhood of the deployers of the security scheme. Hence, since each AS in PA is authenticating to an AS-to-AS relationship between itself and its predecessor and successor ASes, it is clear that the set of acceptable AS_PATHs in PA is a subset of the set of paths acceptable under RTPV. Hence, we can say that, under the set of scenarios we are considering, PA is always at least as strong as RTPV. Clearly, RTPV is at least as strong as OA+1, which is at least as strong as OA. This order of properties under our model implies that each stronger property completely captures the functionality of all the weaker properties—any of our modeled attacks that succeeds against a stronger property will always succeed against the weaker properties. This ordering can be summarized as:

OA $\leq$ OA+1 $\leq$ RTPV $\leq$ PA

RPI is omitted from the ordering since, unlike the other schemes, its security properties vary depending on the attacker's eavesdropping capabilities. Based on the above classification, the goal of our study is to establish the relative quantitative contributions of each of the security properties to the adoptability of a scheme. For clarity, we pick a single well-known security scheme from each class to represent that class of security properties. Hence, we use soBGP to represent RTPV, S-BGP to represent PA, and SPV to represent RPI. Since every OA protocol is meant to be implemented alongside some kind of path security protocol, we do not use any existing OA scheme to represent the security class. Instead, we refer to the class directly as "OA". Likewise, there exists no security protocol that implements only OA+1 — hence, we refer to this class directly as "OA+1". To summarize, the five security classes that we investigate in this paper are denoted by OA, OA+1, soBGP, S-BGP, and SPV respectively.

4. SIMULATION MODEL

In this section we discuss the methodology, models, and assumptions made to develop a viable simulation environment. In later sections we provide sensitivity analysis on the parameters discussed here.

4.1 Simulation Methodology

Our methodology for measuring the adoptability of a given protocol aims to discover the range of possible adoption transition costs (or *switching thresholds*) for which *incentives-compatible deployment scenarios* exist. Recall from Section 1 that an AS adopts the protocol if and only if the immediate security benefit of adopting the protocol is greater than the switching threshold. Hence, the more adoptable a protocol is, the greater the range of switching thresholds for which full adoption eventually occurs.

We measure the adoptability of the protocol for various switching thresholds by simulating the dynamics of the adoption process using a model of the decision-making process of the ASes on the Internet. At the end of the simulation, we consider the final fraction of ASes in the Internet which are adopters of the protocol. If this final fraction is large, then we know that incentives-compatible deployment has succeeded for this particular scenario; if the final fraction is small, then the adoption process has stalled because it is not incentives-compatible for a majority of the ASes to adopt the new protocol.

We assume that an initial set S_0 of ASes have deployed the security protocol prior to the start of the simulation (i.e., at iteration 0). We call our S_0 set the set of *initial adopters*. We consider several possibilities that may account for a particular set of initial adopters. For example, governmental policy could dictate that all military ASes initiate deployment of a secure BGP protocol. Alternatively, large-scale Tier-1 ISPs could coordinate to become initial

adopters via a wide-ranging business agreement. Another possible scenario would be an academic partnership causing a set of university ASes to become initial adopters. We evaluate how the choice of various initial adopter sets affects a security protocol's adoptability. The reason for this is twofold: first, this allows us to check the sensitivity of our results to different initial conditions. Second, we hope that this study will help guide policy decisions on how to best initiate deployment.

After we have selected our initial set S_0, the simulation proceeds in iterations. For each iteration $i \geq 1$, we consider each AS that has not yet adopted the protocol, and we model its adoption decision process as greedily rational and selfishly motivated — hence it will become an adopter of the protocol in the next iteration if and only if the immediate *security benefits* of adopting the protocol is greater than the *switching threshold*, which represents the costs of transitioning to and supporting the new protocol. We explore modeling the security benefit in a variety of ways; we describe these in detail in Section 4.3. The switching threshold is an independent variable (expressed in the same units as the security benefit), which can be arbitrarily varied as a parameter of the simulation. We assume that the switching threshold is a constant value for each AS across the Internet, we justify this assumption in Section 4.4.

Using this method, for iteration i, we use the set of deployed (protocol-speaking) ASes in the previous iteration (S_{i-1}) to determine the set A_i of ASes that will adopt the security protocol in the current iteration i. We then add them to the set of protocol-speaking ASes, i.e., $S_i = S_{i-1} \cup A_i$. The simulation ends when no more ASes have been found to be new adopters of the secure routing protocol in an iteration or the whole of Internet has already adopted the protocol, i.e., $|A_i| = 0$.

4.2 Attacker Model

There are many actions an adversary can take in a partially-secure Internet. We focus on a specific attack and assume a single malicious AS which is attempting to divert legitimate routes towards itself. We chose this general attack since it is a necessary first step for other sophisticated attacks such as eavesdropping, selective packet dropping, and blackholing. This is the most direct form of attack for an adversary whose goal is to gain control of some set of flows on the Internet. Other attacker models which may be analyzed in our framework include multiple-adversary models in which malicious ASes may collude to share information or launch coordinated active attacks. We do not perform these analyses in this paper but hope that they will be the subject of future work.

We base our security analysis on two main attacker models: the *Strong Attacker Model* and the *Weak Attacker Model*. Both models make the assumption that a malicious AS cannot inject new announcements into a non-neighbor AS. This is because typical BGP routers only accept BGP sessions via direct physical links from a small set of neighboring routers, making it difficult for a malicious router to inject false information outside of its immediate AS neighborhood.

We vary the ability of a malicious AS to eavesdrop on BGP announcements from other ASes. In the *Strong Attacker Model*, a malicious AS can eavesdrop on BGP traffic between any two ASes on the Internet. This assumption is somewhat unrealistic since most inter-AS border routers communicate via direct physical links on which remote eavesdropping is impractical if neither AS has been compromised by the adversary. However, the practice in standard security analysis is to assume that all unencrypted communications are known to the attacker. The strong attacker model is thus based on this standard assumption. In the *Weak Attacker Model*, the malicious AS can only access BGP traffic sent directly to it, but cannot eavesdrop on BGP communications elsewhere in the Internet. We will discuss in Sections 5 and 7 how these different attacker models affect the security properties of secure BGP protocols.

4.3 Security Metric

We define the *security benefit* of each AS as being the net difference in its *security metric* between having deployed the protocol and not having deployed the protocol. Intuitively, the security metric for each AS is the expected probability that some uniformly randomly chosen bit passing through the AS cannot be diverted by a single malicious AS somewhere else in the Internet. The *security benefit* is thus the increase in this probability of resistance to diversion due to the AS deploying the secure BGP protocol.

We model the security metric as follows. Let the set of all ASes be V. Let a be the AS deciding on adopting a secure BGP scheme. We assume that a is concerned with all the traffic that passes through itself, that is, every AS-to-AS route that passes through a (or starts or ends at a) has an effect on its security metric. This assumption is motivated by the intuition that ASes are commercial entities which are paid to carry traffic; hence the ability to secure any given bit of traffic should improve the AS's ability to bring in revenue. To enumerate all such routes, we need an AS-level routing model of the entire Internet—we discuss our model for this in Section 4.5. Let r be some route that passes through a; let R be the set of ASes traversed by r. We measure the probability of compromise of r by a single malicious AS $\mathcal{M}$. We define a route r as compromised if $\mathcal{M}$ can successfully cause packets from the source to be routed to $\mathcal{M}$ instead of to the correct destination. In order to do this, $\mathcal{M}$ can hijack the prefix by advertising itself as owning the prefix, or it can advertise an invalid short route to the legitimate destination thus causing packets to be routed to itself. If any of these attacks succeed for a given position of $\mathcal{M}$ on the Internet, then the route r is considered compromised for that position of $\mathcal{M}$. The details on how we determine whether or not an attacker was successful is detailed in Sections 5 and 7. We evaluate the average security s_r of the route r by averaging the binary event variable $E_{r,\mathcal{M}}$ (0=r is compromised by $\mathcal{M}$, 1=r is secure from $\mathcal{M}$) over all possible locations of $\mathcal{M}$ on the Internet not including ASes that are already on the route r. We do not consider malicious ASes already on the route r since in such a case, the attacker has already achieved its goal needing to disrupt the correct operation of BGP.

$$s_r = \sum_{\mathcal{M} \in V, \notin r} E_{r,\mathcal{M}} \cdot P(\mathcal{M})$$

Where $P(\mathcal{M})$ is the probability of $\mathcal{M}$ being the malicious AS. We investigate two probability distributions for $\mathcal{M}$: (1) a uniform distribution, where any AS has an equal chance of being malicious, and (2) a distribution biased towards small ASes, where the probability of an AS being malicious is inversely proportional to its degree—the intuition being that larger ASes are better monitored and administered and hence more secure.

Based on the formula, s_r can take values in $[0,1]$ where 1 means that the route is always secure, and 0 means that the route can always be compromised regardless of which AS happens to be malicious. Let the set of *all* routes passing through a be R_a. We then take the average of the security metrics for each route going through the node a to get the security metric s_a for a, weighted by the estimated traffic w_r for each route r, as shown in Equation 1. We consider several different traffic models which we discuss in Section 4.6.

$$s_a = \frac{\sum_{r \in R_a} s_r w_r}{\sum_{r \in R_a} w_r} \quad (1)$$

4.4 AS behavior model

We assume that each AS adopts the new protocol in some iteration of the simulation if its immediate security benefit is greater than some switching threshold in that iteration. We assume a constant switching threshold for all ASes. This implies that an AS will adopt a secure protocol if adoption secures at least a certain expected fraction of its traffic, regardless of the size, capacity, or position of the AS in the Internet. We believe this is a reasonable approximate model of AS behavior since larger ASes with more traffic capacity will receive greater net benefits from adoption; but at the same time their cost of transition would also be higher due to their larger scale. More precisely, our model assumes that transition costs scale linearly with the traffic carried by an AS, so the natural measure of security benefit is security provided per unit traffic. Clearly, costs in the real world do not scale linearly with traffic, but real-world costs are also affected by unmeasurable factors such as existing infrastructure and business strategies; as a first approximation, the linear assumption will at least allow us to perform tractable simulations and analyses.

We chose the greedy bounded-rational strategy model where each AS only considers its immediate benefit and does not consider the strategies of other ASes. This is because it is the most risk-averse strategy—whenever an AS performs the switch, it is assured that its choice will immediately improve its utility. In a game with significant uncertainty about opponent utilities and strategies and large negative payoffs if beliefs (or predictions) are inconsistent with reality, this is a reasonable approximation to rationality.

4.5 AS Topology

We model the AS topology as a weighted AS-level graph. Each AS is represented as a node, while transit or peering relationships between ASes are represented as edges. The edges are weighted by the number of times a source AS prepends itself in the AS_PATH when it advertises to the neighboring AS. We consider that the weight is symmetric in both directions, i.e., if an AS weighs one edge higher by prepending then it wants both less outgoing and less incoming traffic.

We extract the structure of the AS-level graph from RouteViews data [5]. Since RouteViews is merely a collection of BGP messages at a few limited vantage points on the Internet, it does not reveal the actual AS-level graph. However, this slight inaccuracy is tolerable compared with the strength of our other assumptions. To construct our AS graph, we examine all the paths observed by RouteViews and draw an undirected edge between every pair of ASes that appear consecutively on a path. If the same AS appears consecutively to itself on a path, then this AS is performing AS_PATH prepending on this path. To reflect this, we set the weight of the edge to the number of times the AS prepended itself. For example, for an AS_PATH $AS_1AS_2AS_2AS_3$ in the routing table, we give the AS_1-AS_2 edge an edge-weight of 2 and the AS_2-AS_3 edge an edge-weight of 1.

Since we lack comprehensive policy information about all possible routes on the Internet, in our study we use the least-edge-weight paths to approximate actual routes found on the Internet. If more than one least-edge-weight path exists, one is chosen at random.

In the strong attacker model, due to the universal eavesdropping assumption, any malicious AS receives the same amount of information regardless of its position, and hence the amount of information available to a malicious AS anywhere in the Internet can be precomputed at the beginning of the iteration. Furthermore, since a malicious AS can only inject BGP messages at one point, its ability to attack a given route is dependent solely on its distance from the destination of the route. This means that in each iteration, for each of $O(n^2)$ routes r of length $O(D)$ (where n is the number of ASes and D is the diameter of the AS-graph), we can consider each AS on the route r and check in constant time what the effect is on the security of r when it adopts the protocol. Hence the overall complexity of the computations for the strong attacker model is $O(n^2D)$. Hence, it is possible to perform all the computations for the strong attacker model on the actual Internet AS topology extracted from RouteViews

In the weak attacker model, for each attacker AS M and for every route from AS A to AS B, we have to find the route from A to M that is weakest in security, and consider how this route can be used to attack the route from A to B. Since the information to attack a given route that is available to each malicious AS is now different, for each route, the computation must now iterate over all $O(n)$ possible malicious ASes and revaluate the security benefit of each AS along the route adopting security. This computation is more complex compared with the strong attacker model. Thus for the weak attacker model, we perform the simulation on a smaller generated AS topology. There are a number of topology generators such as BRITE [11], GT-ITM [19] and Inet [17]. We chose Inet as our topology generator because of its close match with the known characteristics of the AS-level graph. This provides us with the AS-level topology graph, however the models do not provide associated IP address origination information for each generated AS, nor do they indicate how AS_PATH prepending could occur in the generated topology. Based on empirical observations, we assume that prepending behavior follows a power-law distribution in the Internet. We replicate this distribution in our generated topologies. To verify our generated topologies, we ran our computations for the strong attacker model on both the full-size AS topology extracted from RouteViews, and the generated topology. The two sets of results were closely matched, indicating that the generated topology was likely to be an adequate approximation to the actual AS topology for our purposes.

4.6 Traffic

To compute Equation 1 in Section 4.3, a model of the traffic load of each possible route on the Internet is needed. Gathering accurate data on actual inter-AS traffic is impractical since this data is usually confidential. Thus for our simulations we approximate the traffic load for each route in four different ways:

- Uniform: Here the traffic is assumed to be uniform between any two pairs of ASes.
- Product of the two endpoint ASes' IP Spaces: An AS's IP space is the number of addresses that it originates to the Internet. If we assume that the Internet's IP space is uniformly populated by hosts, and each host communicates uniformly with all other hosts on the Internet, then the amount of end-to-end traffic between two ASes will be proportional to the product of the two AS's IP spaces.
- Product of the logarithm of the two endpoint ASes' IP Space: This reflects the trend of large IP spaces to be more sparsely populated but the smaller ones are densely populated.
- Gravitational Model: Following the results from Zhang et al. [20], we have also considered the model where the traffic between two hosts is proportional to the product of the IP

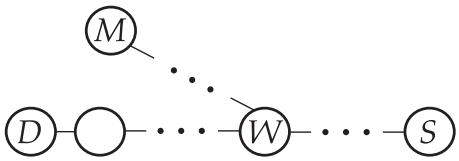

Figure 1: Prefix hijacking

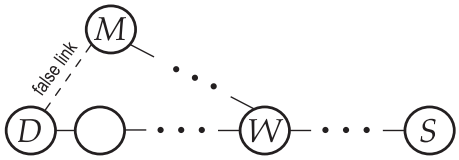

Figure 2: Path spoofing with origin authentication

space and inversely proportional to the square distance between them.

$$\text{Traffic(s,d)} = \frac{IP_s \times IP_d}{hopcount^2}$$

We note that despite the marked differences in each of these metrics, our results in Sections 6 and 8 hold equally well for all of them, indicating that the findings are insensitive to the traffic model.

5. ANALYSIS: STRONG ATTACKER

In this section, we describe the details of how we determine the security level of a given route for each of the five security schemes described in Section 3 under the strong attacker model described in Section 4.2. Under this attacker model, we assume that the attacker is able to read unencrypted BGP traffic anywhere in the Internet. Note that in this section we use the terms *path* and *route* interchangeably.

5.1 Origin Authentication

In this section we describe how we assess the security of a path given that some security scheme with only the Origin Authentication security property is partially deployed on the Internet.

As we explain in Section 4.5, we assume that ASes prefer shortest path routes. Hence, if an adversary is able to falsely advertise a shorter path to the prefix to any AS on the legitimate path, then it is able to divert the legitimate path to itself. It can do so in two ways: *prefix hijacking* or *path spoofing*.

Figure 1 illustrates a prefix hijacking attack using this process. The circles represent ASes, the legitimate originator of the prefix is the destination AS D of the traffic, and the AS at the other end of the path is the source AS S. The malicious AS is denoted by M and it is performing an attack at some given AS in the path denoted by W (W could be any AS along the path). Let $d(A,B)$ denote the distance in AS-hops between A and B. In the figure, M is illegally originating D's prefix. In the absence of authenticating information W is unable to determine which originator is legitimate, and so we assume it simply chooses the closer AS in terms of hops. This attack succeeds whenever one or both of D or W have not deployed origin authentication, and $d(M,W) < d(D,W)$.

If both D and W have deployed origin authentication then the attacker must perform *path spoofing*, i.e., advertise a short path to the originator D instead of originating the prefix directly. Figure 2 describes this attack. The malicious AS falsely advertises itself as being adjacent to the originator D in an attempt to cause W to choose its path over the legitimate path. This attack succeeds if $1 + d(M,W) < d(D,W)$.

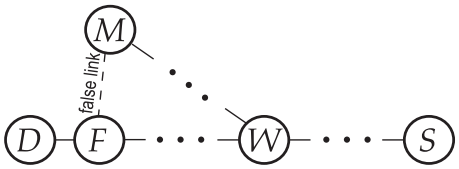

Figure 3: Path spoofing with OA+1

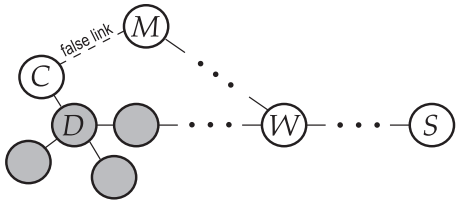

Figure 4: Path spoofing with full security

5.2 First-hop Authentication (OA+1), SPV

Recall from Section 3 that in First-hop Authentication, both the origin and the first hop AS along the path from the origin are authenticated and thus cannot be altered by an adversary. We use the term "Origin Authentication +1" (OA+1) to denote this class of schemes that performs one extra hop of authentication in addition to origin authentication. Note that with OA+1, the path-spoofing attack in Figure 2 fails if both D and W have deployed OA+1. Under the strong attacker model for SPV, the attacker can eavesdrop on the route advertisement as soon as the originating AS sends it to the first-hop AS. This allows it to perform arbitrary alterations to the route after the first hop, thus it has the same properties as OA+1. For brevity, we only discuss OA+1 in this section.

Figure 3 reflects what an adversary M now has to do to subvert the path if both D and W have deployed OA+1. M is no longer able to directly claim a link to the originating AS D since D now signs the identities of each legitimate first-hop AS adjacent to itself. M can, however, illegally advertise a link to the first-hop AS F instead. Hence, this attack succeeds if $2 + d(M,W) < d(D,W)$. Note that directly performing prefix hijacking is still the preferable method of attack if either of D or W have not deployed any security.

5.3 Full Path Security: S-BGP, soBGP

Under the strong attacker model, security schemes with Routing Topology Path Verification and Path Authentication have similar security properties. We call this class of schemes, the schemes with *full path security* under the strong attacker model. We describe the properties of each representative protocol in turn.

S-BGP: In S-BGP, the entire AS_PATH is protected by signatures as far as the nearest non-deploying AS. Once the closest non-deploying AS is encountered, the chain of security is broken and further ASes down the path are unable to provide additional security even if they are deployers of S-BGP, since a malicious attacker could simply strip away any signatures and cryptographic information added on after the first insecure AS.

Figure 4 reflects what an adversary M now has to do to subvert the path if both D and W are deployers of S-BGP. The shaded cir-

cles represent ASes that have deployed S-BGP which form a contiguous area with the originator D. The attacker's best opportunity to present a short path to the originator is to falsely claim a link to C, which is the unsecured AS that is closest to D. Hence, in this case the attack succeeds if $d(D,C)+1+d(M,W) < d(D,W)$. Note that since $d(D,C) \geq 1$, full path security is always at least as secure as OA+1.

soBGP: Under soBGP or any other Routing Topology Path Verification protocol, paths are verified to be consistent against a database of known inter-AS routing information. However, recall from Section 3 that in our analysis, we only allow soBGP to use authenticated topological information from ASes that are soBGP speakers. An attacker is thus free to perform path spoofing attacks as long as any edges in the spoofed path incident to a secure AS are correctly verifiable under soBGP. Referring to Figure 4 once more, it is clear that an attacker's best strategy for creating the shortest possible spoofed path under such a constraint, is to advertise a false link from itself to C, which is the closest insecure AS to D. This is exactly identical to the attack against S-BGP, and hence soBGP has the same security properties as S-BGP under the strong attacker model.

6. RESULTS: STRONG ATTACKER

In this section we present and discuss the results of performing the computations under our simulation-based model for estimating the adoptability of the various protocols under the Strong Attacker model.

6.1 Critical Threshold

Recall from Section 4.1 that the *security benefit* of an AS is a value in the interval $[0,1]$, reflecting the estimated increase in the probability that any given bit passing through an AS cannot be diverted, after this AS adopts a security protocol. The *switching threshold* models the adoption transition cost of the protocol: for any given AS, if the security benefit is below the switching threshold, the AS will not adopt the protocol in the next iteration; otherwise, the AS will become a new adopter in the next iteration. Hence, by varying the switching threshold from 0 to 1, we can run our simulation to determine how far adoption will spread for any given value of the switching threshold.

For each given switching threshold, we measure the final fraction of all ASes which are secure protocol adopters (including the set of initial adopters) when the algorithm has converged. Our simulation terminates when no new ASes are found to be adopters in an iteration. This means that any ASes which are still not adopters must have a lower security benefit than the switching threshold.

Figure 5 shows the final fraction of total adopters as the switching threshold changes. For reference, the fraction of adopters in the first iteration for each threshold level is also indicated. It is clear that the fraction of final adopters exhibits a sharp transition between very low adoption and complete adoption at $c = 0.023$. This sharp transition contrasts with the relatively smooth curve of the fraction of adopters in the first iteration, indicating that it is a characteristic of the adoption process resulting from multiple iterations of our simulation.

Our results indicate a *critical threshold* adoption dynamic where adoption is stalled at a low level when the switching threshold is above the critical threshold but is essentially complete whenever the switching threshold is below the critical threshold. We observed the critical threshold adoption dynamic in *every* simulation regardless of topology, adversary model (weak or strong attacker), path weighting metrics, or various traffic and IP-space ownership distributions. The critical threshold dynamic is due to the positive

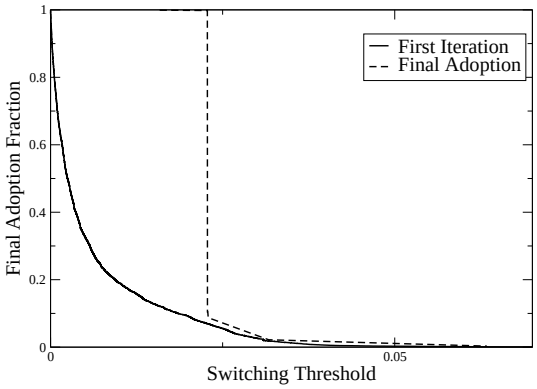

Figure 5: Critical threshold adoption dynamic. With S-BGP Protocol, and the 25 highest degree ASes as initial deployers.

feedback inherent in the system—each AS that adopts the protocol improves the potential benefit of other ASes to adopt the protocol because the benefits of adoption increases as more ISPs adopt the protocol. Hence, as long as the switching threshold of adoption is sufficiently low to sustain a positive rate of adoption for several initial iterations, positive feedback will result in eventual full adoption in the rest of the Internet.

The existence of a critical threshold is significant because we are now able to quantitatively measure adoptability for any secure BGP protocol under a given set of assumptions by a single scalar, i.e., the value of the critical threshold. A scheme with a higher critical threshold can yield full adoption for a larger range of possible switching thresholds, and is thus considered to be more adoptable. For our subsequent analyses, we formally define the following:

Definition 1 *A protocol's adoptability is measured by its* critical threshold*. This is the supremum of the set of switching thresholds for which the final fraction of adopting ASes is greater than* 0.5*.*

6.2 Adoptability of Different Security Schemes

We next investigate the relative adoptability of each of the three classes of security schemes described in Section 5 for the strong attacker model. The three classes are: full security (e.g., S-BGP, soBGP, or SPV), origin authentication only (OA), and first hop authentication (OA+1). Figure 6 shows the final adoption fractions of each scheme as the switching threshold c changes. All three schemes show critical threshold dynamics, switching abruptly from full adoption to almost no adoption when c increases beyond a critical value. The critical threshold for OA is lower than that of OA+1 which is only very slightly lower than that of full path security, indicating a range of switching thresholds where full path security and OA+1 will achieve full adoption while OA achieves little adoption. This reflects the expected result that OA is less adoptable than full security, since OA's security properties are weaker. However, we also observe the surprising result that OA+1 has almost the same adoptability as full security. This is despite the fact that OA+1, which only protects the first hop in its AS_PATH, has significantly weaker security properties than full security, which protects the entire AS_PATH.

In Figure 7, we show the critical thresholds after varying the initial deployment set. We can observe that *T1* deployment yields in a much higher critical threshold than *GOV* or *UNIV* deployment. The critical threshold in *T1* is higher even when the number of ASes is ten times fewer than other deployments. This is expected because

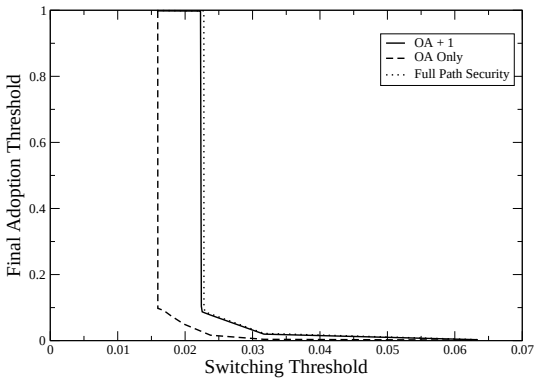

Figure 6: Critical thresholds of various schemes, Tier 1 initial deployers.

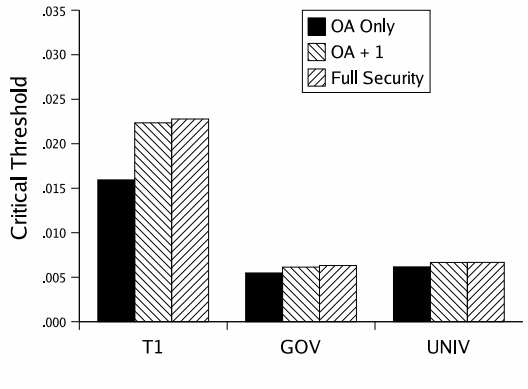

Figure 7: Critical thresholds with different initial deployers.*T1* deployment starts with the 25 highest degree ASes in the Internet, *GOV* deployment starts with all the US governmental institutions, *UNIV* deployment starts with all the educational institutions in the US.

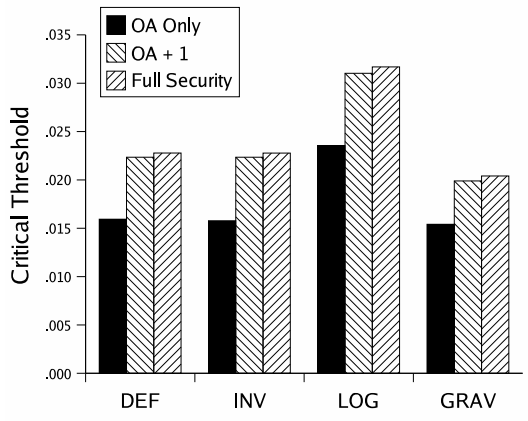

Figure 8: Critical Thresholds with different metrics. Metrics defined in Table 1.

Name	Traffic Metric	Adversary Distribution
DEF	$IP_S \times IP_D$	Uniform
INV	$IP_S \times IP_D$	Inverse Degree
LOG	$log(IP_S) \times log(IP_D)$	Uniform
GRAV	$\frac{IP_S \times IP_D}{distance^2}$	Uniform

Table 1: Metric variations. IP_S is the IP space originated by the source of a path, IP_D is the IP space originated by the destination of a path, *distance* is the distance between them in hop-count. "Uniform" implies any AS has an equal chance of being malicious and "Inverse Degree" implies the the probability of an AS being malicious is inversely proportional to its degree.

the T1 deployers tend to carry most of the traffic in the topology, hence providing security at this set of central points should yield the greatest adoptability.

We investigate how different path metrics and adversary distributions affect the critical threshold values. Table 1 lists the metrics we vary in our simulation and Figure 8 shows the critical thresholds we observe with those conditions. In each experiment we start the adoption by deploying the secure routing protocol on the 25 highest degree ASes in the topology. We observe that the relative values of the critical thresholds for each set of security schemes does not change significantly as we vary our path metric and the adversary distribution.

In the strong attacker model, we observe that using OA results in lower adoptability than full security. On the other hand, OA+1 yields adoptability very close to that of full security. We hypothesize that the similar results of OA+1 and full security are due to the critical threshold dynamics of the adoption process. Since the adoption process experiences positive feedback which drives ASes rapidly to reach full adoption once a sufficient number of ASes have made the decision to adopt the protocol, early stages are particularly crucial. Recalling Figures 3 and 4, the main difference between full path security and OA+1 is that the attacker can spoof a metric of 2 in OA+1 and a metric of $1+d(C,M)$ in full security. However, in the critical early stages of deployment, it is unlikely that D would have completely surrounded itself with secure ASes. Hence, $d(C,M)$ is typically 1. As a result, both OA+1 and full path security share nearly identical properties in early deployment. It is only in mid-deployment, when a significant fraction of ASes have already adopted, that the two schemes begin to diverge, where full path security yields improved security benefits through its full path authentication. However, if adoption is able to proceed to mid-deployment, positive feedback is sufficient to drive both schemes all the way to full adoption regardless of the improved benefits of full security. Hence both schemes show very close adoptability characteristics.

7. ANALYSIS: WEAK ATTACKER MODEL

In this section, we perform the security analysis for each of the five security schemes under the weak attacker model described in Section 4.2. Under this attacker model, we assume that the attacker is only able to read incoming BGP updates at the malicious AS; it is not privy to other BGP messages elsewhere in the Internet even if those messages are unencrypted.

In particular, for any given prefix, a malicious AS M will receive one or more updates for that prefix. Since we assume that BGP uses shortest path routing, only the shortest such path is of relevance to the malicious AS. This path is indicated in Figure 9 as the path

from D to M traversing ASes X, Y and Z. The attacker must now use this information to advertise a short route to some AS W in the path in order to divert the legitimate path through itself.

7.1 Analysis Unchanged: OA, OA+1, soBGP

The security analysis for Origin Authentication (OA), First-hop Authentication (OA+1), and soBGP (Routing Topology Path Verification) remain unchanged under the weak attacker model as compared with the strong attacker model (see Section 5). We discuss each in turn.

OA: The attacker remains free to perform prefix hijacking if either one of D or W do not have origin authentication deployed; otherwise, it claims a direct link to D and performs path spoofing as per Figure 2.

OA+1: The adversary is free to perform path spoofing as in Figure 3, with one minor variation: instead of spoofing a false link to the first AS (F_1) in the legitimate path, it spoofs a false link to the first AS on the path to itself (F_2 in Figure 9). The end result is identical; the attack succeeds if $2 + d(M,W) < d(D,W)$.

soBGP: In soBGP (or any Routing Topology Path Verification protocol), BGP routes are verified against a database of known routing information for all the ASes that have deployed the security scheme. We assume that access to the information in this database is public—AS routing information can already be readily deduced through such mechanisms as RouteViews. Furthermore, whether or not ASes are soBGP speakers can be easily determined by monitoring their route advertisements. Hence, the task of the attacker is identical for both the weak attacker model and the strong attacker model. By querying the database (or through any other information channel), the attacker determines the closest non-deploying AS C to the originator D. The attacker then spoofs a short path from D to itself through C, exactly as in Figure 4.

7.2 S-BGP (Path Authentication)

Path Authentication, as represented by S-BGP, is one of the two classes of protocols that behaves differently under the weak attacker model than the strong attacker model. Under the weak attacker model with S-BGP, the attacker is no longer able to eavesdrop on messages sent to any potential non-deploying ASes that are close to the originator. In particular, the attacker is not always able to perform the attack described in Section 5.3, because this attack requires the attacker to have eavesdropped on the BGP update sent to the closest insecure AS C from the originator AS D, but this AS may not be on the path to the malicious AS M (see Figure 9). In fact, the only useful messages accessible to M are the BGP update messages that are received by M from D. For simplicity, we consider the case where only one such update message was received (indicated in Figure 9 as the path from D to M through ASes X,Y,Z). Extension to the case where multiple messages were received is straightforward. Suppose X is the first non-deploying AS in this path; all previous ASes are S-BGP speakers. Since this represents the first break in the chain of signatures, M can remove any security information appended after X, and spoof a path by claiming a direct link to X. This is the best M can do since M cannot remove any of the authentication information prior to X as all the ASes prior to X are secure. Therefore, attacker can hijack the path from S to D only if $d(D,X) + 1 + d(M,W) < d(D,W)$.

7.3 SPV (Retroactive Path Integrity)

Under SPV (or any similar Retroactive Path Integrity protocol), the adversary's task is further complicated by the fact that subsequent ASes in the path can add cryptographic information to repair the break in the chain of security caused by non-deploying ASes.

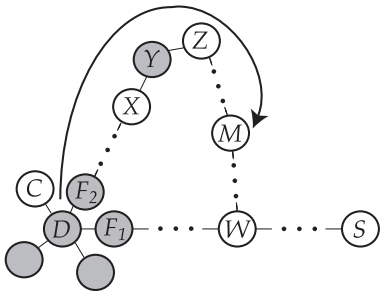

Figure 9: Path spoofing with Weak Attacker Model. Arrow indicates path of BGP update message received by attacker at M.

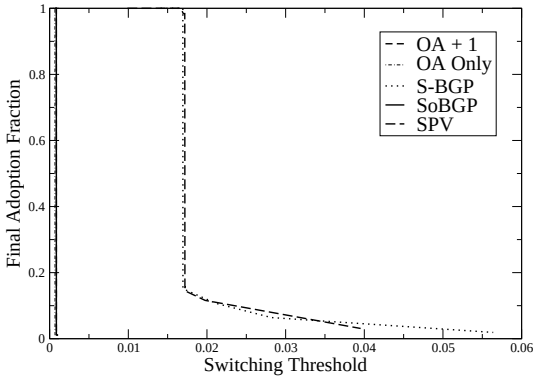

Figure 10: Critical thresholds of various schemes, Weak Attacker Model.

For example, suppose that in Figure 9, X and Z are non-deploying ASes, and Y is the closest SPV-speaking AS to M on the path. Under S-BGP the attacker could spoof a direct link to AS X as described in Section 7.2. However, SPV prevents the attacker from stripping off the cryptographic signatures of AS Y because AS Y will have added the necessary signatures to close the break in the chain (for example, by performing signatures on behalf of AS X) before adding its own signature. This means that the attacker is now restricted to spoofing a direct link to the first non-deploying AS (in this case, AS Z) after the latest secure AS in the path received by the malicious AS (in this case, AS Y). The attack is successful only if $d(D,Z) + 1 + d(M,W) < d(D,W)$.

8. RESULTS: WEAK ATTACKER

In the weak attacker model, we run the simulation on a smaller generated model of 1000 ASes. For cross-validation of the generated model, when we ran our analysis for the Strong Attacker Model on the smaller model topology, we observed the same behavior as in the larger topology extracted from RouteViews, which indicated that the generated model exhibited properties close enough to the actual AS-level topology for our purposes. We chose the 5 highest degree ASes (which represent Tier-1 ASes) from the generated topology as our initial adopters.

Figure 10 shows the critical thresholds of each of the five classes of schemes. The critical thresholds of OA, OA+1, and soBGP remained unchanged from the strong attacker model since their security analyses were identical for both attacker models. On the other hand, SPV and S-BGP both showed significant (approx. $10\times$) in-

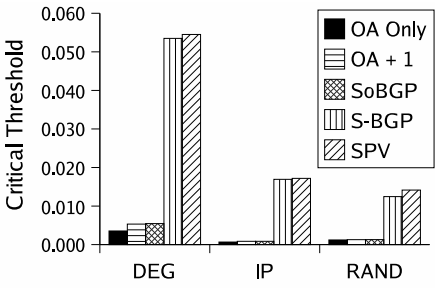

Figure 11: Critical thresholds of different deployments.In *DEG*, the initial deployment set consists of 5 highest degree ASes, *IP*'s initial deployment set is the 5 ASes originating the most IP prefix space, *RAND*'s initial deployment set consists of 10 ASes we chose randomly.

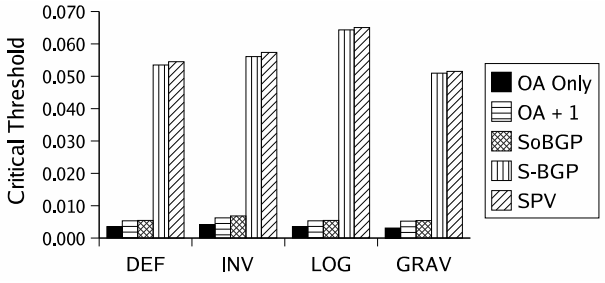

Figure 12: Critical thresholds of different metrics. Metrics defined in Table 1.

creases in adoptability, indicating that full path security is indeed highly valuable in driving protocol adoption in the weak attacker model. We conjecture that the significant difference in the effectiveness of full path security between the strong and weak attacker models is again due to events in the most vital early stages of adoption. Under the strong attacker model, as long as some node close to the originator (e.g., node *C* in Figure 9) is a non-adopter, the attacker retains a very strong path-spoofing ability. Hence, early in the adoption process when not many ASes are adopters, there remain many points of vulnerability. This prevents full path security from being significantly more useful than origin authentication until a large fraction of the Internet has already adopted the protocol. However, in the weak attacker model, it is more likely that a small number of early adopters can significantly improve the subsequent adoption benefits of later adopters—whenever an AS adopts the protocol, as long as the rest of the ASes between the originator and the newly adopting AS are also secure, this will reduce (by one hop) the spoofing capabilities of every malicious AS downstream from it. This improvement in security may be sufficient to cause further adoption in later iterations and thus continue to drive the adoption process for values of the switching threshold which would have stalled adoption under the strong attacker model.

It was expected that SPV, having Retroactive Path Integrity, would be more adoptable than S-BGP; however, the observed difference in the two schemes' critical thresholds was small. This indicates that Retroactive Path Integrity is only slightly more effective in increasing the adoptability of a protocol, and is hence probably not a feature that should be emphasized in future protocol research.

As we do for the strong attacker model, we vary the initial adoption set to verify that the adoptability relationships between each class of schemes holds for different initial conditions. As our alternative initial adoption sets, we chose (1) the top 5 ASes which originated the largest amounts of IP space, and (2) 5 ASes at random from the entire set of ASes. Figure 11 shows that the relative adoptabilities between the schemes hold as we vary the initial conditions.

Similarly, we vary the traffic metric and the adversary distribution to observe their effect on critical threshold. As before, we used the 5 highest degree ASes as our initial adopters. The different traffic metrics and adversary distributions are the same as the ones investigated for the strong attacker model (see Table 1). Again we note that the relative adoptabilities of the five classes of schemes remain stable despite different path metrics and adversary distributions.

9. DISCUSSION

Current research on protocol design focuses on exploring various tradeoffs between security and implementation cost. However, such a limited set of metrics is insufficient to adequately inform researchers as to the most desirable tradeoffs in the design space. For example, there exists no method to quantify how much a given security property contributes to the likelihood of a protocol to be widely adopted. Given two protocols, one with a strong security property and high implementation cost, and another with a weaker security property and lower implementation cost, it is unclear which protocol is in fact the more feasible technology.

Since the "security" provided by a protocol is a set of qualitative properties describing how the protocol is resistant to various attacks, simply examining the security of a protocol does not quantify the relative contributions of each property to the protocol's attractiveness to potential implementors since there is no ordering relation on the set of all possible combinations of security properties. Our methodology for extracting a critical threshold measure of adoptability through simulation allows us to provide one possible ordering relation, allowing researchers to compare the attractiveness of various sets of security properties under any given context. A concise definition of the metric of adoptability is as follows:

Definition 2 *The **adoptability** of a given (security) protocol is a measure of the attractiveness or usefulness of the protocol's (security) properties in terms of how strongly these properties might motivate eventual full adoption of the protocol in the Internet.*

"Security" is parenthesized because the more general definition of adoptability can be used for arbitrary protocol properties, as long as the utility of these properties to the adopter can be mathematically modeled in some way. In this paper, we consider only security protocols. Hence, for a given adoption context, adoptability can be viewed as a **security metric**, i.e., it measures only the strength of the protocol's security properties.

Thus far, the discussion of the simulation results have focused on coarse-grained ordering-based comparisons of the critical thresholds of the various protocols. Even coarse-grained comparisons of adoptability are useful, since an ordering can be derived for a previously unordered set of security properties. Sometimes, the ordering of security properties is trivial, for example, no attack is possible against full path security that was not possible for OA+1. However, this is not necessarily the case. For example, consider the case where soBGP performs routing topology verification where well-known and stable routes between non-adopters of soBGP are considered part of the normal topology against which routes are compared. In such a scenario, the attacks that can be performed

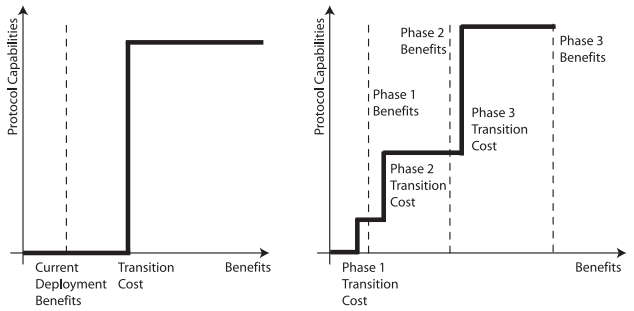

Figure 13: Advantage of Staged Deployment

by an attacker against S-BGP are no longer a subset of the attacks on soBGP: for example a route could be forged involving non-adopters of security that would be accepted by S-BGP but rejected by soBGP. With such an assumption, the relative strengths of the two classes of security properties would be uncertain for different deployment contexts. Examination of their critical thresholds would help inform the researcher as to the relative effectiveness of each class in driving adoption.

The presentation of adoptability as a measurable metric for protocol design allows for a new perspective on protocol design. Researchers may now consider the adoption dynamics of a protocol during the design process, and optimize for protocols that have the highest adoptability for the lowest costs. Adoptability is an attractive metric not only because it allows us to directly compare qualitative properties, but also in terms of its association with feasibility: a protocol with strong properties that does not get adopted is not as desirable as a protocol with weaker properties, but does get adopted.

As an example, one possible way in which specific adoption dynamics may be designed into the protocol is the possibility of protocols that support *staged deployment*. In this setting, a modular protocol is deployed in multiple stages. By breaking up the sharp transition from legacy to new protocol into a series of small stages, a protocol can turn a prohibitively expensive transition scenario into a sequence of incentives-compatible steps, thus greatly enhancing its adoptability. Figure 13 illustrates this process. The graphs show a typical AS's decision process, with benefits on the x-axis, which determine whether or not the protocol is deployed (y-axis). In a typical monolithic protocol, deployment is all-or-nothing, as shown in the sharp step function in Figure 13a—at any point where benefits outweigh the transition cost, the AS will decide to deploy the protocol. However, if the current benefits are below the transition cost, as shown on the diagram, then the AS will decide not to deploy the protocol. If this is the case for all ASes in the Internet then the protocol has stalled in its deployment. Figure 13 shows what might happen in the protocol was designed in a series of mutually-supportive modular phases. Each module has its own incremental transition cost, which provides incrementally increasing capabilities and greater benefits for the next phase. The adoption process is thus facilitated since at every phase, the incremental benefits of moving onto the next phase outweigh the incremental transition costs. Hence such a protocol would provide a solution to the *coordination problem* where social benefit is maximized if all ASes switch to the new protocol, but no AS wants to be the first to commit to the costly switch.

Besides highlighting the importance of adoptability as a design dimension for new protocols and providing new insights into protocol design, our research also highlights the importance of the problem of the selection of the initial adopter set in the deployment of new protocols. This is a crucial step in the adoption process, and yet there has been little quantitative research to date into the best models and heuristics for approaching this problem. We hope that our initial approach will open the field to more focused efforts into this important subproblem.

10. LIMITATIONS AND FUTURE WORK

While our methodology makes necessary assumptions to facilitate tractable simulation and analysis, it remains the first quantitative approach to measure adoptability in Internet protocols. In this section we revisit some important assumptions made earlier in the paper.

We intuit in Section 4, that traffic security is the economic incentive for ASes adopting secure BGP protocols. However, adoption decisions for an AS may involve complex economic and political factors, which are difficult to model. Similarly, accurately modeling the cost of a secure BGP protocol is also difficult; for example, soBGP requires exposing all neighboring information, including peering information, to other adopting ASes. On the other hand, peering information can be highly valued by some ASes [13], thus the cost of adopting soBGP for these ASes may be higher than others. Even with these simplifying assumptions on BGP protocol, we believe that our model considers the important factors affecting the adoption process. Furthermore, the model can be easily extended to other fledgling Internet protocols such as DNSSEC [2], etc.

Another limitation of our methodology is that, although the critical threshold values can be used for qualitative comparisons between protocols, their actual numerical value does not map onto any directly measurable quantity. For example, we are unable to translate a critical threshold to a dollar amount or map it to a meaningful numerical prediction about the absolute likelihood of adoption of either scheme.

One way to make direct quantitative comparisons between protocols is to introduce a (strong) linearity assumption. Suppose protocol A has a critical threshold $10\times$ higher than protocol B. However, protocol B is somehow able to make the claim that, in most contexts, the switching costs to adopt B is k times *lower* than the switching costs for A. In particular assume that the distribution of switching thresholds for B is similar to the distribution of switching thresholds for A, but linearly scaled k times smaller. With such an assumption we can now compare the relative likelihoods of adoption for the two schemes. We know that adoption will proceed if the switching cost is below the critical threshold. If $k = 10$, then, since both the critical threshold and the distribution of switching costs are 10 times lower for B than A, the two schemes have roughly equal likelihoods of adoption. However, if $k < 10$, then the savings in switching costs for B are insufficient to compensate for the reduced adoptability of B's weaker properties, and hence B is less likely to be adopted than A. Conversely, if $k > 10$, then B is more likely to be adopted than A.

This flavor of result, although laden with strong assumptions and hence acceptable only as a very rough estimate, is nonetheless currently the only known method in which qualitative security properties can be quantitatively compared. It is hoped that with refinements of this technique, confidence in the assumptions can be improved, and the significance of the numerical difference between critical threshold can be increased.

11. CONCLUSION

In this paper, we argue that it is important to consider the dimension of *adoptability* in protocol design. We present the following formulation: a protocol's adoptability corresponds to the space of incentives compatible adoption scenarios that yield widespread adoption under given assumptions. We propose a simulation methodology to explore and characterize this space under a range of assumptions and contexts, including multiple attacker models, different initial adopter sets, different network models and different security metrics.

In the process of applying this methodology to the known BGP security schemes to date, we have created a taxonomy for classifying and distinguishing the security properties of a wide range of protocols under partial deployments. Such a taxonomy has also been lacking to date; our new taxonomy enables us to model both existing protocols and variations that do not correspond to any published protocol, yielding interesting design points that have not been explored in the literature, but have good adoptability properties under certain conditions (e.g., OA+1 under the Strong Attacker Model).

Using our methodology, we make the following observations about the adoption dynamics of BGP security: (1) all known BGP security schemes experience *critical threshold* dynamics under all the simulation models we tested; when the switching threshold (transition cost) was above the critical threshold, very little adoption took place. In contrast, when the switching threshold was below the critical threshold, the system eventually converged to full adoption; (2) under the Strong Attacker Model, OA+1 yields comparable adoptability compared with schemes with full AS_PATH security. This is surprising because OA+1 possesses weak security properties and yet drives this high level of adoptability; (3) under the Weak Attacker Model, Path Authentication experiences a significant increase in adoptability, greatly outperforming the lower classes of security schemes. Furthermore, RPI yields only slightly better adoptability than Path Authentication, indicating that the advantages of RPI are not as significant as expected; (4) while it is clear that a larger set of initial adopters always serves to increase the critical threshold, it is unclear how to select this initial set under various constraints. This would be a fruitful area for future research.

Acknowledgments

We wish to thank Yih-Chun Hu, Patrick McDaniel, David McGrew, Jon Peha, Marvin Sirbu, Damon Smith, Brian Weis and the anonymous referees for their invaluable comments and suggestions.

12. REFERENCES

[1] W. Aiello, J. Ioannidis, and P. McDaniel. Origin authentication in interdomain routing. In *ACM Conference on Computer and Communications Security (CCS 2003)*, 2003.

[2] D. Eastlake 3rd. Domain Name System Security Extensions. RFC 2535 (Proposed Standard), March 1999.

[3] G. Goodell, W. Aiello, T. Griffin, J. Ioannidis, P. McDaniel, and A. Rubin. Working around BGP: An incremental approach to improving security and accuracy in interdomain routing. In *Proceedings of Symposium on Network and Distributed System Security (NDSS'03)*, February 2003.

[4] Xinming He, Christos Papadopoulos, and Pavlin Radoslavov. A framework for incremental deployment strategies for router-assisted services. In *INFOCOM*, 2003.

[5] http://www.routeviews.org/. University of Oregon route views project.

[6] Y.-C. Hu, A. Perrig, and M. Sirbu. SPV: Secure path vector routing for securing BGP. In *Proceedings of ACM SIGCOMM 2004*, September 2004.

[7] David Kempe, Jon Kleinberg, and Eva Tardos. Maximizing the spread of influence through a social network. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146, New York, NY, USA, 2003. ACM Press.

[8] S. Kent, C. Lynn, J. Mikkelson, and K. Seo. Secure border gateway protocol (S-BGP) — real world performance and deployment issues. In *Symposium on Network and Distributed Systems Security (NDSS '00)*, pages 103–116, San Diego, CA, February 2000.

[9] S. Kent, C. Lynn, and K. Seo. Secure border gateway protocol (S-BGP). *IEEE Journal on Selected Areas in Communications*, 18(4):582–592, apr 2000.

[10] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur. Topology-based detection of anomalous BGP messages. In *Proceedings of the 6th Symposium on Recent Advances in Intrusion Detection (RAID)*, 2003.

[11] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE: An approach to universal topology generation. In *Proceedings of Workshop on Modeling, Analysis and Simulation of Computer and Telecommunications Systems (MASCOTS '01)*, 2001.

[12] S. Murphy. BGP Security Vulnerabilities Analysis. IETF draft-ietf-idr-bgp-vuln-00, February 2002.

[13] William B. Norton. Internet service providers and peering. In *Proceedings of NANOG 19*, Albuquerque, New Mexico, June 2000.

[14] L. Subramanian, V. Roth, I. Stoica, S. Shenker, and R. Katz. Listen and whisper: Security mechanisms for BGP. In *Proc. of the First Symposium on Networked Systems Design and Implementation (NSDI'04)*, 2004.

[15] T. Wan, E. Kranakis, and P. van Oorschot. Pretty secure BGP (psBGP). In *Proc. of Internet Society Symposium on Network and Distributed System Security (NDSS'05)*, 2005.

[16] R. White. Securing BGP through secure origin BGP. Technical report, Cisco Internet Protocol Journal, September 2003.

[17] J. Winick and S. Jamin. Inet 3.0: Internet topology generator. Technical Report CSE-TR-456-02, University Of Michigan, 2002.

[18] Harlan Yu, Jennifer Rexford, and Edward W. Felten. A distributed reputation approach to cooperative internet routing protection. In *Workshop on Secure Network Protocols*, 2005.

[19] E. Zegura, K. Calvert, and S. Bhattacharjee. How to model an internetwork. In *Proceedings of IEEE Infocom '96*, 1996.

[20] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg. Fast accurate computation of large-scale IP traffic matrices from link loads. In *Proceedings of ACM SIGMETRICS*, June 2003.

[21] Meiyuan Zhao, Sean W. Smith, and David M. Nicol. Aggregated path authentication for efficient bgp security. In *ACM Conferernce on Computer and Communication Security (CCS)*, 2005.

[22] X. Zhao, D. Pei, L. Wang, D. Massey, A. Mankin, S. Wu, and L. Zhang. Detection of invalid routing announcements in the internet. In *IEEE DSN 2002*, 2002.