Inferring BGP Blackholing Activity in the Internet

Vasileios Giotsas
CAIDA/TU Berlin
vasilis@inet.tu-berlin.de

Georgios Smaragdakis
MIT/TU Berlin
gsmaragd@csail.mit.edu

Christoph Dietzel
TU Berlin/DE-CIX
cdietzel@inet.tu-berlin.de

Philipp Richter
TU Berlin
prichter@inet.tu-berlin.de

Anja Feldmann
TU Berlin
anja@inet.tu-berlin.de

Arthur Berger
MIT/Akamai
awberger@csail.mit.edu

ABSTRACT

The Border Gateway Protocol (BGP) has been used for decades as the de facto protocol to exchange reachability information among networks in the Internet. However, little is known about how this protocol is used to *restrict* reachability to selected destinations, e.g., that are under attack. While such a feature, BGP blackholing, has been available for some time, we lack a systematic study of its Internet-wide adoption, practices, and network efficacy, as well as the profile of blackholed destinations.

In this paper, we develop and evaluate a methodology to automatically detect BGP blackholing activity in the wild. We apply our method to both public and private BGP datasets. We find that hundreds of networks, including large transit providers, as well as about 50 Internet exchange points (IXPs) offer blackholing service to their customers, peers, and members. Between 2014–2017, the number of blackholed prefixes increased by a factor of 6, peaking at 5K concurrently blackholed prefixes by up to 400 Autonomous Systems. We assess the effect of blackholing on the data plane using both targeted active measurements as well as passive datasets, finding that blackholing is indeed highly effective in dropping traffic before it reaches its destination, though it also discards legitimate traffic. We augment our findings with an analysis of the target IP addresses of blackholing. Our tools and insights are relevant for operators considering offering or using BGP blackholing services as well as for researchers studying DDoS mitigation in the Internet.

CCS CONCEPTS

• **Networks Routing protocols**; **Denial-of-service attacks**;

KEYWORDS

BGP; Blackholing; DDoS Mitigation.

1 INTRODUCTION

Access to online information, content, services, and social communities has fueled the phenomenal growth of the Internet for decades. To enable such access *global reachability* is imperative [18, 40], i.e., every publicly advertised address should be reachable from any device connected to the Internet. The de facto protocol to achieve global reachability in the Internet is the Border Gateway Protocol (BGP) [62]. BGP enables autonomously operated networks to exchange reachability information with their immediate peers and, eventually, with all the networks in the Internet. Because the Internet is an uncoordinated global communication system [19], it took a substantial effort to achieve stable global connectivity in the face of outages [8, 33] and disasters [15], independent routing decisions [30, 46, 47], network attacks [52, 70], misconfigurations, and security loopholes [36].

Today's network challenges go beyond global connectivity. Distributed denial of service (DDoS) attacks have increased both in terms of attack bandwidth and numbers of compromised machines [1]. The profile of attackers ranges from hackers that try to gain commercial benefits [51], activists that try to protest ("Hacktivism"), e.g., by attacking banks [22] and government Web sites [23], to regimes that attack the network infrastructure of political opponents [6] or neighboring countries [65]. Many high profile sites pay for sophisticated defense mechanisms such as traffic-scrubbing [39] (which filters attack traffic from legitimate traffic) and firewalls [13]. Many content providers use content delivery networks which offer sophisticated security services and are able to absorb large DDoS attacks due to the size and scope of their infrastructure [31, 53]. Operators can have significant difficulties handling DDoS attacks. They may have to transport voluminous attack traffic to mitigation middleboxes or to the targeted server, only to have the traffic dropped or for it to create harm. The larger the DDoS attack volume, the higher the cost for the network operator, including transit costs with upstream providers, or transgression of peering agreements. High volume attacks also impede on the service quality of the legitimate traffic. Such degradation may harm the service provider's reputation or lead to violations of service level agreements.

One mitigation option is *blackholing*, i.e., *dropping* traffic, to a destination, ideally as close to the source as possible. While aggressive, blackholing has the potential to be fast, cheap, and very effective, especially when the attack volume is very high such that alternative mitigation options become more difficult or expensive. Blackholing drops all traffic to a targeted DoS destination, not just the attack traffic, and can have the drawback that this destination becomes *unreachable* – the goal of the DDoS attack. Also, if the

target organization has purchased a traffic-scrubbing service, then this service is degraded if the traffic is discarded prior to reaching the scrubbing center. Hence there is a conflict of interest between organizations that use such services, and intermediate networks through which the traffic would traverse. Such negative impact is less severe if (much of) the attack traffic originates from only a few locations and blackholing can be limited to a small set of associated providers, hence not affecting legitimate traffic from other origins. We note that blackholing is not limited to DDoS mitigation, but can also be used to restrict reachability for other reasons, e.g., censorship, compliance with regulations, or court decisions.

BGP enables blackholing by leveraging the *BGP communities attribute* (RFC 1997 [11]). Networks trigger blackholing requests by sending BGP announcements to their BGP neighbors for specific destination prefixes with the appropriate blackhole community. The neighbor, upon receiving such an announcement, discards at its ingress traffic whose destination address is in the blackholed prefix. Internet exchange points (IXPs) also offer blackholing as a service to their members [12, 25]. Although BGP blackholing has been available for some time, little is known about its adoption and effectiveness. One complication is that BGP community values, in general as well as for blackholing, are not standardized, as underlined in an earlier study of BGP communities (for the period 2004 to 2007) [26]. The authors could extract the semantic meaning for only 22% of BGP communities tags in BGP updates from RIPE and RouteViews route collectors. Of these, only 60 (0.2%) appeared to be related to BGP blackholing. Yet, we know from two prior studies that today BGP-based blackholing services are deployed in the Internet and, in principle, perform as desired. Orsini et al. [54] used blackholing as a case study to highlight the capabilities of BGPstream. They showed that BGPStream provides timely parsing capabilities to capture blackholing events in a set of 30 networks observed at two BGP collectors. Dietzel et al. [25] studied the popularity of the blackholing service offered by a single IXP.

The present paper builds upon this prior work and extends it to systematically assess: (i) the *Internet-wide adoption* of BGP blackholing by different types of network operators over the last years, (ii) current blackholing *practices*, (iii) the network *efficacy* of BGP blackholing, i.e., how much prior to the destination the traffic is dropped, and (iv) the *profile* of destinations targeted by blackholing. Our contributions and findings can be summarized as follows:

- We develop and evaluate a methodology to automatically detect blackholed prefixes in the Internet. We use a number of public and private BGP datasets to assess the visibility of Internet-wide BGP blackholing activity over more than two years (2014-2017).
- We find that an increasing number of networks of different types offer BGP blackholing services to their customers and peers. We identify more than 250 transit, access, and content providers as well as about 50 IXPs around the globe. Moreover, blackholing usage is increasing. Over the last three years the number of blackholed prefixes has increased by a factor of 6, peaking up to 5K concurrently blackholed prefixes by up to 400 Autonomous Systems in recent months.
- We show that BGP blackholing activity can potentially serve as a proxy for identifying high profile attacks in the Internet. Indeed,
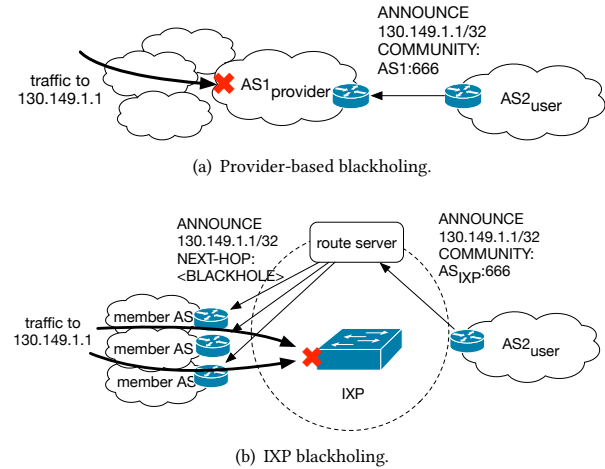

(a) Provider-based blackholing.

(b) IXP blackholing.

Figure 1: Blackholing: Triggering via BGP and effect on the data plane.

increased activity of BGP blackholing correlates with periods of high activity of DDoS attacks.
- We assess the effect of blackholing on the data plane using both targeted active measurements as well as passive datasets, finding that blackholing is in fact highly effective in dropping traffic before it reaches its destination.
- We profile blackholed destinations and find that the most popular blackholed prefixes of the more than 88K prefixes during the last eight months involve content providers. This has implications on the reputation of such networks and for unrelated services that happen to use the same IP.

Our study provides insights to operators, regulators, and researchers. Our evaluation of the deployed BGP blackholing service is of interest to network operators that (i) consider offering BGP blackholing services, or (ii) consider using the offered BGP blackholing service. For regulators, it is important to observe which IP prefixes are not reachable (blackholed) and why. An Internet-wide inference of the BGP blackholing activity either online or based on archived data sheds light on the current trends and can inform policy decisions and best practices. In summary, this study provides the tools and a first analysis of Internet-wide BGP blackholing.

2 BACKGROUND AND DEFINITIONS

Blackholing is a popular DDoS mitigation strategy inside a single network or among multiple networks. Hereby, the victim network (AS) announces the attacked destination IP or prefix to its upstream via BGP, also known as remotely triggered blackholing [16]. The upstream AS can then drop traffic towards this prefix, the *blackholed prefix*, usually at the AS ingress point by rewriting the next-hop address to a null interface. To distinguish regular routing updates from blackholing messages, ASes tag such route updates with a *blackhole community*, which is typically documented in the Internet Routing Registry (IRR) record of the network operator or on the network provider's web page. The *blackholed prefix* is the destination prefix in BGP announcements that is tagged with a blackhole community.

A *Blackholing provider* is a network that offers the BGP blackholing service. It is also the network that specifies which community

to use for BGP blackholing, see Figure 1(a). Historically, blackholing was implemented mainly at the network edge (customer or provider networks). However, over time it has moved to the Internet core, and is now offered by Internet Service Providers (ISPs) and IXPs as well. In recent years, an increasing number of IXPs offer blackholing as a free service to their users [12]. At IXPs the connected IXP members often take advantage of the IXP route server for ease of peering via the joined layer-2 infrastructure [58]. If a member establishes a session with the IXP route server, it can announce a prefix with an appropriate blackhole community tag to the route server. The route server then propagates the announcement to all connected IXP member ASes that do not filter such announcements, see Figure 1(b). These ASes can then drop traffic towards the blackholed prefix to the null interface specified by the IXP. The *Blackholing IP* is the address to which the next-hop of the blackholed prefix is set in order for the traffic to be dropped, effectively a null interface.

A *Blackholing user* is a network that makes use of the BGP blackholing service, i.e., the network that inserts the blackhole community tag in the BGP announcement, offered by a blackholing provider. In the case that the blackholing provider is a peer or an upstream provider, the announcement for the blackholed prefix has to be announced with the associated blackhole community as shown in Figure 1(a). If an IXP is the blackholing provider, the user has to announce the blackholed prefix to the route server with the IXP blackhole community as shown in Figure 1(b).

While blackholing essentially takes the victim destination offline, it prevents saturation and collateral damage in the network resources along the attack path. Therefore, it is beneficial for both the destination network as well as the upstream ASes, at the expense of the target of the attack. According to the recommended best practices, operators do not accept BGP routes for prefixes more-specific than /24, in order to prevent routing table deaggregation [49]. However, in the case of BGP blackholing, /24 or less-specific prefixes to mitigate a DDoS attack on specific hosts would lead to blackholing of all the hosts in the prefix, even those not affected by the attack. To restrict the impact of blackholing, blackholing providers accept routes more-specific than /24, if they are tagged with a blackholing community. Contrarily, and according to the best practice, prefixes less-specific than /24 should not be allowed to be blackholed [45]. A common practice by blackholing providers is to require some type of authentication to accept the blackhole community. Typically, they will accept a blackhole community if the request comes from the originator of the prefix or from a network provider that has this prefix in its customer cone. Some of the blackholing providers will accept announcements only via secure BGP using the RPKI [37] and others will accept blackholing announcements only if the prefix is registered in one of the IRRs.

3 BGP DATASETS

In this section we provide an overview of the BGP datasets that we analyze in this study. A summary of statistics per BGP dataset (for March 2017) is presented in Table 1. Note that we report on the total number of prefixes, however, over 96% of the prefixes across the datasets are IPv4. For our analysis we also rely on non-BGP data, which we introduce in the appropriate sections.

Source	#IP peers	#AS peers	#Unique AS peers	#Prefixes	#Unique prefixes
RIS	425	313	77	712,176	11,876
RV	269	197	42	784,700	87,536
PCH	8,897	1,721	1,175	765,005	38,847
CDN	3,349	1,282	911	1,840,321	1,055,196
Total	12,940	2,798	2,205	2,012,404	1,193,455

Table 1: Overview of BGP dataset for March 2017. IPv4 prefixes account for 96.64% of the total prefixes.

Public BGP Data: We analyze widely-used public datasets from the route collectors of the (i) RIPE Routing Information Service (RIS) [60], (ii) Route Views (RV) [69], and (iii) Packet Clearing House (PCH) [55]. All of these platforms consist of several routers that collect default-free BGP routing information from a multitude of BGP peers. Some BGP peers send full routing tables, others partial views, and even others only their customer routes. The platforms then publicly offer full BGP routing updates. Many IXPs offer route servers as a free value-added service to simplify BGP session management for their members. Route servers collect routing information at the IXP in a centralized manner and redistribute them to connected member routers. As such, they offer BGP routing information for most of the IXP members [58]. PCH maintains route collectors at 111 different IXPs (March 2017) [56] and makes the data available.

Private BGP Data: While the above datasets cover a significant part of the Internet, their scope is biased by where the collectors are placed, which networks participate, e.g., RIS and RV are biased to what is announced by large transit providers in the core of Internet [62], and if a direct peering feed via BGP is available. To overcome some of these limitations we augment the publicly available datasets with BGP updates from a large CDN which receives BGP feeds from about 3,350 BGP peers in about 1,300 networks. The CDN BGP dataset is unique because the CDN collectors also receive customer-specific and internal BGP announcements as the CDN deploys network equipment within many ISPs. This unique view of the CDN is the reason why it receives multiple times more unique prefixes than the collectors of the public datasets. Note that the CDN itself does not offer a BGP blackholing service.

Looking Glasses: We use the Periscope system [32] that gives us access to about 150 BGP looking glasses. Out of them, 30 looking glasses either enable queries for full BGP table dumps or based on community values. We mainly use the looking glass data for validation for those ASes where we do not have a direct peering feed via a BGP collector or to query for a specific prefix/IP.

BGP Data Cleaning: To eliminate obvious misconfigurations in the BGP data we filter out non-routable, private, and bogon prefixes (archived weekly snapshots) that are reported in the Cymru bogon list [21], and eliminate prefixes less-specific than /8.

4 METHODOLOGY

Next, we present our methodology for inferring ISP and IXP BGP blackholing activity in the Internet. The key observation is that BGP announcements are used to restrict reachability if they are tagged with specific communities. Thus, we first outline how to build a blackhole communities dictionary. Then, we describe our inference methodology in detail.

4.1 Blackhole Communities Dictionary

No universally followed convention exists with regards to how ASes use community values [26]. This affects which community values are used by each AS for signaling BGP blackholing. Thus, we have to infer such information from various sources to construct a blackhole communities dictionary.

Inferring Blackhole Communities: To gather lists of BGP blackhole communities we use a similar methodology as was used by Giotsas et al. [33] for extracting BGP location communities. Since ISPs and IXPs offer blackholing as a service for their customers/members, many of them use well-documented communities and publish them either on their Web pages or in their Internet Routing Registry (IRR) records (for our analysis we use the IRR records in Merit RADb [50]). To gather this information we first develop a Web scraper to collect the relevant text from ISP and IXP Web pages as well as IRR records. Then, we apply natural language processing techniques using the Natural Language ToolKit [7] text parser to extract all community values relevant for BGP blackholing by searching for lemmas of certain text patterns, and certain keywords e.g., "blackhole", or "null route". In addition, we can sometimes collect meta-information about these blackhole communities, e.g., the maximum accepted prefix length of blackhole communities, or region-specific blackhole communities. The IRR records contribute the largest fraction of blackhole communities to our dictionary, namely 172 communities for 209 networks. For 93 other ASes we found an additional 130 BGP communities on Web pages that document their BGP policies. We collected additional BGP blackhole communities for 5 networks via private communication. Together this accounts for BGP blackhole communities for 307 ASes, including 49 IXPs and 13 Tier-1 ISPs.

Typically, each blackholing provider only uses a single BGP community for global blackholing. However, there are several blackholing providers that use additional ones for more fine grained control over the scope of the blackholing, e.g., blackhole only in Europe, US, or Asia. Thus, in some cases we have multiple blackhole communities for one blackholing provider. Note, almost all blackholing providers use the same community tags for IPv4 and IPv6. We take all these peculiarities into account while constructing our BGP blackhole community dictionary. The most common convention for blackhole communities (51%) is to use the format ASN:666, where ASN is the AS number of the AS of the blackholing provider. Other popular values are ASN:66 and ASN:999. However, note that the use of the value 666 in a community does not necessarily imply blackholing, as 666 is also used by some ASes to encode other information, e.g., Level3 uses the community 3356:666 to tag peering routes, and the community 3356:9999 for remotely triggered blackholing. As a result we only include communities in our dictionary if we can validate them either via published information by the ASes or private communication, and we refer to them as *documented* communities. We also see cases of blackhole communities where the first 16 bits do not indicate a public ASN, e.g., 65536:666 or 0:666. From the documentation we found that multiple networks share such BGP blackhole communities. Hence, we augment the dictionary of documented communities with information about which networks provide this community.

Network Type	#Networks	#Blackhole communities
Transit/Access	198 (81)	223 (90)
IXP	49 (0)	2 (0)
Content	23 (14)	25 (14)
Education/Research/NfP	15 (1)	20 (1)
Enterprise	8 (3)	9 (3)
Unknown	14 (3)	4 (3)
TOTAL unique	307 (102)	292 (111)

Table 2: Documented blackholed communities distribution used in our study (in parenthesis we report the inferred but undocumented blackhole communities distribution).

The most popular BGP community format is 32 bits, where the first 16 bits refer to the ASN, as specified in RFC1997 [11]. More recent community formats, such as the extended [64] and the large [38] communities, have been introduced to address issues with 32-bit ASNs, but so far their adoption is limited. In our dictionary, only 6 of the 307 networks have adopted the new community format, and only 1 of these use it for blackholing purposes. For the IXPs, we follow a similar approach. Most IXPs that offer BGP blackholing publish their communities on their Web page or IRR records, so that their members can easily access this information. Among the 111 IXPs for which PCH has a BGP collector, we are able to identify 26 IXPs worldwide that offer BGP blackholing service. Using our Web scraper and natural language based analysis, we discovered 23 additional IXPs with BGP blackholing service. The most noticeable among these is MSK IXP, which has locations in 9 Russian cities. The large majority, 47 out of 49, follow the suggestion of RFC 7999 [42] regarding the usage of the BGP blackhole community value 65535:666. In almost all cases they also provide a blackholing IP for IPv4 and IPv6. The most common last octet for IPv4 is .66, while the most common string for IPv6 addresses is dead:beef *[sic]*.

Table 2 provides details on the documented communities per network type. We group the networks that provide these communities according to their declared network type in the PeeringDB database [57]. If the network does not maintain a PeeringDB record, or does not disclose its network type, we use CAIDA's AS classification dataset [9]. CAIDA's classification combines the *NSP* (Network Service Provider) and *Cable/DSL/ISP* types found in PeeringDB in the *Transit/Access* class. To get a consistent classification we also follow the same convention. The network types *Educational/Research*, and *NfP* (Not-for-Profit) are unique to PeeringDB's classification.

The resulting dictionary contains documented community values for 307 ISPs and IXPs[1], i.e., five times as many networks as reported in 2008 in the most detailed study so far [26]. Indeed, our method discovers blackhole communities for all networks included in this prior study in an automated fashion. Since BGP blackhole community values can change, we compare ours with the ones from that study, which are publicly available. We find that 72% are still active and none of the inactive ones have been re-purposed, which indicates that the community usage is relatively stable over time.

---

[1]Each ISP is identified by its ASN. Sibling ASes are counted as separate networks if they use different ASNs.
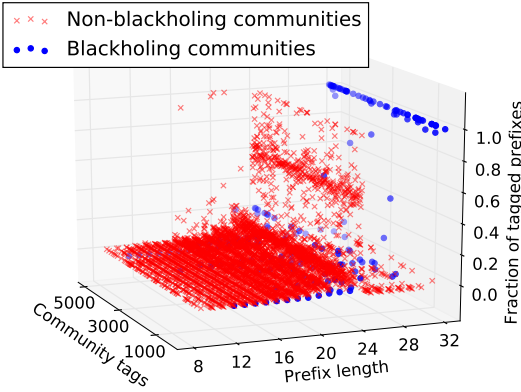
Figure 2: For a given community tag, the fraction of occurrences with the given prefix length.

Possibilities for Extended Dictionary: The above process for creating a BGP blackhole community dictionary is likely to miss some BGP blackhole communities. One possible way for extending the dictionary is to take advantage of the properties of blackholing BGP updates. Indeed, prior work [25] pointed out that most IPv4 blackholing announcements are for /32s, while at the same time host routes should not be part of the global routing tables. Moreover, it should be best practice to not accept blackholing for prefixes less-specific than /24. Thus, communities that are predominantly associated with announcements for more-specific prefixes than /24 are prime candidates for BGP blackhole communities. To confirm this hypothesis we constructed a second dictionary of BGP communities that includes communities for non-blackholing purposes (i.e., relationship tagging, traffic engineering), by parsing again IRR records and operator's Web sites. We then compared the prefix lengths on which the blackholing and non-blackhole communities are applied.

In Figure 2 the axes are respectively a numbered list of the different community tags, the prefix length in the BGP announcement in which the community tag appeared, and on the z-axis, for a given community tag, the fraction of appearances with the given prefix length. Figure 2 shows that the largest fraction of non-blackhole communities are applied on /24 or less-specific prefixes, and where the vertical plane of red cross markers occurs at /24. On the other hand, the blackhole communities in our dictionary (blue dot markers) are applied almost exclusively on /32 prefixes. We could use this observation to extend the blackhole communities dictionary, by collecting communities values which are exclusively applied on prefixes more-specific than /24. Moreover, to increase our confidence that these communities are indeed used for BGP blackholing, we require them to be used together with other known blackhole communities at least once. Many of the *inferred* communities also follow the pattern ASN:666. We ignore communities for which the first 16 bits do not encode a public ASN since without documentation it is not possible to map such communities to blackholing providers. Overall, we found 111 such inferred communities in 102 ASes that are not present in the dictionary of the documented communities. Since these communities are not documented, we decided not to include them in the community dictionary.

4.2 BGP Blackholing Inference

To infer BGP blackholing activity, we identify related BGP announcements using the blackhole community dictionary from Section 4.1 (with documented blackhole communities). We also perform additional checks to eliminate false positives and collect useful data to later characterize BGP blackholing activity.

Identifying Blackholed Prefixes in BGP Datasets: Each prefix that is tagged with a community from our blackhole community dictionary is a potential candidate. However, before we conclude that it is part of a blackholing event, we check if the blackhole community can be used by multiple blackholing providers (with different ASNs), e.g., 0:666. In the case of such ambiguous communities we further check if any potential ASN is on the AS path. If it is not we do not consider the update any further. We infer the blackholing user as the AS before the blackholing provider along the AS path (after removing AS path prepending).

When we encounter IXP blackholing communities, we first check if the ASN of the IXP's route server appears in the AS path, or if the `peer-ip` attribute of the BGP message belongs to the address space of the IXP's peering LAN, according to PeeringDB. In these cases we can confirm that the IXP is indeed traversed, and we infer the IXP as the blackholing provider. If the route server ASN appears in the AS path, we infer the AS hop before it as the blackholing user. When we determine the IXP blackholing provider based on the `peer-ip` attribute (in this case is the blackholing IP), the blackholing user is inferred as the AS encoded in the `peer-as` attribute, which is the IXP member that announces the blackholed prefix using the corresponding IXP blackhole community.

Note that blackholing users often want to blackhole a prefix at multiple providers. In this case they may either send a different prefix advertisement to each provider, with the blackhole community of each respective blackhole community attached, or bundle the blackhole communities together for all of the intended blackholing providers, and send the same prefix advertisement to all of their external BGP neighbors. In the case of such *blackhole community bundling*, our methodology is able to detect blackholing requests at providers, even if these providers are not forwarding the blackholed prefix outside their network. Figure 3 illustrates this detection process. Blackholing user AS_{C1} announces the blackholed prefix $Prefix_{B1}$ with different blackhole communities (IXP:666 and P1:666) for each respective blackholing provider. In this case we can infer only the IXP blackholing provider but not AS_{P1}, since AS_{P1} does not propagate the announcement to the BGP collector. On the other hand, blackholing user AS_{C2} bundles the blackhole communities P1:666 and P2:666 to prefix $Prefix_{B2}$, even at BGP neighbors which do not provide blackholing such as AS_{peer}. Even though neither AS_{P1} nor AS_{P2} propagate the blackholed prefix to a collector, we are able to infer the blackholing at both providers by getting the BGP feed from AS_{peer}. We show in Section 9 that bundling contributes about half of our inferences and helps us discover a large number of blackholings, despite the inherent restrictions in the propagation of blackholed prefixes (due to prefix lengths).

Initialization Based on BGP Table Dump: To initiate our analysis we use the oldest BGP dump table from the BGP dataset. For each prefix that can be identified as a blackhole prefix, we store: the blackhole communities, the associated blackholing providers/IXPs,
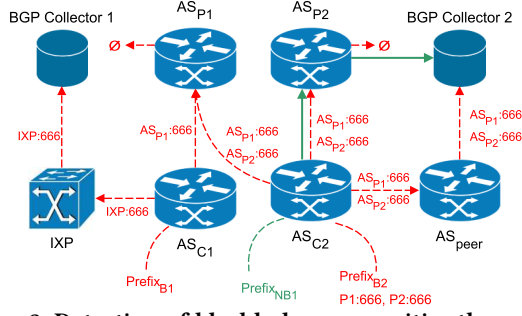
Figure 3: Detection of blackhole communities through passive BGP monitoring.

BGP path, next hop, and the peer IP and ASN. At this point we cannot accurately pinpoint the start time as we can only conclude that the blackholing event started before the BGP dump was stored. To account for this we use an initial starting time of zero. Once we have initialized our inference we move into continuous monitoring mode by monitoring all BGP updates, which includes BGP announcements as well as withdrawals.

Continuous Monitoring of BGP Announcements: When a BGP collector receives a new BGP announcement, we parse the BGP communities attribute for possible blackhole values in the same manner as during the initialization phase. We consider two cases where we have to update our initial record of blackholed prefixes:

- The announced prefix has blackhole communities attached and it was not previously blackholed. In this case we consider the BGP announcement the start of a new blackholing event for the BGP peer that received the announcement, and we add the prefix in the list of blackholed prefixes to monitor.
- The announced prefix has no blackhole communities attached, but the prefix was previously observed as blackholed at that particular BGP peer. In this case we infer an *implicit withdrawal* of blackholing event at the time of the BGP announcement.

Note that a de-activation of a prefix blackholing may be observed only at a subset of the BGP peers that observed the initial blackholing. Therefore, we track all blackholing events at the granularity of individual BGP peers. Then, we correlate the observed activation and de-activation for a given blackholed prefix across all the BGP peers. We do this to decide on the exact starting and ending of a blackholing activity for a given blackholed prefix.

Continuous Monitoring of BGP Withdrawals: To estimate the end of a blackholing event for a blackholed prefix we also track BGP withdrawals on a per BGP peer basis for each prefix. Whenever we detect an *explicit withdrawal* of a previously blackholed prefix, we mark the end of the blackholing event at the time of the withdrawal message.

5 BGP BLACKHOLING VISIBILITY

In this section we assess to what extent our analysis can infer blackholing activity using a number of public and private datasets and we discuss factors that determine visibility into BGP blackholing.

5.1 Visibility across Datasets

To assess the visibility of blackholing activity of the 307 ISPs and IXPs that are included in our blackhole communities dictionary, we rely on the four large BGP datasets introduced in Section 3. For RIPE RIS and Route Views we utilize the BGPStream API [54] and for the PCH and the CDN datasets we develop our own parser. We focus on the period spanning August 2016 through March 2017, since it includes time periods with high DDoS activity (see Section 6).

Table 3 provides summary statistics for each individual dataset as well as the combined datasets. We find activity by more than 1,100 users, blackholing 88,381 prefixes. Out of these prefixes, 88,209 are IPv4, which we focus on in the remainder of this analysis. It is worth mentioning that 98% (86,834 of 88,209) IPv4 prefixes are /32 prefixes. We find active blackholed prefixes for some 242 (79%) out of the 307 blackholing providers in our BGP communities dictionary. We do not see blackholing activity for 24 small IXPs (out of the 49 IXPs in our dictionary) and some small/regional ISP blackholing providers with known blackholing communities.

Not all BGP datasets contribute evenly to the visibility of BGP blackholing activity, as shown in Table 3. The CDN BGP dataset contributes most blackholing activity and observes almost all active blackholing providers. This is to be expected since it receives BGP feeds from the largest number of peers (around 1K networks). Notice that the number of visible blackholing providers is almost twice as large as compared to any other BGP dataset, with 111 blackholing providers that are only visible in the CDN dataset. With respect to blackholing users, however, we note that the majority of them are visible in multiple datasets. Nevertheless, each dataset observes some 5-15% unique users. Here, the CDN again contributes the largest share of unique users. In terms of visibility of blackholed prefixes, both the CDN and the PCH dataset provide significantly better coverage when compared to RIPE RIS and Route Views (some 70K+ blackholed prefixes in CDN/PCH versus some 25K prefixes in RIPE RIS/Route Views). The underlying reason here is that the CDN and PCH have deployed BGP monitors at various IXPs, where they peer directly with the IXP's route servers and hence provide direct visibility into blackholing offered by IXPs. When combining all the datasets, PCH contributes the largest number of unique prefixes. Nevertheless, RIS also contributes a large number of unique prefixes due to a blackholing provider that peers directly with a RIS collector, but with no other collector. Another factor that contributes to the large diversity of visibility across and within our datasets is the percentage of BGP feeds that individual BGP collectors receive directly from a BGP blackholing provider, i.e., when a blackholing provider has a direct BGP session with a BGP collector (See Table 3, Blackholing providers with direct BGP feed). BGP datasets with a large fraction of direct peerings, i.e., CDN and PCH datasets with 21% and 44% respectively, provide the best visibility.

In Table 4 we look across datasets and explore the business types of networks associated with the blackholed entity, following the same approach as with Table 2. A majority of the blackholed prefixes are tagged with transit/access blackhole communities, which is expected since this type of networks serve the largest number of customers and therefore have the largest pool of potential blackholing users. IXPs are the second largest type of blackholing providers,

Source	#Blackholing providers	#Unique bh. providers	#Blackholing users	#Unique bh. users	#Blackholed prefixes	#Unique bh. prefixes	Blackholing providers with direct BGP feeds
CDN	231	111	894	94	73,400	5,908	20.8%
RIS	113	0	739	57	24,637	6,217	4.42%
RV	116	2	729	27	24,420	417	17.2%
PCH	119	5	831	63	74,709	7,224	43.6%
ALL	242	118	1112	241	88,209	19,766	33.05%

Table 3: Blackhole dataset overview for IPv4 prefixes (August 2016 – March 2017).

Network Type	#Bh. prov.	#Bh. users	#Bh. pref.	Direct feed
Transit/Access	184	986	80,262	28%
IXP	25	673	20,824	100%
Content Provider	19	90	2,428	21%
Enterprise	5	127	4,144	20%
Educ./Res./NfP	5	40	1,244	20%
unknown	4	19	882	0%
Total (unique)	242	1,112	88,209	33.8%

Table 4: Blackhole visibility of IPv4 prefixes (August 2016 – March 2017).

mainly due to a few very large IXPs - such DE-CIX, Equinix, and HK-IX - that serve many hundreds of AS members.

5.2 Limitations

We are well aware that our methodology relies on datasets that have some limitations.

First, although we do our best to maintain and update the BGP blackhole community dictionary, it is possible that we miss community values for some blackholing providers. Moreover, the validation of discovered BGP blackhole communities is a slow process if we have to rely on private communication rather than official documentation, e.g., the official web page or the IRR records. Finally, it is possible (but unlikely) that ASes change their BGP blackholing communities or add additional ones. Maintaining an up-to-date dictionary of communities presents a challenging task.

Second, we only observe prefixes whose blackholing announcements are visible in the public Internet. Note that according to RFC7999 [42] and RFC5635 [45], blackhole announcements must not be propagated outside the local AS and should carry the no-export community. Nonetheless, our findings suggest that these recommendations are not respected by a number of networks, as we present in Section 9. Some networks do not rely on BGP to trigger blackholing, but provide specialized interfaces to their customers to request a prefix to be blackholed. For example, Cogent uses an interface that requires login and password by other networks or authorities to blackhole prefixes [20]. In February 2017, the blackholing of two CloudFlare IP addresses hosting Pirate Bay content received some attention [67]. The corresponding prefixes were not visible in any of our BGP datasets. However, by querying a BGP looking glass within Cogent, we were able to observe it. Hence, looking glasses have the potential to reveal additional blackholing that is invisible in BGP data.

Third, while we are using multiple BGP datasets and a large number of collectors (see Section 3), it is possible that a blackholing announcement that was in fact advertised in the public Internet is not observed at any of our collectors. This is the case if the collector does not directly peer with a BGP blackholing provider, or if a BGP blackholing provider strips blackhole communities or does not propagate the route to its peers, customers, or providers.

Summary: We conclude that even though our datasets have some limitations with regards to where the BGP collectors are located, and with whom and how they peer, they unveil blackholing activity for a large number of blackholing providers and users that are included in the blackhole communities dictionary. With regards to blackholed prefixes, the numbers are multiple times higher than reported previously [25, 54]. Nevertheless, this study provides a *lower bound* regarding the number of blackholing providers, blackholing users, and blackholed prefixes.

6 THE RISE OF BGP BLACKHOLING

To explore the extent to which BGP blackholing has been adopted by network operators, we apply our methodology to several large BGP datasets (see Section 3) from December 2014 till March 2017. We only report IPv4 data since the number of blackholed IPv6 prefixes is less than 1% of the total during this period.

Blackholing Adoption: To understand the longitudinal trends with regards to adoption, we show in Figure 4(a) the number of active blackholing providers per day that are visible across all datasets. We find that the number of blackholing providers has more than doubled during this period (from 40 in December 2014 to up to 100 in early 2017). At the same time, the number of routed transit ASes, i.e., ASes that carry traffic between at least two different other ASes and are hence possible blackholing providers, has increased only by 18%. Next, we turn our attention to blackholing users, see Figure 4(b). Here, the increase is even higher. Indeed, the daily blackholing user population has increased by a factor of 4 since December 2014, peaking up to 400 in early 2017. This increase again outpaces the 21% increase of all the visibly routed ASes in the public routing data. Even more striking is the increase in the number of prefixes over time. As shown in Figure 4(c), the daily number of prefixes increased from about 500 at the end of 2014 to over 3,000 in early 2017 and peaking over 5,000. The increased usage underlines the value of BGP-based blackholing for ISPs. Over the entire measurement period we were able to identify 161,031 blackholed prefixes i.e., a factor of 20 times higher than previously reported blackholing activity [25]. This is possible since our methodology does not depend on private vantage points but can be applied on publicly available BGP feeds from many different BGP peers.

Blackholing Activity Spikes: A closer look at the daily BGP blackholing activity shows that in addition to the constant increase in all three usage metrics, namely, number of blackholing providers,
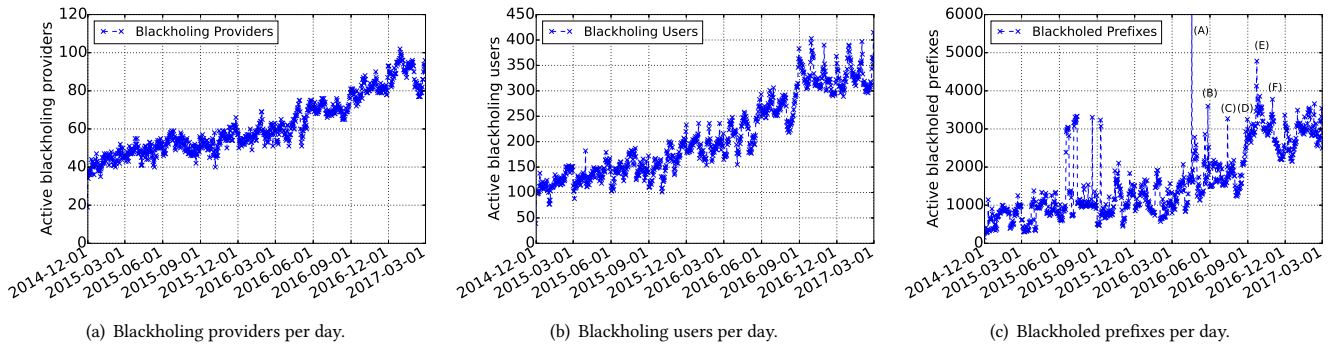
(a) Blackholing providers per day.

(b) Blackholing users per day.

(c) Blackholed prefixes per day.

Figure 4: Longitudinal growth of blackholing usage.

blackholing users, and blackholed prefixes, there are significant spikes. Manual investigation of the most noticeable spikes shows that they correlate well with large-scale cyber-attacks, especially with DDoS attacks, which made headlines in the press. This suggests that such attacks contribute to the rapid adoption of BGP blackholing as a mitigation technique.

In particular, we were able to correlate some of the most noticeable spikes in blackholing activity with well-documented DDoS attacks, which we annotate in Figure 4(c). Spike (B) on 2016/05/16 coincides with a large amplification attack against NS1, a major DNS provider with global footprint [5]. Spike (C) on 2016/07/15 occurred during the Turkish coup attempt which was accompanied by DDoS attacks against several news sites [63]. During spike (D) on 2016/08/22, DDoS attacks that peaked at 540 Gbps targeted the Rio Olympic games [68]. At the beginning of September 2017 we noticed a significant increase in the level of blackholing activity that lasted for months and correlated with the operation of the Mirai botnet [4]. The spike (E) correlates with the "Krebs on Security" attack [44] that started around 2016/09/20 and lasted for days. Spike (F) correlates with the massive attack against Liberia's Internet infrastructure on 2016/10/31 that is also related to the Mirai botnet [66]. We confirm that most of the remaining spikes in the first nine months of 2015 also correlate well with large scale DDoS attacks on corporate websites. An important observation is that the magnitude of the spike in blackholing activity depends on the number of blackholed prefixes, as well as the blackholing strategy of the users, i.e., whether they decide to blackhole individual IP addresses or entire prefixes. Thus, the magnitude of the spike alone can not be an indicator of how large an attack is.

We note that while we observe correlations across the number of blackholed prefixes and DDoS attacks, we can not readily determine whether the increase in blackholed prefixes is indeed the result of DDoS attacks taking place at the same time. We also find instances of accidental blackholing events, such as the spike (A) on 2016/04/18. This spike seems to be unrelated to a DDoS attack, but rather the result of a misconfiguration where a European academic network accidentally blackholed its entire routing table for less than 2 minutes. In future work, we plan to assess the causality between DDoS attacks and blackholing events.

7 BGP BLACKHOLING PROVIDERS

Next, we ask who and where are the BGP blackholing providers again focusing on the time period August 2016 to March 2017 (Table 4) . We start by grouping the observed blackholing providers according to their network type following the same technique as Section 4.1. As one might expect, most blackholing providers, 184 out of 242, are transit/access providers. This group has significant incentives for providing such services to their customers. They can reduce their own traffic by not forwarding the "unwanted" customer traffic. This group includes many large transit and Tier-1 providers. Consequently, the number of blackholing users using these blackholing providers is also the largest, 986 combined users out of 1,112. Similarly, the number of IPv4 prefixes that are blackholed via the transit/access providers accounts for about 90% of the total number of prefixes (80,262 out of 88,209).

IXPs are the second largest group of blackholing providers, offering blackholing to 60% of the total observed users and 25% of the total observed blackholed prefixes, although they account for "only" 25 out of 242 (10.3%) observed blackholing providers. That IXPs have a relatively greater proportion of blackholing users and prefixes is due, at least partially, to the IXPs in our dataset all contributing direct feeds to an available BGP collector, and therefore they offer good visibility in their blackholing activity. Moreover, most IXPs that offer blackholing service have a significant number of member ASes, often in the order of hundreds, which explains why the number of blackholing users that rely on the IXP blackholing service is the second largest among all the types of blackholing providers.

Although blackholing providers that are content providers are about as numerous as IXPs, they have a relatively small number of blackholing users and blackholed prefixes. Enterprise and unknown type blackholing providers are a small group with relatively small numbers of blackholing users and blackholed prefixes. For these we have limited direct feeds to our collectors, which also limits their visibility. To better understand how blackholing providers are used, we plot in Figure 5(a) the CDF of the number of blackholed prefixes for each of the blackholing providers separately for transit/access network provider and IXPs. We observe that the number of blackholed prefixes associated with transit/access networks ranges from a few to more than 1,000, and only 20 are associated with more than 1,000 blackholed prefixes. The CDF for the IXP group (recall it only has 25 members) roughly follows that for the transit/access
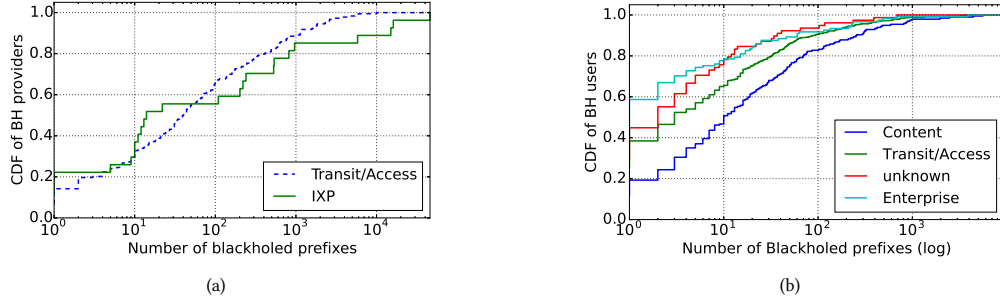
(a)

(b)

Figure 5: CDF: #blackholed prefixes per (a) blackholing provider and (b) blackholing user type.
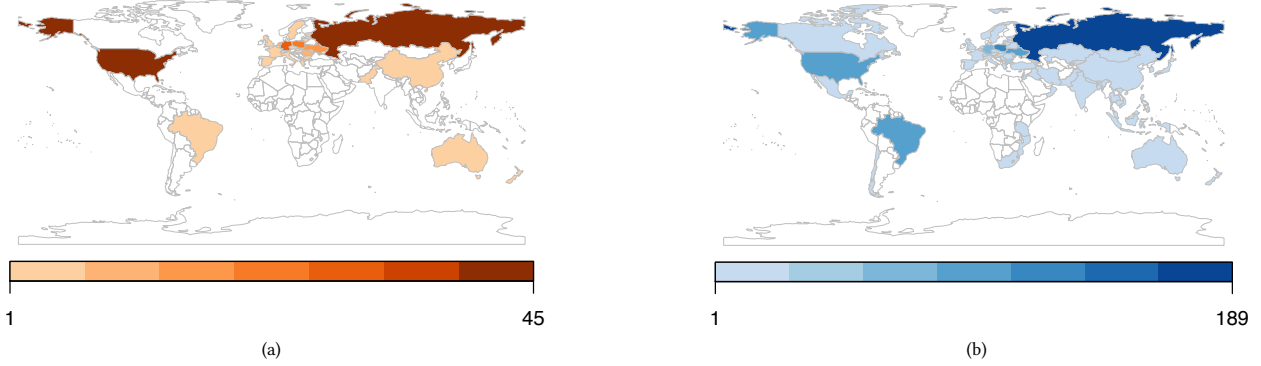

(a)

(b)

Figure 6: #blackholing (a) provider ASes and (b) user ASes per country.

providers. But is more extreme at each end: about 20% have just one blackholed prefix, as compared with 15%; and 14% have more than 1,000 blackholed prefixes. Next, we identify to which region each blackholing provider belongs to, see Figure 6(a), using the country registered in the RIR of each AS. Most blackholing providers are in Russia, USA, and Germany. IXPs that provide blackholing services are in major cities which are also telecommunication hubs, particularly in Europe, USA, and Asia.

8 BGP BLACKHOLING USERS

Next, we focus on the observed blackholing users and the profile of blackholed prefixes. Figure 6(b) is the counterpart of Figure 6(a) for the blackholing users. It highlights similarities and differences. Again the largest group is in Russia, US, and Germany, but other countries including Brazil and Ukraine are in the top-5. Figure 5(b) shows the CDF of the blackholing users vs. the number of blackholed prefixes (in log scale), for the period August 2016 to March 2017. Content providers are by far the most active group in terms of blackholed prefixes. While content providers account only for some 18% of the total blackholing users, they originate 43% of the blackholed prefixes. This confirms expectations, since content providers host servers that can potentially targets of attacks. Manual investigation shows that the large majority of these providers are small cloud providers and hosters and their top-5 locations are: Russia (46), US (30), Germany (21), Ukraine (18) and Poland (10).

Services/Applications on Blackholed IPs: Next, we are interested in application-layer characteristics of the blackholed prefixes or, more specifically, the IP addresses covered by blackholed prefixes.

Initially, we focus on March 2017 with 20,948 blackholed prefixes, which include 20,088 host routes (/32s) and cover a total of 5.2M unique IPv4 addresses. To determine which applications/services are offered by blackholed hosts, we take advantage of the Internet-wide scanning data as provided by *scans.io* [10, 27] for several protocols/port numbers, including TCP-SYN scans for HTTP(S), SSH, Telnet, FTP, SMTP(S), POP3(S), and IMAP(S) and UDP scan data for DNS and NTP.

We identified offered services, or more precisely, open host ports for more than 60% of the blackholed prefixes, see Figure 7(a). The classes are not mutually exclusive, since a host can offer multiple services. Overall, HTTP is the dominant service (53% of prefixes) and more than 90% (79%) of the FTP (SSH) servers are co-located with HTTP servers. This corresponds to the *default* configuration of hosters offering preconfigured virtualized Web servers. Moreover, we find that hosts in some 10% of the blackholed prefixes offer all 6 mail-related protocols. Hosts in 845 blackholed prefixes (some 4%) accept TCP connections on all of the 10 tested protocols, suggesting that these hosts might be tarpits [3].

Web Servers and Content: Given the large number of possible blackholed Web servers we next explore which content they host based on the results of an `HTTP GET` request to the respective IP addresses – data also provided by *scans.io*. We find that only 61% of the blackholed IP addresses reply with an HTTP response, while the general response rate is roughly 90% (for the entire population of HTTP servers). Thus, a significant share of blackholed hosts run *some* service on port 80 other than Web. In future work, we plan to explore this in more detail. Next, we explore which content the blackholed Web servers may host using the DNS lookups

for Alexa top 1M domain names, also from *scans.io*. While useful, the dataset is limited by the fact that the vantage point is a single location at the University of Michigan and, thus, we might miss some Domain-to-IP address mappings, e.g., for CDNs and other distributed infrastructures. In total, we find that only 334 black-holed prefixes (about 3% of HTTP hosts) host Web sites that are among the Alexa top 1M. This suggests that many of these Web servers are active and reachable, but do not host the typical popular content. This has to be expected since many Alexa top sites rely on CDNs and other large content hosters to provide reliable scalable services. Among the top domains are 51auto.com (rank 2282), ebrun.com (rank 3044), imooc.com (rank 3148), and gdz.ru (rank 5199). Among the most dominant TLDs are .com (38% of domains), followed by .ru (16%), .org (11.9%), .net (6%), and .se (3%). We confirmed our findings using data from DNSDB [29], a proprietary database with a list of active domains per IP for a specified period.

Malicious Activity of Blackholed IPs: Typically, blackholing is used when a blackholed prefix is the victim of a DDoS attack. However, we find incidents when blackholed addresses–a minority of the overall–engage in suspicious activity. For this analysis we use proprietary information that characterizes the source activity of IPs. In particular, we use IP-level information for (i) port scanners, i.e., IPs that perform port scans against a major CDN, (ii) vulnerability probes, i.e., IPs that scan multiple CDN servers for a specific port, and (iii) IPs that participate in repeated login attempts against CDN customers. This information is used, in part, in features of the Kona Site Defender service [2]. On a daily basis we find between 400 to 900 matches for blackholed prefixes in this database. More than 90% of these IPs are probers. The remaining ones are scanners and about 2% did both. Moreover, on a daily basis about 500 to 800 IPs in the blackholed prefixes participated in unauthorized login attempts. The union of all of the above IPs belong to about 2% of the blackholed prefixes.

9 BGP BLACKHOLING PRACTICES

Next, we explore common practices on the use of BGP blackholing for the period August 2016 to March 2017.

Global vs. Local Blackholing: Today, ASes peer at multiple locations and/or have multiple upstream providers, either for resilience or for cost efficiency [28, 34]. Many networks are also members at multiple IXPs [35, 43]. Thus, ASes can choose to use one or more of them as blackholing provider. Indeed, we can expect that during an attack a prefix is blackholed using multiple blackholing providers. We consider a *blackholing event* to be the blackholing of a prefix at one or at multiple providers concurrently (or within a small time window).

Figure 7(b) shows a histogram of the number of blackholing providers associated with a blackholed prefix for each blackhole event (y-axis in log scale). Note: a given prefix could be counted in multiple bins, for example if it ever were blackholed by just one provider, and another time was blackholed by, say, three providers. A significant fraction of prefixes, 28%, are associated with multiple blackholing providers, 2% are associated with more than 10 blackholing providers. The largest number of blackholing providers across our BGP datasets is 20. Note that the actual number of blackholing events with multiple blackholing providers may be even higher

given the visibility restrictions of blackholed prefixes that are not advertised with bundled communities, as explained in Section 4.2.

BGP Blackholing Propagation: The majority of the blackhole announcements are for host routes (/32s). Prefixes that are more specific than /24 are typically not accepted and not propagated by BGP. However, according to RFC7999 and RFC5635 [42, 45], blackhole announcements for more-specific prefixes should be accepted to enable blackholing for specific hosts, without affecting the reachability of large address spaces. Nonetheless, as explained in Section 5.2, the same RFCs require that the propagation of these prefixes outside the local AS must be suppressed.

We investigate the propagation patterns of blackholed prefixes by checking the position of the blackholing provider in the AS path in the BGP data[2]. Figure 7(c) shows that in the most common case (about 50%) the blackholing provider does not appear in the path and we are able to detect the blackholing activity thanks to the bundling of blackholing communities, see Section 4.2. In about 20% of the blackholing events the collector is on an IXP that is a blackholing provider, we annotate this as 0 AS distance. In more than 10% of the blackholing events the collector is directly peering with the blackholing provider, denoted 1 AS distance. Nevertheless, in 30% of the cases we observe that the blackholed prefix is propagated at least one hop away from the blackholing provider, 1 - 6 AS distance.

BGP Blackholing Duration Patterns: To understand the temporal dynamics of BGP blackholing, Figure 8(a) shows the CDF of the duration of each blackhole event (ungrouped) for August 2016 to March 2017. Hereby, the blackhole duration refers to the time between the start (announcement) and end (implicit or explicit withdrawal, see Section 4.2) of a blackholed prefix. Over 70% of the blackhole events have a duration of one minute or less, which may seem surprisingly short. However, when we group together events for the same prefix using a 5 minute timeout (assuming that the collector infrastructure is sufficiently closely synchronized), then just 4% of the (grouped) events are this short. This shows that a significant number of events, by these accounts more than 70% (5 minutes bin aggregation), have an ON/OFF pattern. While this looks strange at first, private communication with a number of blackholing users confirms this is a best practice. An awkwardness of using blackholing is the inability to know when an attack is over. Thus, it is a common practice to blackhole, e.g., a host, observe a drop in traffic, then withdraw the blackholing to check if the attack is over. If not, the process is repeated. This allows operators to reduce the impact of an attack while limiting its operational disruption.

Figure 8(a) also highlights that a minority of the events are rather long: 2% of the ungrouped events and 30% of the grouped events are longer than 16 hours. Figure 8(b) shows the histogram of the duration of the ungrouped events (x-axis in hours, y-axis log scale). We observe roughly three event regimes: short-lived (minutes), long-lived (weeks), and very long-lived (months). Manual investigation shows that some long and very long-lived events are either due to misconfiguration or intentional blocking of IPs with bad reputation. We do not (yet) have a general methodology to explain all instances of long-lasting blackholed prefixes.

---

[2]We only test this for BGP blackhole communities that are associated with a single AS.
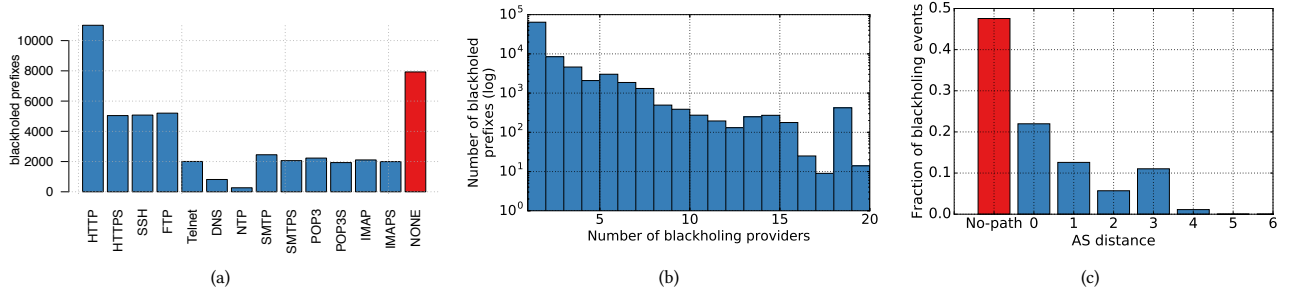
(a)

(b)

(c)

Figure 7: Distribution of (a) services run on blackhole hosts (aggregated per blackhole prefix), (b)# blackholing providers per BGP blackholing event, and (c) histogram of AS-distance between BGP collector and blackholing provider (No-path means that the blackholing provider is not in the AS path).
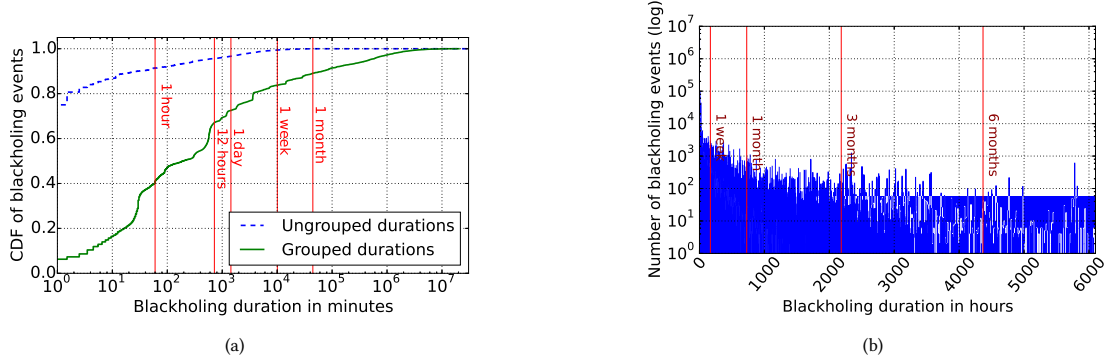

(a)

(b)

Figure 8: (a) CDF: Durations of blackholing events (Ungrouped) and periods (Grouped, using 5 minute timeout) and (b) histogram of blackholing events duration.

10 BGP BLACKHOLING EFFICACY

To assess the impact of BGP blackholing on network reachability and traffic, we use active and passive measurements collected during and after blackholing events.

Assessment using Active Measurements: To assess the data plane reachability of the blackholed prefixes, we orchestrated traceroute measurements to the blackholed prefixes. We utilize BGP-Stream [54] to obtain a near real-time BGP stream from all the RIPE RIS and RouteViews collectors, and upon detecting a new blackholing event, we select 40 probes from the RIPE Atlas platform [61].

To obtain a diverse set of probes, for each blackholing event we request ten probes for each one of the following four groups: probes in the downstream cone of the blackholing user, probes in the upstream cone, probes accessible through peering links and probes inside the blackholing user AS, according to CAIDA's AS relationship inferences [48]. We then select 4 probes (uniformly at random) from each group. If a group doesn't have enough probes we select the remaining probes randomly. For the traceroute target we select a host inside each blackholed prefix. We also select a non-blackholed target from the most specific prefix that includes the blackholed prefix, in order to compare their respective data plane reachability. For instance, if the blackholing prefix length is 32, we select another target in the same /31 if possible[3]. For each target we execute two traceroute queries from the same probe, one while

[3]Sometimes there are separate BGP announcements for blackholed /32 prefixes that are in the same /31, or /30. In this case the neighbor is chosen from the next less-specific prefix that includes the blackholed host and at least one non-blackholed one.

the blackholing event is active, and a follow-up measurement one hour after we detect the withdrawal of the blackholing. In total we collected traceroute paths for 2,967 blackholing events involving 337 blackholing users during March 2017. Our findings indicate that blackholing is effective. The reachability of the blackholed host decreases significantly, by an average of 5.9 IP-level hops when comparing the paths to the blackholed host during and after the blackholing.

Next, we study where on the path, during blackholing, the traffic is dropped in the sense of how far away from the destination, see Figure 9(a). Here, we only consider events where the destination was reachable again after the blackholing event to eliminate artifacts due to route changes, misconfigurations, and ICMP blocking which limits traceroute reachability [59]. We include two comparisons: (1) The first (red solid line) shows the difference in traced path lengths *after* vs. *during* the blackhole event for a given RIPE Atlas probe. The "path length" is the number of hops to the last responding interface on the trace. (2) Likewise, the second (blue-dashed line) shows the path length difference to the neighboring non-blackholed host vs. the blackholed host during the blackholing event. Both comparisons highlight that blackholing is successful. More than 80% of the paths to the blackholed hosts terminate earlier during the blackholing event than after the blackholing event. For about 15% of the cases the path to the blackholed host is of equal length or shorter. This can occur if (a) the prefix is not blackholed by utilizing all the providers, (b) not all providers offer blackholing, (c) BGP

(a) IP-level paths

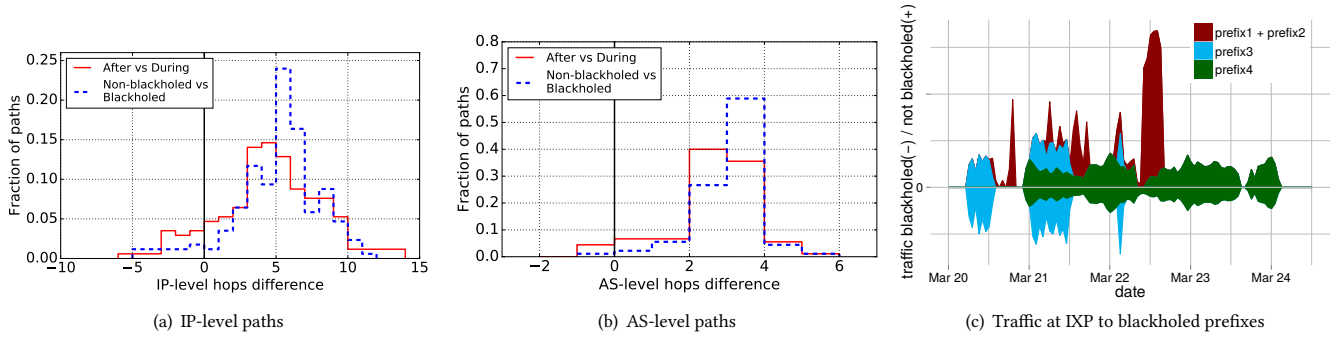(b) AS-level paths

(c) Traffic at IXP to blackholed prefixes

Figure 9: Histogram of impact of blackholing on (a) IP-level and (b) AS-level paths for /32 blackholed hosts, and (c) levels for blackholed (negative y-axis) and non-blackholed traffic (positive y-axis) to blackholed prefixes.

misconfigurations, (d) route changes, or even (e) inconsistency in the responding interface, e.g., due to load balancing.

We use the same data to study the impact on the length of the AS-level path, where, likewise, the "AS-level path length" is the number of ASes on the path to the last responding interface on the trace, see Figure 9(b). Overall, we see a significant shortening of the AS-level path, on average 2 to 4 AS-hops, during vs. after the blackholing event. This finding confirms our observation from the analysis of the control-plane paths, that in many cases the blackholing request is propagated along the AS path. We also observe that in 16% of the cases the traffic to the destination is dropped at the destination AS or the upstream provider. Therefore, if blackholing is used and deployed at scale it is possible to significantly reduce unwanted/attack traffic in the Internet, especially for long lasting large-scale DDoS attacks. Lastly, we observe that for blackholed prefixes less specific than /24 we find virtually no difference between the path lengths, when comparing the paths during and after the blackholing events and between the blackholed and non-blackholed hosts. This finding indicates that operators respect, to a large extent, the requirement to blackhole prefixes only more specific than /24.

Assessment using Passive Measurements: To validate our inferences we collaborate with major blackholing providers: IXPs in the USA, Central and South Europe. First, we validate the completeness of the visibility of our BGP datasets regarding the blackholed prefixes at each of the IXPs. We confirm that we have 99.5% visibility of all blackholing events that involve the IXPs route server. We also confirm that the large majority of blackholed prefixes are IPv4 /32s. Next, we check the impact on traffic. Therefore, we use IPFIX [17] traffic traces collected from the switching fabric of a major European IXP, which offers a blackholing service. The traces, sampled at a rate of 1 out of 10K packets, provide flow-level information about the source/destination IP, port, and exchanged traffic volume via the IXP infrastructure on a per IXP member basis. To understand if and how much traffic is discarded, we focus on the blackholed prefixes with the largest traffic volumes at the IXP (these prefixes are blackholed throughout the week). Figure 9(c) shows the traffic volume in two stacked plots across one week. The values below the zero line are the traffic that is dropped at the IXP while the ones on top of the zero line show the traffic still traversing the IXP to its destination. We observe that a significant amount, i.e., more than 50% of traffic for some of the successfully announced /32s is dropped (negative part). Nevertheless, some traffic still is forwarded

via the IXP towards the destination[4]. A closer analysis regarding the traffic sources reveals that 80% comes from less than ten member ASes. Apparently, they do not honor the correctly announced BGP blackhole route. Looking into the IXP's route server logs points out two reasons: (a) some ASes do not accept /32 announcements because they have not yet changed their router configurations appropriately, (b) other ASes do not use the route server and, thus, miss the announcement.

To complement our findings we focus on prefixes that are blackholed on the control plane (announced to the route server) but for which we observe no reduction on the data plane (red region in Figure 9(c)). Our analysis shows that a common reason is misconfiguration at the blackholing user: (a) the blackholing user must maintain proper entries within the RIR/verification database that is used to filter incoming announcements at IXPs—the blackholing provider. The route servers will only redistribute prefixes to other peers if the advertising AS is authorized to announce this prefix, (b) some announcements are misconfigured in the sense that they either carry invalid next hop IPs or wrong BGP communities.

We confirm this finding for other ASes by analyzing all blackholed /32 prefixes for one day. Of 508 ASes that send traffic to these IPs, for only about one third of these ASes, traffic is dropped for at least one of the blackholed IPs. Note, this is a lower bound. During large-scale DDoS attacks even networks that do not use the route server discard traffic to DDoS victims, indicating that blackholing is also used over bilateral peering sessions. Overall, we confirm that blackholing can be efficient if configured according to best common practices. However, we also find room for improvement. We strongly encourage operators to update their BGP configurations to accept advertised prefixes more specific than /24 and blackholing users to maintain proper RIR/verification database entries.

11 IMPLICATIONS

Implications to Network Troubleshooting: Our methodology of studying the BGP unreachable signals due to blackholing allows us to infer reachability problems of a destination solely based on control plane-data, including the root cause—useful for troubleshooting reachability. This is in contrast to previous work which combined active and passive measurements for investigation of unreachability of prefixes in the Internet [40, 41]. Our methodology

---

[4]We do not know if it reaches the destination.

can also accurately estimate when a prefix starts to be unreachable via the data plane, again by monitoring the control plane only. Such insights are also important for regulatory authorities which need to understand reachability issues. Since our methodology can also identify unconventional use of BGP blackholing, it has the potential to reveal censorship by entities that restrict access to information.

Implications for Reputation: An important asset of network operators, cloud/hosters, and cloud tenants is the reputation of their IP address space. Our methodology enables potential customers to check to which degree the IPs of a provider are repeatedly blackholed. However, our results differ from other reputation systems as we have information about the targets rather than the sources of problematic traffic. On the one hand the use of blackholing indicates that the provider is able to tackle DDoS attacks. On the other hand it might be indicative of customers with frequently attacked services, mis-configurations, or other mis-management. Moreover, the blackholing may also be caused by a third party. In the case of unauthorized blackholing (via prefix hijacking) it is possible to determine the blackholing provider and start an investigation. For cloud services, BGP blackholing may be problematic as all applications that use the same shared IP will be blocked. Thus, potential customers may shy away from providers with IP address space that is too frequently blackholed. Indeed, ongoing work investigates fine-grained blackholing, where additional restrictions, such as a given port number, are imposed [14, 24]. Thus, our tools and analysis of blackholing activity can be used to enhance existing reputation systems, e.g., IP reputation based on previous malicious activity.

Need for Standardization: Unfortunately, there is no strict convention for BGP communities. Yet, a recent initiative for standardizing a BGP blackhole community [42] is already adopted by a large majority of the IXPs and some of the ISPs. This points out that standardization is feasible and that operators do adopt. With "standardized" BGP blackhole communities, the corresponding BGP configuration is simplified and, thus, usage should increase while the risk of misconfiguration decreases. It is important to establish best common practices, e.g., strategies on targeted use of blackholing that mitigates the negative impact on legitimate traffic, or on how to validate BGP blackholing announcements, and guidelines regarding the size of the blackholed prefix. The latter is needed to reduce misuse and unintentional mis-configuration of BGP blackholing which can result in large-scale reachability problems.

12 CONCLUSION

In this paper we provide the first Internet-wide study of the state and adoption of a popular attack-mitigation technique, BGP blackholing. We develop a methodology to infer BGP blackholing based on BGP announcements and BGP community tags, which enables a significant increase in the visibility of the provider and user networks of this capability.

Our study shows that the documented increase of cyber-attacks and threats in the Internet has significantly increased the adoption of BGP blackholing. The number of blackholing providers has more than doubled the last three years. Our analysis shows that, today, more than 300 networks, including about 50 IXPs, worldwide offer blackholing as a service to their customers, peers, and members

respectively. Even more impressive is the rise in the number of blackholing users and the number of blackholed prefixes. During the last three years, the number of blackholing users has increased fourfold and the number of blackholed prefixes has increased sixfold, peaking up to 400 blackholing users and up to 5K blackholed prefixes per day, respectively, in recent months. Along with the steady increase of BGP blackholing adoption, our study identifies spikes that correlate well with large-scale DDoS attacks. We also show that BGP blackholing delivers on the promise of dropping the unwanted traffic early saving, typically by two or three AS-hops and multiple IP-hops. Thus, blackholing is readily available, cheap, and effective at attack-mitigation and reducing traffic on intermediate networks, and has already been widely adopted. However, blackholing also discards legitimate traffic, and can have the drawback that the target becomes unreachable – the goal of the DDoS attack. Also, if the target organization has purchased a traffic-scrubbing service, then this service is degraded if the traffic is discarded prior to reaching the scrubbing center. Hence there is a conflict of interests between organizations that use such a service, and intermediate networks through which the traffic would traverse. The negative impact of blackholing is less severe if (much of) the attack traffic originates from only a few locations and blackholing is limited to the associated providers. We find from our passive measurements that blackholing can be even more effective if all operators would follow best common practices. Best common practices are in need to minimize the negative externalities of BGP blackholing, i.e., mitigate the impact on legitimate traffic. We believe that our study provides insights to potential providers and users of BGP blackholing and develops tools to better study reachability issues in the Internet, improve troubleshooting, increase transparency, and inform peering and hosting decisions.

ACKNOWLEDGMENTS

REFERENCES

[1] Akamai. 2015–2017. State of the Internet; Quarterly Security Reports. https://www.akamai.com/us/en/our-thinking/state-of-the-internet-report/global-state-of-the-internet-security-ddos-attack-reports.jsp. (2015–2017).

[2] Akamai. 2017. Kona Site Defender. https://www.akamai.com/us/en/products/cloud-security/kona-site-defender.jsp. (2017).

[3] L. Alt, R. Beverly, and A. Dainotti. 2014. Uncovering Network Tarpits with Degreaser. In *ACSAC*.

[4] M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou. 2017. Understanding the Mirai Botnet. In *USENIX Security Symposium*.

[5] Ars Technica. 2016. Major DNS provider hit by mysterious, focused DDoS attack. https://arstechnica.com/information-technology/2016/05/major-dns-provider-hit-by-mysterious-focused-ddos-attack. (May 2016).

[6] BBC. 2012. 'Hacking attacks' hit Russian political sites. http://www.bbc.com/news/technology-16032402. (2012).

[7] S. Bird. 2006. NLTK: The Natural Language Toolkit. In *COLING-ACL*.

[8] C. Partridge, P. Barford, D. D. Clark, S. Donelan, V. Paxson, J. Rexford, and M. K. Vernon. 2003. *The Internet Under Crisis Conditions: Learning from September 11*. The National Academy Press.

[9] CAIDA. 2014-2017. AS Classification. https://www.caida.org/data/as-classification/. (2014-2017).

[10] Censys Team at the University of Michigan. 2017. Internet-Wide Scan Data Repository. https://scans.io/. (2017).

[11] R. Chandra, P. Traina, and T. Li. 1996. BGP Communities Attribute. IETF RFC 1997. (1996).

[12] N. Chatzis, G. Smaragdakis, A. Feldmann, and W. Willinger. 2013. There is More to IXPs than Meets the Eye. *ACM CCR* 45, 5 (2013).

[13] W. R. Cheswick, S. M. Bellovin, and A. D. Rubin. 2003. *Firewalls and Internet Security: repelling the wily hacker*. Addison-Wesley.

[14] M. Chiesa, C. Dietzel, G. Antichi, M. Bruyere, I. Castro, M. Gusat, T. King, A. Moore, T. Nguyen, P. Owezarski, S. Uhlig, and M. Canini. 2016. Inter-domain Networking Innovation on Steroids: Empowering IXPs with SDN Capabilities. *IEEE Communications Magazine* 54, 10 (2016), 102–108.

[15] K. Cho, C. Pelsser, R. Bush, and Y. Won. 2011. The Japan Earthquake: the impact on traffic and routing observed by a local ISP. In *ACM CoNEXT SWID workshop*.

[16] CISCO. 2005. Remotely Triggered Black Hole Filtering - Destination Based and Source Based. Cisco White Paper, http://www.cisco.com/c/dam/en_us/about/security/intelligence/blackhole.pdf. (2005).

[17] B. Claise, B. Trammell, and P. Aitken. 2013. RFC 7011: Specification of the IPFIX Protocol for the Exchange of Flow Information. (2013).

[18] D. D. Clark. 1988. The Design Philosophy of the DARPA Internet Protocols. In *ACM SIGCOMM*.

[19] D. D. Clark, J. Wroclawski, K. Sollins, and R. Braden. 2002. Tussle in Cyberspace: Defining Tomorrow's Internet. In *ACM SIGCOMM*.

[20] Cogent. 2017. Customer User Guide. http://www.cogentco.com/files/docs/customer_service/guide/global_cogent_customer_user_guide.pdf. (2017).

[21] Cymru. 2017. BGP Bogon Refence. http://www.team-cymru.org/bogon-reference-bgp.html. (2017).

[22] Daily Mirror. 2016. Hackers attack the Stock Exchange: Cyber criminals take down website for more than two hours as part of protest against world's banks. (2016).

[23] Deutsche Welle. 2015. Anonymous hacktivist explains why group is targeting Saudi Arabian government. (2015).

[24] C. Dietzel, G. Antichi, I. Castro, E. Fernandes, M. Chiesa, and D. Kopp. 2017. SDN-enabled Traffic Engineering and Advanced Blackholing at IXPs. In *Proceedings of the ACM Symposium on SDN Research*.

[25] C. Dietzel, A. Feldmann, and T. King. 2016. Blackholing at IXPs: On the Effectiveness of DDoS Mitigation in the Wild. In *PAM*.

[26] B. Donnet and O. Bonaventure. 2008. On BGP Communities. *ACM CCR* 38, 2 (Mar 2008), 55–59.

[27] Z. Durumeric, E. Wustrow, and J. A. Halderman. 2013. ZMap: Fast Internet-Wide Scanning and its Security Applications. In *USENIX Security Symposium*.

[28] P. Faratin, D. D. Clark, S. Bauer, W. Lehr, P. Gilmore, and A. Berger. 2008. The Growing Complexity of Internet Interconnection. *Communications and Strategies* 72 (2008).

[29] Farsight Security. 2017. DNSDB. https://www.dnsdb.info/. (2017).

[30] L. Gao and J. Rexford. 2001. Stable Internet Routing Without Global Coordination. *IEEE/ACM Trans. Networking* 9, 6 (2001), 681–692.

[31] D. Gillman, Y. Lin, B. Maggs, and R. K. Sitaraman. 2015. Protecting Websites from Attack with Secure Delivery Networks. *IEEE Computer Magazine* 48, 4 (2015).

[32] V. Giotsas, A. Dhamdhere, and k. claffy. 2016. Periscope: Unifying Looking Glass Querying. In *PAM*.

[33] V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben. 2017. Detecting Peering Infrastructure Outages in the Wild. In *ACM SIGCOMM*.

[34] V. Giotsas, M. Luckie, B. Huffaker, and k. claffy. 2014. Inferring Complex AS Relationships. In *ACM IMC*.

[35] V. Giotsas, G. Smaragdakis, B. Huffaker, M. Luckie, and k. claffy. 2015. Mapping Peering Interconnections at the Facility Level. In *CoNEXT*.

[36] S. Goldberg. 2014. Why is It Taking So Long to Secure Internet Routing? *Comm. of the ACM* 57, 10 (2014).

[37] E. Heilman, D. Cooper, L. Reyzin, and S. Goldberg. 2014. From the Consent of the Routed: Improving the Transparency of the RPKI. In *ACM SIGCOMM*.

[38] J. Heitz, J. Snijders, K. Patel, I. Bagdonas, and N. Hilliard. 2017. BGP Large Communities Attribute. IETF RFC 8092. (2017).

[39] M. Jonker, A. Sperotto, R. van Rijswijk-Deij, R. Sadre, and A. Pras. 2016. Measuring the Adoption of DDoS Protection Services. In *ACM IMC*.

[40] E. Katz-Bassett, H. V. Madhyastha, J. P. John, A. Krishnamurthy, D. Wetherall, and T. Anderson. 2008. Studying Black Holes in the Internet with Hubble. In *NSDI*.

[41] S. Khattak, D. Fifield, S. Afroz, M. Javed, S. Sundaresan, V. Paxson, S. J. Murdoch, and D. McCoy. 2016. Do You See What I See? Differential Treatment of Anonymous Users. In *NDSS*.

[42] T. King, C. Dietzel, J. Snijders, G. Doering, and G. Hankins. 2016. BLACKHOLE Community. IETF RFC 7999. (2016).

[43] R. Kloti, B. Ager, V. Kotronis, G. Nomikos, and X. Dimitropoulos. 2016. A Comparative Look into Public IXP Datasets. *ACM CCR* 46, 1 (2016).

[44] KrebsOnSecurity. 2017. KrebsOnSecurity Hit With Record DDoS. https://krebsonsecurity.com/2016/09/krebsonsecurity-hit-with-record-ddos. (2017).

[45] W. Kumari and D. McPherson. 2009. Remote Triggered Black Hole Filtering with Unicast Reverse Path Forwarding (uRPF). IETF RFC 5635. (2009).

[46] C. Labovitz, A. Ahuja, A. Bose, and F. Jahanian. 2000. Delayed Internet Routing Convergence. In *ACM SIGCOMM*.

[47] C. Labovitz, G. R. Malan, and F. Jahanian. 1999. Origins of Internet routing instability. In *IEEE INFOCOM*.

[48] M. Luckie, B. Huffaker, A. Dhamdhere, V. Giotsas, and kc claffy. 2013. AS Relationships, Customers Cones, and Validations. In *ACM IMC*.

[49] A. Lutu, M. Bagnulo, and O. Maennel. 2013. The BGP Visibility Scanner. In *IEEE Infocom Workshops*.

[50] Merit Network. 2017. Merit RADb. http://radb.net/. (2017).

[51] New York Times. 2017. Hackers Hit Dozens of Countries Exploiting Stolen N.S.A. Tool. https://www.nytimes.com/2017/05/12/world/europe/uk-national-health-service-cyberattack.html. (2017).

[52] O. Nordstrom and C. Dovrolis. 2004. Beware of BGP attacks. *ACM CCR* 34, 2 (2004).

[53] E. Nygren, R. K. Sitaraman, and J. Sun. 2010. The Akamai Network: A Platform for High-performance Internet Applications. *SIGOPS Oper. Syst. Rev.* 44, 3 (2010).

[54] C. Orsini, A. King, D. Giordano, V. Giotsas, and A. Dainotti. 2016. BGPStream: a software framework for live and historical BGP data analysis. In *ACM IMC*.

[55] Packet Clearing House. 2014–20017. Routing archive. https://www.pch.net/resources/data.php. (2014–20017).

[56] Packet Clearing House. 2017. PCH PoPs. https://www.pch.net/about/pops. (2017).

[57] PeeringDB. 2014–2017. PeeringDB website. https://www.peeringdb.com. (2014–2017).

[58] P. Richter, G. Smaragdakis, A. Feldmann, N. Chatzis, J. Boettger, and W. Willinger. 2014. Peering at Peerings: On the Role of IXP Route Servers. In *ACM IMC*.

[59] P. Richter, G. Smaragdakis, D. Plonka, and A. Berger. 2016. Beyond Counting: New Perspectives on the Active IPv4 Address Space. In *ACM IMC*.

[60] RIPE. 2014–2017. Routing Information Service. http://www.ripe.net/ris/. (2014–2017).

[61] RIPE. 2014–2017. RIPE Atlas. https://atlas.ripe.net/. (2014–2017).

[62] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush. 2011. 10 Lessons from 10 Years of Measuring and Modeling the Internet's Autonomous Systems. *IEEE J. on Sel. Areas in Comm.* 29, 9 (2011).

[63] Russia Today. 2016. RT targeted by massive DDoS attack during attempted Turkey coup. https://www.rt.com/news/351645-rt-massive-ddos-attack/. (2016).

[64] S. Sangli, D. Tappan, and Y. Rekhter. 2006. BGP Extended Communities Attribute. IETF RFC 4360. (2006).

[65] R. Stapleton-Gray and W. Woodcock. 2011. National Internet Defense – Small States on the Skirmish Line. *Comm. of the ACM* 54, 3 (2011).

[66] Telegraph. 2017. Unprecedented cyber attack takes Liberia's entire internet down. http://www.telegraph.co.uk/technology/2016/11/04/unprecedented-cyber-attack-takes-liberias-entire-internet-down/. (2017).

[67] TorrentFreak. 2017. Internet Backbone Provider Cogent Blocks Pirate Bay and other Pirate Sites. https://torrentfreak.com/internet-backbone-provider-cogent-blocks-pirate-bay-and-other-pirate-sites-170209/. (2017).

[68] Tripwire. 2016. How a Massive 540 Gb/sec DDoS Attack Failed to Spoil the Rio Olympics. https://www.tripwire.com/state-of-security/security-data-protection/cyber-security/how-a-massive-540-gbsec-ddos-attack-failed-to-spoil-the-rio-olympics/. (2016).

[69] University of Oregon. 2014–2017. Routeviews Project. http://www.routeviews.org/. (2014–2017).

[70] L. Wang, X. Zhao, D. Pei, R. Bush, D. Massey, A. Mankin, S. F. Wu, and L. Zhang. 2002. Observation and Analysis of BGP Behavior under Stress. In *ACM IMW*.