DISCO: Sidestepping RPKI's Deployment Barriers

Tomas Hlavacek[†] Italo Cunha[‡♮] Yossi Gilad[♯] Amir Herzberg[⋆]
Ethan Katz-Bassett[♮] Michael Schapira[♯] Haya Shulman[†]

[†] Fraunhofer SIT [‡] Universidade Federal de Minas Gerais [♯] Hebrew University of Jerusalem
[⋆] University of Connecticut [♮] Columbia University

***Abstract*—BGP is a gaping security hole in today's Internet, as evidenced by numerous Internet outages and blackouts, repeated traffic hijacking, and surveillance incidents. To protect against prefix hijacking, the Resource Public Key Infrastructure (RPKI) has been standardized. Yet, despite Herculean efforts, ubiquitous deployment of the RPKI remains distant, due to RPKI's manual and error-prone certification process. We argue that deploying origin authentication at scale requires substituting the standard requirement of certifying *legal* ownership of IP address blocks with the goal of certifying *de facto* ownership. We show that settling for de facto ownership is sufficient for protecting against hazardous prefix hijacking and can be accomplished without requiring *any* changes to today's routing infrastructure. We present DISCO, a readily deployable system that *automatically* certifies de facto ownership and generates the appropriate BGP-path-filtering rules at routers. We evaluate DISCO's security and deployability via live experiments on the Internet using a prototype implementation of DISCO and through simulations on empirically-derived datasets. To facilitate the reproducibility of our results, we open source our prototype, simulator, and measurement analysis code [30].**

I. Introduction

The Border Gateway Protocol (BGP) glues together the organizational networks, called "Autonomous Systems" (ASes), that make up the Internet. It is thus, arguably, the most crucial component of the Internet's infrastructure. Unfortunately, BGP was designed decades ago, when security was not the foremost consideration, and, consequently, BGP is disastrously vulnerable to configuration errors and attacks [47]. Indeed, BGP's insecurity is the cause for repeated major Internet outages [51], [60], [61] and traffic hijacking incidents [1], [5], [63].

The most common and devastating attack on BGP is prefix hijacking, where an attacker advertises in BGP IP addresses that belong to another AS and thereby attracts traffic destined for that AS (for the purpose of monitoring, eavesdropping on, or manipulating traffic; blackholing traffic; masquerading as the legitimate destination; etc.). In fact, today's BGP routing infrastructure is so fragile that misconfigurations of BGP routers often result in inadvertent prefix hijacks.

To protect against prefix hijacks, the Internet Engineering Task Force (IETF) is promoting the deployment of the Resource Public Key Infrastructure (RPKI), which binds IP address blocks to "owner" ASes via cryptographic signatures [35]. RPKI enables ASes to validate that an AS advertising IP addresses in BGP is authorized to do so and thus to detect and discard prefix hijacks. This form of RPKI-based filtering is termed Route-Origin Validation (ROV). Beyond its significance for thwarting prefix hijackers, RPKI is also the first step towards combating more sophisticated attacks on BGP, namely, path manipulation attacks [40].

Unfortunately, despite its critical role, RPKI suffers from significant drawbacks, both in terms of security and in terms of deployability.

Security concerns. RPKI violates the fundamental principle of "do no harm". Due to human error, over 5% of the records in RPKI repositories conflict with *legitimate* long-lived BGP announcements and would cause ROV-enforcing ASes to discard *legitimate* BGP route-advertisements, thus disconnecting from thousands of legitimate destinations [21], [31]. In addition, almost a third of the records are misconfigured in a way that leaves the issuer unprotected from prefix hijacks [24].

Deployability concerns. RPKI adoption has been frustratingly slow [21], [29], [45], [49]. With few notable exceptions (e.g., AT&T [44]), almost no AS uses ROV [21], [29], [49], and the very few that do may enforce ROV only partially [44]. In addition, RPKI certification of IP-prefix ownership is quite limited [45]. One reason for this dismal state of affairs is the classic chicken and egg problem: both certifying ownership of IP addresses and using ROV require nontrivial effort, yet each is largely ineffective without the other being widely deployed. Certifying ownership of IP address blocks in RPKI is a manual, bureaucratic process that requires coordination between ASes, yet certification is effective only when many ASes use ROV to discard bogus BGP route-advertisements. Indeed, as shown in [21], even high ROV adoption rates amongst the largest ISPs which comprise the core of the Internet still leave the issuers of RPKI certificates vulnerable to prefix (and especially subprefix) hijacks. Moreover, using ROV requires carefully analyzing the expected effect on actual traffic of different ROV strategies (to avoid disconnecting from legitimate destinations) [44].

Certifying de facto ownership with DISCO. We argue that (1) today's limited adoption of RPKI/ROV leads to a less secure Internet than a hypothetical one with widespread adoption of a scheme that provides some but not all of the protections of today's RPKI, and (2) this tradeoff is realizable by moving away from today's manual and error-prone RPKI certification process. To this end, we propose revisiting the conventional requirement of binding IP address blocks to their *legal* owners and consider the more modest goal of certifying *de facto* possession of IP addresses. De facto ownership, in this context, means that the AS being certified as the owner of a block of IP addresses controls those addresses in BGP.

We present Decentralized Infrastructure for Securing and Certifying Origins. DISCO *automatically* certifies de facto

Network and Distributed Systems Security (NDSS) Symposium 2020
23-26 February 2020, San Diego, CA, USA
ISBN 1-891562-61-4
https://dx.doi.org/10.14722/ndss.2020.24355
www.ndss-symposium.org

ownership of IP addresses, populates public repositories, and generates filtering rules for ROV. DISCO is designed to provide reliable security guarantees while being easier and safer to deploy than today's RPKI. First, by automating the certification process, DISCO avoids the costs and risks associated with the traditional manual and error-prone RPKI certification. Second, to deploy DISCO, an AS need not coordinate with any other AS (as opposed to today's hierarchical dependencies [21]). Finally, DISCO maintains consistency between certificates and the BGP control plane to avoid incidents resulting from human neglect in synchronizing the two.

We designed DISCO to be compatible with today's Internet architecture. An AS deploying DISCO need only configure iBGP sessions between its BGP routers and a local machine running a DISCO agent (DISCO requires no changes to the routing hardware or software). We present a prototype implementation of DISCO and show that DISCO is readily deployable on today's Internet and can provide significant security benefits through live (control-plane and data-plane) experiments on the PEERING platform [52], [53], as well as extensive simulations on empirically-derived datasets. We make all the code and artifacts from our study available [30].

We regard de facto ownership certification with DISCO as a practical way to circumvent the obstacles facing the adoption of today's RPKI, leading to a more secure Internet. We also regard DISCO as a means for enhancing RPKI's security and driving its deployment forward. By comparing DISCO's automatically-generated certificates with operator-issued RPKI certificates and flagging inconsistencies, operators can identify and fix misconfigured RPKI records, thus mitigating the adverse effects of human error in issuing RPKI records on RPKI's security and deployability.

II. RPKI and Its Adoption Challenges

RPKI associates public keys with network resources such as IP prefixes [38]. After certifying their IP prefixes, owners can use their private keys to authorize specific AS numbers to advertise these prefixes in BGP. Authorizations are cryptographically signed and published in public repositories, which enables other ASes to verify the authorization and filter routes with invalid origins so as to protect against prefix hijacks. The RPKI certification and validation system consists of:

- **Resource Certificates (RC):** certificates for ownership of IP prefixes (and AS numbers) mapping owner public keys to IP prefixes (and AS numbers).
- **Route Origin Authorizations (ROAs):** signed statements using the certified private key of the owner of an IP prefix to specify an AS number authorized to originate this prefix in BGP. The ROA might also permit the AS to advertise subprefixes, up to a specified maximum prefix length.
- **Route Origin Validation (ROV):** filtering rules to be applied by BGP routers to discard or depreference a BGP advertisement whose origin AS does not match the information in the prefix's ROA.[1]

[1]Depreferring invalid routes reduces some of the risks from errors, but leaves the AS completely vulnerable to subprefix hijacking [16], [28].

Despite the importance of RPKI for Internet security and extensive efforts to push its deployment forward, RPKI adoption is sluggish. The vast majority of prefixes advertised in BGP are still not in the system [45] (including most IP addresses for popular web-services [65]), and very few ASes filter BGP advertisements based on the information recorded in RPKI [21], [29], [49], although there is some progress on both these fronts [13]. We next discuss obstacles to RPKI's ubiquitous adoption [31], [24], [21].

Certification can be challenging. RPKI certification is manual and hierarchical. Network operators need to request whoever allocated the IP addresses to them to issue them a resource certificate. Since many organizations do not yet have resource certificates for their IP address blocks, in many cases certification requires other organizations (higher up in the hierarchy) to first be certified themselves [21].

Consider the following example. Level 3 owns prefix 8.0.0.0/9 and allocated subprefixes to hundreds of organizations. However, suppose that Level 3 did not yet obtain a resource certificate. With today's RPKI, all these organizations must wait for Level 3 to certify ownership over 8.0.0.0/9 before they can ask Level 3 to sign a certificate for them.

Worse yet, if Level 3 had a certificate and issued a ROA for 8.0.0.0/9, then any organization holding a subprefix of 8.0.0.0/9 without a ROA for this subprefix would appear as an attacker to ROV-enforcing ASes. This constitutes a real problem for some of the world's largest Internet service providers [21].

Human error. Issuing ROAs requires network operators to manually authorize origin ASes and to specify the maximum permissible length for advertised subprefixes. However, an operator might inadvertently not authorize all origins, or restrict the maximum length to be too short, and so advertise BGP prefixes that violate their ROAs. About 6% of the BGP announcements that are covered by ROAs are invalid [45], with the vast majority of these attributed to human error [21], [31], [64]. ASes that perform ROV would unwittingly discard legitimate BGP advertisements for those prefixes, and thus disconnect from legitimate destinations.

Another common type of human error is issuing ROAs with an unnecessarily long maximum prefix length [24]. Such ROAs void RPKI's benefits and leave the issuer completely vulnerable to devastating subprefix hijacks [24]. Recent measurements reveal that almost 30% of the prefixes covered by RPKI are, in fact, not protected [21], [20], [24].

Human error impacts ROV enforcement. Only a few ASes actually use ROV [21], [29], [49]. According to a recent survey of network operators [21], the most common reason for not filtering invalid routes is fear of "being disconnected from legitimate destinations" due to erroneous RPKI records. Even some of the few that perform ROV (e.g., AT&T [44]) do so only partially, and must have sufficient expertise and measurement capabilities to first carefully analyze the expected implications of doing ROV for their traffic and to periodically re-evaluate this over time [44].

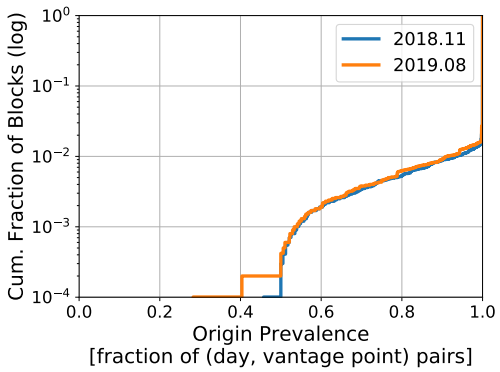


Fig. 1. Origin stability for globally routed prefixes over the course of entire months (Nov. 2018 and Aug. 2019). More than 97% of IP blocks have a de facto owner (prevalence = 1).

III. Introducing De Facto Ownership

To overcome the obstacles facing RPKI adoption, we argue that the classic desideratum of certifying *legal* ownership of IP prefixes should be substituted for certification of *de facto* ownership over IP addresses. We use the term de facto ownership of an IP block to indicate that the AS being certified for ownership is the AS to which traffic destined for the relevant IP addresses is forwarded to over a considerable amount of time. Traffic sent to almost all routed addresses on the Internet reaches a single destination AS (with the exception of multi-origin addresses, which we discuss in §III-B).

Because IP prefixes (of different lengths) can overlap, we define an IP *block* as a non-contiguous non-overlapping portion of the IP address space forwarded to an AS, for which we can establish de facto ownership. For example, if AS3356 (Level3) announces $P_1 = 8.0.0.0/9$, AS15169 (Google) announces a subprefix $P_2 = 8.8.8.0/24$, and considering no other prefixes overlap, we define two blocks: $B_1 = P_1 \setminus P_2$ and $B_2 = P_2$. In the rest of the discussions in this paper, we refer to IP address blocks, rather than prefixes, as the objects in DISCO's certificates to reflect the nature of non-contiguous certification.

We introduce a system, DISCO, for automated certification of de facto ownership and generation of filtering rules for ROV. We show below that certifying de facto ownership can be executed in a manner that guarantees that: (1) Any AS can certify its own IP addresses. (2) An attacker capable of fooling DISCO's certification either has no point in doing so (because they serve as a sole upstream provider of the victim and can already intercept their traffic) or must launch a highly visible, long-running attack.

A. Almost All Routable Addresses Have De Facto Owners

To leverage de facto ownership for certification purposes, we must first establish that most routed IP blocks have persistent de facto owners. That is, we aim to show that most IP blocks are advertised by a single origin AS for extended periods of time.

To investigate the stability of origin AS numbers (ASNs) for IP blocks on the Internet, we examine routing tables from all RouteViews and RIPE RIS collectors. The routing tables were collected midnight (UTC) each day throughout November 2018 and August 2019. We define the *origin prevalence* of a specific AS A with respect to a specific IP block B to be the fraction of collected routing tables in which A is the origin AS (receives traffic) for B. We call the origin AS with the highest prevalence with respect to an IP block B its *prevalent origin*.

Blocks	Nov. 1st, 2018			Aug. 1st, 2019		
	Total	De facto	Multi-origin	Total	De facto	Multi-origin
IPv4	778894	97.5%	0.78%	825063	97.0%	0.98%
IPv6	66751	98.8%	0.47%	79454	98.6%	0.57%

TABLE I. Number and fraction of multi-origin prefixes

Figure 1 plots the distribution of the prevalence values for the prevalent origins across all IP blocks, with a logarithmic scale on the y-axis.[2] Results are consistent across both months, with most blocks announced by a single origin AS (prevalence = 1), a clear case of de facto ownership. Table I summarizes the results. Additionally, the fraction of individual IP addresses with a de facto owner is higher than for blocks (e.g., 98.1% for IPv4 addresses vs. 97.0% for blocks in Aug. 2019), possibly because large blocks are often older allocations with stable owners and/or because traffic engineering or delegations are more common at finer granularities (e.g., /24s).

The above measurement results indicate that certifying de facto ownership applies to the vast majority of IP blocks advertised in BGP. All these blocks can benefit from DISCO.

B. Inherent Limitations

De facto ownership is not without its limitations. In particular, IP prefixes not advertised in BGP may have a legal owner, but not a de facto owner (as no AS is advertising itself as the owner of these prefixes in BGP). The absence of a de facto owner implies that de facto ownership certification alone is unsuitable for such prefixes.

Another limitation of de facto ownership is when multiple ASes originate the same prefix. As our results in §III-A show, BGP advertisements observed by RouteViews and RIPE RIS vantage points on Nov. 1st, 2018 and Aug. 1st, 2019 reveal that only a small fraction of prefixes are advertised simultaneously by multiple ASes. Table I summarizes the results for both IP protocols.

In §VIII, we describe extensions to DISCO to address these issues.

IV. Overview

In this section we present an overview of DISCO's goals, components, and threats it protects against.

A. Goals

As we will show, "settling" for de facto ownership enables DISCO to meet the following design goals:

Security against prefix hijacks. DISCO is designed to protect against the most common and alarming prefix-hijacking attacks

[2]When constructing IP blocks for this analysis, we ignore announcements for IPv4 prefixes longer than /25 and IPv6 prefixes longer than /64 prefixes (e.g., announcements to blackhole traffic during DDoS attacks [32]). We also ignore IP prefixes that are announced less than 20% of the time.

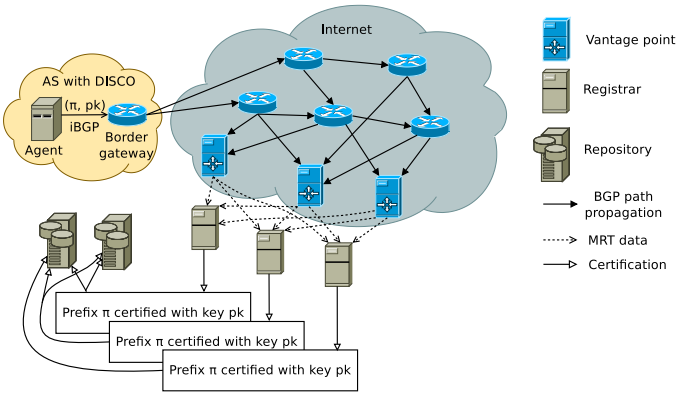


Fig. 2. Certification through DISCO. The agent associates public key pk with prefix π.

against BGP (like [1], [5], [63]). DISCO's design targets both *safety*—an attacker that is not the de facto owner of a prefix should not be able to certify ownership—and *liveness*—the legitimate owner should be able to certify ownership.

Do no harm (security-wise and performance-wise). DISCO should be safe to deploy. To accomplish this goal, DISCO's design is aimed at avoiding the ill effects of human error by automating certification. In particular, DISCO enforces *consistency* between certificates and the BGP control plane during initialization and delegation to avoid incidents resulting from human neglect in synchronizing the two. In addition, DISCO is carefully designed to not harm the performance of today's routing system (e.g., by slowing down the processing of route-advertisements, or by causing route-flapping).

Be readily deployable. We design DISCO to minimize the operational costs entailed in deploying and running the system. In addition, to facilitate deployment, DISCO is compatible with today's routing infrastructure and, in particular, it does not involve changes to BGP routers, and eliminates the need for an adopter to coordinate with other ASes.

B. System Components

Figure 2 illustrates DISCO's main components.

The agent. A software-implemented agent, installed within the AS, initiates the certification process for address blocks belonging to that AS by attaching the AS's public key to its BGP advertisements using BGP attributes. Once certified, the agent issues a ROA associating an AS number with its address block. By automating ROA issuance, DISCO eliminates the possibility of human error, which is common in RPKI's ROAs (§II). Given DISCO's certificates and ROAs, the agent computes the filtering rules for using ROV on advertisements from other ASes.

Registrars and vantage points. A registrar continuously monitors BGP advertisements from a distributed set of vantage points to obtain a global view of Internet routing. This information is used to generate and sign certificates associating owners' public keys with their IP address blocks. We envision registrars as deployed by different organizations that are distributed across geographical and political boundaries and use independent sets of vantage points, so as to avoid centralized control over global Internet routing. DISCO's design decouples the certifying entities (the registrars) from the vantage points that monitor BGP routing information. This separation allows DISCO to leverage available public sources of routing information, such as RouteViews and RIPE RIS, for initial bootstrapping of DISCO (as evidenced by our implementation and experiments).

Repositories. Similarly to RPKI, DISCO uses public repositories to store and distribute certificates and ROAs (also illustrated in Figure 2).

C. Threat Model

Alongside its original goal of protecting ASes from prefix hijacks, DISCO must also protect from attackers that seek to exploit DISCO itself to adversely influence Internet routing. Thus, DISCO is designed to protect against three threats:

- An attacker in control of one or more ASes, which may choose to advertise in BGP a prefix it does not own from these ASes.
- An attacker that compromised a fraction of the DISCO registrars, which may attempt to falsely certify IP address blocks in DISCO. We assume that most of the registrars are available (for liveness) and honest (for safety). A registrar is honest if it follows DISCO's protocol and receives feeds from a set of vantage points that is mostly honest (i.e., most vantage points that feed the registrar provide correct reports of BGP announcements that they observe).
- An attacker in control of a fraction of the DISCO repositories. The system's repositories only store cryptographically signed objects (by a quorum of registrars or the owner of an address block) and are therefore trustless. We only assume that some honest repositories are available to guarantee liveness (so agents can receive DISCO's certificates and route authorizations).

We assume that agents run correct implementations of DISCO's protocol, that private keys of honest participants remain secret, and that standard cryptographic primitives such as signatures are secure.

V. Design

We next dive into the mechanics of DISCO's certification and its protection against routing attacks.

A. Ownership Certification

We describe the ownership certification procedure following Figure 2. The agent connects to the AS's border routers through iBGP sessions. By using iBGP, DISCO avoids changing the router software or network infrastructure, only requiring changes to router configuration to set up iBGP sessions. The agent uses iBGP to initiate the certification process by attaching the origin AS's public key to the AS's route advertisements. Specifically, DISCO uses a 32 byte BGP optional transitive attribute. Using an optional transitive

attribute makes DISCO compatible with the current Internet, since BGP speakers are meant to send such attributes to neighbors even if they do not recognize them [48]. We verify that this indeed occurs on the Internet in §VII. (Although BGP communities could, in principle, be an alternate means of propagating DISCO keys, they are sometimes filtered by intermediate networks [57].)

A registrar certifies an IP block owner when more than a predetermined fraction (the *certification threshold*) of the vantage points that provide it information observe the same public key attached to the BGP advertisements for an IP prefix for a predetermined period of time (the *certification interval*). Normally, an announcement carrying DISCO's attribute would propagate globally and be received at all vantage points.[3] The certification threshold prevents an attacker from obtaining a certificate by hijacking traffic to a few of the vantage points that feed a registrar, while providing robustness against availability errors (e.g., due to temporary connectivity issues or vantage point unavailability). The certification interval ensures that short-lived prefix hijacks (that propagate through the global Internet) cannot be used to obtain a certificate, but otherwise has minimal impact as legitimate owners can permanently announce prefixes with DISCO's attribute. When a registrar approves an owner for some block, it creates a certificate, signs it, and sends it to a public repository.

The repository collects signatures from registrars regarding IP-block-to-public-key associations. When more than a predetermined threshold of registrars approves the same association, the repository publishes the aggregate of the registrars' signatures on that association; this aggregate is the DISCO certificate. This threshold reflects a trade-off between liveness and security: a certificate should be generated even if a few registrars are down or refuse to sign, but should not be generated if only a few registrars sign. In practice, we expect most registrars to be up most of the time since they do not serve public requests and are hence less vulnerable to DoS attacks. We therefore believe that this threshold should be high.

1) Initial certification: DISCO can automatically generate certificates for the vast majority of the IP blocks, namely blocks that have *de facto* owners (about 97%, see Figure 1); this immediately allows the owners of these prefixes to issue ROAs. The remaining 3% of the IP blocks include (i) prefixes that recently changed ownership, which will undergo the certification process without issue; (ii) prefixes announced by multiple ASes, which require the multiple owners to coordinate deployment and announce an attribute with the same public key (§VIII-B); and (iii) prefixes with ongoing hijacks, which require the legitimate owner to coordinate with other operators to mitigate the ongoing hijacks before the certification process can succeed.

2) Continuous (re-)certification: DISCO's automated certification procedure has very little overhead (BGP advertisements only include an additional short optional attribute). This enables the owner to continuously run the certification procedure by incorporating DISCO's attribute in its BGP messages. Thus, DISCO registrars can issue short-lived certificates, which are renewed often (for example, every few weeks). Short-lived certificates have the advantage of avoiding long-term commitment to public keys, adapting quickly to changes in ownership, and simplifying revocation procedures (discussed in §V-A6).

[3]Some ASes export *partial feeds* to RIPE RIS and RouteViews route collectors (i.e., only export routes from customers) [49]. These ASes will not export routes received from peers and providers. They can be identified by characterizing their exports or by RIPE RIS and RouteViews (while coordinating with the AS to establish the BGP session) and thus not considered when verifying the certification threshold.

PARAMETER	DEFAULT	DISCUSSION
Certification threshold	95% of VPs	§V-A
Certification interval	2 weeks	§V-A
Owner control during interval	80%	§V-A
Registrar consensus	80% of registrars	§V-A, §VI-B

TABLE II. DISCO PARAMETERS, DEFAULT VALUES, AND REFERENCES

3) Certification under attack: DISCO cannot protect against hijacks when a prefix is uncertified or not covered by a ROA, similarly to RPKI. This is similar to the *initial certification* of prefixes without *de facto* owners (§V-A1). In cases of initial certification and partial ROV adoption, ASes close to the hijacker and that have not deployed ROV need to be contacted to mitigate the hijack—similar to how prefix hijacks are dealt with in the Internet today.

Prefixes with certificates and active ROAs are harder to attack, as continuous recertification happens under the protection of ROV. When ROV is partially adopted, it partially mitigates prefix hijacks, and recertification is more likely to succeed as ROV adoption increases (incremental benefit during partial deployment). When ROV is widely adopted, prefix hijacks are mitigated, and recertification is guaranteed to succeed.

DISCO must ensure that an attacker cannot prevent issuance of a certificate by hijacking the prefix. To accommodate initial certification and the interim where most ASes do not perform ROV, registrars certify an organization who receives traffic for an IP address block throughout *almost* all of the certification interval (*owner control* in Table II).

4) Setting certification parameters: Table II summarizes DISCO's certification parameters. To set the certification threshold, we use real-world experiments and simulations in §VII. To set the certification interval and what portion of it an announcement must be visible for to establish de facto ownership, we consider the length of past widespread prefix hijacking incidents, which typically last up to a few hours [63]. Finally, the fraction of system registrars that should approve a certificate reflects the high availability we expect from these services and expected security against malicious registrars.

5) Certificates with exclusions: Since DISCO validates de facto ownership, we must handle the case where a prefix is allocated to organization A, but its sub-prefix belongs to organization B. In that case, A should not be certified as the de facto owner of the entire prefix it advertises in BGP, but rather as the owner of all IP addresses in that prefix except those in the sub-prefix announced by B. DISCO supports this scenario by extending RPKI's resource certificate format to specify excluded sub-prefixes, making it possible for A's certificate to exclude B's IP addresses. This encoding is efficient; the number of IP prefixes specified in all certificates is no more than twice that of all prefixes announced through BGP (every prefix can only be included once in an owner's certificate and once as an exclusion). DISCO registrars generate the

exclusions automatically, based on sub-prefixes visible at their vantage points. A prefix and list of subprefix exclusions define an IP block in DISCO certificates.

6) Revocation: If a private key is exposed, an AS may need to revoke its certificate and not wait until it expires so as to prevent the attacker from issuing ROAs. In this case, the owner issues a "revocation request" signed with the private key associated with the certificate being revoked. Repositories store revocations until the certificate that is being revoked expires, and the short-lived certificates mean that revocation lists do not need to be stored for long. They distribute revocations to agents along with DISCO's certificates and ROAs, so any agent that syncs with a repository would discard revoked certificates it received in the past.

The distributed nature of trust in DISCO allows registrar keys to be revoked without invalidating certificates. Similarly to certificate revocation, a registrar can also sign a message that revokes its public key, and store the revocation at the system's repositories. Importantly, revocation of a registrar's key does not imply invalidation of DISCO-issued certificates. If a certificate still has a number of signing registrars higher than the certificate threshold, it can still be considered valid. The short certificate lifetime also means that the owner will refresh its certificate soon after a registrar changes its key (obtaining a new signature and updated registrar keys). So DISCO allows to gradually replace registrar keys (routinely or in case of compromise).

7) Prefix transfers and delegations: DISCO supports secure prefix transfers and delegations by allowing the previous prefix owner (with a DISCO certificate) to generate a ROA allowing the new owner to announce the transferred prefix or delegated subprefix. This ROA would allow the new owner to announce the prefix and obtain a DISCO certificate. In the case of transfers, a revocation request (§V-A6) against the previous owner's certificate is possible but optional as certificates expire quickly. In the case of subprefix delegations, DISCO ultimately issues a new certificate for the delegator with a hole for the delegated subprefix (§V-A5).

8) Illegitimate certificates for IP prefixes that are advertised in BGP: Our simulation results in §VII show that for IP prefixes that are advertised in BGP by legitimate owners, an attacker is highly unlikely to succeed in falsely certifying these prefixes (an expected success rate of 3%). Moreover, in the vast majority of successful attacks (81%), the attacker is the sole upstream provider of the victim, and can thus observe and intercept all of the victim's traffic without attacking DISCO (and risking exposure). In scenarios where an attacker that is not the sole ISP of the victim succeeds in false certification (0.6% in our simulations), measures such as those discussed in §VIII for revoking certificates for IP prefixes not advertised in BGP can be applied.

9) Addressing conflicts between DISCO *and RPKI certificates:* To accommodate fast and incremental deployability, DISCO, as described here, does not rely on RPKI certification and can be deployed in parallel as an independent certification system. We believe that the question of which policy to apply when conflicts arise between DISCO-issued and existing RPKI certificates merits further discussion. This question is particularly relevant in scenarios in which *de facto* certification is inherently limited (see §III-B). In such scenarios, DISCO adopters could possibly automatically ignore ROAs that are incompatible with existing RPKI certificates.

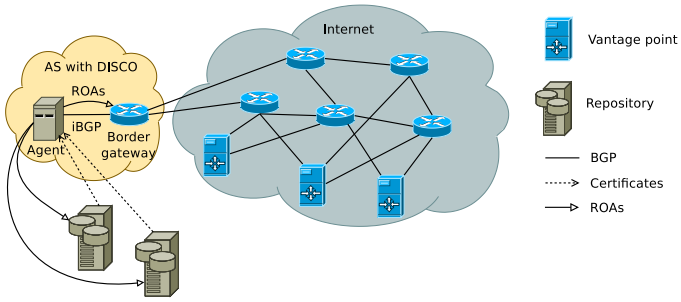


Fig. 3. Origin validation with DISCO. The agent issues a ROA, and fetches ROAs issued by others.

B. Origin Authorization

DISCO uses its certificates to create route origin authorizations (ROAs), as illustrated in Figure 3. DISCO's ROAs are conceptually similar to those of RPKI: they are signed by the owner's private key, and include a list of approved IP prefixes with maximum length and an authorized origin AS number for each prefix. The ROAs in DISCO extend RPKI's format to allow for the exclusion of sub-prefixes which belong to others (as indicated by corresponding exclusions defining the IP block in the certificate). These exclusions conform with the "wildcard-ROA" format previously suggested to address some of RPKI's deployment problems [21]. Exclusions are necessary since in many cases an RPKI ROA issued for one organization may invalidate legitimate BGP advertisements by other organizations (§II).

1) Reducing human involvement: The DISCO agent can operate under two modes. In the first mode the agent does not have access to the private key (e.g., the network operator may want to keep the private key offline). In this mode, the agent automates ROA generation to the extent possible (without signing it using the private key). When the agent observes a new certificate in DISCO's repositories for one of its IP blocks, it computes the ROA that the network operator then needs to sign using the certified private key. The agent creating the ROA configures the maximum length for each prefix automatically to avoid the pitfalls of manual configurations. More precisely, the agent gathers all prefixes that the AS advertises and computes the shortest list of prefixes and maxLength combination that exactly covers the IP block (by running Algorithm 1 from [24]). The agent specifies this prefix list in the ROA, and each prefix in the list is potentially followed by excluded sub-prefixes (not owned by the AS). The agent then sends the proposed ROA to the operator (e.g., via email), and the operator may edit it before signing with the private key and returning to the agent. For example, the operator may edit the ROA to allow for a not-yet-announced subprefix which supports traffic engineering. DISCO also supports an automated mode, where the agent stores the private key and signs the ROA without involving the operator at all.

Under both modes of operation, DISCO's agent automates ROA generation to the largest extent possible. Compared to RPKI, DISCO frees the operator from deciding which prefixes

should be listed, which maxLength to allow for each prefix, and what are the origin ASes. Once the agent has the signed ROA, it stores the ROA on DISCO's public repositories which verify that the ROAs are valid by checking the signature.

2) Route origin validation: The agent periodically checks for new ROAs and certificates at DISCO's repositories. Since DISCO repositories are untrusted, the agent also validates that the new ROAs are valid, corresponding to registrar certificates. For each prefix in a ROA, the agent configures the corresponding new filter at the AS's border routers. We describe the implementation details involved in configuring these rules into existing routers in §VI.

VI. Implementation

We built a prototype implementation of DISCO. Our implementation consists of less than 200 lines of non-library Python code for each of the agent and registrar. The prototype code is available online [30]. Our prototype attaches to BGP announcements an optional transitive attribute which consists of a 32-byte ED25519 public key. The number identifying this attribute is configurable in our implementation. Our tests running DISCO (§VII) use the 0xFF attribute which is reserved for experimental use to avoid interference with another standardized or squatted attribute types (see [56], [55]) during experiments. We acknowledge the need for standardizing the use of this BGP attribute type before DISCO can be widely adopted.

A. Agent

The system's design offloads the certification logic from the border routers to a local machine running a software agent. A network operator installs the agent to certify ownership over its prefixes through DISCO and to enforce validation of other prefix origins. We next discuss how the agent communicates with the AS's border routers to achieve these goals.

Running certification using iBGP. Our implementation uses iBGP to interface the agent with the AS's border routers. The network operator configures the agent with the AS's public key (the network operator may keep the corresponding private key offline for security). Through iBGP, the agent takes over generating the announcement from the router and appends DISCO's attribute. This approach decouples the existing routing architecture from DISCO, which saves complexity at the router and makes DISCO readily deployable.The following example shows a configuration on a border router of `AS65535` that runs the DISCO agent on a server with IP address `192.168.10.10`, configured to announce prefix `172.16.0.0/24` with the proper key in the optional transitive attribute:

```
router bgp 65535
  // DISCO agent connection
  neighbor 192.168.10.10 remote-as 65535
  // disable local announcement
  no network 172.16.0.0 mask 255.255.255.0
```

Enforcing origin validation. In addition to certifying the AS's IP blocks, the agent is also responsible for configuring the AS's border routers to enforce route origin validation. The agent has a list of public repositories from which it periodically syncs. Once it observes a new DISCO ROA, it validates its certificate and then creates a filtering rule for each prefix in the ROA. The agent configures the border routers to enforce ROV using a standard access list interface. For example, enforcing a ROA for IP prefix `10.0.0.0/22` with maxlength 24 and origin AS number 1 is achieved through the following access list:

```
// allow 10.0.0.0/22 maxlen 24, enforce last AS number is 1
ip as-path access-list rov seq 1 permit 10.0.0.0/22 le 24 1$
```

After all of DISCO's rules (for all prefixes), the agent adds another rule that denies advertisements for `10.0.0.0/22` and its subprefixes that were not captured by earlier rules:

```
// deny other advertisements for 10.0.0.0/22
// or its subprefixes
ip as-path access-list rov seq 100 deny 10.0.0.0/22 le 32
```

Supporting address blocks. Assume that, in addition, the ROA covering `10.0.0.0/22` has an exclusion for subprefix `10.0.2.0/23`, and no ROA covers that subprefix. DISCO's filtering rule would exclude that prefix by adding the following "permit" access control entry between the permit and deny filters above:

```
// allow 10.0.2.0/23 from any origin and set maxlen 32
// to allow any subprefix of any length
ip as-path access-list rov seq 2 permit 10.0.2.0/23 le 32
```

We view the above implementation through the BGP router's access control list interface as a temporary bridge that allows using DISCO with today's routers. For the longer term, we expect a protocol similar to RPKI-to-Router [11] to allow configuring filters according to DISCO's ROAs. Such a protocol provides a simple interface allowing the operator to configure a local machine that provides origin validation rules. RPKI's RTR protocol only needs a modest extension for subprefix exclusions in order to support DISCO.

B. Registrars

DISCO's registrars are machines operated by different organizations, e.g., RIRs or reputable network providers, to decentralize trust in the system. A registrar receives BGP feeds from vantage points at different locations on the Internet. The vantage points that each registrar uses are configurable and may reflect the registrar's administrator trust assumptions. DISCO requires a threshold of registrars to agree on a certificate, so, even if a few registrars are malicious or make a bad decision choosing their vantage points, the certified information recorded in DISCO's repositories maintains integrity. We note that if a registrar discloses its list of vantage points and certification threshold, and if the vantage points provide public feeds, then anyone can verify the certificates issued by that registrar. This effectively limits a malicious registrar's capability to issue multiple incorrect certificates, as a malicious (or misbehaving) registrar can be identified immediately after it incorrectly issues a certificate. Lying about the set of vantage points used or the certification threshold is futile, as it still defines (and allows verification of) what certificates the registrar should issue. Finally, if the list of vantage points or the certification threshold used by a registrar are untrustworthy, it may be removed from the list of registrars.

Our implementation's vantage points. We focused on creating a readily deployable system. Therefore, we opted to use as vantage points 262 routers that peer with RouteViews and RIPE RIS [50], [62], two publicly accessible BGP advertisement collection systems. RouteViews and RIS provide the content of the full BGP advertisements they receive (including the optional transitive attributes) in the MRT [9] data format. The list of vantage points is configurable in our implementation, so operators running the registrars can add new vantage points to improve the visibility of the Internet.

Incentives for adoption. Our implementation runs on commodity hardware, relies entirely on open source components, and uses only public data. We argue that the low adoption cost, large number of interested parties, and potential positive impact may be sufficient to drive deployment of registrars. We note that even a handful of registrars would be enough to provide reliability against registrar failures in DISCO (as long as the fraction of honest registrars is higher than the registrar consensus threshold).

VII. Evaluation

We evaluate DISCO in two respects. First, we evaluate DISCO's compatibility with today's Internet (§VII-A). We perform this evaluation by deploying DISCO and advertising a prefix with its attribute through the PEERING platform [52], [53]. Second, we evaluate DISCO's security using simulations of different types of attacks on an empirically derived dataset of the Internet AS-level topology (§VII-B). The code for our simulations and measurement analysis is available online [30].

A. Compatibility with today's Internet

DISCO's strategy for distributing public keys through optional transitive attributes aims to be compatible with today's Internet. Routers need not understand these (optional) attributes but are still expected to forward them onwards to other networks (transitive); hence they provide a useful mechanism for extending BGP [48]. In practice, however, BGP implementations might violate the protocol's specification and filter unknown optional transitive attributes (as indeed happened during the standardization of BGP large communities [55]). We describe below how we evaluate the propagation of DISCO's new public key BGP attribute and provide evidence that such a new BGP attribute would reliably propagate to the global Internet.

1) Experiment setup: We performed two experiments to quantify the propagation of BGP announcements carrying DISCO's attribute from the PEERING platform [52], [53]. We shared our experiment proposal with a few operators for feedback to make sure it had community support and was seen as safe and useful, integrating their feedback into the experiment design. We then submitted the proposal to the PEERING operations team, who approved it. PEERING did not previously support the ability to use such attributes. We extended the platform to support per-experiment capabilities, such that our approved experiment could use an optional transitive BGP attribute but other experiments could not, in keeping with the principle of least privilege. We announced a "DISCO" prefix with an optional transitive attribute carrying DISCO's 32-byte owner's public key. We used attribute `0xFF`, which is reserved by the BGP specification for experimental use. We also announced a "control" prefix without unknown attributes to track the routes ASes use to reach PEERING in normal circumstances.

The first experiment announced the prefixes from PEERING's point of presence at UFMG, Brazil, and the second announced the prefixes from PEERING's point of presence at the University of Washington (UW), USA. We limited the duration of the announcements to 15 minutes each.

DISCO's BGP announcements conform with the specification but are unusual compared to the BGP announcements that typically propagate the Internet. We attempted to limit the potential adverse impact on the global routing system in the case that a BGP implementation does not comply with the BGP RFCs, or contains a bug triggered by our BGP announcements. We disseminated the experiment plan on the NANOG operator mailing lists (which was then forwarded to other operator mailing lists), giving advance notice of the experiment schedule. We gave operators ample time between the experiments to identify and report any issues. Finally, we tested successful propagation of our announcements in controlled environments using Cisco IOS-based routers running versions 12.2(33)SRA and 15.3(1)S, Quagga 0.99.23.1 and 1.1.1, as well as BIRD 1.4.5 and 1.6.3.

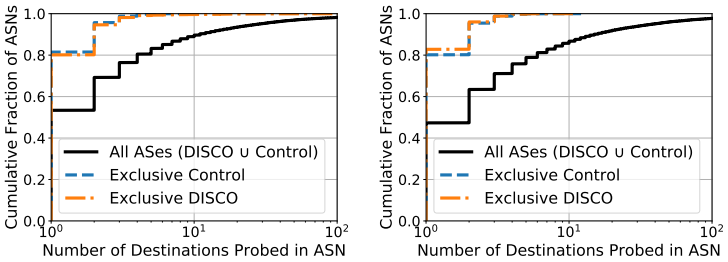
2) Measuring DISCO*'s effect on reachability:* We first evaluate whether the presence of the attribute on the DISCO prefix announcement impacts whether Internet destinations can reach the prefix. That would be the case if many routers filtered announcements with unusual attributes, leaving destinations "behind" them without a route to the prefix. During our experiments, we used `zmap` [17] to send ICMP Echo Requests to a destination hitlist built from ISI's Internet census data from Nov. 2018 [18]. We target a list of 5,651,501 IP addresses which includes the IP addresses with the highest response rate in responsive /24 prefixes during ISI's census measurements (we ignore /24 prefixes without any IP address with at least a 10% response rate in ISI's census). We identify the set of ASes responding to pings from the DISCO and control prefixes (denoted $\mathcal{A}_{\text{DISCO}}$ and $\mathcal{A}_{\text{control}}$) by mapping responding targets to their ASes. We consider an AS as responsive if at least one address in the AS responded.

Table III presents results indicating that the DISCO and control prefixes have equivalent (global) reachability. The average AS-level response rate during the UFMG and UW experiments is around 47% and 66%, respectively. This low response rate is a result of scattering probes across a large set of targets which may have gone offline, turned unreachable, or stopped responding since ISI's Internet census [18]. The increase in response rate from the UFMG to the UW experiment is explained by `zmap` overloading our VMs during the UFMG experiment and dropping response packets; we reconfigured the VMs for the UW experiment.

Although the overall response rate is low, it is similar for the DISCO and control prefixes. The number of ASes that responded only to the DISCO prefix is roughly equal to the number that responded to the control prefix only (the *exclusive ASes* columns). Two factors outside attribute propagation can cause responsiveness to one prefix and not another. First, `zmap` probes in a random order, which can combine with delayed

	UFMG announcements			UW announcements		
	AS Response Rate	Responding ASes	Exclusive ASes	AS Response Rate	Responding ASes	Exclusive ASes
DISCO pings	47.10%	27794	7727	64.65%	38150	7542
Control pings	47.01%	27738	7671	66.80%	39418	8810

TABLE III. SUMMARY OF ZMAP PING MEASUREMENTS DURING PEERING EXPERIMENTS. *Exclusive ASes* INDICATES THE NUMBER OF ASES THAT RESPONDED TO PINGS FROM THE DISCO PREFIX BUT NOT THE CONTROL PREFIX, OR VICE VERSA.



(a) UFMG Experiment (b) UW Experiment

Fig. 4. Comparison of reachability of test DISCO prefix vs control prefix as a function of the number of targets. Most of the ASes that respond to one prefix but not the other host few destinations and are thus more susceptible to measurement errors.

route convergence to cause a target to be probed from one prefix before the route converges and from the other prefix after the route converges. Second, the measured responsiveness of an individual destination can vary due to ICMP rate limiting and packet loss. Finally, the BGP session used to announce the DISCO prefix during the UW experiment was temporarily down due to a flap of the OpenVPN tunnel used to connect to the PEERING router. This temporary disconnection and the subsequent BGP session reset may have negatively impacted propagation of announcements for the DISCO prefix and data plane reachability. This may explain (at least part of) why reachability on the DISCO prefix during the UW experiment is 2.15% lower than on the control prefix.

In Figure 4 we present our reachability results as a function of the number of target destinations in each AS in our dataset as further support to the conclusion that the DISCO and control prefixes are similarly reachable. The "all ASes" line in Figure 4 shows the distribution of the number of targets probed in all ASes with responsive destinations. Most ASes own only a small number of /24 prefixes and thus host only a few destinations in our hit list, but some ASes own tens or hundreds of /24s and host a large number of destinations. The figure compares this line against the distributions of the number of targets probed in ASes where we observe responses on only one of the prefixes (i.e., ASes in the *exclusive ASes* sets). We notice that these distributions are heavily skewed to the left, indicating that, usually, the AS sets differ in ASes where the number of targets for our probes was low. A low number of targets implies a higher probability for measurement errors (e.g., when a target does not respond to an ICMP Echo Reply probe or a packet is lost) due to convergence delay (e.g., when a target is probed from one of the prefixes before convergence has completed). Such errors lead to responses for one prefix but not the other, even when both DISCO and control prefixes are utlimately reachable from the AS.

3) Identifying filtering on the control plane: We check the propagation and attributes of routes used to reach the control and DISCO prefixes by downloading BGP updates from routers around the world collected by the RouteViews and RIPE RIS projects. We compare routes toward the control prefix and the DISCO prefix to identify sets of candidate ASes that could be filtering the announcement of the DISCO prefix (denoted $\mathcal{F}$) or discarding DISCO's attribute and forwarding the announcement without it (denoted $\mathcal{D}$). We apply the following rules, in order, to estimate $\mathcal{F}$ and $\mathcal{D}$ for each experiment:[4]

1) For each router R that exports a route toward the control prefix to a BGP collector, we add to $\mathcal{A}$ all ASes in R's route toward the control prefix. These represent the set of ASes we might expect to see in routes to the DISCO prefix. We initialize $\mathcal{F}$ to this full set $\mathcal{A}$, and the following rules remove the ASes we do see in routes to the DISCO prefix, leaving candidates that may be filtering.
2) For each router R that exports a route toward the DISCO prefix without the DISCO attribute to a BGP collector, we add to $\mathcal{D}$ all ASes in R's route toward the DISCO prefix. We do not remove ASes from $\mathcal{F}$ as we cannot know if ASes after the one discarding the attribute would have filtered the announcement.
3) For each router R that exports a route toward the DISCO prefix with the DISCO attribute, we remove from $\mathcal{D}$ and $\mathcal{F}$ all ASes in R's route toward the DISCO prefix.

Although route convergence in the Internet's control plane usually takes on the order of 3 minutes, some cases can take more than 15 minutes to converge [34]. To avoid considering transient routes observed during route convergence, we require that a route remains stable for 5 minutes during the execution of the experiment. If a router exports two routes that satisfy this condition during an experiment, we consider both routes.

For ASes in the converged routes to the control prefix (step 1), we check whether we observe them in any route to the DISCO prefix (steps 2 and 3). If we observe an AS on a route (transient or otherwise) to the DISCO prefix, it is proof that the AS does not filter the prefix. And if we observe an AS on a route to the DISCO prefix with the DISCO attribute, it is proof that the AS does not discard the attribute.

The columns under "BGP updates" in Table IV summarize our results. We show $|\mathcal{F}|$ and $|\mathcal{D}|$, and the number of routers that export paths to the DISCO and control prefixes. The results indicate that the great majority of ASes receive DISCO's prefix with its attribute. If we observe an AS in a route with the attribute in one experiment, it indicates that the AS does

[4]The rules assume that all routers in an AS behave uniformly. Although this is not always true for routing decisions [43], we expect filtering to be more uniform across an AS.

Announcement Locations	BGP updates (§VII-A3)						Traceroute (§VII-A4)				
	$\lvert\mathcal{A}\rvert$	$\lvert\mathcal{F}\rvert$	$\lvert\mathcal{D}\rvert$	number of routes			$\lvert\mathcal{A}\rvert$	$\lvert\mathcal{F}\rvert$	number of traces		
	control			DISCO	control	$\neq$ paths	control		DISCO	control	$\neq$ paths
UFMG	295	4	2	359	364	36	1656	14–24	1196	1198	172
UW	292	5	2	328	330	8	1684	8–27	1196	1196	58
Joined	—	3	2	—	—	—	—	3–3	—	—	—

TABLE IV. EVALUATION OF FILTERING RESULTS DURING PEERING EXPERIMENTS.

not filter on the attribute, even if we did not observe the AS in the other experiment.

A single AS exports routes toward the DISCO prefix without DISCO's attribute. This route traversed 2 ASes that are candidates for having discarded the attribute (i.e., $\mathcal{D} = \{\text{AS16150}, \text{AS48285}\}$). As other routes never traverse these 2 ASes, we cannot identify which one is discarding the DISCO attribute. AS48285 peers directly with a route collector, and hosts BGP and DNS security-related databases. AS16150 is a transit provider for AS48285 and has been bought and is currently being merged into a larger transit provider AS12552, which we observe forwarding announcements with the DISCO attribute, suggesting the DISCO attribute may not be dropped after the transit merger is complete.

By combining the results from our two experiments in this way (the "Joined" row in Table IV), we flag the two ASes in $\mathcal{D}$ (discussed above) and a single additional AS as a candidate for filtering the DISCO prefix. We have confirmed with the operators of the additional network that they do *not* employ any filtering of optional-transitive attributes; instead, the network uses the FRR software router and, due to a bug in the FRR software, experienced instabilities during our experiment (§VII-A5).

4) Identifying filtering on the data plane: We use RIPE Atlas to run traceroute measurements to the DISCO and control prefixes and identify filtering ASes. We convert traceroutes to AS-level routes by mapping IP addresses to AS numbers using Team Cymru's IP-to-AS database [59]. We then use the inferred AS-level route to repeat the analysis in §VII-A3. Before the experiments, we run measurements from all RIPE Atlas vantage points to each PEERING announcement location, then greedily choose vantage points to use during the experiment so as to maximize the total number of ASes traversed on routes to the PEERING location. We run traceroutes from 1600 locations over 20 minutes to conform with the rate limit that RIPE Atlas enforces.

Columns under "Traceroute" in Table IV summarize our results. We show the number of traceroutes collected, the number of ASes covered on traceroute measurements toward the control prefix ($|\mathcal{A}|$), and the number of ASes that may be filtering ($|\mathcal{F}|$). Since we lack information about BGP attributes in the traceroute measurements, we cannot measure which ASes discard DISCO's attribute when relaying the advertisement (i.e., we cannot compute $|\mathcal{D}|$); these ASes maintain connectivity and are therefore counted toward the $|\mathcal{A}|$ column. We observe measurements from approximately 1200 RIPE Atlas probes (in the columns under "number of traces"), which is what we expect given the measurement period and experiment duration ($1200/1600 = 15/20$).

We compute $\mathcal{F}$ as before, and then remove from $\mathcal{F}$ ASes which replied to our zmap data-plane measurements, since we know they had a route to the DISCO prefix. ASes with a default route may filter the DISCO prefix and still respond to pings [12]. To correctly account for these cases, we infer ASes that use default routes (Appendix A) and report a range: the lower end assumes that no AS inferred to use a default route filters the prefix, and the higher end assumes that all ASes inferred to use a default route filter the prefix.

We find that the number of ASes seen on routes toward each prefix is similar and that $\mathcal{F}$ is small, providing further indication that only a few ASes are likely to be filtering the announcement to the DISCO prefix. The number of ASes in $\mathcal{F}$ grows in comparison to the previous control-plane based experiment (§VII-A3) since unresponsive routers and incorrect IP-to-AS mappings may increase the number of candidates for filtering ASes flagged by our analysis, and the broader coverage will accumulate more errors.

The "$\neq$ paths" column shows the number of ASes that choose different routes to the DISCO and control prefixes, with the condition that the chosen route toward the control prefix *does not* intersect $\mathcal{F}$. This column counts ASes that choose different routes to our prefixes, but for reasons unrelated to filtering. For example, ASes can choose different routes for different prefixes due to tie-breakers in BGP's best path algorithm, like preferring the oldest among multiple equally-preferred routes [4]. ASes that choose different routes to our prefixes could cause false positives when inferring candidates for filtering. We expect that our inferred candidates are unlikely to be performing filtering (and instead have been labelled as candidates due to routing decisions of other ASes).

Finally, the "Joined" row shows results when we consider AS-paths from both the UFMG and UW experiments when computing $\mathcal{F}$. The decrease in $|\mathcal{F}|$ indicates that most ASes flagged as candidates for filtering in one experiment appear in a route to the DISCO prefix on the other experiment.

5) Router support for DISCO *and the FRR incident:* The results of our experiments, summarized in Table IV, show that the DISCO prefix had disseminated across the network to a similar extent as the control prefix, indicating that today's border gateway routers can support the DISCO protocol.

One notable exception is the case of FRR software routers. These routers had used the 0xFF attribute (reserved by the BGP standard for experiments) to communicate internal state among several FRR routers in the same AS. Our experiment used the 0xFF attribute to carry the DISCO prefix, which is protocol compliant but did not conform with the encoding an FRR router expects. As a result, FRR routers reset their sessions upon receipt of the attribute. FRR is the only reported router distribution affected by our experiments.

Upon receiving notification of the issue after our UFMG

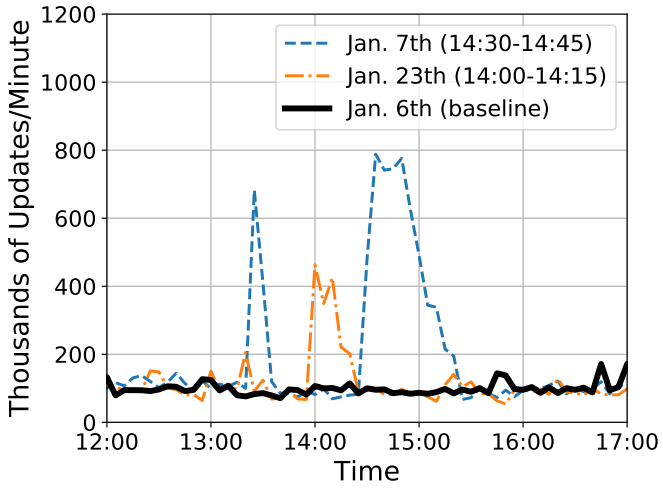


Fig. 5. Aggregate rate of prefix updates over time during experiments with a BGP unassigned attribute. Update peaks in the periods indicated in the legend correspond to disruption. The spike around 13:30 is unrelated to our experiment, but of similar magnitude. The reduced spike on Jan. 23rd indicates operators deployed software patches.

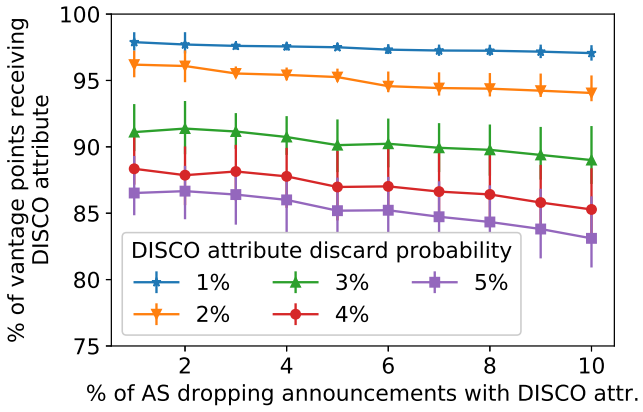


Fig. 6. Propagation of DISCO's attribute to the vantage points. DISCO's attribute propagates to most vantage points for a wide array of filtering scenarios.

announcement, we immediately put our experiments on hold, added FRR to our controlled testing environment, and coordinated with FRR developers and network operators. The FRR developers issued a patch within two days of this incident (CVE-2019-5892 [46]), and we confirmed that it fixes the bug triggered by our experiment. We postponed the experiments to allow for a two-week upgrade window after the release of the FRR updates. We received new reports of disruption after resuming the experiment from UW and decided to cancel more experiments from other PEERING locations. As time passes, we believe that this patch will disseminate to the vast majority of FRR deployments.

Figure 5 shows the number of BGP updates received by RIPE RIS and RouteViews route collectors during the execution of our experiments. The UFMG experiment ran on Jan. 7, 2019, between 14:30 and 14:45 GMT and the UW experiment ran on Jan. 23, 2019 between 14:00 and 14:15 GMT. We also show lines for Jan. 6, 2019, as a baseline. We note that the spike in the number of updates around 13:30 on Jan. 7 is unrelated to our experiment but of a similar magnitude. We can see an increase in the number of updates during our experiments, which we attribute to FRR routers. However, we also see that our second experiment had a much smaller effect, indicating that the patch was getting adopted. We believe that this incident does not mean that DISCO is incompatible with today's Internet but rather reflects a bootstrapping cost that will diminish as operators using the FRR routing daemon upgrade and adopt the existing patch.

After the incident, most operators who sent messages to the NANOG mailing list expressed support for continuing the experiment, arguing that the announcements comply with BGP standards and that operating routers with known remotely exploitable bugs is a severe vulnerability. A small number of operators expressed concerns that research should be careful not to impact Internet operations, which we agreed with and tried to implement by coordinating with operators, executing tests in controlled environments, and changing the experiment schedule upon receiving the report of the problem. We discuss safe re-execution of our experiments and, more generally, similar experiments extending BGP in Appendix B.

6) Summary: Despite experimental challenges (BGP session resets, VM overload, and FRR failure) and a limited number of experiments, the results indicate no significant difference between the propagation of a plain announcement and an announcement carrying a custom BGP attribute. Our announcements with the DISCO BGP attribute propagated widely, and possibly globally, evidence that DISCO may be readily deployable. The main conjecture we make is that a standardized BGP attribute would allow prompt deployment of DISCO. However, our results are based on limited measurements, and it is desirable to further confirm them using additional experiments and analyses, in particular additional measurements exploring the topology from more additional announcement locations.

B. DISCO *Security Evaluation*

Our security evaluation focuses on DISCO's certification mechanism. For DISCO to provide a reliable source of information for filtering malicious announcements, we need to show that, (i) under normal conditions (no attack) the de facto owner can obtain a certificate with high probability, and that (ii) an attacker's attempt to obtain a certificate for a prefix it does not control is likely to fail. For case (ii), we consider initial certification as a worst-case scenario as prefixes with a certificate and active ROAs are harder to attack.

To perform our evaluation we build on an existing BGP simulator [15] that uses the BGP route-computation framework by Gill et al. [25]. We extend the simulator to mark ASes as DISCO's vantage points and provide statistics to the routes they observe. Our simulator uses the 262 AS vantage points afforded by RouteViews and RIPE RIS which feed our implementation (§VI). We run simulations on the empirically-derived CAIDA AS-level graph from August 2019 [2].

1) Setting the certification threshold: We evaluate how different rates for ASes discarding DISCO's attribute or filtering

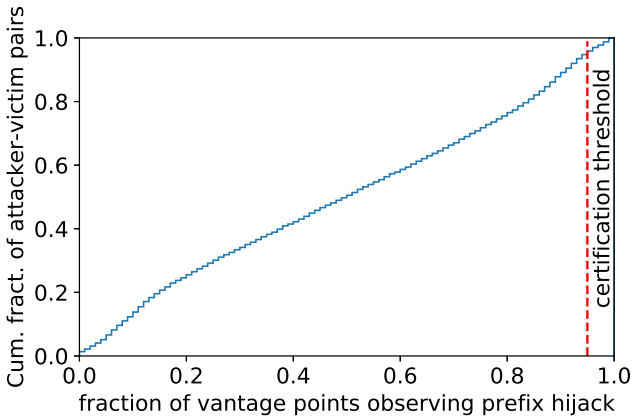


Fig. 7. Certification under prefix hijack (using vantage points provided by RIPE RIS and RouteViews). With 95% as the certification threshold, a registrar will not certify ownership under most hijacks.

its announcement affect the certification mechanism, assuming that a registrar feeds from the 262 vantage points mentioned earlier. In each iteration of the following simulation, we randomly select a different origin AS for the announcement, $1\% - 10\%$ of ASes chosen at random to drop announcements with the DISCO attribute (shown on X axis), and $1\% - 5\%$ of ASes chosen to discard the attribute (shown as different lines). Figure 6 shows the percent of vantage points that would observe DISCO's attribute (each data point is the average of 10^5 random iterations). Our Internet measurements in §VII-A indicate that less than 1% filter DISCO's announcement or discard its attribute. Under these conditions, Figure 6 shows that a certification threshold of 95% allows certifying ownership in all cases (topmost line).

2) Attacks on DISCO*:* DISCO issues certificates based on the public key attached to announcements that vantage points observe. To launch a successful attack, an attacker needs to hijack routes to the victim's prefix from many vantage points (the certification threshold) for the certification interval. Our simulations assume that the victim is not already protected by DISCO (or it would be protected from the attack).

Prefix hijack. We first consider a prefix hijacker who announces the victim's prefix. To evaluate the attacker's success rate, we select 10^5 random attacker-victim pairs where both attacker and victim announce the same prefix. We select 1% of ASes at random to filter DISCO announcements and 2% of ASes to discard the attribute. Figure 7 shows the fraction of vantage points whose routes to the victim the attacker succeeds in hijacking. A certification threshold of 95% as suggested earlier means that in about 3% of cases the prefix hijacker would succeed. The reason that the prefix hijacker typically fails is that the victim also announces the same prefix, so traffic splits between the victim and attacker and typically neither party achieves the required level of visibility at vantage points to certify ownership. Only 3% of hijacks successfully reach 95% of vantage points and obtain an illegitimate certificate. For 81% of illegitimate certificates, the successful attacker is the upstream provider of the victim. However, an upstream provider is in a position to intercept its customer's traffic (even without launching a BGP hijack) and manipulate their customers' BGP announcements,[5] and have no incentives to launch attacks against customers.

In case an attacker controls the vantage points that feed a registrar, crossing the 95% certification threshold is still very challenging. For example, even if the attacker controlled half of the vantage points that feed every registrar and had them report announcements with its public key for DISCO's attribute, the other half of vantage points would still report the announcement it received (as reported in Figure 7). The attacker would only succeed in certifying a block when less than 10% of the honest vantage points observe the legitimate route (in other words, when 90% of the honest vantage points choose the hijacked route), which happens for about 8% of the cases (Figure 7).

Safety is not very sensitive to the choice of vantage points. We evaluate the attacker success rate in simulations where the registrar uses a random subset of only 50 out of the 262 vantage points afforded by RouteViews and RIPE RIS. For each of 13 combinations of vantage points, we verify the probability that an attacker hijacks a prefix and successfully obtains a certificate by running 10^5 simulations varying the victim and attacker ASes. We find that, on average, the attacker fails to cross the certification threshold 99.3% of the time, and the standard deviation across the 13 combinations of vantage points is 0.5%. (We assumed all ASes propagate announcements with DISCO's attribute during these simulations.)

Subprefix hijack. The attacker may also launch a subprefix hijack. In this case, the attacker is the only one announcing the subprefix. Therefore, the attacker will hijack the routes from all vantage points to the subprefix. However, running such an attack will also disconnect the victim's subprefix from the Internet: the only route for the subprefix is to the attacker; hence even the attacker does not have a route to the victim AS and cannot relay to the victim intercepted traffic. Running a subprefix hijack for a long certification interval is highly visible and allows the victim to prevent certification by announcing the same subprefix which would create a scenario similar to the prefix hijack discussed above. As a result, DISCO significantly raises the bar for a successful attack. In §VIII we extend DISCO to limit an attacker's ability to certify, even if they launch a subprefix hijack.

Temporary hijacks. An attacker may aim to simply prevent DISCO (re-)certification by the legitimate owner (instead of obtaining an illegitimate certificate) by launching a temporary hijack during the certification period. This is a less damaging attack, but would prevent issuance of ROAs and undermine the benefits of DISCO. However, such an attack still requires a concurrent widely-visible hijack of the prefix that lasts long enough to reduce the owner's control of the prefix below the threshold (Table II, §V-A3). Such an attack does not limit deployability of DISCO and is impossible against certified prefixes when ROV is widely adopted.

[5]In particular, a provider can drop a client's DISCO attribute from BGP updates, preventing certification of that client's prefixes.

VIII. Handling the Limitations of De Facto Ownership Certification

A. Certifying Non-advertised Prefixes

Relying on de facto ownership makes it challenging to certify ownership over prefixes not advertised in BGP. An attacker can claim ownership of an unannounced prefix by starting to advertise it in BGP and initiating the DISCO certification procedure. Once certified, the attacker can issue itself a ROA for that prefix. Since such prefixes do not receive any traffic, rogue certificates would not threaten the Internet's immediate connectivity. But they allow the attacker to disconnect the owner later or hijack all their traffic at will, since, even if the legitimate owner starts advertising the prefix, ASes that use DISCO for ROV will discard its advertisements, violating the "do no harm" principle.

To prevent the issuance of such rogue certificates, we extend DISCO to leverage information recorded in Regional Internet Registry (RIR) databases about prefix allocations and the corresponding controlling organizations. RIR databases contain a reduced set of mostly static information maintained by the RIR itself (e.g., the allocations) and are only partially editable by network operators (e.g., organization and point-of-contact information). The RIR itself already specifies in its database which organization controls a prefix, establishing a link between org and inetnum objects.

Due to their authoritative nature and recurring resource allocation maintenance costs, inetnum and org objects are kept more up-to-date than Internet Routing Registries (IRR) objects used for specifying routing policies (e.g., route, route6, and AS-SET objects). IRRs are known for being incomplete, out-of-date, or outright incorrect.

Although RIPE maintains a single database that combines all these functionalities, operators have limited control over allocation (inetnum) objects, and so it is safe to use these objects as a basis for preventing rogue DISCO certificates. Specifically, an owner specifies the public key that it intends to certify for its prefixes in its organization's description in the RIR database. Operators can currently include the DISCO public key as part of their organization's address, which would be sufficient for the purposes of this extension, although database schemas could be updated with a dedicated field for storing DISCO public keys.

DISCO then compares the public key in the RIR database to the public key seen in an announcement. To claim a previously unadvertised prefix, the attacker would first need to obtain the credentials to modify the owner's organization entry in the database, then successfully advertise the prefix while attaching the DISCO attribute during the certification period. This raises the bar for false certification, allows time for the legitimate owner to react, and guarantees more visible traces of attacks. In addition, the legitimate owner of the IP prefix can void the false certificate by editing their information at the RIR database for that prefix at any time. Conceptually, the above extension extends the notion of de facto ownership to control of the authoritative information recorded in RIR as well as the routes propagating on the Internet.

As a security measure, DISCO uses only the information recorded in the database of a prefix's authoritative regional (RIR), local (LIR), or national (NIR) Internet registry; this requires querying multiple databases to identify which one is authoritative for a given prefix, particularly for legacy resources. If a prefix is not assigned to an organization in any database (e.g., not allocated), then DISCO does not issue a certificate for that prefix. Although this extension does not protect prefixes missing from databases or defend against compromised point-of-contact and organization records (e.g., point-of-contact e-mail addresses whose domains have expired and been reregistered by third-parties), neither does RPKI.

B. Certifying Multiple Origins for the Same IP Address Block

As discussed in §III-B, on fairly rare occasions, the same IP prefix is advertised by multiple ASes. To validate de facto ownership of a prefix with multiple origin ASes, the ASes should use the same public key when advertising the prefix. Thus, DISCO vantage points will observe this key on all BGP route-advertisements for that prefix, and de facto ownership will be verified. This can be accomplished *without* sharing the corresponding *private* key; after de facto ownership is verified, the owner of the private key can create ROAs authorizing all ASes to announce the prefix.

IX. Discussion

We next discuss alternative approaches to performing de facto ownership validation and explain why they fall short in meeting DISCO's goals (§IV). We conclude with a discussion on using DISCO as a basis for further validation of BGP paths.

Validating ownership of the entire IP prefix *vs.* validating ownership of a small set of IP addresses. As discussed in §III, de facto ownership is a strictly weaker desideratum than the traditional goal of binding IP prefixes to their legal owners. Settling for this weaker goal is intended to enable the design of deployable solutions while not losing "too much" in terms of security. A natural question is thus whether de facto ownership of the certified address block can be further weakened to achieve similar goals. Consider, for instance, a certification scheme in which, when certifying an IP prefix, the party being certified needs only prove control over a small set of IP addresses within the IP prefix, or even just a single IP address (e.g., of a Web, DNS or mail server). Such a scheme would be easy to deploy but would be more vulnerable to manipulation. Someone with control over few IP addresses, e.g., by 'renting' them from a hosting provider, or taking over an end-host, may be able to abuse such a mechanism to claim ownership of an entire address block.

Control plane *vs.* data plane certification of de facto ownership. DISCO could, in principle, be replaced by a certification mechanism that validates de facto ownership of an IP prefix by requiring the alleged owner to respond to "challenges" sent to addresses in the prefix from multiple locations over a sufficiently long time period.

One shortcoming of data plane validation of de facto ownership is the need to intercept (e.g., at a firewall or border router) all challenges sent to IP addresses in the address block being certified and forward these to a location from which the responses are sent. DISCO, in contrast, adds an attribute

to existing BGP advertisements, thus avoiding the need to actively intercept and forward challenges.

Data plane approaches are also vulnerable to *stealthy manipulations*. Someone capable of intercepting the challenge packets, e.g., via BGP prefix hijacking, could respond to them and so establish ownership over another's IP address block. A sophisticated attacker could also ensure that all other traffic to the legitimate owner of the address block safely reaches its destination (and so not arouse suspicions). DISCO, by design, uses many vantage points, and so attacks on it cannot be successful unless they are widely visible.

Beyond origin validation. BGPsec [40] was proposed to prevent path-manipulation attacks, in which an AS advertises bogus BGP routes to influence other ASes' path selection, by cryptographically authenticating the links between ASes. Unfortunately, BGPsec requires widespread adoption of RPKI as a prerequisite, involves nontrivial changes to the Internet infrastructure to support on-path cryptography, and provides limited security benefits until universal adoption [39]. For these reasons, the adoption of BGPsec is far more difficult than that of RPKI [26]. BGPsec relies on certificates of ownership over AS numbers, which DISCO does not support.

DISCO is a good match for path-end validation, a recently proposed alternative to BGPsec that does not require modifications to the Internet's infrastructure and is effective in partial deployment [14], [15]. In path-end validation, a prefix's owner uses its private key to approve neighboring ASes for relaying BGP advertisements it originates. Since DISCO assigns IP addresses to public keys, it naturally supports authenticating neighbors through path-end validation records. The combination of DISCO and path-end validation provides a tangible defense against path manipulation attacks.

X. Related Work

Previous research has pointed out complementary approaches to RPKI, and operators continuously refine deployment approaches. Regardless, RPKI adoption has been very slow, and known deployments are not without serious challenges: AT&T's deployment, for example, requires constant monitoring by experienced administrators, is partial (ROV is enforced on peering links only), and still is far beyond what most networks have done. We believe that DISCO's automated certification mechanism significantly lowers the bar for adoption compared to previously proposed approaches. We also believe DISCO certification is an advancement relative to today's monitoring and alert systems.

Similar to DISCO's *de facto* ownership, PGBGP maintains a history mapping which origins announce what prefixes [33]. When a prefix is announced by a different origin AS, a PGBGP router quarantines the announcement for a predefined period (e.g., 24h) before installing the routes. The quarantine period allows network operators to check the quarantined announcement and take action before damage is done. Left unchecked, hijacks (due to misconfigurations or attacks) would be propagated as regular announcements after the quarantine period. Furthermore, quarantining routes requires changes to existing routers and may accidentally quarantine legitimate changes in announcement configuration (e.g., for traffic engineering). DISCO instead uses *de facto* ownership to create certificates, which allows flexible, deterministic, permanent filtering without the involvement of network operators. Another fundamental difference is that DISCO uses a global view of the Internet to establish ownership, while PGBGP routers operate in isolation and are thus more easily subverted by targeted, localized attacks.

Some studies proposed using anomaly detection to identify attacks on BGP [27], [36], [54], [58]. Like DISCO, these proposals rely on a global view of Internet routes afforded by vantage points. Although anomaly detection is useful for identifying suspicious routes, it does not bind IP prefixes to owners (through public keys) and so does not enable the owner to publish filtering rules pertaining to its prefixes like ROAs.

Our recent workshop paper proposes using BGP advertisements to establish de facto ownership over prefixes [23]. However, (1) the work presents only preliminary results from one Internet measurement and did not advertise a control prefix, limiting the conclusions that can be drawn; (2) no results regarding the global effects on reachability or data-plane measurements are presented; (3) no measurements of the fraction of announced prefixes with de facto owners are given so as to evaluate how useful the certification approach would be in practice; (4) important implementation details such as how to integrate with border routers are not explained; and (5) alternative approaches to establishing de facto ownership are not discussed.

Other proposals advocate establishing ownership by checking for control of a single machine (such as the reverse DNS server [19]). As discussed in §IX, such approaches might not provide sufficient security.

Human involvement in the configuration of cryptographic protocols induces errors and limits adoption. Consequently, automating configuration has been investigated in the context of other protocols as well, including IPsec and TLS [6], [22]. De facto ownership for establishing security proved useful in bootstrapping TLS, as reflected by the popular Let's Encrypt service for issuing X.509 TLS certificates [3], [6], [41]. Nevertheless, this approach was shown vulnerable to MitM attacks in the data plane [7], [8] and was later improved by utilizing control plane information for performing validation [10].

DISCO's decentralized approach for certifying ownership over IP address blocks resembles the design of the Convergence system for validating the correctness of TLS certificates [42]. In contrast to DISCO, Convergence relies on the data plane to validate ownership (discussed as an alternative approach to DISCO's design in section IX). DISCO's repositories publish a list of all issued certificates, allowing anyone to identify whether their prefix was certificated to someone else, an idea resembling certificate transparency for TLS [37].

XI. Conclusion

We presented DISCO, a system for certifying ownership of IP address blocks that yields substantial security benefits while circumventing the obstacles to adoption facing RPKI and ROV. We evaluated the security and deployability of DISCO through a combination of extensive simulations on empirically-derived datasets and live (control-plane and data-plane) experiments using the PEERING platform.

We view DISCO as the first step towards a broader agenda for securing BGP routing. Beyond protecting against prefix hijacking attacks, DISCO certification is sufficient to support path-end validation [14], [15], a recently proposed alternative to BGPsec that achieves comparable security benefits in a deployable manner. Combined, DISCO and path-end validation constitute a feasible path to BGP security.

ACKNOWLEDGEMENTS

We thank our shepherd Brad Reaves and the NDSS reviewers for valuable feedback. We appreciate the support and feedback from Job Snijders and others in the network operator community. Donald Sharp and others in the FRR community fixed the FRR bug triggered by our announcements, enabling further experiments. Michael Schapira is supported by an ERC Starting Grant. Ethan Katz-Bassett and Italo Cunha were partially supported by NSF grants CNS-1740883 and CNS-1835252, as well as a Google Faculty Research Award. Italo Cunha is additionally funded by RNP project 2955, CNPq award 311049, and CAPES award 88881.17164. Amir Herzberg was partially supported by an endowment from the Comcast corporation and by NSF grant 1840041. Yossi Gilad is supported by the Alon fellowship, the Hebrew university cybersecurity research center, and Mobileye. This research work has been funded in part by the German Federal Ministry of Education and Research and the Hessen State Ministry for Higher Education, Research and the Arts within their joint support of ATHENE – National Research Center for Applied Cybersecurity, and co-funded by the DFG as part of project S3 within the CRC 1119 CROSSING. The opinions expressed in the paper are those of the researchers themselves and not of their universities or sources of funding.

REFERENCES

[1] "The New Threat: Targeted Internet Traffic Misdirection," http://www.renesys.com/2013/11/mitm-internet-hijacking/.

[2] "The CAIDA AS Relationships Dataset," http://www.caida.org/data/as-relationships/, Feb. 2019.

[3] M. Aertsen, M. Korczynski, G. C. M. Moura, S. Tajalizadehkhoob, and J. van den Berg, "No domain left behind: Is Let's Encrypt democratizing encryption?" in *ANRW*. ACM, 2017, pp. 48–54. [Online]. Available: http://doi.acm.org/10.1145/3106328

[4] R. Anwar, H. Niaz, D. R. Choffnes, I. Cunha, P. Gill, and E. Katz-Bassett, "Investigating Interdomain Routing Policies in the Wild," in *Proc. ACM Internet Measurement Conference*, 2015.

[5] H. Ballani, P. Francis, and X. Zhang, "A Study of Prefix Hijacking and Interception in the Internet," in *Proceedings of ACM SIGCOMM*, 2007.

[6] R. Barnesm, J. Hoffman-Andrews, and J. Kasten, "Automatic Certificate Management Environment (ACME)," https://tools.ietf.org/html/draft-ietf-acme-acme-08, October 2017, internet-Draft.

[7] H. Birge-Lee, Y. Sun, A. Edmundson, J. Rexford, and P. Mittal, "Bamboozling certificate authorities with BGP," in *USENIX Security Symposium*, 2018, pp. 833–849.

[8] ——, "Using BGP to Acquire Bogus TLS Certificates," Hotpets, 2017.

[9] L. Blunk, M. Karir, and C. Labovitz, "Multi-threaded routing toolkit (MRT) routing information export format," October 2011, RFC6396. [Online]. Available: http://tools.ietf.org/rfc/rfc6396.txt

[10] M. Brandt, T. Dai, A. Klein, H. Shulman, and M. Waidner, "Domain validation++ for MitM-resilient PKI," in *Proc. ACM Conference on Computer and Communications Security*. ACM, 2018, pp. 2060–2076.

[11] R. Bush and R. Austein, "The resource public key infrastructure (RPKI) to router protocol," September 2017, RFC8210. [Online]. Available: http://tools.ietf.org/rfc/rfc8210.txt

[12] R. Bush, O. Maennel, M. Roughan, and S. Uhlig, "Internet Optometry: Assessing the Broken Glasses in Internet Reachability," in *Proc. ACM Internet Measurement Conference*, 2009.

[13] T. Chung, E. Aben, T. Bruijnzeels, B. Chandrasekaran, D. Choffnes, D. Levin, B. M. Maggs, A. Mislove, R. v. Rijswijk-Deij, J. Rula *et al.*, "RPKI is coming of age: A longitudinal study of RPKI deployment and invalid route origins," in *Proc. of the Internet Measurement Conference*, 2019.

[14] A. Cohen, Y. Gilad, A. Herzberg, and M. Schapira, "One hop for RPKI, one giant leap for BGP security," in *Proc. ACM HotNets*, 2015.

[15] ——, "Jumpstarting BGP Security with Path-End Validation," in *Proc. ACM SIGCOMM*, 2016, pp. 342–355. [Online]. Available: http://doi.acm.org/10.1145/2934872

[16] D. Cooper, E. Heilman, K. Brogle, L. Reyzin, and S. Goldberg, "On the risk of misbehaving RPKI authorities," in *HotNets*. ACM, 2013, pp. 16:1–16:7. [Online]. Available: http://dl.acm.org/citation.cfm?id=2535771

[17] Z. Durumeric, E. Wustrow, and J. A. Halderman, "ZMap: Fast Internet-wide Scanning and Its Security Applications," in *Proc. USENIX Security*, 2013.

[18] X. Fan and J. Heidemann, "Selecting Representative IP Addresses for Internet Topology Studies," in *Proc. ACM Internet Measurement Conference*, 2010.

[19] J. Gersch and D. Massey, "ROVER: Route Origin Verification Using DNS," in *ICCCN*, 2013, pp. 1–9.

[20] Y. Gilad, S. Goldberg, K. Sriram, and J. Snijders, "The Use of Maxlength in the RPKI," https://tools.ietf.org/html/draft-ietf-sidrops-rpkimaxlen-02, April 2019, proposed Best Current Practice.

[21] Y. Gilad, A. Cohen, A. Herzberg, M. Schapira, and H. Shulman, "Are We There Yet? On RPKI's Deployment and Security," in *NDSS*, 2017.

[22] Y. Gilad and A. Herzberg, "Plug-and-Play IP Security: Anonymity Infrastructure instead of PKI," in *ESORICS*. Springer, 2013, pp. 255–272.

[23] Y. Gilad, T. Hlavacek, A. Herzberg, M. Schapira, and H. Shulman, "Perfect is the Enemy of Good: Setting Realistic Goals for BGP Security," in *Proceedings of ACM Workshop on Hot Topics in Networks*, 2018, pp. 57–63.

[24] Y. Gilad, O. Sagga, and S. Goldberg, "Maxlength considered harmful to the RPKI," in *CoNEXT*, 2017, pp. 101–107. [Online]. Available: http://doi.acm.org/10.1145/3143361.3143363

[25] P. Gill, M. Schapira, and S. Goldberg, "Modeling on Quicksand: Dealing with the Scarcity of Ground Truth in Interdomain Routing Data," *Computer Communication Review*, vol. 42, no. 1, pp. 40–46, 2012.

[26] S. Goldberg, "Why is it taking so long to secure internet routing?" *Communication of the ACM*, vol. 57, no. 10, pp. 56–63, 2014. [Online]. Available: http://doi.acm.org/10.1145/2659899

[27] G. Goodell, W. Aiello, T. Griffin, J. Ioannidis, P. D. McDaniel, and A. D. Rubin, "Working around BGP: An Incremental Approach to Improving Security and Accuracy in Interdomain Routing," in *NDSS*. The Internet Society, 2003. [Online]. Available: http://www.isoc.org/isoc/conferences/ndss/03/proceedings/papers/5.pdf

[28] E. Heilman, D. Cooper, L. Reyzin, and S. Goldberg, "From the consent of the routed: Improving the transparency of the RPKI," in *SIGCOMM*, 2014, pp. 51–62. [Online]. Available: http://doi.acm.org/10.1145/2619239.2626293

[29] T. Hlavacek, A. Herzberg, H. Shulman, and M. Waidner, "Practical experience: Methodologies for measuring route origin validation," in *International Conference on Dependable Systems and Networks (DSN)*, June 2018, pp. 634–641.

[30] T. Hlavacek, I. Cunha, Y. Gilad, A. Herzberg, E. Katz-Bassett, M. Schapira, and H. Shulman, "Code artifacts," https://github.com/yossigi/disco.

[31] D. Iamartino, C. Pelsser, and R. Bush, "Measuring BGP Route Origin Registration and Validation," in *PAM*, ser. LNCS, vol. 8995. Springer, 2015, pp. 28–40. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-15509-8

[32] M. Jonker, A. Pras, A. Dainotti, and A. Sperotto, "A First Joint Look at DoS Attacks and BGP Blackholing in the Wild," in *ACM IMC*, 2018.

[33] J. Karlin, S. Forrest, and J. Rexford, "Pretty Good BGP: Improving BGP by Cautiously Adopting Routes," in *ICNP*. IEEE Computer Society, 2006, pp. 290–299. [Online]. Available: http://doi.ieeecomputersociety.org/10.1109/ICNP.2006.320179

[34] E. Katz-Bassett, C. Scott, D. R. Choffnes, I. Cunha, V. Valancius, N. Feamster, H. V. Madhyastha, T. Anderson, and A. Krishnamurthy, "LIFEGUARD: Practical Repair of Persistent Route Failures," in *Proc. ACM SIGCOMM*, 2012.

[35] S. Kent and K.seo, "An Infrastructure to Support Secure Internet Routing," Internet Requests for Comments, RFC 6480, February 2012. [Online]. Available: http://tools.ietf.org/html/rfc6480

[36] M. Lad, D. Massey, D. Pei, Y. Wu, B. Zhang, and L. Zhang, "PHAS: A Prefix Hijack Alert System," in *USENIX Security Symposium*, 2006.

[37] B. Laurie, A. Langley, and E. Kasper, "Certificate transparency." *ACM Queue*, vol. 12, no. 8, pp. 10–19, 2014.

[38] M. Lepinski and S. Kent, "An infrastructure to support secure internet routing," February 2012, RFC6480. [Online]. Available: http://tools.ietf.org/rfc/rfc6480.txt

[39] R. Lychev, S. Goldberg, and M. Schapira, "BGP Security in Partial Deployment: Is the Juice Worth the Squeeze?" in *SIGCOMM*. ACM, 2013, pp. 171–182. [Online]. Available: http://dl.acm.org/citation.cfm?id=2486001

[40] E. M. Lepinski and K. Sriram, "BGPsec Protocol Specification," RFC, Internet Engineering Task Force, Apr. 2017. [Online]. Available: https://tools.ietf.org/html/draft-ietf-sidr-bgpsec-protocol-23

[41] A. Manousis, R. Ragsdale, B. Draffin, A. Agrawal, and V. Sekar, "Shedding Light on the Adoption of Let's Encrypt," *CoRR*, vol. abs/1611.00469, 2016. [Online]. Available: http://arxiv.org/abs/1611.00469

[42] M. Marlinspike, "SSL And The Future Of Authenticity," https://www.youtube.com/watch?v=Z7Wl2FW2TcA, 2011.

[43] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig, "Building an AS-topology Model that Captures Route Diversity," in *Proc. ACM SIGCOMM*, 2006.

[44] Nimrod Levy, "Lightning Talk: Dropping RPKI invalid routes in a service provider network ," https://www.youtube.com/watch?v=DkUZvlj1wCk.

[45] NIST, "RPKI Monitor," http://rpki-monitor.antd.nist.gov/, 2015.

[46] ——, "CVE-2019-5892," https://nvd.nist.gov/vuln/detail/CVE-2019-5892, 2019.

[47] R. O'Donnell, "A survey of bgp security issues and solutions," *Proceedings of the IEEE*, vol. 98, no. 1, pp. 97–99, 2010.

[48] Y. Rekhter, T. Li, and S. Hares, "A border gateway protocol 4 (bgp-4)," January 2006, RFC4271. [Online]. Available: http://tools.ietf.org/rfc/rfc4271.txt

[49] A. Reuter, R. Bush, Í. Cunha, E. Katz-Bassett, T. C. Schmidt, and M. Wählisch, "Towards a rigorous methodology for measuring adoption of RPKI route validation and filtering," *ACM Computer Communication Review*, 2018.

[50] RIPE NCC, "Routing Information Service (RIS)," https://www.ripe.net/analyse/internet-measurements/routing-information-service-ris.

[51] ——, "YouTube Hijacking: A RIPE NCC RIS case study," March 2008.

[52] B. Schlinker, T. Arnold, I. Cunha, and E. Katz-Bassett, "PEERING: Virtualizing BGP at the Edge for Research," in *Proc. ACM CoNEXT*, 2019.

[53] B. Schlinker, K. Zarifis, I. Cunha, N. Feamster, and E. Katz-Bassett, "PEERING: An AS for Us," in *Proc. ACM HotNets*, 2014.

[54] P. Sermpezis, V. Kotronis, P. Gigis, X. Dimitropoulos, D. Cicalese, A. King, and A. Dainotti, "ARTEMIS: Neutralizing BGP Hijacking within a Minute," *IEEE/ACM Transactions on Networking*, vol. 26, no. 6, pp. 2471–2486, Dec 2018.

[55] J. Snijders, "Deprecation of bgp path attribute values 30, 31, 129, 241, 242, and 243," February 2017, RFC8093. [Online]. Available: http://tools.ietf.org/rfc/rfc8093.txt

[56] ——, "BGP Large Communities," May 2017, SINOG 4, Ljubljana, Slovenia. [Online]. Available: http://largebgpcommunities.net/presentations/SINOG2017_Snijders_Large_Communities.pdf

[57] F. Streibelt, F. Lichtblau, R. Beverly, A. Feldmann, C. Pelsser, G. Smaragdakis, and R. Bush, "BGP Communities: Even More Worms in the Routing Can," in *IMC*, 2018.

[58] L. Subramanian, V. Roth, I. Stoica, S. Shenker, and R. H. Katz, "Listen and Whisper: Security Mechanisms for BGP," in *NSDI*. USENIX, 2004, pp. 127–140. [Online]. Available: http://www.usenix.org/events/nsdi04/tech/subramanianListen.html

[59] Team Cymru, "IP-ASN-mapping," http://www.team-cymru.com/IP-ASN-mapping.html.

[60] A. Toonk, "Hijack Event Today by Indosat," http://www.bgpmon.net/hijack-event-today-by-indosat/.

[61] ——, "Turkey Hijacking IP Addresses for Popular Global DNS Providers," BGPMon.

[62] University of Oregon, "Route Views Project," http://www.routeviews.org.

[63] P.-A. Vervier, O. Thonnard, and M. Dacier, "Mind Your Blocks: On the Stealthiness of Malicious BGP Hijacks," in *NDSS*. The Internet Society, 2015. [Online]. Available: http://www.internetsociety.org/events/ndss-symposium-2015

[64] M. Wählisch, O. Maennel, and T. C. Schmidt, "Towards Detecting BGP Route Hijacking Using the RPKI," *SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 4, pp. 103–104, Aug. 2012.

[65] M. Wählisch, R. Schmidt, T. C. Schmidt, O. Maennel, S. Uhlig, and G. Tyson, "RiPKI: The Tragic Story of RPKI Deployment in the Web Ecosystem," in *Proceedings of ACM Workshop on Hot Topics in Networks (HotNets)*. New York: ACM, 2015.

APPENDIX A
INFERRING NETWORKS USING DEFAULT ROUTES

We used measurements to identify ASes using default routes. To do so, we kept a /24 prefix in PEERING address space permanently withdrawn and issued traceroutes from our set of RIPE Atlas vantage points to the withdrawn prefix. If the network hosting the probe does not have default route, it will drop the packets, and the traceroute will terminate. If the network does have a default route, the traceroute packets can be forwarded along the default route until they reach a router in a network that does not have a default route. We infer that a network has a default route if the traceroute reaches (at least) a router outside that network. Although the coverage of this approach is limited to ASes at the beginning of routes from the selected RIPE Atlas probes toward PEERING, it overlaps exactly with the traceroute measurements used in our analysis of DISCO, which we measured from the same set of RIPE Atlas vantage points.

We identify 768 RIPE Atlas probes hosted in networks using default routes, and a total of 941 networks using default routes (some traceroutes traverse multiple ASes using default routes). Figure 8 shows the distribution of the customer cone sizes of ASes using default routes, and the distribution of customer cone sizes of the ASes where these default routes terminate. As observed in previous work [12], we find that most cases are small networks employing default routes toward larger provider networks.

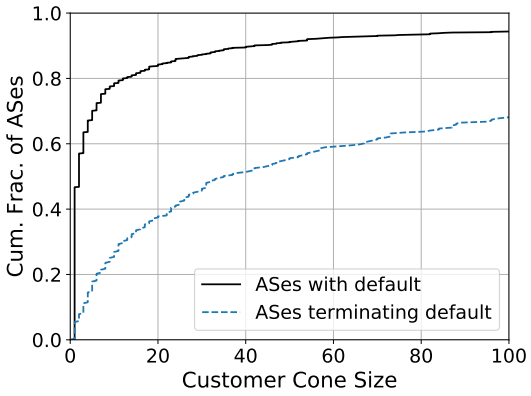


Fig. 8. Characterization of ASes inferred to be using default routes.

APPENDIX B
EXPERIMENTING WITH BGP IN THE WILD

We did not repeat the experiments in §VII-A for multiple reasons: (1) we anticipate that the qualitative results would remain very similar, even if the exact numbers would change; (2) there is a chance of a small number of operators not applying the FRR patch; (3) while we have safely tested against multiple routers and configurations in a controlled environment (including the patched FRR), we cannot know the full set of configurations and deployments that exist in the world to test against, and (4) new router bugs may have been introduced since the last experiment and be triggered by a new experiment.

Our experiments and recent experience with BGP large communities indicate that extending BGP is a slow, iterative process. The safest course of action is to go through the IETF/IRTF to request a BGP attribute (IANA does issue temporary allocations). However, even a BGP attribute allocation from IANA is not a guarantee of disruption-free deployments, as implementations may be using the attribute for other ends or mishandle it.