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

NANOG 80

Cecilia Testart  
MIT

Philipp Richter  
MIT

Alistair King  
CAIDA, UC San Diego

Alberto Dainotti  
CAIDA, UC San Diego

David Clark  
MIT

Originally presented at the ACM Internet Measurement Conference 2019
BGP hijacking is pervasive in the Internet

How Pakistan knocked YouTube offline (and how to make sure it never happens again)
YouTube becoming unreachable isn’t the first time that Internet addresses were hijacked. But if it spurs interest in better security, it may be the last.

How 3ve’s BGP hijackers eluded the Internet—and made $29M
3ve used addresses of unsuspecting owners—like the US Air Force.

Suspicious event hijacks Amazon traffic for 2 hours, steals cryptocurrency
Almost 1,300 addresses for Amazon Route 53 rerouted for two hours.

Criminals, Nation-States Keep Hijacking BGP and DNS
While Exploitable Protocols and Processes Persist, Adoption of Secure Fixes Lags

Why BGP Hijacking Remains a Security Scourge
Cyber criminals are stepping up their attacks against routing protocols, creating new problems for enterprise security
BGP hijacking is pervasive in the Internet

The problem of BGP hijacking is still far from solved.
Hijack disclosure in mailing lists

OmanTel hijacking of IP space

Jared Mauch
Wed Jan 11 15:50:49 UTC 2017

- Previous message (by thread): Advice re network compromise and "law enforcement" (PCI certification)
- Next message (by thread): OmanTel hijacking of IP space
- Messages sorted by: [ date ] [ thread ] [ subject ] [ author ]

There is an ongoing pattern of OmanTel hijacking IP space and advertising it to many of their peers.

IPv4 and IPv6 hijacking by AS 6

Matt Harris
Thu Apr 12 16:34:31 UTC 2018

- Previous message (by thread): 198.154.60.0/22 bogon/hijacked?
- Next message (by thread): IPv4 prefix hijack by INDOSAT AS4795 / AS4761
- Messages sorted by: [ date ] [ thread ] [ subject ] [ author ]

AS 6 is now announcing a prefix like I'm not alone. Does anyone know what's going on? The phone is non-functional. I've seen some of this before. But I don't know what's going on.

Jeremy Parsons
Mon Nov 14 00:49:29 UTC 2016

198.154.60.0/22 bogon/hijacked?

AS9498 Bharti BGP hijacks

George William Herbert
Sat Apr 1 18:19:55 UTC 2017

- Next message (by thread): AS9498 Bharti BGP hijacks
- Messages sorted by: [ date ] [ thread ] [ subject ] [ author ]

Hey, Bharti, knock that off. It's just not on.

Prefix hijack by INDOSAT AS4795 / AS4761

Randy amps at dijab.com
Thu Mar 26 14:08:20 UTC 2015

- Previous message: booster to gain distance above 60km
- Next message: Prefix hijack by INDOSAT AS4795 / AS4761
- Messages sorted by: [ date ] [ thread ] [ subject ] [ author ]

AS3266: BitCanal hijack factory, courtesy of Cogent, GTT, and Level3

Ronald F. Guilmette
Tue Jun 26 04:49:15 UTC 2018

- Previous message (by thread): Call for presentations RIPE 77
- Next message (by thread): AS3266: BitCanal hijack factory, courtesy of Cogent, GTT, and Level3
- Messages sorted by: [ date ] [ thread ] [ subject ] [ author ]

AS3266: BitCanal hijack factory, courtesy of Cogent, GTT, and Level3

Ronald F. Guilmette
tue Jun 26 04:49:15 UTC 2018

- Previous message (by thread): Call for presentations RIPE 77
- Next message (by thread): AS3266: BitCanal hijack factory, courtesy of Cogent, GTT, and Level3
- Messages sorted by: [ date ] [ thread ] [ subject ] [ author ]
Hijack disclosure in mailing lists

Serial hijackers: ASes that repeatedly hijack over long periods of time.
Bitcanal: an infamous serial hijacker

**1.** September 2014: Blog post

**2.** January 2015: Blog post

**3.** June 25, 2018: Email in NANOG

**Disconnection**

July 10, 2018

It took **4 years** to disconnect this serial hijacker.
Research goals

Find serial hijackers in the Internet

(i) Identify hijackers distinctive routing characteristics

(ii) Build a machine learning system to flag suspicious ASes

(iii) Evaluate our results

What can we learn about serial hijackers?
Ground truth: serial hijackers

23 serial hijackers:
• 10+ hijacks
• Most have been active over a year
• Up to 30,000 originated prefixes
Ground truth: legitimate ASes

230 Legitimate ASes:

• 191 MANRS ASes
• 26 ASes manually selected
BGP dataset and processing

- RIPE RIS and RouteViews collectors (~40 col., ~1400+ col. peers)

- We process all **BGP updates** to reconstruct peer routing tables

- We extract (**prefix**, **origin AS**) pairs and the number of peers with each pair in their routing table (**visibility**)

- Data from Jan. 2014 to Dec. 2018

  **(prefix, origin AS, visibility, timestamp)** every 5 min.
BGP origination behavior: legitimate vs. serial hijacker

Legitimate ASes mostly show **stable** BGP behavior. Serial hijackers BGP activity is **visually different**.
BGP origination behavior: legitimate vs. serial hijacker

We need features that capture this behavioral difference.
Variability of BGP behavior: serial hijackers
Variability of BGP behavior: legitimate ASes
Expected serial hijacker behavior

• Repeated AS absence from the global routing table.
• Short prefix origination times.
• More multi-origin conflicts (MOAS).
• Volatile count of concurrently advertised prefixes.
• Broad geographical distribution of address space originated.
Expected serial hijacker behavior

- Repeated AS absence from the global routing table.
- Short prefix origination times.
- More multi-origin conflicts (MOAS).
- Volatile count of concurrently advertised prefixes.
- Broad geographical distribution of address space originated.
Expected serial hijacker behavior

• Repeated AS absence from the global routing table.
• Short prefix origination times.
• More multi-origin conflicts (MOAS).
• Volatile count of concurrently advertised prefixes.
• Broad geographical distribution of address space originated.

We derived 52 features to capture differences.
Challenges of applying ML to find more potential serial hijackers

• Heavy-tailed and skewed data:
  Monthly prefix changes [0,2600], Gini in [0,0.8]

• Very small ground truth:
  240 AS for 19,000 ASes

• Class Imbalance:
  23 serial hijacker vs. 217 legitimate networks
Our ML approach

• Tree based classifier.

• Voting ensemble of extremely randomized forests.

• 3 over-sampling techniques.

• All 52 features with positive median drop column importance.

79% precision and 100% recall
(in ground-truth using out-of-bag score)
Putting our classifier to work

• **Goal:** Find ASes exhibiting similar BGP behavior to serial hijackers in our ground truth.
Putting our classifier to work

• **Goal:** Find ASes exhibiting similar BGP behavior to serial hijackers in our ground truth.
Putting our classifier to work

• **Goal:** Find ASes exhibiting similar BGP behavior to serial hijackers in our ground truth.
Putting our classifier to work

• **Goal:** Find ASes exhibiting similar BGP behavior to serial hijackers in our ground truth.

Flagged ASes are:
- **4.9%** of ASes originating 10+ prefixes
- **1.2%** of all ASes.
BGP behavior of flagged ASes
What are ASes flagged by our classifier?

- Indication of malicious behavior
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Block listed ASNs: 934
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Block listed ASNs: 84/290 ASes in Spamhaus ASN DROP list
  
  Flagged ASes are 10x more likely to be block listed
What are ASes flagged by our classifier?

- Indication of malicious behavior

- Block listed ASNs: **84/290** ASes in *Spamhaus ASN DROP list*

- Spammer ASNs:
What are ASes flagged by our classifier?

- Indication of malicious behavior
  - Block listed ASNs: 84/290 ASes in *Spamhaus ASN DROP list*
  - Spammer ASNs: 33% ASes have a prefix in UCE-PROTECT level 2 spam blacklist
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Indication of misconfigurations

- Block listed
  - ASNs: 84

- Spammer
  - ASNs: 304
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Indication of misconfigurations
- Private ASNs: 12%

Block listed ASNs: 84
Spammer ASNs: 304
Private ASNs: 114
What are ASes flagged by our classifier?

- Indication of malicious behavior
- **Indication of misconfigurations**
  - Private ASNs **12%**
  - Fat-finger error ASNs **1%**

![Bar chart showing the proportion of ASes flagged by various criteria.](chart.png)
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Indication of misconfigurations
- Known false positives
What are ASes flagged by our classifier?

• Indication of malicious behavior

• Indication of misconfigurations

• **Known false positives**
  
  • DDos protection ASNs: **2%**

<table>
<thead>
<tr>
<th>Category</th>
<th>ASNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block listed</td>
<td>84</td>
</tr>
<tr>
<td>Spammer</td>
<td>304</td>
</tr>
<tr>
<td>Private</td>
<td>114</td>
</tr>
<tr>
<td>Fat-finger error</td>
<td>9</td>
</tr>
<tr>
<td>DDoS protection ASNs</td>
<td>18</td>
</tr>
</tbody>
</table>
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Indication of misconfigurations
- Known false positives
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Indication of misconfigurations
- Known false positives

<table>
<thead>
<tr>
<th>Category</th>
<th>ASNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block listed</td>
<td>84</td>
</tr>
<tr>
<td>Spammer</td>
<td>304</td>
</tr>
<tr>
<td>Private</td>
<td>114</td>
</tr>
<tr>
<td>Fat-finger error</td>
<td>9</td>
</tr>
<tr>
<td>DDoS protection</td>
<td>18</td>
</tr>
</tbody>
</table>

AS 134190

- Block listed: 84
- Spammer: 304
- Private: 114
- Fat-finger error: 9
- DDoS protection: 18
What are ASes flagged by our classifier?

- Indication of malicious behavior
- Indication of misconfigurations
- Known false positives

Block listed ASNs: 84
Spammer ASNs: 304
Private ASNs: 114
Fat-finger error ASNs: 9
DDoS protection ASNs: 18

53% of flagged ASes are in known categories.
Many interesting ASes are in the other 47%.
What our classifier is not...

• A bulletproof identifier of malicious ASes.

• A system that exhaustively captures hijackers.
Key takeaways

• **First** longitudinal analysis of **serial hijacker** ASes.

• Features offer **state of affairs** of AS-wide **BGP behavior**.

• Classifier outcome provides **new data for network reputation** scoring systems.

• Effectively **narrows the focus on suspicious networks**, with much future work to be done.
Key takeaways

- **First** longitudinal analysis of serial hijacker ASes.
- Features offer **state of affairs** of AS-wide BGP behavior.
- Classifier outcome provides **new data for network reputation** scoring systems.
- Effectively **narrows the focus on suspicious networks**, with much future work to be done.