Towards Characterizing International Routing Detours

Anant Shah
Colorado State University
akshah@cs.colostate.edu

Romain Fontugne
IIJ Research Lab
romain@iij.ad.jp

Christos Papadopoulos
Colorado State University
christos@cs.colostate.edu

ABSTRACT

There are currently no requirements (technical or otherwise) that routing paths must be contained within national boundaries. Indeed, some paths experience *international detours*, i.e., originate in one country, cross international boundaries and return to the same country. In most cases these are sensible traffic engineering or peering decisions at ISPs that serve multiple countries. In some cases such detours may be suspicious. Characterizing international detours is useful to a number of players: (a) network engineers trying to diagnose persistent problems, (b) policy makers aiming at adhering to certain national communication policies, (c) entrepreneurs looking for opportunities to deploy new networks, or (d) privacy-conscious states trying to minimize the amount of internal communication traversing different jurisdictions.

In this paper we characterize international detours in the Internet during the month of January 2016. To detect detours we sample BGP RIBs every 8 hours from 461 RouteViews and RIPE RIS peers spanning 30 countries. We use geolocation of ASes which geolocates each BGP prefix announced by each AS, mapping its presence at IXPs and geolocation infrastructure IPs. Finally, we analyze each global BGP RIB entry looking for detours. Our analysis shows more than 5K unique BGP prefixes experienced a detour. 132 prefixes experienced more than 50% of the detours. We observe about 544K detours. Detours either last for a few days or persist the entire month. Out of all the detours, more than 90% were *transient* detours that lasted for 72 hours or less. We also show different countries experience different characteristics of detours.

Categories and Subject Descriptors

C.2.3 [**Computer-Communication Networks**]: Network Operations—*Network monitoring*

AINTEC '16, November 30-December 02, 2016, Bangkok, Thailand
©2016 ACM ISBN 978-1-4503-4552-1/16/11...$15.00
DOI: http://dx.doi.org/10.1145/3012695.3012698 .

General Terms

Measurement, Security

Keywords

Routing Detours, AS Geolocation, MITM

1. INTRODUCTION

We define an international detour (detour for short) as a BGP path that originates in an AS located in one country, traverses an AS located in a different country and returns to an AS in the original country. Detours have been observed in the Internet, for example, cities located in the African continent communicating via an external exchange point in Europe [7]. Many autonomous systems are also multinational, which means that routes traversing the AS may cross international boundaries. There have also been suspicious cases of detours. In November, 2013, Internet intelligence company Renesys published an online article detailing an attack they called Targeted Internet Traffic Misdirection [6]. Using `Traceroute` data they discovered three paths that suffered a man-in-the-middle (MITM) attack. One path originated from and destined to organizations in Denver, CO was passing through Iceland, prompting concern and uncomfortable discussions with ISP customers. Each of these anecdotes, while interesting in its own right, does not address the broader question about how prevalent such detours are, their dynamics and impact. Characterizing detours is important to several players: (a) for network engineers trying to diagnose problems; (b) policy makers aiming at adhering to potential national communication policies mandating that all intra-country communication be confined within national boundaries, (c) entrepreneurs looking for opportunities to deploy new infrastructure in sparsely covered geographical areas such as Africa, or (d) privacy-conscious states trying to minimize the amount of internal communication traversing different jurisdictions.

In this paper we first develop methodology to detect detours from BGP data, and validate our results with traceroute data. We also employ this methodology to characterize detours at a global scale on historical BGP data of January 2016 from RouteViews and RIPE RIS.

We provide novel insights on detour duration and characteristics. 90% of the detours are short-lived which last for less than 72 hours; some detours appear only once. We also discover that a few ASes cause most of the detours and detours affect a small fraction of prefixes. We show US, Brazil and Russia start more than 90% of the detours and stability of these varies by country. We make our tool, `Netra`, publicly available[1] so network operators can monitor the Internet routing system and obtain alerts in near real-time.

The rest of this paper is organized as follows. In Section 2 we present related work and point out key areas where our work differs from them. In Section 3 we describe our data sources and motivation to use them. Section 4 explains in detail detour detection process and corresponding terminologies used throughout the paper. In Section 5 we validate our methodology with data plane measurements. In Section 6 we characterize detours seen in January 2016. In Section 7 we discuss value additions of our work, summarize and present future work.

2. RELATED WORK

Detour detection:
In November 2013 Renesys reported a few suspicious paths [6]. One went from Guadalajara, Mexico to Washington, D.C. via Belarus; another went from Denver, CO through Reykjavik, Iceland, back to Denver. They used mostly data plane information from traceroute for their analysis. In [7] the authors focus on ISP interconnectivity in the continent of Africa. They searched for paths that leave Africa only to return back. The goal, however, was to investigate large latencies in Africa and ways to reduce it. The premise was that if a route crosses international boundaries it would exhibit high latency. The work pointed to cases where local ISPs are not present at regional IXPs and IXP participants don't peer with each other. Similar to Renesys, they also use traceroute measurements, this time from the BISmark infrastructure (a deployment of home routers with custom firmware) in South Africa. Our study extends beyond Africa and investigates transient in addition to long-lasting detours. In *Boomerang* [11], the authors again use traceroute to identify routes from Canada to Canada that detour through the US. In this work the motivation was concerns about potential surveillance by the NSA. This work differs from ours in a number of ways: we characterize detours not just for one but 30 countries using control plane information rather than data plane. We use data plane measurement only for validation purposes.

To detect detours we only use only control plane data. This has a number of advantages: 1) Collecting data plane information at an Internet scale is hard. It needs infrastructure and visibility provided by Atlas probes or Ark monitors is limited. 2) Small footprint of our methodology makes it easily reproducible. Any network operator can pull a RIB dump from his/her border router and run `Netra` to detect detours for prefixes they own. Our goal is to not only detect detours but show characteristics about them which previous work does not present.

Data plane vs Control plane Incongruities:
In [5] authors focus on routing policies and point out cases where routing decisions taken by ASes do not conform to expected behavior. There are complex AS relationships, such as, hybrid or partial transit which impact routing. Such relationships may lead to false positives in our results. However, the paper points out that most violations of expected routing behavior caused by complex AS relationships are very few and most violations were caused by major content providers. Our work identifies detours for variety of ASes, including both large content providers and small institutions. Moreover, in [13] authors argue that such incongruities are caused due to incorrect IP to AS mappings. About 60% of mismatches occur due to IP sharing between adjacent ASes. Authors here show that 63% to 88% of paths observed in control plane are valid in data plane as well. The work in [8] also analyzes the control plane (RIBs and AS paths) to construct a network topology and then uses traceroute to construct country-level paths. The goal of this work was to understand the role of different countries that act as hubs in cross-country Internet paths. Their results show that western countries are important players in country level internet connectivity.

3. DATA SOURCES

The methodology and evaluation presented in this paper rely on the five data sources listed in Table 1. The detection methodology is based on three data sources: BGP, ASMap, and CAIDA AS Relationship.
BGP: Traffic routes are obtained from BGP RIBs. The sampling rate is 3 RIBs per day (one every eight hours, as provided by RIPE RIS) for a total of 38,688 RIBs from 416 peers. This spans 30 countries, which amounts to about 55GB of compressed MRT data.
ASMap: To geolocate ASes observed in the AS path we use the AS to country mapping provided by ASMap BGPmon GeoInfo API [1]. ASMap provides the most comprehensive information of the geographic presence of ASes at the country level. For example: {`"ASN": 12145,"ASNLocation":"`{`US`}`"`}.
ASMap detects presence of an AS in a country by detecting a) prefixes announced from that country, b) infrastructure geolocation and 3) IXP participation from peering DB [4], Packet Clearing House (PCH) [3] and

[1] `https://github.com/akshah/netra`

Table 1: Dataset Description

Name	Usage	Date	Sources	Info
BGP	Detour Detection	2016-01	RouteViews, RIPE RIS	38,688 RIBS, 416 peers, 30 countries, 55GB
ASMap	AS Geolocation	2016-01	ASMap: BGPmon GeoInfo API	52K AS to country set mappings
AS Relationship	Filtering peered paths from detection	2016-01	CAIDA AS Relationship	482,657 distinct relationships
Traceroute	Detour Validation	2016-05-01	RIPE Atlas	Used by Netra, 163 traceroutes
MaxMind	Detour Validation	2016-01, 2016-03	MaxMind GeoLite City (paid)	Detour validation

358 IXP websites. 11.6% ASes out of a total of 52,984 geolocate to multiple countries. The correlation between CAIDA's AS Rank and number of countries in AS geolocation is high. ASes with higher rank have larger customer cones, hence many countries in their geolocation set.

AS Relationship: We use CAIDA's AS Relationship dataset to detect paths that may show peering relationships and affect the detours detected. Such paths are ignored.

Traceroute and MaxMind: We validate detected detours using traceroutes from RIPE Atlas probes to detoured prefixes. The IP addresses found in traceroutes are geolocated with the Maxmind GeoLite City DB paid version.

4. NETRA: DETOUR DETECTION

We define a path as having a detour if the origin and destination is country *'A'* but the path unambiguously includes some other country *'B'*. Note that this approach examines paths where the prefix origin AS and the AS where the BGP peer is located are in the same country. We geolocate all the ASes along the AS path using ASMap [1]. To analyze the AS path, we provide the following definitions:

- **Prefix Origin**: The AS that announces the BGP prefix.
- **Detour Origin AS**: The AS that starts a detour in country *'A'* that diverts the path to foreign country *'B'*.
- **Detour Origin Country**: The country where we approximate location of Detour Origin AS, country *'A'*.
- **Detour Destination AS**: The AS in foreign country *'B'*.
- **Detour Return AS**: The AS where detour returns back in country *'A'*.

Figure 1 illustrates detours. *AS0* announces prefix *a.b.c.0/24* to AS1, AS2 and AS3. AS1 geolocates to JP whereas AS0, AS2 and AS3 are in the US. In this case, data traversing from AS3 to AS0 will contain a detour from AS2 (Detour Origin) to AS1 (Detour Destination).

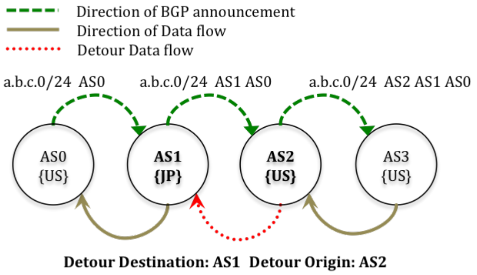

Figure 1: Example showing direction of BGP announcement and direction of observed detour

We do not include sub-paths in our analysis; other portions of the path that may experience a detour. For example, in path AS1{US}-AS2{IN}-AS3{CN}-AS4{IN}-AS5{US}, we only count the detour US-IN-US. We do not count the detour IN-CN-IN.

There are some cases where we need to approximate detour origin and country. In a path such as AS1{US}-AS2{US, BR}-AS3{CN}-AS4{US}. We resolve the uncertainty of the detour origin by assuming that it starts in AS2, since there is a likely path to AS2 from AS1 through the US and AS2 starts the detour from US, not BR. We do not characterize ***possible*** detours. For example, a path that geolocates to {US}-{US,IN}-{US} may in fact stay within the US and never visit India. In this work we only focus on paths that contain ***definite*** detours, such as {US}-{IN}-{US} or {US}-{IN,CN}-{US}. Again, we emphasize that in this work we only look at paths that confidently start and end in the same country; paths like {US,BR}-{IN}-{US} or {US}-{IN}-{BR} are not considered. We discard paths where we see an AS whose geolocation is unknown and a detour is not certain. For example, paths like AS1{US}-AS2{}-AS3{US} are discarded. However, if we see the detour occurring before the AS that could not be geolocated we do count it as a valid detour i.e., in AS1{US}-AS2{BR}-AS3{US}-AS4{}-AS5{US}, AS4 does not have geolocation information but the US-BR-US detour occurred earlier. We treat this path as definite detour.

Filtering peered AS paths:
It is possible that the detour origin and the detour return ASes have a peering relationship and in reality traffic was not detoured at all. This, however, is hard

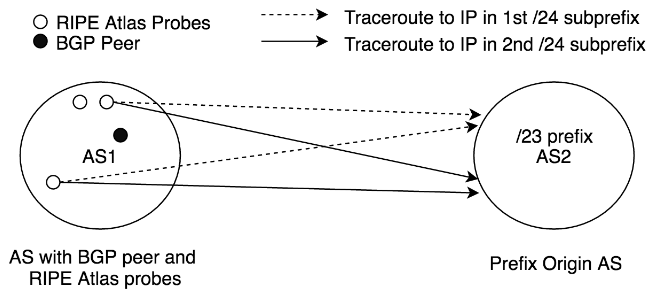

Figure 2: Data plane measurements: Example showing selection of RIPE Atlas probes and target IPs

to determine with certainty since peering relations and policies are not public. What we can do is provide an upper bound on how many detours may be eliminated due to peering. To detect such cases we use CAIDA's AS relationship dataset [2]. This dataset provides information of provider to provider (p2p) and provider to customer (p2c) relationship between ASes. We count cases where p2p link might be used, i.e., data originates from the peer itself or from a downstream customer. In case of p2c link we assume this link is always chosen. We eliminate such paths from our analysis and revisit this issue in Section 6.1.

5. DETOUR VALIDATION

In this section we validate detours in near real time using *traceroutes* from RIPE Atlas probes. Our validation comprises of four steps:

1. Run `Netra` with live BGP feeds from 416 peers to detect detours.
2. When a detour is detected, run corresponding traceroutes (from same country and same AS) using RIPE Atlas.
3. Check if the traceroute and detour see similar AS path.
4. Validate using traceroute IP hops and RTT.

5.1 Data-plane Measurements

We ran `Netra` from May 2^{nd} 2016 noon to midnight (using BGP feeds from 416 peers). When a detour was detected in control plane we selected RIPE Atlas probes in the same country and same AS which we detected detour from and ran traceroute (ICMP Paris-traceroute [10]) to IP addresses in the detoured prefix. The methodology to run data plane measurements is shown in Figure 2. There are a few cases where more than two Atlas probes are present in selected AS; in this case we selected 2 probes that are geographically farthest from each other. By doing this we aimed to account for cases where routes seen from geographically distant vantage points within the same AS are different. To select target IPs from detoured prefix, we break the prefix into its constituent /24s and randomly select an IP from each /24. For example, in a /23 prefix we select 2 IPs belonging to different /24s. By doing this we account for cases where a large prefix, even though in the same country, has different connectivity via different upstream provider. During this live run detours were seen from only 63 ASes. We then select ASes that also have active RIPE Atlas probes; there were only 10 ASes that both saw a detour and host a RIPE Atlas probe. 169 detours were seen from these 10 ASes corresponding to 6 countries: {Brazil, Italy, Norway, Russia, United States, South Africa}.

5.2 Selecting Congruent Paths

For the detours detected we initiated 169 traceroutes to detoured prefixes, we discard 6 traceroutes where less than 3 hops responded since drawing detour conclusion from these is not possible. Finally, we are left with 163 traceroutes that can be used for validation. Running `Netra` for more hours does not necessarily increase the number of usable traceroutes for validation by a lot, we are limited by the number of ASes that have RIPE Atlas probes which also see a detour and detour-origin and detour-destination have no peering.

In total we detected 85 prefixes (corresponding to the 163 traceroutes) that suffered a detour that was visible from an AS which has RIPE Atlas probes. Note that some detoured prefixes were larger than /24, so we traceroute multiple IPs within it as explained in Section 5.1. As previous work [12] has pointed out, we found many cases where AS path seen in control plane and AS path seen in data plane do not match. However, these paths can still show detour if the detour origin AS and the detour destination AS are still present in the traceroute observed AS path. We call such AS paths *congruent*. More specifically, we consider the detoured AS path congruent only if detour origin AS and detour return AS both are present in the traceroute-observed AS path in the same order (detour origin first). For example, if an AS path '`A B C D E`' in control plane changed to '`A X B C E`' in data plane where '`B`' was detour origin and '`C`' was detour destination, we consider it as a congruent path.

To resolve traceroute path to AS path we used CAIDA ITDK and iPlane IP to AS mappings and in cases where no match was found we use longest prefix match on the global routing table for the hop IP. Then we map the longest prefix match to the AS that originated it. Out of all the IPs we saw in 163 traceroutes, only 44 could be mapped to an AS using the IP to AS datasets. All other IPs were mapped using longest prefix match.

We observed 113 congruent AS paths. This includes 3 cases, insertions, deletions and mix of both. We borrow nomenclature of these paths from [12]. We saw 73 deletions, 29 insertions, 4 mix of insertion and deletions. The remaining 7 AS paths were exact matches. Note that these insertions and deletions occurred only

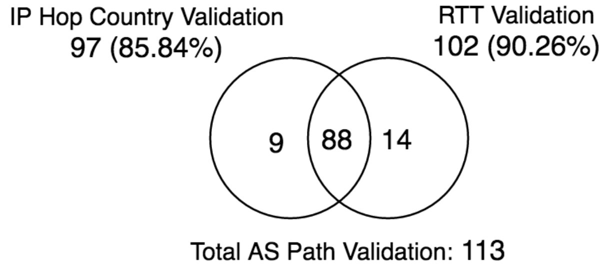

Figure 3: Validation Results: Live traceroutes using RIPE Atlas

for ASes that were not involved in the detour.

5.3 Validation

Now we validate detours detected by our methodology by comparing it with detours seen in data plane. For the 113 congruent AS paths, we evaluate if a data plane detour was seen. We chose to perform two tests. First, we resolve IPs observed in the hops of traceroute to country level geolocation using Maxmind (paid version). We detect data plane detour if a path traversed foreign country and returned. We make sure that country visited (detour destination country) in data plane is present in the set of destination countries expected for this particular detour by `Netra`. We do this filtering to avoid false positives like: `Netra` detected detour {US}–{GB,DE}–{US} and traceroute detected detour {US}–{IT}–{US}. Although still a detour, since it was not accurately captured we count it as a miss. However, no such case was found. Second, we validate using RTT measurements. We detect RTT based detour if a hop in the traceroute showed increase in RTT by an order of magnitude (at least 10 times increase). The results of this analysis are shown in Figure 3. We observed accuracy of about 85% (97 out of 113) in country-wise method and 90% (102 out of 113) by RTT measurements. The overlap between these two different tests was also large. 88 detours were detected in both (77.8%).

We investigate further the 9 detours that were seen in country-wise method but not in RTT. These detours covered small geographic area; 4 from Italy to France, 2 Norway to Sweden, 2 from Brazil to US and 1 from Russia to Sweden. RTTs between these countries have been previously reported to be low. Next we investigate 14 cases which were captured in RTT measurements but not in country-wise method. All of these do cross international boundaries. For 12 of these cases, due to large number of traceroute hops (especially towards the end of the traceroute) not responding we do not see the route returning to the origin country, hence not detected by country-wise method. We attribute remaining 2 cases as false positives due to inaccurate AS geolocation.

6. RESULTS

In this section we quantify detours detected in January 2016. First, Section 6.1 presents an overview of all the detours detected in our dataset. Section 6.2 defines metrics and classifies detours based on their stability and availability. Section 6.3 focuses on transient detours.

6.1 Aggregate Results

We begin by characterizing aggregate results, namely all detours seen by all peers. We count an incident every time an AS path in a RIB contains a detour. As expected, we observe that detours are not generally common. Only 79 peers, out of 416, saw one or more detours. Table 2 details the number of detours seen. We analyzed about 14 billion RIB entries and about 659k entries showed a detour.

Impact of peering:
We now estimate the effect of peering links on detours. Specifically, we are interested in cases where a peering relationship exists between the *Detour Origin AS* and the *Detour Return AS* as described in Section 4 using CAIDA AS relationship dataset. If such a link exists, it is possible that traffic traverses that link instead of the detour. We have found 115,085 detours between ASes that also have peering relations compared to total number of detours without filtering peered paths. Thereby, we discard 17.4% of detours to avoid ambiguities due to peering relations. We do not count these as detours in the rest of our analysis.

Aggregate characteristics:
We are left with 544K detour entries. On an average we find about 17.5K detoured entries per day and they are evenly spread out throughout the month. But many of these incidents are duplicates, thus, we also compute the number of unique detours ({peer,prefix,aspath} tuple), and obtain only 18.9K unique detours (most detours re-appear during the month).

Next we examine the visibility of detours, where we observe an uneven distribution among ASes. Just 9 ASes originate more than 50% of the detours. Similarly, some prefixes experience detours more than others. 132 prefixes experienced more than 50% of the total detours. Looking at the average length of a detour, we see that a detour visits 1 to 2 foreign ASes before returning to its origin country.

Country-wise analysis:
To provide an understanding on number of detours per peer in each country we normalize the data by dividing the number of detours by number of peers in the country. The reason to normalize data is simple, RouteViews and RIPE RIS peers are not evenly distributed

Table 2: Aggregate number of detours detected

#Total RIB entries	#Total Detours without filtering peered paths	#Detours with peered paths	#Analyzed detours	#Unique detours
14,366,653,046	659,569	115,085	544,484	18,995

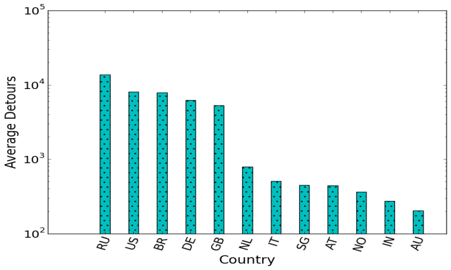

Figure 4: Average number of detours per country

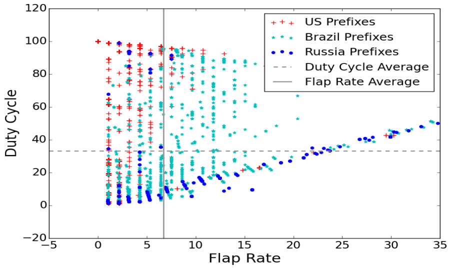

Figure 5: Flap Rate vs DC for US, RU and BR prefixes

among different countries. Therefore it is possible that more detours are seen in countries that have more peers due to more visibility. An average number of detours per peer per country provides better insight. Out of 30 countries, only 12 countries observed a detour. Figure 4 shows average number of detours per country. Russia showed most number of average detours. Understanding the total number of detours in different countries is important but it does not reflect if detours seen in different countries have different characteristics. In the next section we focus on characterizing these detours.

6.2 Characterizing Detours

To characterize detours we define two metrics:

1. **Detour Dynamics**
   (a) **Flap Rate**: Measure of *stability* of a detour; how many times a detour disappeared and reappeared.
   (b) **Duty Cycle**: Measure of *uptime* of a detour throughout the month measurement period.
2. **Persistence**: Total number of continuous hours a prefix was seen detoured.

Before using the above metrics to characterize the detours, we perform data pruning to avoid skewing of data towards ASes that have more peers that provide BGP feeds to RouteViews and RIPE RIS. Also, ASes with multiple peers and similar views can contribute duplicate detours to our dataset. We follow a simple approach to deal with this problem: if an AS contains more than one peer we select the peer that saw the most detours as the representative of that AS. This may potentially undercount detours since some peers in same AS may see different detours. After selecting a representative we are left with 36 (out of 79) peers. We now continue our characterization of detours by looking at **detour dynamics**. Specifically we focus on flap rate and duty cycle, defined as follows:

$$FlapRate = \frac{TotalTransitions}{TotalTime} \times 100$$

$$DutyCycle = \frac{TotalUptime}{TotalTime} \times 100$$

To understand if country where detours occur plays a role in detour dynamics, next we drill into country specific detours. Figure 5 shows a scatter plot of flap rate vs. duty cycle for various detours in US, Brazil and Russia. We selected these three countries because they show the most detours in our dataset; they account for 93% of detours. We see a triangular pattern with some outliers. Large number of detours show high duty cycle and low flap rate. We divide Figure 5 into 4 quadrants based on average flap rate and average duty cycle of all detours. We name quadrants anti-clockwise starting from top right. US detoured paths appear more stable (lower flap rate and higher duty cycle) in II^{nd} quadrant. On the other hand, Russian and Brazilian detoured paths fall mostly in the I^{st}, III^{rd} and IV^{th} quadrant. Russian detours in general showed lower duty cycle than US and Brazil. We also studied non US, BR and RU detours separately, in this case we observed detours mostly in extreme ends on II^{nd} and III^{rd} quadrant indicating two categories of detours, either long lasting or very rare events.

A network operator can use information like this and decide which quadrant detours are more interesting to focus on. While all of detours may need attention, we believe detours with low duty cycle and low flap rate may need immediate attention. We talk more about this in Section 6.3.

Next, we examine the **persistence** of detours. Figure 6 shows the number of consecutive days a detour was visible by any peer. Note that persistence is measured in number of consecutive hours hence captures different characteristics than duty cycle which measures uptime throughout the dataset. We see a U-shaped pattern in

Table 3: Top Transient Detour Origin ASNs

Transient Detour Origin AS	Total %	Frequent Detour Destination AS	% to frequent destination
9002 (RETN-AS RETN Limited,RU)	22.64%	2914 (NTT America)	99.07%
6939 (Hurricane Electric,IT)	10.94%	8551 (Bezeq International)	100%
1299 (TELIANET,IT)	10.87%	8708 (RCS-RDS)	100%

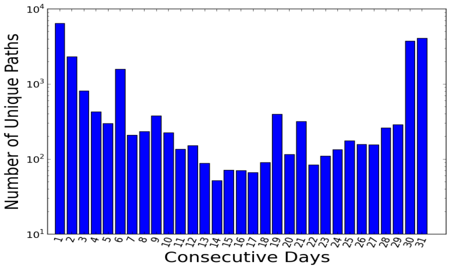

Figure 6: Persistence of definite detoured paths as seen by all peers

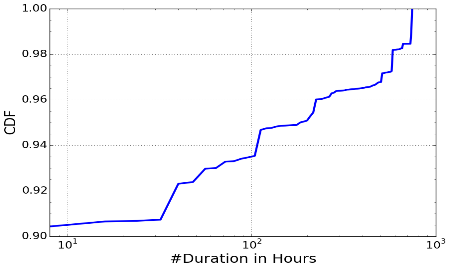

Figure 7: Distribution of detour duration

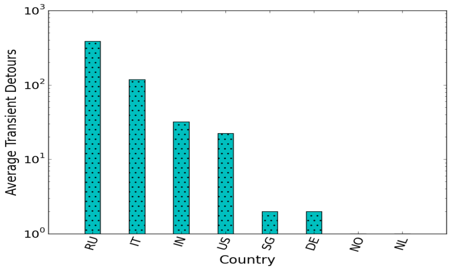

Figure 8: Average number of transient detours per country

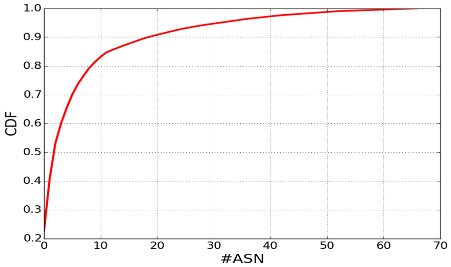

Figure 9: Distribution of ASes that originated a transient detour. The top 4 Detour Origin ASes account for 50% of all transient detours

Figure 6, meaning that many detours are either short lived (one day) or they persist for entire month. We take a different view at persistence in Figure 7 by plotting CDF of duration in hours. We see that most detours are short-lived, with about 92% lasting less than 72 hours, defined as *transient* detours. Finally, we examine a specific case of a transient detour, namely *flash detours* which appeared only once and never appeared again during the month.

6.3 Transient and Flash Detours

We first present an understanding of the transient detours on per-country basis. Since there are more than one peers in some countries and different peers see varying number of transient detours, we calculate an average number of transient detours per country by dividing total number of transient detours in a country by number of peers in the given country. This average value per country is presented in Figure 8. We detected transient detours in only 8 countries where Russia topped the list. In comparison to Figure 4 Italy and India showed more average number of transient detours than US. Figure 9 shows a distribution of ASes that initiate detours. We observe that 4 ASes originate 50% of the transient detours. We also studied the distribution of prefixes that suffer a detours, here only 30 prefixes account for 50% of the transient detours. Table 3 shows the most common transient detour origins and country where the AS was approximated to origin the detour from. Next is the percentage of detours out of the total that started from given origin. Following the percentage, is the most frequent destination that was visited from the origin, and lastly is the percentage of detours that went to most common destination from the said origin. AS9002, RETN-AS, started the most number of transient detours in our dataset. We note that in *ASWatch* [9] authors gathered ground-truth data from security blogs which enlisted AS9002 as a malicious AS. Another previously know malicious AS that appeared in our findings was AS49934 as a detour destination for 7 Russian prefixes. AS49934 is currently unassigned. It was assigned in Ukraine between 2009-10-14 and 2016-01-03 and was known to announce bogus prefixes and host bots.

Finally, we look at *flash* detours. These are detours that appeared only once and were observed in only one RIB of a peer. Flash detours account for 26% of the transient detours, 328 prefixes (6% of all prefixes that suffered detour) experienced at least one flash detour.

Owners of the prefix which suffered flash detours might be interested to know such findings. While 328 prefixes suffered flash detours in our dataset, due to space limitation we point out a few interesting ones in Table 4.

The list in Table 4 raises serious concerns. Data from

Table 4: Some prefixes affected by flash detours

Prefix Affected	Owner	Detour Destination
170.61.199.0/24	Mellon Bank, US	28513 (Uninet, MX)
192.230.0.0/20	Washington State Department of Information Services, US	7660(Asia Pacific Advanced Network, JP)
212.11.152.0/21	Moscow Mayor Office, RU	2603(NORDUnet, NO)
208.79.7.0/24	Security Equipment Inc, US	53185(William Roberto Zago, BR)
161.151.72.0/21	The Prudential Insurance Company of America, US	2510(Infoweb Fujitsu, JP)

government agencies, banks, insurance companies can easily be subject to wiretapping once it leaves national boundaries. Based on our control-plane only data, it is not possible to verify if these institutions were attacked or not. Nevertheless, we believe our findings will motivate network operators to look more closely into why their prefix detoured and if they intended it to happen.

7. CONCLUSION AND FUTURE WORK

In this paper we present a first attempt to characterize detours in the Internet. We sampled BGP routing tables from 416 peers around the world over the entire month of January 2016. We found about 18.9K distinct entries in RIBs that show a detour. More than 90% of the detours last less than 72 hours. We also discover that a few ASes cause most of the detours and detours affect a small fraction of prefixes. Some detours appear only once. Our work is the first to present different types of detours, namely, persistent and transient. We also present novel insights on their characteristics such as detour dynamics in different countries, top impacted prefixes and detour origins. Our work raises interesting questions that span multiple research directions. Detected detours can be studied more to understand internet routing better. False positives (about 10%) of our work could be a result of publicly unknown peering and IXP relationships on ASes. Mining these relationship will be useful to our work as well as other internet measurement researchers. We plan to enhance our methodology to learn common cases of modifications to AS path in data plane and appropriately detect detours. With this paper we aim to fetch for participation from service providers to deploy our tool `Netra`, validate AS geolocation and detours to improve its detection capabilities.

In the future we plan to continue to build a system that detects international detours in real time. It is very apparent that we need to include both control and data plane measurements and study algorithms that take input from both. Our first goal is to provide ISPs with a tool to alert when a detour has taken place, followed by information about it (origin and destination AS, duration, source and amount of data in the ISP that followed the detour). We also plan to study emerging regulatory requirements and provide feedback about the challenges they pose.

Acknowledgements

We would like to thank Randy Bush (IIJ) for helpful discussion on AS geolocation; Emile Aben (RIPE NCC) for helpful insights on using OpenIPMap; Roya Ensafi (Princeton) for help in geolocating IP addresses with Maxmind; Kensuke Fukuda (NII/Sokendai) and Johan Mazel (NII/JFLI) for their valuable comments.

The work in this paper is partially sponsored by the Department of Homeland Security (DHS) Science and Technology Directorate, HSARPA, Cyber Security Division, via SPAWAR Systems Center Pacific under Contract No. N66001-13- C-3001 (Shah, Papadopoulos), and via BAA 11-01-RIKA and Air Force Research Laboratory, Information Directorate under agreement numbers FA8750-12-2-0344 and FA8750-15-2-0224 (Papadopoulos). The U.S. Government is authorized to make reprints for Governmental purposes notwithstanding any copyright. The views contained herein are those of the authors and do not necessarily represent those of DHS or the U.S. Government.

8. REFERENCES

[1] Bgpmon geoinfo api. `http://geoinfo.bgpmon.io`.

[2] Caida as relationships. `http://www.caida.org/data/as-relationships/`.

[3] Packet clearing house ixp datasets. `https://prefix.pch.net/applications/ixpdir/menu_download.php`.

[4] Peering db 2.0 api. `https://prefix.pch.net/applications/ixpdir/menu_download.php`.

[5] Ruwaifa Anwar, Haseeb Niaz, David Choffnes, Ítalo Cunha, Phillipa Gill, and Ethan Katz-Bassett. Investigating interdomain routing policies in the wild. In *Proceedings of the 2015 ACM Internet Measurement Conference*, IMC '15, pages 71–77, New York, NY, USA, 2015. ACM.

[6] Jim Cowie. The new threat: Targeted internet traffic misdirection, Nov 2013. `http://www.renesys.com/2013/11/mitm-internet-hijacking/`.

[7] Arpit Gupta, Matt Calder, Nick Feamster, Marshini Chetty, Enrico Calandro, and Ethan Katz-Bassett. Peering at the internets frontier: A first look at isp interconnectivity in africa. In Michalis Faloutsos and Aleksandar Kuzmanovic, editors, *Passive and Active Measurement*, volume 8362 of *Lecture Notes in Computer Science*, pages 204–213. Springer International Publishing, 2014.

[8] Josh Karlin, Stephanie Forrest, and Jennifer Rexford. Nation-state routing: Censorship, wiretapping, and BGP. *CoRR*, abs/0903.3218, 2009.

[9] Maria Konte, Roberto Perdisci, and Nick Feamster. Aswatch: An as reputation system to expose bulletproof hosting ases. *SIGCOMM Comput. Commun. Rev.*, 45(5):625–638, August 2015.

[10] Matthew Luckie, Young Hyun, and Bradley Huffaker. Traceroute probe method and forward ip path inference. In *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement*, IMC '08, pages 311–324, New York, NY, USA, 2008. ACM.

[11] Jonathan A. Obar and Andrew Clement. Internet Surveillance and Boomerang Routing: A Call for Canadian Network Sovereignty. *SSRN Electronic Journal*, 2013.

[12] kc claffy Young Hyun, Andre Broido. Traceroute and bgp as path incongruities.

[13] Yu Zhang, Ricardo Oliveira, Hongli Zhang, and Lixia Zhang. Quantifying the pitfalls of traceroute in as connectivity inference. In *Proceedings of the 11th Passive and Active Measurement Conference*, PAM'10, pages 91–100, Berlin, Heidelberg, 2010. Springer-Verlag.