Profiling BGP Serial Hijackers: Capturing Persistent Misbehavior in the Global Routing Table

Cecilia Testart
MIT
ctestart@csail.mit.edu

Philipp Richter
MIT
richterp@csail.mit.edu

Alistair King
CAIDA, UC San Diego
alistair@caida.org

Alberto Dainotti
CAIDA, UC San Diego
alberto@caida.org

David Clark
MIT
ddc@csail.mit.edu

ABSTRACT

BGP hijacks remain an acute problem in today's Internet, with widespread consequences. While hijack detection systems are readily available, they typically rely on a priori prefix-ownership information and are reactive in nature. In this work, we take on a new perspective on BGP hijacking activity: we introduce and track the long-term routing behavior of *serial hijackers*, networks that repeatedly hijack address blocks for malicious purposes, often over the course of many months or even years. Based on a ground truth dataset that we construct by extracting information from network operator mailing lists, we illuminate the dominant routing characteristics of serial hijackers, and how they differ from legitimate networks. We then distill features that can capture these behavioral differences and train a machine learning model to automatically identify Autonomous Systems (ASes) that exhibit characteristics similar to serial hijackers. Our classifier identifies ≈ 900 ASes with similar behavior in the global IPv4 routing table. We analyze and categorize these networks, finding a wide range of indicators of malicious activity, misconfiguration, as well as benign hijacking activity. Our work presents a solid first step towards identifying and understanding this important category of networks, which can aid network operators in taking proactive measures to defend themselves against prefix hijacking and serve as input for current and future detection systems.

CCS CONCEPTS

• **Networks** → **Network measurement**; **Network security**.

KEYWORDS

Internet security, BGP, routing, route hijacks.

ACM Reference Format:
Cecilia Testart, Philipp Richter, Alistair King, Alberto Dainotti, and David Clark. 2019. Profiling BGP Serial Hijackers: Capturing Persistent Misbehavior in the Global Routing Table. In *Internet Measurement Conference (IMC '19), October 21–23, 2019, Amsterdam, Netherlands.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3355369.3355581

IMC '19, October 21–23, 2019, Amsterdam, Netherlands

ACM ISBN 978-1-4503-6948-0/19/10…$15.00
https://doi.org/10.1145/3355369.3355581

1 INTRODUCTION

BGP's lack of route authentication and validation remains a pressing problem in today's Internet. The lack of deployment of basic origin validation of route announcements in BGP not only makes the Internet more susceptible to connectivity issues due to misconfigurations, but also opens the door for malicious actors. While a long-standing problem, its severity becomes clear in numerous recent reports of widespread connectivity issues due to BGP misconfiguration [14], as well as hijacking events of popular destinations in the Internet [38]. Episodes range from simpler attacks with the goal of using blocks to send spam emails [56] to more sophisticated misuse of BGP to intercept traffic or steal crypto currencies [9].

While the operator and research communities have devoted substantial resources to improve the state-of-the-art of BGP security (*i.e.,* the RPKI [12]), little has changed in production environments. Today, operators can use monitoring services [2] to automatically detect potential hijacks of their prefix announcements. Current hijack detection systems typically rely on assumptions of prefix ownership and track origin changes in the global routing table. If an event is detected, the victim network can react and attempt to get in contact with the perpetrator or its upstream networks to solve the problem. However, many times this contact is not fruitful or not even possible. At that point, victims of hijacks are only left with the option of publicly disclosing the event in network operator mailing lists in the hope that peer pressure and manual interventions by other networks, such as filtering announcements or refusing to provide transit, will remediate the situation.

What most BGP hijack detection systems have in common is that *(i)* they are *reactive* in nature, *i.e.,* they identify hijacking events only after they occurred, and *(ii)* they are event-based, *i.e.,* they track individual hijacking events. However, malicious BGP behavior by an actor is sometimes consistent over time, creating opportunities for methods based on longitudinal analysis, potentially informing *proactive* approaches (*e.g.,* scoring systems) and providing situational awareness. We indeed find that many hijacking events disclosed in operator mailing lists and network security blogs involve malicious Autonomous Systems (ASes) that repeatedly hijack prefixes, *i.e.,* originate prefixes allocated to and routed by other networks. In fact, some of these ASes show malicious activity in the global routing table for *multiple years*, and we refer to networks of this type as *serial hijackers*. Serial hijackers pose an ongoing threat, yet they have received surprisingly little attention in terms of empirical assessment.

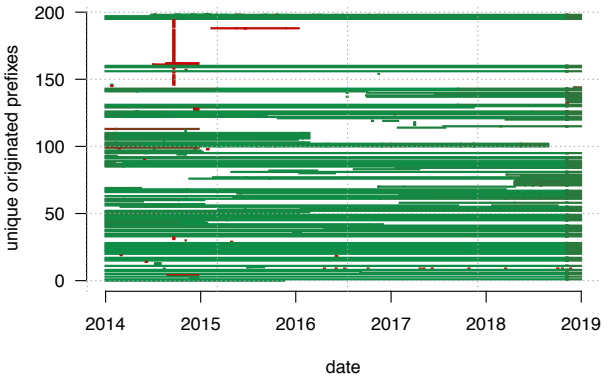
In this paper, we provide a systematic empirical analysis of the behavioral characteristics of serial hijacker ASes. We take on a new perspective on illicit BGP activity: instead of looking at individual BGP hijacking events, we study the long-term prefix advertisement dynamics in the global routing table in space and time. Our analysis leads to a set of key attributes that broadly capture the behavior of serial hijacker ASes, highlighting several interesting and previously undocumented cases. Our main contributions are:

- We provide a detailed and longitudinal study of BGP announcement dynamics of serial hijacker ASes over the course of 5 years. We develop hypotheses on prefix origination behavior (announcement stability, visibility, growth, address space fragmentation, origin conflicts) and identify dominant characteristics of serial hijackers and how they differ from legitimate ASes. We show that some of these behavioral patterns are clearly visible when studying announcement dynamics of networks over long time periods.
- Based on these behavioral patterns, we propose a set of metrics and we use a machine-learning model to evaluate their applicability to the problem of automatically identifying ASes with BGP origination patterns similar to serial hijackers. Our classifier flags ≈ 900 ASes that exhibit characteristics similar to our ground truth serial hijackers. We provide a detailed analysis of these preliminary results, revealing insight into false positives, actual malicious activity, as well as ASes appearing as illegitimately originating prefixes because of third-party misconfigurations.
- We illuminate behavioral patterns of serial hijackers in the wild with three case studies featuring a known serial hijacker, and two newly identified ones: a detailed analysis of multiple years of hijacking activity by AS197426, a glaring case of a serial hijacker from our ground truth dataset; AS19529, a hijacker network that was detected by our classifier, where we found corroborating evidence of hijacks; and AS134190, the network that, among the ASes we identify, shows the most recent indications of potential serial hijacker behavior.

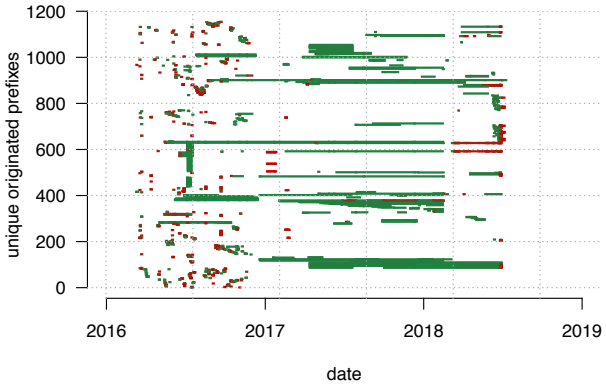
This work shows that, through analysis of readily available public BGP data—without leveraging blacklists or other indicators—it is possible to identify dominant patterns of serial hijackers. Our preliminary results suggest that these patterns can be leveraged in automated applications, potentially revealing undetected behavior or generating a novel category of reputation scores. Our findings have thus relevance for the operator community, since they can aid network operators to identify suspicious ASes *a priori*, potentially allowing for preventive defense. Our findings are also of relevance to the broader research community, since they provide viable input for new prefix hijacking detection systems, as well as for the development of AS reputation metrics and scoring systems.

To the best of our knowledge, this is the first work focusing on the long-term characteristics of this important category of networks, serial hijacker ASes. We make the feature dataset and the results of this work publicly available to allow both for reproducibility and for other works to leverage our list of identified ASes.[1] While the majority of hijacker ASes only target the IPv4 space, we show metrics both for IPv4 and IPv6.

[1] Auxiliary material can be found at https://github.com/ctestart/BGP-SerialHijackers.



(a) Legitimate AS: Prefix origination of AS5400 (British Telecom) over the course of 5 years. This AS originates prefixes consistently over long time periods.



(b) Serial Hijacker AS: Prefix origination of AS3266 (Bitcanal) over the course of almost 3 years. This AS announces a large number of prefixes over short time periods.

Figure 1: Long-term prefix announcement behavior for a regular AS, and a serial hijacker AS. We visualize each originated prefix as a row on the y-axis and color prefixes by their normalized visibility in the global routing table. We sort prefixes numerically and show time (3-5 years) on the y-axis.

2 BACKGROUND

To bootstrap our analysis, we first introduce the *serial hijacker* network type, and illustrate some of its pertinent characteristics by example. We review related work in the field of hijack detection and network profiling, and present a roadmap for this paper.

2.1 Introducing Serial Hijackers

Since as of today, no reliable and widely deployed system to automatically discard illegitimate BGP route announcements exists, the network operator community frequently relies on mailing lists (*e.g.,* NANOG [6]) to exchange information about illegitimate BGP announcements and to coordinate efforts to limit their propagation and impact by blocking announcements from networks originating such prefixes. The key observation that motivates this work came from studying 5 years of threads from operator mailing lists: many

reported hijacks are not "one-off" events, where a previously unknown AS number starts to advertise prefixes. Instead, we often find reports of the very same ASes repeatedly carrying out prefix hijacks. In fact, some of these networks continue to hijack different prefixes over the course of multiple *years*. Figure 1b shows a visualization of the origination activity of AS3266, a network that was repeatedly reported to hijack address space. We see that, over the course of 3 years, this AS originated almost 1,200 unique prefixes, and we observe a highly irregular pattern of short-lived origination of disparate address blocks. To put this behavior in contrast, we show the origination activity of AS5400 (British Telecom) in Figure 1a. This network, a large British residential and mobile ISP, shows a much more steady pattern, longer prefix announcement times, and an overall constant, and monotonically increasing number of advertised prefixes. We note, however, that also legitimate ASes can exhibit irregular patterns (see the white space between lines indicating a prefix was not originated at that time), often due to configuration issues of the network in question or of third-party ASes. Thus, metrics and systems attempting to isolate ASes with potentially malicious behavior must be chosen and evaluated carefully to allow for robustness. From Figure 1 it becomes clear that these two networks show wildly different long-term behavior in the global routing table. The goal of this paper is to identify and scrutinize the dominant prefix origination characteristics of this important class of networks: serial hijackers.

2.2 Related Work

BGP vulnerabilities and hijacks have been studied for a long time [10, 13, 36, 37, 52]. However, proposals to secure BGP have not gained widespread traction. Even though the Internet Engineering Task Force (IETF) standardized BGP prefix origin authorization and validation many years ago [29, 30, 47], deployment in production networks is still limited [16, 18]. As a result, BGP hijacks are a prevalent threat and concern for network operators [48]. There have been many efforts in the research community to characterize BGP hijacking events [28, 56] and to develop hijack detection systems using different approaches, metrics, and vantage points [22, 27, 42, 43, 46, 49, 50, 57]. While most systems focus on detecting individual BGP hijacking events, some attempt to identify the source of the cause and a few even tackle mitigation and remediation [7].

In contrast to most earlier works on BGP hijacks, our approach works by profiling the network-wide BGP prefix origination behavior of ASes. Few previous works study network-wide behavior of malicious actors. In [45], the authors study BGP announcements dynamics of prefixes found in email spam blacklists. They find that some spammers use short-lived (a few minutes long) BGP route announcements of large address blocks to send spam from IP addresses scattered throughout the advertised prefix. In [51], the authors study ASes that are over-represented in blacklists of phishing, scam, spam, malware and exploited hosts. Analyzing a month of BGP data, they find that these ASes are more likely to become unreachable and that they have more changes in their connectivity than most ASes in the Internet. Konte et al. [25] developed a system to identify bulletproof hosting ASes, leveraging features such as frequent re-wiring of transit interconnections. Our work is complementary in that we do focus on a specific group of malicious ASes, *serial hijackers*. We focus exclusively on behavioral characteristics related to their BGP origination patterns (*i.e.,* we do not leverage any data other than BGP for our classification), and specifically study *long-term* behavior of networks.

2.3 Roadmap

The rest of the paper is organized as follows: in § 3 we first describe how we build a ground-truth dataset of serial hijacker ASes, as well as a control set of legitimate ASes. We also introduce our longitudinal dataset that covers 5 years of BGP activity at a 5-minute granularity. We introduce necessary data cleaning and preprocessing steps in § 4. In § 5, we first introduce a set of behavioral characteristics and pose hypotheses on how the behavior of serial hijacker ASes might differ from legitimate ASes. For each category, we introduce different metrics to capture AS behavior and study in detail how serial hijackers' BGP origination behavior differs from legitimate ASes in our ground-truth dataset and how our metrics capture these differences. With our metrics in hand, in § 6 we proceed and train a machine-learning model to identify networks in the global routing table exhibiting similar behavior to serial hijacker ASes. In § 7, we present a broad and detailed study of the ≈ 900 networks flagged by our classifier "in the wild". Finally, we feature three networks in case studies in § 8, and discuss implications and limitations of our work as well as avenues for future work in § 9.

3 DATASETS

In this section we first describe the datasets we leverage for identifying serial hijackers and a control group of legitimate ASes. We then introduce our longitudinal BGP dataset.

3.1 Legitimate ASes and Serial Hijackers

Legitimate ASes: We start our selection of legitimate ASes using the participants to the Mutually Agreed Norms for Routing Security (MANRS) initiative [5]. MANRS is a global initiative started by network operators and supported by the Internet Society, which proposes a set of actions, such as filtering and global validation of Internet resources, that network operators can implement to foster routing security. Since MANRS participants voluntarily agree to implement a set of proactive security measures in BGP, it is unlikely that they would repeatedly—and willingly—engage in repeated BGP misbehavior or malicious activities. 272 ASes[2] are part of the MANRS initiative. Additionally, we manually select 35 ASes that represent the full spectrum of routed ASes: major end-user ISPs, enterprise networks, content/cloud providers, and academic networks. For these ASes, we are reasonably certain that the administrators do not willingly engage in repeated hostile activity.

Serial Hijacker ASes: Finding ground truth on serial hijacker ASes is a more difficult task: we process 5 years worth of email threads on the NANOG [6] mailing list and extract 23 AS numbers for which network operators repeatedly disclosed hijacking events. We note that for each of these ASes the email threads included several address blocks that had recently been (or were being) hijacked. Furthermore, in 4 cases, hijacker ASes were mentioned in connection to hijacking events spanning multiple years.

[2]Later in § 6 we only leverage MANRS ASes that have originated at least 10 prefixes in the 5 years considered in our study.

Start date	Jan 1, 2014 00:00:00 UTC
End date	Dec 31, 2018 23:55:00 UTC
Snapshot files	525,888
Unique prefixes	6,044,333
Unique ASNs	76,769
Prefix-origin pairs	7,351,829

Table 1: Dataset properties.

	IPv4	IPv6
Snapshot files	524,556	524,290
Unique prefixes	1,907,397	196,136
Unique ASNs	75,261	22,248
Prefix-origin pairs	2,317,168	196,137

Table 2: Dataset properties after removal of incomplete snapshots and very low visibility prefix-origin pairs.

In the remainder of this paper, we use the set of *Legitimate ASes* and *Serial Hijacker ASes* to first study the dominant characteristics of serial hijackers in § 5, and to later train a classifier to identify these characteristics in the larger AS population in § 6.

3.2 Longitudinal BGP Dataset

We base our study on snapshots taken from the global routing table computed every 5 minutes over a time period of 5 years, leveraging historical BGP data from all available RIPE and RouteViews collectors. Starting on January 1st, 2014 and ending in December 31, 2018, we build an individual routing table for each peer (network that feeds into any of the collectors) of each collector every 5 minutes using RIB dumps and BGP updates received over the respective peer-collector BGP sessions. For each of these routing tables, we extract prefix and origin AS numbers to generate 5 minute snapshots listing prefix-origin AS pairs (*prefix-origins* in the following) together with the count of peers observing them. Each snapshot file contains between 560,000 and 1,240,000 prefix-origin pairs. We obtain 288 files per day, 525,888 snapshot files in total. Across the entirety of our dataset covering 5 years, we find 7,370,019 unique prefix-origins to be advertised by at least one peer. We find a total of 76,769 unique ASes and 6,044,333 unique prefixes. Table 1 summarizes the main properties of the dataset.

4 DATA PREPROCESSING

In this section, we describe the necessary steps to de-noise our dataset, and to convert individual snapshots into aggregated prefix-origin timelines for further analysis.

4.1 Dataset De-Noising

Variability of BGP peer availability: We leverage the count of peers that see and propagate an individual prefix-origin pair as a proxy for the prefix-origin visibility in the global routing table. Figure 2a shows the maximum visibility of IPv4 and IPv6 prefix-origin pairs in each snapshot file, *i.e.,* the maximum number of peers that reported the same prefix-origin pair to any of the RIPE or RouteViews collectors. Over the course of 5 years, the maximum visibility increases from the 250-300 range for IPv4 and 160-210 range for IPv6 in 2014 to 400-500 (IPv4) and 300-400 (IPv6) in 2018, mainly a result of increasing participation of networks in the BGP collection infrastructure. However, we see constant variability, *e.g.,* caused by lost BGP sessions between peers and collectors, or outages of individual collectors. Indeed, we find a number of episodes of significant reduction in the number of peers with active connections to collectors. During the 5 year period, the lowest maximum peer count is 83 for IPv4 and 102 for IPv6. In order to reduce the impact of significant peer disconnections and other BGP collector infrastructure problems, for IPv4 and IPv6, we do not consider a snapshot file if the maximum peer count drops below 20% of the median maximum peer count of the previous week for the same protocol. In total, for the 5 year period, we ignore 1332 (for IPv4) and 1598 (for IPv6) snapshot files, representing 0.25% and 0.30% of all available files respectively.

Highly localized BGP advertisements: In every snapshot file, we find prefix-origin pairs with very low visibility. These BGP advertisements can either be the result of highly localized traffic engineering efforts or related to misconfigurations and errors of the collector infrastructure itself or of a single, or a few, of their connected peers (recall that the total number of peers ranges between 300 and 500 for IPv4 during our measurement period). We remove prefix-origin pairs that were seen by 5 or less peers. While we specifically track both low-visibility and high-visibility prefix advertisements in this work, these cases of very low visibility are unlikely to represent actual routing events of interest for this study. We find that, on average, of all prefix-origin pairs of a snapshot file, less than 20% of IPv4 and 15% of IPv6 prefix-origin pairs are seen by 5 or less peers, but point out that they represent only 0.09% of IPv4 and 0.1% of IPv6 prefix-origins found in the routing tables of BGP collectors' peers at the time of the snapshot. Two thirds of the low-visibility IPv4 prefix-origins are announcements more specific than /24, and three quarters of IPv6 prefix-origins more specific than /48. Table 2 summarizes the properties of the cleaned routing dataset for IPv4 and IPv6. We note that although filtering very low visibility prefix-origins reduces the overall number of prefix-origin pairs from some 7.4M to 2.5M, it only represents ≈ 0.1% of all BGP collectors' peers routing table data during the time of the study.

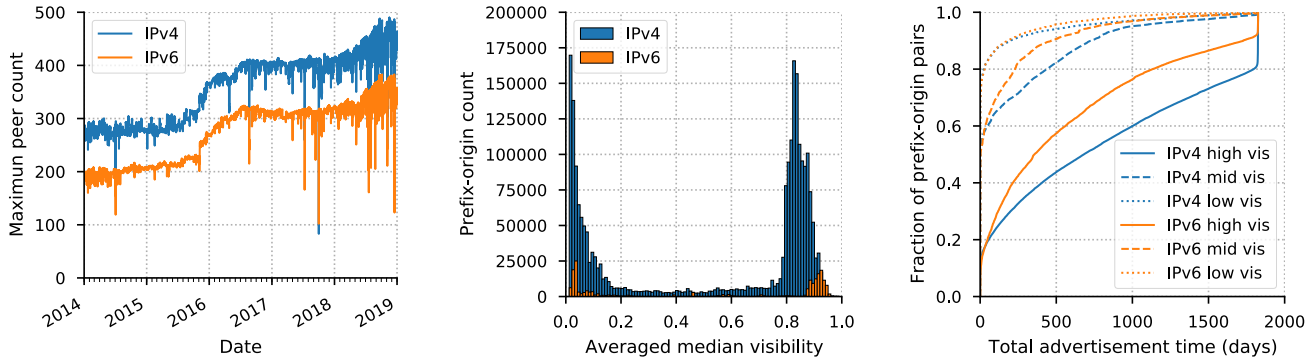
4.2 Aggregating Snapshots to Timelines

Our methodology to go from individual snapshot files to a suitable data representation for longitudinal analysis of prefix-origin characteristics consists of 3 steps:

***(i)* Normalizing visibility:** To deal with absolute changes in peer count when evaluating prefix-origin visibility, we normalize the raw prefix-origin peer count from *each snapshot* by dividing the absolute visibility of a prefix-origin pair by the maximum peer count seen in each snapshot for the respective protocol (IPv4 or IPv6). Our normalized visibility thus is in the $(0, 1]$ interval for each prefix-origin pair.

***(ii)* Building prefix-origin timelines:** We next create *timelines* for each prefix-origin aggregating the 5-minutes-apart snapshot files, requiring *(i)* constant existence of the prefix-origin pair in consecutive snapshot files,[3] and *(ii)* a steady level of visibility of the prefix-origin pair. We find that prefix-origin visibility is overall

[3]Since some snapshot files are not considered due to low BGP peer availability (see § 4.1), consecutive files can be more than 5 minutes apart.



(a) Variability and growth of the maximum visibility (max. number of peers) in RouteViews and RIPE RIS combined.

(b) Distribution of prefix-origin averaged median visibility. Most prefixes have either high (> 0.75) or low (< 0.15) visibility.

(c) Prefix-origin total advertisement time for different visibility levels for IPv4 and IPv6.

Figure 2: Visibility of prefix-origin pairs in the global routing table.

relatively stable, but we want to capture significant changes. For each prefix-origin timeline, we require that the visibility range (maximum visibility minus minimum visibility) of the prefix-origin pair in all contained snapshots does not exceed 0.1, that is 10%.[4]

***(iii)* Classifying prefix-origin pairs by visibility level:** We next tag each prefix-origin pair with its aggregated visibility, *i.e.,* the median visibility of all contained timelines, weighted by their duration. Figure 2b shows a histogram of the visibility for all prefix-origin pairs. Here, we observe a bi-modal behavior: for IPv4, 65.3% of prefix-origin pairs show visibility greater than 0.75, while 26.1% show visibility lower than 0.25 (55.9% and 32.6% for IPv6 respectively). To better understand the relationship of prefix-origin visibility and the total time they are originated by an AS, we leverage this bi-modal behavior of visibility and classify prefix-origins according to 3 levels of visibility as follows:

- Low visibility: prefix-origin pairs with an averaged median visibility of less than 15% of active peers.
- Medium visibility: prefix-origin pairs with an averaged median visibility of less than 75% but more than 15% of active peers.
- High visibility: prefix-origin pairs with an averaged median visibility of 75% of active peers.

Figure 2c shows the total time that prefix-origin pairs are visible in the global routing table for high, mid and low visibility, for IPv4 and IPv6. We note that, generally, high visibility prefix-origins are present in the global routing table for longer time periods when compared to medium visibility prefix-origins, and low visibility prefix-origins. Note that in Figure 2c, the maximum duration is naturally constrained by our measurement window of 5 years.

In the next section, we leverage our generated prefix-origin timelines from step *(ii)* and the visibility and total advertisement distribution from step *(iii)* to compute features at the prefix-origin and AS level to scrutinize the prefix origination behavior of serial hijackers in the global routing table.

[4]We note that for a single snapshot file, visibility of prefix-origins is strictly bi-modal, *i.e.,* visibility is either close to 1 or close to 0. Our threshold of 0.1 thus works well to capture significant changes.

5 DOMINANT ORIGIN AS CHARACTERISTICS

Since little is known about BGP behavior of serial hijacker ASes other than the anecdotal evidence that these networks are repeatedly involved in BGP hijacks, we start with a mental exercise of describing how origination behavior of a network dedicated to malicious activity might look like in our BGP data. We identify five main characteristics:

- **Intermittent AS presence:** BGP activity of hijackers might be intermittent. We expect some serial hijackers to have offline periods, during which they do not originate any prefix and are thus not present in the global routing table.
- **Volatile prefix origination behavior:** We expect hijackers to show higher variability in terms of the number of originated prefixes over time than legitimate ASes. Further, we expect serial hijackers to change prefixes more frequently, resulting in a higher number of *unique* prefixes originated by serial hijackers when compared to the average number of originated prefixes.
- **Short prefix origination duration:** We expect that serial hijackers originate prefixes for shorter time periods than legitimate ASes. However, we also expect to see short-term origination of prefixes from legitimate ASes due to misconfigurations (*cf.* Figure 1a). We expect that different visibility levels of such events might help to disambiguate hijacks from misconfiguration events.
- **Fragmentation of originated address space:** We expect that serial hijackers originate prefixes allocated to different RIRs (Regional Internet Registries), whereas most legitimate ASes originate prefixes allocated to a single RIR, reflecting geographic boundaries of ASes. Further, we expect that some serial hijackers originate unassigned address space.

- **Multi-Origin conflicts (MOAS) of originated prefixes:** Since hijackers originate address space routed by other ASes, we expect to see a significantly higher share of MOAS conflicts for prefixes originated by hijackers, when compared to legitimate ASes. We note, however, that there are also benign cases of MOAS conflicts that are not indicative of hijacks. We take the behavioral characteristics, *i.e.,* duration and frequency, of MOAS conflicts into account to disambiguate such cases.

In the remainder of this section, we elaborate and test each of these assumptions, introduce metrics that can capture these behavioral patterns, and contrast the behavior of our ground truth *serial hijackers* against our manually selected 35 *legitimate ASes* (*cf.* § 3.1). We test the relevance of our metrics using the broader set of ground truth ASes in § 6 using a machine-learning classification algorithm. The features used to train the algorithm are based on the properties described in this section.[5]

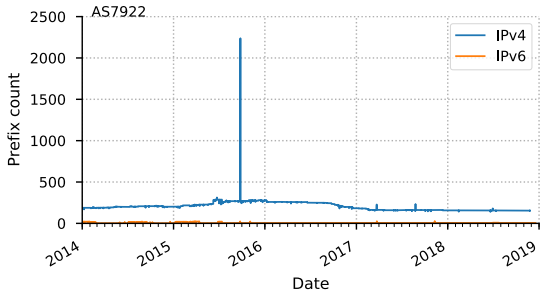
5.1 Inconsistency and Volatility of AS Activity

To exemplify differences in AS activity, Figures 3a and 3b show the number of originated IPv4 and IPv6 prefixes over time for a legitimate AS (AS7922, top), and a serial hijacker AS (AS133955, bottom). Here, we see a strong contrast: while the legitimate AS is present in the global routing table 100% of the time, we see that the serial hijacker AS showed activity in 2015, no activity in 2016, and then again higher levels of activity starting in mid-2017. Although the number of prefixes originated by both ASes varies over time, the legitimate AS shows an overall much more stable origination pattern. We note, however, that also legitimate ASes can show high levels of short-term variability, as evidenced in Figure 3a. This peak is the result of AS7922 de-aggregating large prefixes for localized traffic engineering purposes to handle an infrastructure problem in 2015.[6]

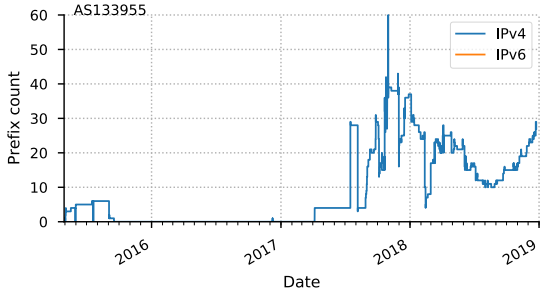
Intermittency of AS presence: To investigate the length and frequency of AS offline periods, we compute two metrics: the number of times an AS stops originating prefixes (offline drop count), and the percentage of time an AS originates prefixes during its entire lifetime (active time), where the active time is the range between the first and the last visible prefix origination of an AS. Figure 4a shows the distribution of these two metrics for legitimate and hijacker ASes. We find that all legitimate ASes cluster in the lower right corner, *i.e.,* once they start originating prefixes they are almost always seen originating prefixes, being active close to 100% of the time. In contrast, a large share of the serial hijacker ASes have lower overall activity times and we see multiple offline drops, *i.e.,* instances where an AS ceased to originate any prefix.

We also compute these metric for ASes originating IPv6 and obtain similar results (not shown). However, we find a few legitimate ASes that show a low activity-time percentage and high count of offline drops. Possible explanations include the fact that some networks may have originated IPv6 prefixes for testing purposes (recall that we cover a period of 5 years) before starting to steadily announce IPv6 prefixes and thus have experienced offline periods in IPv6.

[5]The full feature list can be found at https://github.com/ctestart/BGP-SerialHijackers.
[6]A contact in AS7922 confirmed this incident.



(a) Prefixes originated over time by a legitimate AS (AS7922).

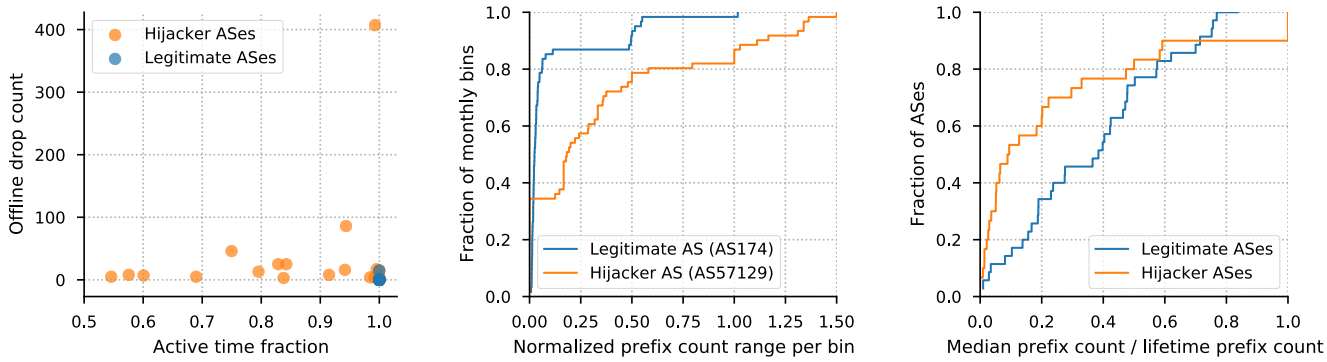


(b) Prefixes originated over time by a hijacker AS (AS133955).

Figure 3: Example of changes in prefix origination over time.

Volatility in the number of originated prefixes: To quantify volatility in the number of originated prefixes over time (*e.g.,* as shown in Figure 3b), we partition our dataset into different time bins: one day, one week and one month. Then, for each AS and bin we compute statistics over the number of originated prefixes: range, median, and the absolute number of prefix changes. We normalize both the range and the number of prefix changes by the median number of advertised prefixes. This is to allow for more variability for large ASes, as compared to small ones. Figure 4b shows the distribution of the normalized range of originated prefixes for monthly bins for a legitimate AS (AS174) and a serial hijacker (AS57129). In a legitimate AS (AS174 in Figure 4b), we see that their normalized range is small for most time bins, since the number of prefixes originated during a typical month does not vary much. AS57129, a serial hijacker, on the other hand, shows a higher number of bins with higher normalized ranges.

Volatility in the set of originated prefixes: So far, we developed metrics that can capture volatility in the number of originated prefixes over time. Next, we are interested in the stability of the *set* of originated prefixes. In particular, we want to capture if an AS typically advertises a fixed set of prefixes (the legitimate case) or if it "hops" through a large number of unique prefixes. To this end, we compute the median number of originated prefixes per AS, and we divide this median by the total number of unique prefixes this AS ever announced over the course of 5 years. The distribution of this ratio for legitimate and hijacker ASes (Figure 4c) suggests that serial hijackers tend to show a lower ratio compared to legitimate ASes, which indicates that they have a higher turnover of prefixes. Note however, that some legitimate ASes also show a low ratio,



(a) Fraction of active time and offline drop count per AS. Many hijacker ASes are only intermittently visible in the global routing table, resulting in an active time < 1 and multiple instances of offline drops.

(b) Example ASes: Monthly prefix count range normalized by median prefix count. The hijacker AS shows higher volatility in the number of advertised prefixes resulting in larger prefix count range values.

(c) Median prefix count divided by lifetime prefix count per AS. Hijacker ASes originate a smaller share of their lifetime prefixes at a given time, i.e., they have a higher turnover rate of prefixes.

Figure 4: Volatility metrics of prefix origination behavior for serial hijackers and legitimate ASes.

if, *e.g.,* a network had a route leak or misconfiguration problem that significantly increased the number of prefixes it advertised for a short period of time. Nonetheless, these types of events do not occur frequently in our set of legitimate ASes and our metric separates our two classes well.

5.2 Prefix-origin Longevity and Visibility

In this section, we study the dynamics of individual prefixes originated by ASes, in particular how hijackers' prefix total duration and visibility in the global routing table differ from prefixes originated by legitimate ASes.

Longevity of prefix announcements: Our hypothesis is that hijackers originate prefixes for a shorter period of time than legitimate ASes. While we find this clear distinction when looking at aggregate data, *i.e.,* hijackers' median prefix-origin duration is 27.25 days v.s. 264.17 days for legitimate ASes, we found it challenging to identify a threshold that separates short-term and long-term prefixes and hence separates our two categories of ASes well. To sharpen the picture, we next take the visibility of announcements into account.

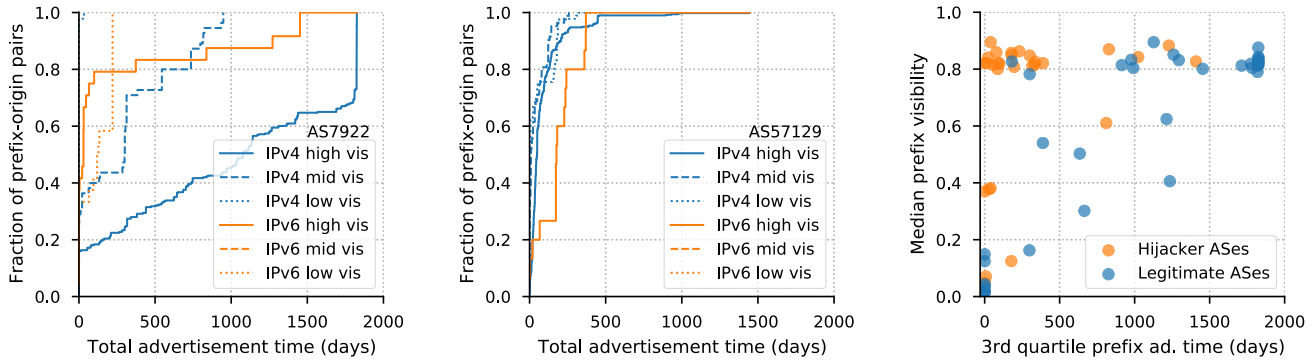
Longevity vs. visibility level: Figures 5a and 5b show the distributions of the total advertisement time of prefix-origin pairs, for different levels of visibility, for a legitimate AS and a serial hijacker AS. AS7922, a legitimate AS, has a large fraction of long-term originated prefixes, *i.e.,* more than 50% of high visibility IPv4 prefixes it originates are advertised for over 1,000 days. On the other hand, the lower the visibility the larger the share of short-term prefixes. We notice that most of the low visibility prefixes that AS7922 originates have a very short total advertisement time. Indeed, a large share of the prefixes advertised by AS7922 for only a short period of time come from highly localized traffic engineering efforts used to handle infrastructure problems and hence have very limited visibility in the global routing table (*cf.* § 5.1). AS57129, a serial hijacker, however, shows vastly different behavior: some 50% of high visibility IPv4 prefixes originated by AS57129 have less than 50 days of total advertisement time, and the share of short and long-term prefixes it originates is very similar for all levels of visibility.

When plotting ASes by median prefix visibility and total advertisement time (3^{rd} quartile shown, Figure 5c), a large portion of serial hijacker ASes cluster in the high visibility, low advertisement time corner (upper left). In contrast, legitimate ASes are spread out and high visibility is correlated with longer advertisement time for these networks. Thus, we find that the longevity of prefix origination can only be meaningfully leveraged to separate our two classes of ASes when qualified by their visibility level.

5.3 Address Space Properties

In this section, we study different properties of the IP addresses that ASes originate. We take into account the Regional Internet Registry (RIR) that assigned originated IP addresses, whether ASes originate bogon or unassigned IP space, and if originated prefixes were originated by other ASes at the same time (MOAS conflicts).

Address space fragmentation: Our hypothesis is that legitimate ASes only originate address blocks that were allocated to them by a respective Regional Internet Registry (RIR). Since most networks are limited in geographic scope, and individual RIRs cover individual geographic regions, we expect most legitimate ASes to either originate addresses from a single RIR, or, if they originate prefixes from different RIRs, they would still be concentrated in one of them. Since we do not expect serial hijackers to originate address space allocated to them, nor respect RIR boundaries, we expect them to originate prefixes from multiple RIRs, and show much less concentration on any particular RIR. To express concentration of originated address space across RIRs, we compute the Gini coefficient of ASes' RIR distribution using the percentage of prefixes ASes originate from each of the five RIRs. A Gini of 0.8 means all IP resources come from one RIR, whereas a Gini index closer to 0 means resources are uniformly distributed across the 5 RIRs. Figure 6a depicts the distribution of serial hijackers and

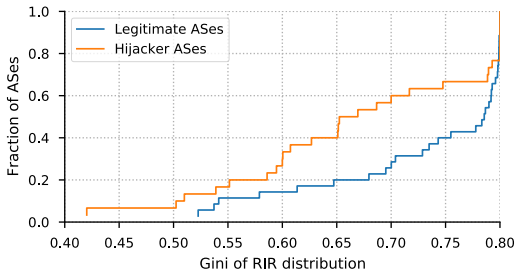


(a) Legitimate AS example: Total prefix advertisement time. Over 50% of prefixes are originated for more than 1,000 days.

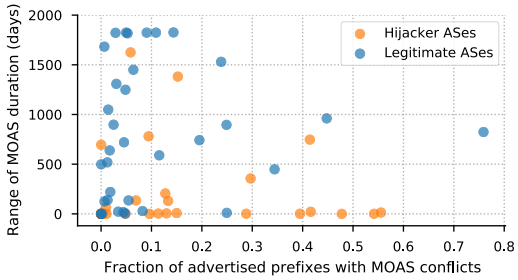
(b) Hijacker AS example: Total prefix advertisement time. Over 50% of prefixes are originated for less than 50 days total.

(c) Advertisement time and visibility per AS. Hijacker ASes show shorter, high-visibility announcements.

Figure 5: Advertisement longevity and visibility of prefixes originated by legitimate and serial hijacker ASes.



(a) Gini coefficient of originated prefix RIR concentration per AS. Serial hijackers' prefixes are more spread out over different RIRs when compared to legitimate ASes.



(b) Fraction of prefixes with MOAS conflicts and range of MOAS duration per AS. Some hijacker ASes show a higher fraction of prefixes with MOAS conflicts with a low duration range of MOAS conflicts.

Figure 6: Specific address space characteristics example for legitimate and serial hijacker ASes.

legitimate ASes with respect to the Gini coefficient over the RIR distribution. We observe that many serial hijackers show a lower Gini coefficient compared to legitimate ASes, meaning that the prefixes they originate are comparably more uniformly distributed among RIRs. This is in contrast to legitimate ASes, which typically show high RIR concentration.

Multiple Origin AS prefixes: We compute the number of prefixes and the share of address space an AS originates that is also originated by another AS at the same time, *i.e.,* the prefix has Multiple Origin ASes (MOAS) in the global routing table. Figure 6b shows per AS the fraction of advertised prefixes with MOAS conflicts (x-axis) and the range of the duration of the MOAS announcements (y-axis). We chose to show the range of the MOAS duration, since we found that serial hijackers have almost exclusively short-term MOAS announcements, resulting in a small MOAS duration range, whereas legitimate ASes show variable MOAS durations, with many short-term and long-term prefix originations with MOAS conflicts, resulting in a large MOAS duration range. Many serial hijacker ASes have a very short range of MOAS duration and a significant share of the address space they originate are MOAS prefixes, which is what we would expect for illegitimate MOAS events (*e.g.,* replaced by new ones as they are detected). We note that, as expected, some legitimate ASes show MOAS conflicts, but that these MOAS events typically last much longer than those of serial hijackers.

6 TOWARDS SCALABLE CLASSIFICATION OF BGP MISBEHAVIOR

Next, we describe how we build a classifier to identify more ASes in the global routing table that exhibit a prefix origination behavior similar to serial hijackers. We start by explaining the main challenges faced when training a model with our dataset, and elaborate on our resulting choices for our model and its main parameters. We then discuss the features we use, their importance, and present the final ensemble classifier and its accuracy metrics. We present the results of the classification based on our trained classifier in § 7.

6.1 Challenges Faced

We face three main challenges when applying machine learning algorithms to classify whether ASes show behavioral patterns of serial hijackers: *(i)* heavy-tailed and skewed data, *(ii)* limited ground truth, and *(iii)* class imbalance.

Heavy-tailed and skewed data: The routing data on which our analysis is based is extremely heterogeneous. In almost all dimensions, individual prefixes and ASes are heavily concentrated at some level but then there is a long tail of outliers, making the data difficult to normalize. In addition, some of our features range from zero to one (*e.g.,* the Gini coefficient expressing concentration of address space across RIRs described in § 5.3), while other features, such as the total advertisement duration (described in § 5.2) ranges from 5 minutes to 5 years.

Small ground truth: As discussed in § 3.1, building a ground truth dataset including serial hijackers and legitimate ASes is challenging. In total, our ground truth dataset consists of 230 labeled ASes. We only select ASes originating at least 10 prefixes in the 5-year dataset. This includes all hijackers but only 217 ASes from our legitimate AS group described in § 3.1. Therefore, we must carefully select a model to avoid *overfitting*.

Class imbalance: We do not expect that a large share of routed ASes exhibiting serial hijackers' behavior. The true share of such ASes is unknown, and if we were to make an educated guess, we would only expect to find this behavior for a small number of ASes, *i.e.,* less than 1% of routed ASes (over 75,000 ASes are routed in our dataset in the 5-year period). Class imbalance is also present in our ground truth dataset: we only have 23 serial hijacker ASes vs. 217 ASes in the legitimate group of our labeled ground truth.

6.2 Our Classifier

Choice of Classifier: We choose a tree-based classifier since decision trees do not require normalized data and work well with large dimensions and heavy-tailed data such as the features we built to capture different aspect of BGP origination behavior. More specifically, we use Extremely Randomized Trees (Extra-Trees) classifiers [17]. An Extra-Trees classifier is an ensemble (forest) of decision trees that picks feature thresholds to split nodes at random, instead of fitting the threshold to the training data like in a common random forest classifier. This added randomness greatly reduces overfitting, another of our main challenges as discussed in § 6.1.

Model accuracy for parameter selection: To properly select model parameters (sampling methods, forest size, feature selection) without reducing the training data by doing an n-fold cross-validation, we use bootstrapping samples (subset samples) in the training phase of the individual trees and compute the classifier Out-Of-Bag (OOB) error estimate. OOB error estimation is a method to measure the prediction error of random forests, where a lower OOB error indicates higher accuracy of the model. The OOB error estimate is the average error for each data point p in the training sample computed averaging the prediction of trees trained on a bootstrapping sample (bag) not including p [11]. The OOB score has been shown to converge almost identically as the n-fold cross-validation test error and is an established method to validate random forest classifiers [21].

Sampling techniques: To address class imbalance, we try different under- and over-sampling methods to create balanced training sets for our classifier, by either under-sampling the majority class (selecting only a few legitimate ASes) or over-sampling the minority class (artificially expanding the set of serial hijackers) in our original ground truth. Figure 7 shows the mean OOB scores (and error bars) of sets of 100 Extra-Trees classifiers trained using 6 different sampling technique for different forest sizes. We observe that techniques that are purely based on under-sampling perform worse than techniques that include an over-sampling step. In addition, over-sampling techniques use different rules and randomness to expand the serial hijacker set and thus no two synthetic training sets are equal. We therefore decide to use a mixture of over-sampling techniques for the training of our classifier, so that it leverages the different distributions of misclassified points to improve its generalization ability [54].

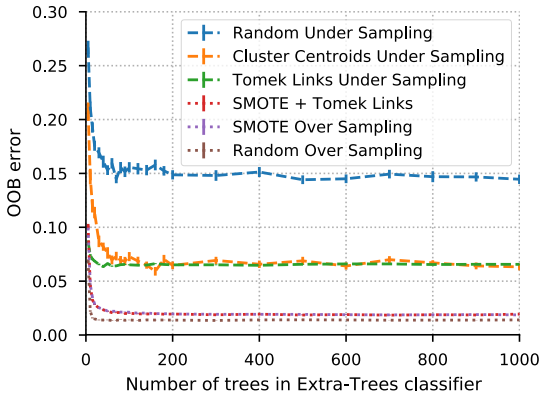


Figure 7: Mean Out-of-bag accuracy scores and error bars of sets of 100 Extra-Trees classifiers trained using different sampling techniques for increasing forest sizes.

Feature selection and importance: Based on the extensive manual analysis described in § 5, we select 52 features that capture BGP behavior according to 8 categories: ASN presence in the global routing table, prefix origination behavior, longevity of individual prefix advertisements, prefix visibility, longevity vs. visibility level, prefix set stability, address space fragmentation, and MOAS statistics. The features capture different characteristics and statistical behavior of the properties discussed in § 5, such as the median origination time of high visibility prefixes and 90^{th} percentile of the distribution of daily changes in prefix origination.

To assess feature importance, we compute the *drop column feature importance* for each feature.[7] The drop column importance captures how the classifier accuracy actually varies when a feature is not considered in the training phase [39]. We learn that all categories have positive median drop column importance, *i.e.,* they all add to the accuracy of the model. We thus proceed to feed all 52 features to train our final classifier.

The trained classifier: Our final ensemble classifier is based on the vote of 34 Extra-Trees classifiers of 500 extremely randomized trees each, and each trained on a different balanced synthetic training set computed using one of the 3 over-sampling algorithms we selected. The model OOB error estimate is 2.5%. We program our classifier using the `sklearn` and `imblearn` libraries [40] in Python,

[7] Given most of our features are computed from the same raw BGP data, selecting features by usual random forest feature importance ranking or information gain is not adequate [8, 19, 53].

which have the Extra-Trees classifiers and sampling algorithms pre-programmed.

False positives from the training set: Using the OOB predictions for the training set, the ensemble classifier precision and recall are 79.3% and 100% respectively. Although our serial hijacker set is small, the high recall rate supports our hypothesis that our small group of serial hijacker have distinctive characteristics in their BGP prefix origination behavior. We note however that the classifier precision is only about 80% — a strong reminder that the behavior of ASes selected by the classifier is not *necessarily illegitimate*. Even in our legitimate group, there are a few ASes that present similar characteristics to serial hijackers. Indeed, throughout all the different classifiers we tested, there are 6 ASes in our legitimate group that get consistently misclassified. Looking in more detail at these ASes, we find that two of them are from Verisign, an organization that offers DDoS protection, and are hence benign cases of serial hijackers, which we discuss in § 7.3. Two other ASes have only originated prefixes for a short period of time and are not currently being routed, which could have adversely affected our metrics and classification. The last two ASes are hosting organizations showing irregular BGP behavior of which the cause is unclear to us.

7 INVESTIGATING BGP MISBEHAVIOR IN THE WILD

In this section, we describe the output from our ensemble classifier. We feed the classifier with features based on IPv4 prefix-origin routing data of ASes that originate at least 10 prefixes in the 5 years of our dataset. Of the 19,103 ASes in our prediction set, our ensemble classifier finds 934 ASes having similar behavior to serial hijackers, we refer to them as*flagged ASes*. We note that the group of flagged ASes is fairly consistent across classifiers trained using different combinations of sampling methods and forest sizes. For models with an OOB error score of 4% at most, at least 95% of the ASes flagged by that classifier where also flagged by the final classifier. In the next sections, we first describe general characteristics of flagged ASes and compare them to *non-flagged* ASes. Then, we further scrutinize flagged ASes, breaking them into sub-categories.

7.1 Behaviors Captured by the Classifier

Table 3 provides summary statistics of some representative metrics for the two classes of ASes identified by the ensemble classifier: ASes flagged as having similar BGP origination behavior to serial hijackers and non-flagged ASes. For each metric, its distribution in flagged ASes is considerably different from its distribution in non-flagged ASes.

Volatile overall BGP behavior: The ASes flagged as having similar behavior to serial hijackers show more sporadic and volatile BGP activity: the 1^{st} quartile of ASN active time is 65.9%, compared to 99.9% for non-flagged ASes. Most prefixes originated by flagged ASes are shorter-lived than those of non-flagged ASes—50% of flagged ASes have a median prefix-origin duration of less than 48.2 days *vs.* only 17.9% of non-flagged ASes.

Large ASes: On average, ASes flagged by our classifier originate more prefixes than the rest—with a median prefix count of 41 compared to 23 for non-flagged ASes. Furthermore, 34 flagged ASes have originated over a thousand prefixes, representing 3.64% in the group, compared to only 1.37% of networks in the Internet announcing more than a thousand prefixes.

Diverse IP sources: ASes flagged by our classifier use IP space spread out across the RIRs—with a median RIR Gini index of 0.675 compared to 0.8 for non-flagged ASes (an RIR Gini index of 0.8 means all prefixes originated by that AS come from only one of the five RIRs). Flagged ASes also exhibit a larger share of MOAS address space than non-flagged ASes, resulting in a median MOAS prefix share of 22.9% *vs.* 6.9%, respectively.

7.2 Indications of Misconfiguration

We find that some ASes were likely flagged as a result of misconfiguration issues in BGP.

Private AS numbers: Per RFC 6996 [35], ASNs [64512, 65534] are reserved for private use. In the group of flagged ASes, we found 114 private ASNs that appear to have very volatile prefix origination behavior with relatively low visibility. A possible explanation is that due to router misconfiguration, these AS numbers appear at the origin of BGP AS-paths. As many ASes filter out prefixes originated by known reserved AS numbers, the spread and visibility of these misconfigurations is often limited. Some of the serial hijackers in our ground truth dataset exhibit lower visibility too, which is likely why these behavior got captured by the classifier.

Fat finger errors: Our classifier flagged all of the single-digit AS numbers. Indeed, the origination behavior of these ASes appears to be extremely volatile using the longitudinal routing data. We note however, that apparent origination of prefixes by theses ASes does not necessarily reflect actual routing decisions by the owner or network with given AS number. The prefix originations by these single digit ASes are likely mere results of misconfigurations, where an origin network accidentally adds an additional AS number (behind its own) to its BGP advertisements. These so-called "fat finger errors" [15] commonly occur when configuring a router to perform AS path prepending, a traffic engineering technique that artificially lengthens the AS path in order to make the advertised path less desirable in the BGP decision process [44]. A notable example of an AS flagged by our classifier is AS5, an AS whose registered company went out of business 20 years ago, periodically revived through router misconfiguration.

Removing private and single digit ASes from our group of flagged ASes, 811 remain.

7.3 Benign Serial Hijackers

In our dataset, we find prefixes originated by 29 DDoS protection networks (*e.g.,* DDoSGuard).[8] 18 of these ASes are flagged by our classifier. We find that a significant share of the address space originated by these networks has MOAS conflicts, representing over 30% of the prefixes they originate in most cases. The DDoS mitigation they perform includes originating prefixes of their customers when a DDoS attack is detected, in order to attract all the traffic destined to the network under attack, "scrub" it (to remove DDoS traffic), and tunnel it to the intended final destination [23]. Thus, DDoS protection networks present a case of "legitimate", or benign, serial hijacking behavior.

[8] Our list of AS numbers of DDoS protection services is manually compiled and hence not necessarily complete.

	Flagged ASes			Non-flagged ASes		
	1^{st} quartile	median	3^{st} quartile	1^{st} quartile	median	3^{st} quartile
Count		934			18,169	
Prefix count	18	41	101	14.0	23.0	53.0
Active time	65.9%	99.2%	100%	99.9%	100%	100%
Prefix origination median time (days)	1.8	48.2	176.9	144.6	598.0	1,217.9
Prefix-origin median visibility (%)	51.1%	80.8%	84.2%	79.7%	82.9%	85.3%
Median origination time of high visibility prefixes (days)	3.4	79.4	227.2	289.7	754.2	1,386.0
Originated/unique prefixes	0.017	0.089	0.222	0.213	0.435	0.684
RIR Gini index from address concentration	0.575	0.675	0.743	0.80	0.80	0.80
MOAS prefix share	6.7%	22.9%	52.7%	0.00%	6.9%	24.0%

Table 3: Summary statistics of selected metrics for ASes *flagged* as having similar BGP origination behavior to serial hijackers ASes and *non-flagged* ASes. Only ASes originating 10 or more prefixes in our dataset (N=19,103) are fed into our classifier. For each metric, we show the median value across ASes in each group, as well as the 1^{st} and 3^{rd} quartile.

7.4 Indications of Malicious Behavior

After removing private AS numbers, single digit ASes, and DDoS protection ones, a total of 793 publicly routable ASes flagged by our classifier remain. Next, we assess if our identified ASes show indications of malicious behavior, *e.g.,* spam or probing activity.

Flagged ASes in Spamhaus DROP list: First, we leverage snapshots of the Spamhaus *Don't Route Or Peer* (DROP) ASN list [41], a list of ASes controlled by "spammers, cyber criminals, and hijackers". We have access to 6 snapshots taken between January 1^{st} 2017 until early 2019, containing a total of 451 unique ASes, and we note that 266 of these ASes appear in all snapshots. We compared the ASes flagged by our classifier with those listed in any of the 6 snapshots of the Spamhaus DROP list we have available, finding that 84 (10.6%) of our flagged ASes are present in the Spamhaus DROP list. For comparison, we find only 206 (1.1%) ASes from the non-flagged group are present in at least one snapshot of the blacklist. Thus, flagged ASes are almost 10 times more likely to be in this list of spammers, hijackers and cyber criminals. Of the 266 ASes that are blacklisted in *all* snapshots of the Spamhaus DROP list, 133 originate more than 10 prefixes during our measurement window, and are thus in the set of ASes we classified. Our classifier flags 50 of them as exhibiting serial hijacker characteristics. In other words, based on our feature set, our classifier detects some 38% of all the ASes with enough BGP activity that repeatedly appear on this blacklist, an indicator of persistent malicious activity in this group of ASes.

Spam activity of flagged ASes: We also check for indications of spam activity in our group of flagged ASes. To this end, we leverage 2.5 years of snapshots taken 4 times a day from the UCEPROTECT [55] Level 2 spam blacklist. Attributing prefix ranges from the UCEPROTECT blacklist to ASes is challenging in our case, since our identified ASes are by definition highly volatile and might only temporarily originate prefixes that are otherwise routed by different ASes. We first load all prefixes and their origination time ranges into a prefix trie. We then process the blacklist snapshots, where we *(i)* perform a lookup in our trie to see if the particular blacklisted address block was ever originated by one (or multiple) flagged AS(es), and *(ii)* tag a given prefix-origin as blacklisted, if the prefix was originated by the respective AS at the time it appeared in the blacklist.[9]

We find indication of spam activity for more than a third of ASes flagged by our classifier. Specifically, for 38.3% of our flagged ASes, we find at least one address block originated and simultaneously blacklisted. Note that while ASes that are *victims* of hijacking for spamming purposes might also appear in spam blacklists, we do not expect them to consistently appear in multiple blacklist snapshots. Indeed, We find that when blacklisted, prefixes originated by flagged ASes tend to be blacklisted for a larger share of their advertisement time, *i.e.,* 27% are blacklisted during more than 50% of their advertisement time, compared to 12% for prefixes originated by ASes not flagged by our classifier.

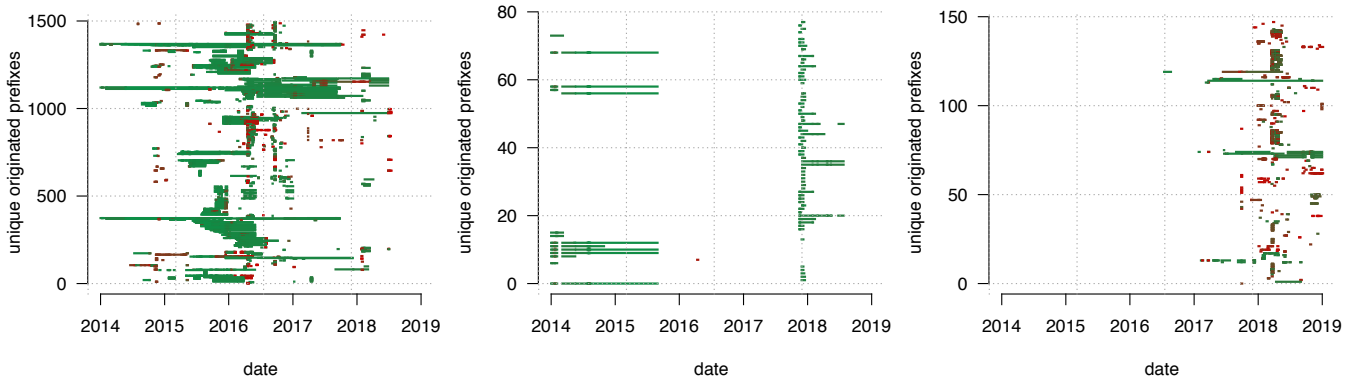
7.5 Big Players

To find possible false positives, we inspect large ASes flagged by our classifier. Using data from CAIDA AS-Rank [4, 31], we find that 4 flagged ASes are in the top 500 ASes by customer cone size, and 21 ASes are in the top 1000. Since it is unlikely that a large prominent transit provider performs serial hijacking, these are probably false positives. Nonetheless, the BGP origination behavior of these large ASes appears to be highly volatile, similar to false positives from the training sample (cf. § 6). As an example, the median of these ASes' median prefix-origin duration is only 69 days compared to 411 for large non-flagged ASes, and they show higher levels of prefix changes—the rate of normalized monthly prefix changes is 1.0 for large flagged ASes vs. only 0.35 for large non-flagged ASes.

8 CASE STUDIES

In this section, we illustrate three cases of ASes actually misbehaving, two of which are not in our ground truth dataset but are instead in the group of ASes identified by our classifier. We picked: AS197426, a serial hijacker from our ground truth dataset that was essentially "kicked off the Internet" in July 2018 because of their repeated malicious behavior [26]; AS19529, an AS flagged by our classifier for which we subsequently found hijacking complaints in a RIPE forum; AS134190, another flagged AS, which only recently started to show characteristics of a potential serial hijacker.

[9] We allow for 24 hours leeway before and after prefix origination.



(a) AS197426, a known serial hijacker, part of our ground truth dataset.

(b) AS19529, a hijacker identified by our classifier for which found corroborating evidence of hijacking activity.

(c) AS134190, the most recent detected case of a potential serial hijacker.

Figure 8: Prefix origination behavior for our selected case studies.

8.1 The Quintessential Serial Hijacker

Bitcanal, the "hijack factory", a Portuguese Web hosting firm, has been featured in several blog posts [32–34], since it represents a glaring case of serial hijacking, and one of the few cases in which prolonged coordinated action among network operators, ISPs, and IXPs, finally resulted in complete disconnection of the company's ASes. Bitcanal leveraged several ASNs: in this case study we focus on AS197426, the most active ASN used by Bitcanal.[10] While multiple incidents of hijacks carried out by Bitcanal were featured in numerous blog posts [32–34], we provide a first comprehensive data-driven assessment of their long-term behavior in the global routing table, revealing the full extent of persistent hijacking activity of this network, *i.e.,* an upwards of 1,500 originated prefixes over the course of 4 years.

Figure 8a provides a graphical representation of their prefix origination activity, each row represents a different prefix that AS197426 has originated. In the first snapshot file of our dataset in January 2014, AS197426 originates only 4 prefixes, but its origination activity soon ramps up. Already in February 2014, the same AS starts originating 15 prefixes and by October 2014 it originates almost 50 prefixes. The first post about hijacking activity by AS197426 appeared as early as September 2014 stating that it originated unrouted IP addresses that were allocated to a diverse set of organizations [32]. And yet, this was only the start of their serial hijacking spree. Starting in early 2015, we see AS197426 progressively increasing the number of prefixes it originates, and in January 2015, another blog post described recent hijacks by AS197426. Origination activity peaks at $\approx$ 300 prefixes in the second trimester of 2016, see vertical structures in late 2016 in Figure 8a. During this time, this AS makes an average of 2.5 changes per day in the set of prefixes they originate. Sometime in 2017, AS197426 was expelled from the German IXP DE-CIX because of their bad behavior. DE-CIX collected and analyzed evidence before contacting the AS and finally suspending their services [3, 24]. On June 25, 2018, a detailed email thread on the NANOG mailing list described multiple hijacks carried out by AS197426 and explicitly called out Cogent, GTT, and Level3 to act, since they provided transit to AS197426 [20]. Reportedly, GTT and Cogent quickly suspended their services to Bitcanal. Then, early in July 2018, Bitcanal appeared using other European transit providers (see sporadic activity in 2018 in Figure 8a), who terminated their relationship with Bitcanal only a few days later. Bitcanal was also present in other European IXPs, including the large LINX and AMS-IX, who terminated services with Bitcanal shortly after. The last transit provider disconnected Bitcanal on July 9, 2018. AS197426 has not been visible in the global routing table since that day.

From 2014 until its disconnection in 2018, our data shows AS197426 originating a total of 1,495 different prefixes. While hijacking activity was reported as early as September 2014, coordinated measures only showed effect and resulted in eventual disconnection in 2018.

[10] Figure 1b features another Bitcanal AS.

8.2 A Recent Hijacker

AS19529, originates about a dozen prefixes in our first snapshot in 2014. As Figure 8b shows, 7 of these prefixes were steadily originated for over a year. In April 2016, we see AS19529 withdrawing these prefixes and disappearing from (our proxy for) the global routing table (white gap in Figure 8b). Although the ARIN WHOIS record [1] for AS19529 has not been updated since 2012, our dataset shows it returns originating prefixes (31 this time) in November 2017. Then, AS19529 quickly increases the number of prefixes it originates, reaching almost 60 prefixes by the end of 2017. This spike in activity is clearly visible in Figure 8b. During these months, new RIPE RIR entries appeared, listing AS19529 as origin of IPv4 blocks owned by a different institution and registered in the ARIN region. At the same time, the legitimate owner of these prefixes raised complaints in a RIPE forum, stating that such RIPE RIR records were incorrect and that the respective address blocks were hijacked [20]. The complaints continued until April 2018 and the result, as of today, is unclear. In our data, we see AS19529 stopping to originate prefixes in July 2018: in its last 9 months of activity, it originates a total of 63 different prefixes, 20 of which are MOAS.

8.3 An Ongoing Potential Hijacker

We see AS134190, for the first time in our data on July 14, 2016, originating only a single prefix for about a month, after which it disappears from the global routing table. In early 2017, AS134190 starts repeatedly originating different prefixes for very short time periods (about a day). Starting in July 2017, AS134190 originates a few prefixes on and off—the small dots in Figure 8c—with some burst of activity reaching over 30 prefixes being simultaneously originated. In this period, AS134190 averages almost 10 changes per day in terms of originated prefixes. In November 2018, BGPmon, a widely known BGP hijack detection system [2], detected a potential hijack from AS134190 and 10 additional potential hijacks in early 2019. As of today, we have not found further evidence in the form of public complaints about potential hijacks carried out by AS134190.

9 DISCUSSION

Our study was motivated by repeated complaints in the operational community about reiterated, even persistent, prefix hijacking activities carried out by certain ASes. On the one hand, BGP's native lack of validation mechanisms exposes it not just to one-off or stealthy attacks but also to routinely executed, in-the-open, forms of abuse. On the other hand, BGP's inherent transparency, combined with the availability of pervasive and "public" BGP measurement infrastructure (*e.g.,* RouteViews, RIPE RIS) provides the opportunity to uncover systematic malicious behavior, also through the application of automated methods.

In this work, we analyzed the origination behavior of a small set of manually identified serial hijacker ASes, finding that they show distinct origination patterns, separating them from most benign ASes. We further showed that, in spite of limited ground truth and severe class imbalance, it is possible to train a machine-learning classifier that effectively narrows our focus to a set of networks exhibiting similar behavior to serial hijackers: this set accounts for 5.5% ($\approx$ 900) of the examined ASes, 1.4% of all ASes visible in IPv4 BGP. Our analysis also reveals clear potential and specific directions to further reduce this set, to the point that fully automated detection approaches and scoring systems can be envisioned in the future.

Practical relevance: To the best of our knowledge, this is the first work that examines the BGP origination behavior of serial hijackers, a category of networks that has received surprisingly little attention in terms of broad and detailed empirical assessment. We argue that serial hijacking behavior needs attention from both operators and the broader research community to allow for faster mitigation or even prevention of hijacking events.

While, as expected, not all ASes flagged by our classifier are serial hijackers, we note that all such networks do show a highly distinctive origination pattern. Scrutinizing these networks, we found widespread indications of malicious behavior, with flagged ASes being more likely to be in blacklists associated with malicious behavior, as well as different indicators of misconfiguration. Since our system is *orthogonal* to commonly deployed reputation systems (*e.g.,* event-based hijack detection), and works out-of-the-box using readily available public BGP data, we believe that, after refinement, the output of our classifier might be used to provide additional scoring data, *e.g.,* in scoring-based reputation systems.

Even after disclosure, hijack reports and discussions on mailing lists typically focus on isolated incidents (*i.e.,* usually the prefixes of the network operator issuing the complaint), and the case of Bitcanal shows that it took *years* to effectively cap hijacking activity and disconnect Bitcanal. Our metrics can compactly, and yet comprehensively, capture the dominant origination characteristics of misbehaving networks. Thus, even after initial disclosure on mailing lists, our metrics and analysis provide an instant picture of the Internet-wide "state-of-affairs" of the networks in question, which can help operators to readily assess the full extent of hijacking activity, and thus inform the process of coordinated mitigation.

Limitations: We note that our classifier is solely based on the routing activity of ASes. We focus on identifying routing characteristics of serial hijackers, which present one particular case of hijacking activity. Our detection mechanism does, naturally, not cover the space of hijacking activity exhaustively. While we find that serial hijackers do show distinct announcement patterns, our classifier does falsely tag some legitimate ASes as having BGP behavior similar to serial hijackers, as reflected in the precision of our classifier of $\approx$ 80%. We hence want to stress that our classifier, while effective in narrowing down the set of flagged ASes to $\approx$ 900 ASes, can and should *not* be deployed, as is, to generate, *e.g.,* filtering rules. Furthermore, if deployed at any point in the future, there is a potential risk that hijackers could craft their BGP announcements to not exhibit the characteristics captured by our classifier and thus evade detection. Another limitation of our work is that we focus solely on distinct features of the BGP *origination* patterns of networks and therefore on BGP origin hijacks. Hijacks which modify the AS path leaving the legitimate origin AS unaltered are therefore not captured in our data. Our work constitutes an initial view into the properties of serial hijackers with much future work to be done.

Future work: In the future, we plan to extend the features we leverage for classification. Potential additional features include more BGP-derived properties, such as AS-path characteristics of hijacked prefixes, as well as sub- and super-MOAS events. We believe that such features could not only further improve separation of ASes, but also shed light on topological properties of hijackers, *e.g.,* upstream networks and peering facilities leveraged by serial hijackers. We further plan to cross-evaluate our findings with other external datasets. In a first step, we correlated our identified ASes against blacklists, finding indications of persistent malicious behavior.

Our work is based on 5 years of historic BGP routing data, and we point out that some of the dominant characteristics of serial hijackers only become visible when studying routing data at longer timescales. We note, however, that our features to capture advertisement volatility are in fact computed over much shorter timescales *i.e.,* bins of weeks and months, and our address space features might well yield distinctive results when applied to shorter timescales. This suggests that early detection of systematic misbehavior might be indeed possible. We plan to further study the time-sensitivity of our approach to assess closer-to-real-time detection possibilities.

Acknowledgments

We thank our shepherd Olaf Maennel and the anonymous reviewers for their thoughtful feedback. We are also thankful for the feedback provided by the members of the Shadow PCs. This work was

partially supported by the MIT Internet Policy Research Initiative, William and Flora Hewlett Foundation grant 2014-1601. We acknowledge funding support from the NSF Grants CNS 1423659, OAC 1848641, CNS 1705024, and OAC 1724853. This material is based on research sponsored by Air Force Research Laboratory under agreement number FA8750-18-2-0049. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions in this paper are those of the authors and do not necessarily reflect the opinions of a sponsor, Air Force Research Laboratory or the U.S. Government.

REFERENCES

[1] American Registry for Internet Numbers ARIN WHOIS. http://whois.arin.net/ui/.
[2] BGPmon Network Solutions Inc. https://bgpmon.net/.
[3] BitCanal hijack factory, courtesy of Cogent, GTT, and Level3. https://seclists.org/nanog/2018/Jun/370.
[4] CAIDA - AS Rank. http://as-rank.caida.org.
[5] Mutually Agreed Norms for Routing Security (MANRS). https://www.manrs.org/.
[6] NANOG mailing list and archives. https://www.nanog.org/list/archives.
[7] Bahaa Al-Musawi, Philip Branch, and Grenville Armitage. BGP Anomaly Detection Techniques: A Survey. *IEEE Communications Surveys & Tutorials*, 19(1):377–396, 2017.
[8] Andre Altmann, Laura Tolosi, Oliver Sander, and Thomas Lengauer. Permutation importance: a corrected feature importance measure. *Bioinformatics*, 26(10):1340–1347, May 2010.
[9] Maria Apostolaki, Aviv Zohar, and Laurent Vanbever. Hijacking Bitcoin: Routing Attacks on Cryptocurrencies. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 375–392, San Jose, CA, USA, May 2017. IEEE.
[10] Hitesh Ballani, Paul Francis, and Xinyang Zhang. A study of prefix hijacking and interception in the Internet. In *ACM SIGCOMM Computer Communication Review*, volume 37, pages 265–276. ACM, 2007.
[11] Leo Breiman. Out-Of-Bag Estimation. Technical report, Dec 1996.
[12] Randy Bush and Rob Austein. The Resource Public Key Infrastructure (RPKI) to Router Protocol, Version 1. RFC 8210, IETF, Sep 2017.
[13] Kevin Butler, Toni R. Farley, Patrick McDaniel, and Jennifer Rexford. A Survey of BGP Security Issues and Solutions. *Proceedings of the IEEE*, 98(1):100–122, Jan 2010.
[14] RIPE Network Coordination Centre. YouTube Hijacking: A RIPE NCC RIS case study. https://www.ripe.net/publications/news/industry-developments/youtube-hijacking-a-ripe-ncc-ris-case-study, Mar 2008.
[15] Shinyoung Cho, Romain Fontugne, Kenjiro Cho, Alberto Dainotti, and Phillipa Gill. BGP hijacking classification. In *Network Traffic Measurement and Analysis Conference (TMA)*, 2019.
[16] National Institute for Stantards and Technology. RPKI Deployment Monitor. https://rpki-monitor.antd.nist.gov/.
[17] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely Randomized Trees. *Mach. Learn.*, 63(1):3–42, Apr 2006.
[18] Yossi Gilad, Tomas Hlavacek, Amir Herzberg, Michael Schapira, and Haya Shulman. Perfect is the Enemy of Good: Setting Realistic Goals for BGP Security. In *Proceedings of the 17th ACM Workshop on Hot Topics in Networks - HotNets '18*, pages 57–63, Redmond, WA, USA, 2018. ACM Press.
[19] Baptiste Gregorutti, Bertrand Michel, and Philippe Saint-Pierre. Correlation and variable importance in random forests. *Statistics and Computing*, 27(3):659–678, May 2017. arXiv: 1310.5726.
[20] Ronald Guilmette. RIPE Forum. https://www.ripe.net/participate/mail/forum/anti-abuse-wg/PDU2ODUzLjE1MzM4NTk2NzhAc2VnZmF1bHQudHJpc3RhdGVsb2dpYy5jb20+, Aug 2018.
[21] Trevor Hastie, Robert Tibshirani, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer series in statistics. Springer, New York, NY, 2nd ed edition, 2009.
[22] Xin Hu and Z. Morley Mao. Accurate Real-time Identification of IP Prefix Hijacking. In *2007 IEEE Symposium on Security and Privacy (SP '07)*, pages 3–17, May 2007.
[23] Mattijs Jonker, Anna Sperotto, Roland van Rijswijk-Deij, Ramin Sadre, and Aiko Pras. Measuring the Adoption of DDoS Protection Services. In *ACM IMC*, 2016.
[24] Thomas King. We Care About Data Quality at IXPs. https://ripe75.ripe.net/presentations/54-20171016-TKJS-RIPE-We_Care_About_Data_Quality_at_IXPs.pdf, Oct 2017.
[25] Maria Konte, Roberto Perdisci, and Nick Feamster. ASwatch: An AS Reputation System to Expose Bulletproof Hosting ASes. In *ACM SIGCOMM*, 2015.
[26] Krebs on Security. Notorious 'Hijack Factory' Shunned from Web. https://krebsonsecurity.com/tag/bitcanal/.
[27] Mohit Lad, Dan Massey, Dan Pei, Yiguo Wu, Beichuan Zhang, and Lixia Zhang. PHAS: A Prefix Hijack Alert System. In *Proceedings of the 15th Conference on USENIX Security Symposium - Volume 15*, USENIX-SS'06, Berkeley, CA, USA, 2006. USENIX Association.
[28] Mohit Lad, Ricardo Oliveira, Beichuan Zhang, and Lixia Zhang. Understanding Resiliency of Internet Topology against Prefix Hijack Attacks. In *37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'07)*, pages 368–377, Jun 2007.
[29] Matt Lepinski, Richard Barnes, and Stephen Kent. RFC 6480: An Infrastructure to Support Secure Internet Routing, Feb 2012.
[30] Matt Lepinski, Derrick Kong, and Stephen Kent. RFC 6482: A Profile for Route Origin Authorizations (ROAs), Feb 2012.
[31] Matthew Luckie, Bradley Huffaker, Amogh Dhamdhere, Vasileios Giotsas, and kc claffy. AS Relationships, Customers Cones, and Validations. In *ACM IMC*, 2013.
[32] Doug Madory. Sprint, Windstream: Latest ISPs to hijack foreign networks | Dyn Blog. https://dyn.com/blog/latest-isps-to-hijack/, Sep 2014.
[33] Doug Madory. The Vast World of Fraudulent Routing | Dyn Blog. https://dyn.com/blog/vast-world-of-fraudulent-routing/, Jan 2015.
[34] Doug Madory. Shutting down the BGP Hijack Factory | Dyn Blog. https://dyn.com/blog/shutting-down-the-bgp-hijack-factory/, Jul 2018.
[35] Jon Mitchell. RFC 6996: Autonomous System (AS) Reservation for Private Use, Jul 2013.
[36] Asya Mitseva, Andriy Panchenko, and Thomas Engel. The state of affairs in BGP security: A survey of attacks and defenses. *Computer Communications*, 124:45–60, Jun 2018.
[37] Martin O. Nicholes and Biswanath Mukherjee. A survey of security techniques for the border gateway protocol (BGP). *IEEE Communications Surveys Tutorials*, 11(1):52–65, 2009.
[38] Pierluigi Paganini. BGP hijacking - Traffic for Google, Apple, Facebook, Microsoft and other tech giants routed through Russia. https://securityaffairs.co/wordpress/66838/hacking/bgp-hijacking-russia.html, Dec 2017.
[39] Terence Parr, Kerem Turgutlu, Christopher Csiszar, and Jeremy Howard. Beware Default Random Forest Importances. http://explained.ai/decision-tree-viz/index.html, Mar 2018.
[40] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in python. *J. Mach. Learn. Res.*, 12:2825–2830, Nov 2011.
[41] The Spamhaus Project. DROP - Don't Route or Peer lists - The Spamhaus Project. https://www.spamhaus.org/drop/.
[42] Jian Qiu and Lixin Gao. Hi-BGP: A Lightweight Hijack-proof Inter-domain Routing Protocol. Technical report, 2006.
[43] Tongqing Qiu, Lusheng Ji, Dan Pei, Jia Wang, Jun Xu, and Hitesh Ballani. Locating Prefix Hijackers Using LOCK. In *Proceedings of the 18th Conference on USENIX Security Symposium*, SSYM'09, pages 135–150, Berkeley, CA, USA, 2009. USENIX Association.
[44] B. Quoitin, C. Pelsser, L. Swinnen, O. Bonaventure, and S. Uhlig. Interdomain traffic engineering with bgp. *Comm. Mag.*, 41(5):122–128, May 2003.
[45] Anirudh Ramachandran and Nick Feamster. Understanding the network-level behavior of spammers. In *ACM SIGCOMM Computer Communication Review*, volume 36, pages 291–302. ACM, 2006.
[46] Johann Schlamp, Ralph Holz, Quentin Jacquemart, Georg Carle, and Ernst W. Biersack. HEAP: Reliable Assessment of BGP Hijacking Attacks. *IEEE Journal on Selected Areas in Communications*, 34(6):1849–1861, Jun 2016.
[47] John Scudder, Randy Bush, Pradosh Mohapatra, David Ward, and Rob Austein. RFC 6811: BGP Prefix Origin Validation, Jan 2013.
[48] Pavlos Sermpezis, Vasileios Kotronis, Alberto Dainotti, and Xenofontas Dimitropoulos. A Survey among Network Operators on BGP Prefix Hijacking. *ACM SIGCOMM Computer Communication Review*, 48(1):64–69, Apr 2018.
[49] Pavlos Sermpezis, Vasileios Kotronis, Petros Gigis, Xenofontas Dimitropoulos, Danilo Cicalese, Alistair King, and Alberto Dainotti. ARTEMIS: Neutralizing BGP Hijacking within a Minute. *arXiv:1801.01085 [cs]*, Jan 2018. arXiv: 1801.01085.
[50] Xingang Shi, Yang Xiang, Zhiliang Wang, Xia Yin, and Jianping Wu. Detecting Prefix Hijackings in the Internet with Argus. In *ACM IMC*, 2012.
[51] Craig A. Shue, Andrew J. Kalafut, and Minaxi Gupta. Abnormally Malicious Autonomous Systems and Their Internet Connectivity. *IEEE/ACM Transactions on Networking*, 20(1):220–230, Feb 2012.
[52] Muhammad S. Siddiqui, Diego Montero, Rene Serral-Gracia, Xavi Masip-Bruin, and Marcelo Yannuzzi. A survey on the recent efforts of the Internet Standardization Body for securing inter-domain routing. *Computer Networks*, 80:1–26, Apr 2015.
[53] Carolin Strobl, Anne-Laure Boulesteix, Achim Zeileis, and Torsten Hothorn. Bias in random forest variable importance measures: Illustrations, sources and a

solution. *BMC Bioinformatics*, 8(1):25, Jan 2007.
[54] Yanmin Sun, Andrew K. C. Wong, and Mohamed S. Kamel. CLASSIFICATION OF IMBALANCED DATA: A REVIEW. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04):687–719, Jun 2009.
[55] UCEPROTECT. Blacklist Policy LEVEL 2. http://www.uceprotect.net/en/index.php?m=3&s=4.
[56] Pierre-Antoine Vervier, Olivier Thonnard, and Marc Dacier. Mind Your Blocks: On the Stealthiness of Malicious BGP Hijacks. In *Proceedings 2015 Network and Distributed System Security Symposium*, San Diego, CA, 2015. Internet Society.
[57] Z. Zhang, Y. Zhang, Y. C. Hu, Z. M. Mao, and R. Bush. iSPY: Detecting IP Prefix Hijacking on My Own. *IEEE/ACM Transactions on Networking*, 18(6):1815–1828, Dec 2010.