Investigating Characteristics of Internet Paths

KHALID BAKHSHALIYEV*, University of Nevada, Reno
MUHAMMED ABDULLAH CANBAZ*, Indiana University Kokomo
MEHMET HADI GUNES, Stevens Institute of Technology

Interactive and multimedia applications depend on the stability of end-to-end paths for predictable performance and good quality of service. On the other hand, network providers depend on multiple paths to ensure fault tolerance and use load balancing between these paths to enhance the overall network throughput. In this study, we analyze path dynamics for both end-to-end paths and path segments within service providers' networks using 2 months of measurement data from the RIPE Atlas platform, which collects path traces between a fixed set of source and destination pairs every 15 minutes. We observe that 78% of the end-to-end routes have at least two alternative Autonomous System (AS) paths with some end-to-end routes going through hundreds of different AS paths during the 2 months of analysis. While AS level paths are often prevalent for a day, there are considerable changes in the routing of the trace packets over the ASes for a longer duration of a month or longer. Analyzing end-to-end paths for routing anomalies, we observe that 4.4% of the path traces (involving 18% of the ASes) contain routing loops indicating misconfiguration of routers. Some of the ASes had over 100 routers involved in loops in path traces through their networks. We observe a much higher rate of anomalies in the AS level, with 45% of path traces containing an AS loop. Additionally, we discovered that few of the ASes bounce-back packets where some traces through their network traverse routers in both forward and backward directions. Determining path segments belonging to each AS, we further explore ingress to egress paths of ASes in addition to the source to destination paths within the AS. Analyzing trace segments between ingresses and egresses of an AS, we realized more than half of the ASes have the same router level path between any ingress-egress pair for the 2 months, but others implement load balancing. These results are different from earlier studies that indicated a high level of path dynamism. Our results indicate that the end-to-end path dynamism is due to the Border Gateway Protocol level rather than the router level within ASes.

CCS Concepts: • **Networks** → **Network measurement**; **Network dynamics**; Network layer protocols; Routers; Network management;

Additional Key Words and Phrases: Autonomous system, end-to-end paths, Internet measurement, path stability, routing anomalies

ACM Reference format:
Khalid Bakhshaliyev, Muhammed Abdullah Canbaz, and Mehmet Hadi Gunes. 2019. Investigating Characteristics of Internet Paths. *ACM Trans. Model. Perform. Eval. Comput. Syst.* 4, 3, Article 16 (August 2019), 24 pages.
https://doi.org/10.1145/3342286

*Equal first authors.

This material is based upon work supported by the National Science Foundation under Grants No. CNS-1321164 and No. EPS-IIA-1301726.

Authors' addresses: K. Bakhshaliyev, University of Nevada, Reno, 1664 N Virginia St. Reno, NV 89557; email: khalidb@nevada.unr.edu; M. A. Canbaz, Indiana University Kokomo, 2300 S. Washington St. Kokomo, IN 46902; email: mcanbaz@iuk.edu; M. H. Gune, Stevens Institute of Technology, Castle Point on Hudson Hoboken, NJ 07030; email: mgunes@stevens.edu.

2376-3639/2019/08-ART16 $15.00
https://doi.org/10.1145/3342286

1 INTRODUCTION

The volume of network traffic has been increasing as Internet use encompasses our daily lives. At the same time, users expect a better quality of service from their network providers. To address increasing demand and service expectations, network operators look to better utilize the underlying topology and deploy traffic engineering mechanisms to maximize throughput and minimize latency [1–3]. In particular, load balancing network flows between multiple paths is a common practice [4–7]. Such multiple paths also allow network operators to enhance reliability by switching to a different path in case of failures and congestion [8, 9]. However, many applications such as network games and multimedia rely on predictable end-to-end communication performance [10, 11], which could be affected by altering paths in the network.

Understanding path dynamics is crucial to understand network characteristics and improve communication performance. Researchers have studied the end-to-end route prevalence and diversity [12–14], investigated the effect of routing policies over network topologies [15, 16], analyzed the effect of routing failures on path persistence [17–19], and compared the IPv6 path prevalence with the IPv4 paths [20]. While earlier studies stated that end-to-end paths are mostly stable for a couple of hours or even days [13], others indicated that the end-to-end route persistence is generally unpredictable and highly correlated to the routing dynamics [21–23].

In an earlier work, we analyzed the Internet topology data provided by public active measurement platforms [24]. When we analyzed the Border Gateway Protocol (BGP) datasets from different platforms, we observed that there are some differences between the datasets where some IP address prefixes are announced by different ASes. However, when we analyzed our measurement dataset in this study, only a few of the IP addresses mapped to different ASes. We believe as the measurement data is focused on the Internet backbone, we do not observe transient network domains that are often at the network periphery. We observed different data sources provide unique topological perspectives. For instance, the edges discovered by different platforms are often unique. We also found routing anomalies, as well as highly dynamic path traces which led us to this particular study.

In this article, we extend [25] and perform an in-depth analysis of Internet path dynamics for both end-to-end paths and path segments within each Autonomous System (AS). Additionally, we analyze path anomalies such as trailing repetitions, bounce-back of path traces, and loops at AS and router levels. In our analysis, we utilize 2 months of path traces collected between RIPE Atlas nodes and RIPE anchors [26]. We choose the RIPE Atlas platform over other measurement platforms as it consistently provides traces every 15 minutes between 16,577 nodes and 183 anchors, at the time of data collection. In the collected data, we have over 433 million path traces that span 3,380 ASes. Obtained path traces along with the resolution results are provided at `https://im.cse.unr.edu/data`.

While earlier studies explored end-to-end paths across the Internet, we focus on paths within each individual AS to measure the path dynamics of each network separately. Additionally, earlier studies ignored IP aliases [27, 28] and hence could incorrectly classify path traces that revealed different IPs of the same router as different paths. Each IP address captured in path traces represents an interface of a router on a path. As routers have multiple interfaces and path traces contain different IP addresses of a router, IP alias resolution is crucial to determine the router level connectivity [29].

Our primary contribution in this work is to perform a large-scale end-to-end path dynamics analysis. The key findings of our analysis are listed below:

(1) While paths are stable at AS level for a short duration of an hour or a day, there is considerable change over the longer duration as end-to-end paths traverse through different ASes (see Sections 3 and 5).
(2) Some end-to-end paths are highly unstable as they frequently fluctuate across different ASes. We observe that most of the path traces traverse 5 to 7 ASes while the median AS path length is six hops (see Section 3.2).
(3) We observe that there is a very high number of AS level loops as only 55% of path traces do not revisit the same AS after arriving at another AS. We identified 1,225 unique AS pairs causing these loops (see Section 4.2).
(4) We observe that only 4.4% of router-level path traces have a loop in it. Also, 89% of ASes did not have any path segment with a routing loop (see Section 4.3).
(5) We report a new type of anomaly, called bounce-back, where a trace backtracks over the same set of subnetwork IPs it has traversed (see Section 4.4).
(6) Overall, we observe AS level loops (caused by BGP misconfigurations) more than 10 times as the router level loops (caused by routing inconsistencies) (see Section 4).
(7) When analyzing the shortest paths between ingress to egress paths within ASes, we observe that non-shortest paths are much more prevalent (see Section 5.1).
(8) Persistence of the intra-AS paths ranges considerably where some are persistent for the whole duration of our measurement period, while others are persistent for less than a minute (see Section 5.2).
(9) While load balancers alternate traffic packets between multiple paths, intra-AS networks have more prevalent and persistent paths than end-to-end paths (see Section 5.3).

In the rest of the article, we first provide an overview of the Internet measurement issues, a summary of our data sources and the data processing we performed to obtain AS and router level graphs in Section 2. Then, we analyze the trace dataset characteristics in Section 3, explore path trace anomalies in Section 4, and investigate path stability within individual ASes in Section 5. We compare our findings with the prior work in Section 6. Finally, we conclude with Section 7.

2 METHODOLOGY

In this section, we present the methodology of the experimental analysis and discuss issues involving accurate path measurements. Figure 1 presents the overall flow of analysis performed in this article.

2.1 Internet Topology Measurements

In this section, we provide a brief background on Internet topology measurements. Internet is a connection of AS, independently operated network domains. Each AS has a network of routers to carry traffic and peers with other ASes. ASes exchange information about their networks using the BGP and use internal policies to determine which ASes to transfer data for a particular destination.

Internet topology measurement approaches utilize tools such as *ping* to directly probe a destination and *traceroute* to obtain information about routers on the path toward a destination. Ping relies on the `ICMP Echo request` mechanism to obtain an `ICMP Echo reply` message. When traceroute is initiated toward a destination, the source host sends a probe packet (often using the `ICMP Echo request` but could also probe a random UDP or TCP port) with the time-to-live (TTL) value starting from 1. The TTL value is incremented until the destination is reached, which elicits an `ICMP Echo reply`, or a certain TTL threshold is reached. These incremental probes cause routers on the path toward the destination to send back an `ICMP TTL exceeded` message, from which an IP address of the router is obtained.

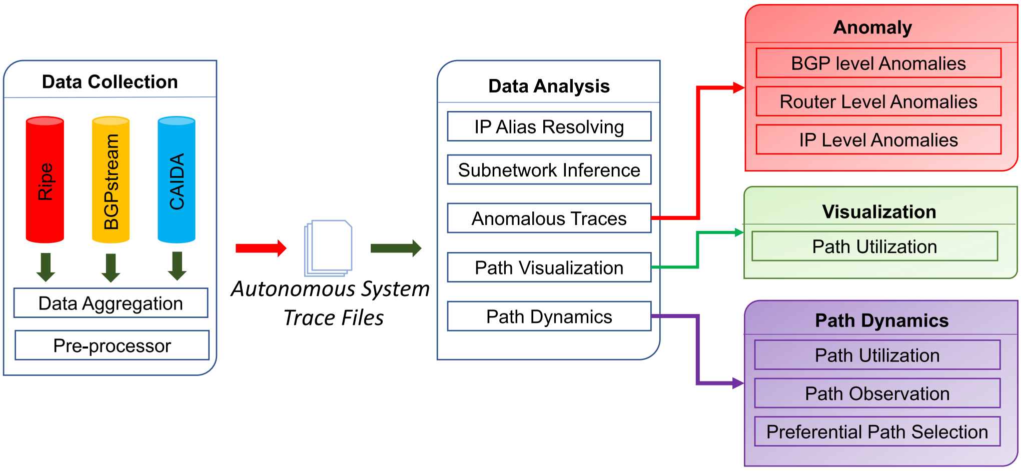

Fig. 1. Overview of the data collection and analysis.

While there are various middleboxes that could impact the Internet traffic, they are typically deployed at the network edge [30]. Some middleboxes could (i) drop a packet, which would be reflected as an unresponsive router (see Section 2.7), (ii) re-route a packet, which would be observed as load balancing (see Section 2.4), or (iii) rewrite the packet, which could happen within an IPv6 or MPLS tunnel. As we perform measurements at IPv4 networks, any IPv6 only network would be observed as a single link between the ingress and egress of the IPv6 network.

Multiprotocol Label Switching (MPLS) utilizes labels rather than IP addresses in forwarding data packets. Labeling allows one to create end-to-end tunnels across a network and may not directly be observed via traceroute probes. When traced, MPLS tunnels might reveal the sequence of links with *explicit tunnels* using the TTL-propagate and the RFC4950 mechanisms and *implicit tunnels* using the ttl-propagate mechanism, or appear as a single link with *opaque tunnels* using the RFC4950 mechanism and *invisible tunnels* using neither options. While the presence of opaque tunnels could be detected in the traceroute, invisible tunnels are not revealed. Sommers et al. [31] investigate the effect of explicit MPLS deployments on the Internet routes using a couple of years of trace data from Ark [32]. Relying on the direct observation of MPLS label information in the ICMP extensions, they show that the total number of tunnels observed varies widely over the time with the largest deployments being in tier-1 providers. Donnet et al. [33] further present the analysis of implicit and opaque tunnels and show that over 30% of the end-to-end paths traverse an MPLS tunnel. They show that a significant fraction of MPLS tunnels are not explicit, but most of these tunnels are traceroute friendly and hence could be discovered.

2.2 Data Source

As the Internet lacks an integrated measurement mechanism, researchers have developed ad-hoc methods to uncover its underlying connectivity. Several research groups have developed public platforms that share data with network practitioners and researchers. Using different deployment approaches, measurement campaigns conducted by these platforms provide different perspectives into the Internet topology. While some platforms provide router level IP measurements (e.g., USC/ISI Ant census [34], CAIDA Archipelago (Ark) [32], Measurement Lab (MLab) [35], RIPE NCC Atlas [26], and UNR IM [36]), others provide AS level BGP measurements (e.g., CAIDA BGPStream [37], CIDR [38], Internet Research Lab (IRL) [39], and Team Cymru [40]). Path traces provide connectivity of routers by identifying a sequence of IP addresses on the path from a source system toward a given destination. Additionally, BGP announcements provide AS level mapping of these IP addresses.

In this study, we utilize path traces collected by the RIPE Atlas measurement platform [26]. The RIPE Atlas measurement platform consists of over 10,000 probes and 400 anchors capable of performing ping, traceroute, DNS, SSL, NTP, and HTTP measurements. They are developed by the Network Coordination Centre (RIPE NCC), and deployed in 3,586 ASes (5.6% of all) by volunteers covering 93% of the countries across the globe. While Ark [32] and M-Lab [35] also provide continuous path traces across the Internet, their traces between a particular source-destination pair are not as frequent or periodic as the RIPE Atlas traces. Also, other platforms shuffle target set for each source in every run, whereas RIPE periodically traces the same targets, i.e., Anchors, from all sources, i.e., Atlas Probes. We harvested trace dataset from RIPE Atlas for 2 months during May-June, 2017 and obtained 433,155,232 path traces which revealed IP level connectivity information of 3,380 ASes. Traces are obtained with the Paris traceroute [41], which addresses load balancing (see Section 2.4).

We also utilized CAIDA BGP Stream to map IP addresses to ASes. BGP Stream provides both historical and real-time BGP measurement data. It enables users to query a number of world-wide BGP routers to analyze BGP tables and prefix announcements. As of August 2017, BGP Stream provides querying capability to 22 Route Views, 24 Ripe NCC BGP, and 83 CAIDA openBMP routers.

2.3 Sampling Bias

One of the challenges in the Internet topology measurements is the sampling bias [42, 43]. Since there are a limited number of vantage points and a large number of destinations, the topological data collected may be biased and not accurately reflect the underlying topology [44]. For instance, Lakhina et al. [42] showed that as traceroutes sample the network region around the vantage point more frequently than the network close to the destination, the degree distribution of routers near the vantage point could differ from the routers that are closer to the destination. This would yield a sampling bias with respect to the degree distribution characteristic. There can be multiple causes for biased samples such as (i) an insufficient number of vantage points [45–47], (ii) limited geolocation of vantage points [48], and (iii) an insufficient number of destinations [49]. As this study focuses on the path characteristics, the main sampling issue would be an imbalance between the number of sources and destinations and their geographic distribution. As shown in Figure 23 in Section 5, the significant majority of the traces are between the 183 RIPE Anchors, which are known to be geographically diverse [26].

2.4 Load Balancing

While TCP performance degrades with unordered packet arrivals [50], traffic engineering over multiple paths enhances network robustness, resiliency, and efficiency [51]. Load balancing routers forward packets through different routes based on certain parameters [52, 53]. Augustin et al. show that routers determine the paths based on (i) destination, (ii) flow characteristics (i.e., based on the first 28 octets), or (iii) randomly [41]. As load balancing routers could transmit measurement probes over different paths, their presence on a path would yield incorrect links in the traceroute output. Paris traceroute fixes the flow identifiers between multiple probes of a traceroute to ensure that packets traverse the same path with per-flow load balancers and obtain a consistent snapshot of a path toward the given destination. However, it cannot control the per-destination load balancing routers but can detect their presence. Multipath detection algorithm (MDA) can identify all paths between a source and a destination pair using *scholastic probing* that generates a large number of probes and analyzes the likelihood of individual links based on the distribution of replies but involves considerable probing overhead [54]. A measurement campaign can also detect the presence of per-destination load balancing. Note, however, that per-destination load balancers would not yield incorrect path traces as traceroute packets would be traversing the same path even if they differ for different destinations.

2.5 IP Alias Resolution

Routers have multiple interfaces, where each interface has a unique IP address. In a given set of path traces, a router may appear with different IP addresses [27]. In IP alias resolution, the goal is to identify nodes that appear to be separate in collected path traces and combine them into one single node (i.e., to detect IP addresses that belong to the same router). Without IP alias resolution the resulting topology may be significantly different from the underlying topology [29].

After slicing path traces into AS regions, we performed IP alias resolution to obtain the router level graph of each AS. We utilized midar [28], which combines probing approaches of the source address based method [55] and the IP identification based resolution [56], as well as kapar [57] which utilizes the analytical alias resolution [27]. We then clustered alias IP addresses to obtain the underlying topology of each AS. Resolving IP aliases helped us correctly classify paths revealing different IP addresses of the same router, which would be marked as different nodes without IP alias resolution.

2.6 Subnetwork Inference

A subnetwork is a set of hosts that share the same physical connection medium and can communicate with each other at the link layer [58, 59]. Interfaces connected to the same subnetwork have a common IP prefix. Each IP address belongs to a subnetwork where all interfaces on the subnetwork have IP addresses with the same maximal x bit prefix, and interfaces on the other subnetworks have different x bit prefixes. Subnetwork inference looks for the distances of IP addresses per vantage point and determines IP address ranges that have the same distance to infer the link-level connectivity. A couple of studies have focused on the efficient inference of the subnetwork [44, 60, 61]. In this work, we only identified /30 or /31 subnetworks for bounce-back analysis as we lack an efficient analytic approach to process 433 million traces to detect larger prefixes. Also, as we utilized historical data, we could not probe the IP addresses to ensure subnetworks in path traces. Hence, there might be a higher number of bounce-backs than the reported ones in Section 4.4 due to larger subnetworks.

2.7 Unresponsive Router Resolution

Unresponsive routers are routers that are passive to measurement probes and are represented by a "*" in a traceroute output [62]. Yao et al. formulate the unresponsive router resolution as an optimization problem where the goal is to build the minimum size topology that does not reduce the distances between known nodes [63]. While unresponsive routers may distort the constructed network topology [64], our focus is on the individual path traces between a source and a destination pair. During edge classification, we ignored unresponsive routers, i.e., nodes that are represented as a "*" in the trace output [64, 65]. They were removed from path segments if they were between two ASes. If they were between IP addresses of the AS, they were kept as a "*" and treated as a match to any IP address in the corresponding trace segment.

2.8 IP to AS Mapping

We performed IP to AS mapping of traces to identify trace segments belonging to each AS using BGPstream data [37]. We sliced path segments into corresponding AS graphs and merged sister ASes. We obtained the sibling AS relationship from the CAIDA AS Organization data [66]. When

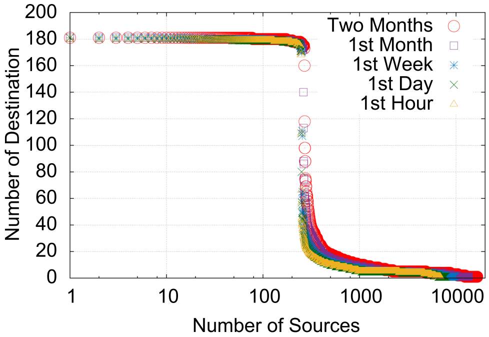

Fig. 2. Number of destinations per source (x:log scale).

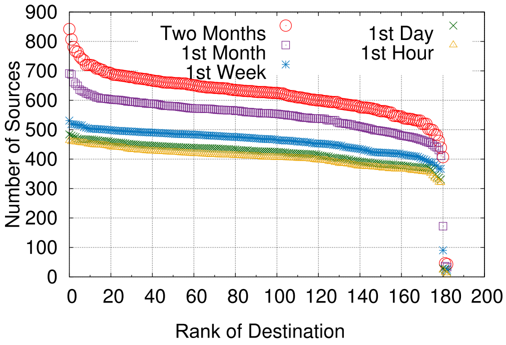

Fig. 3. Number of sources per destination.

we compared AS mapping of our IP dataset with Team Cymru [40], only 18 IP addresses out of 60 thousand IP addresses mapped to a different AS. Path segments within an AS can be in the form of (i) *source-to-destination* if both the source node and the destination of path trace are within the AS, (ii) *source-to-egress* if only the source node is within the AS and the trace exits to another AS, (iii) *ingress-to-destination* if only the destination is within the AS where the trace arrived from another AS, and (iv) *ingress-to-egress* if the AS was an intermediate AS on the path between a source node and a destination that were in different ASes.

3 END-TO-END PATH ANALYSIS

In this section, we first analyze the path trace dataset collected for 2 months (May 1–June 30, 2017) from the RIPE Atlas platform, and then provide an AS level analysis of the measured paths.

3.1 Trace Dataset Characteristics

We observed 0.3, 7.1, 49.5, 221.4, and 433 million traces for the first hour, the first day, the first week, the first month, and 2 months, respectively. We discovered 113,027 source-destination pairs in the whole data while there were 72,967 pairs in an hour. The reason for the increase in the source-destination pairs is due to the Atlas Probes becoming online at different times. While few of the Atlas Probes trace the same pair more frequently than every 15 minutes, some of the probes are not connected to the Internet all the time and hence provide fewer observations. 97%, 89%, 76%, 54%, and 44% of the source-destination pairs are traced *exactly every 15 minutes* for an hour, a day, a week, a month, and 2 months, respectively. Overall, we see that the source-destination pairs are traced consistently over the2-month period.

Furthermore, we analyzed the trace matrix between sources and destinations. Figure 2 presents the number of destinations that are traced by each source in the dataset. We observe that while 183 RIPE Anchor trace toward all other anchors, other Atlas Probes trace toward a few of the Anchors. Similarly, Figure 3 presents the number of sources that trace toward each of the destinations. We observe that almost all of the destinations are traced by a similar number of sources. These results indicate that the dataset is suitable for analysis of path characteristics as it has consistent measurements between the sources and destinations for a long duration.

3.2 AS Path Characteristics

In this section, we analyzed the AS paths observed by each source-destination pair over time. Figure 4 presents the probability distribution function (PDF) of the number of source-destination pairs with a given number of alternative AS paths between them. The figure presents the probability

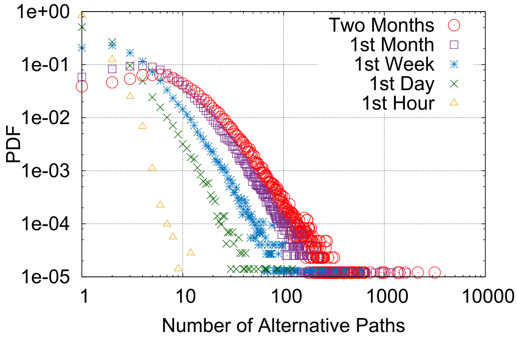

Fig. 4. Source-destination pairs with a given number of alternative AS paths (log-log scale).

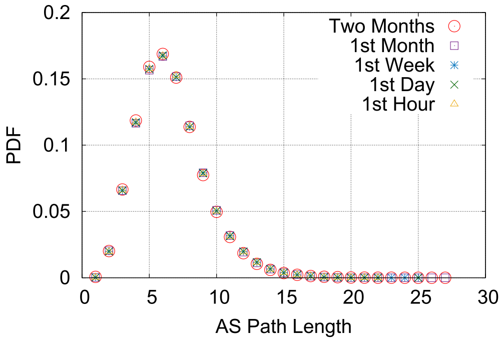

Fig. 5. AS path length distribution.

of a particular number of alternative paths for the first hour, the first day, the first week, the first month, and 2 months. We observed similar distributions for the other time periods with the same duration. The majority of pairs have only one AS path between the source-destination pair in an hour and the same holds for a day's data but with a smaller percentage. As the duration increases to a week, the majority of the source-destination pairs have two AS paths; and most pairs have three to five paths between them with a months' data. We see that the majority of pairs have four to eight alternative AS paths over the 2 months. Out of 113,027 source-destination pairs, there are only 3,333 pairs with a single AS path over the 2 months indicating only 3.9% of traces always traversed the same ASes between the RIPE Atlas probes and the Anchors.

These results indicate that while paths are stable at AS level for a short duration of an hour or a day, there is considerable change over longer durations as end-to-end paths traverse through different ASes. We observe a long-tailed distribution indicating that there are more than thousands of different AS paths for a small number of source-destination pairs when 1 or 2 months data is analyzed. While these could be due to a firewall or such a middlebox that generates random IP addresses in response to traceroute probes at the end of path trace, RIPE Anchors are supposed to be beyond such middleboxes. Additionally, traces with trailing repetitions (see Section 4.1) are filtered and would not contribute to this dynamism. Even during an hour, we observe two paths that crossed through 12 different AS paths. Hence, this indicates that *some end-to-end paths are highly unstable as they fluctuate across different ASes.*

Figure 5 displays the AS path length distributions of the end-to-end traces. The figure shows the PDF of the number of paths with a given AS path length. The figure indicates that the paths are consistent across different time frames in terms of the number of ASes they traverse to reach their destination. We observe that most of the traces pass through five to seven ASes and the median AS path length is six hops. An AS path length of 1 indicates both the source and the destination are within the same AS and is observed in 0.03% of the pairs. In the extreme, there are some paths that cross more than 15 ASes to reach their destinations even in an hour's data and some that traverse over 25 ASes in a month's data. Further analysis of such long AS paths indicated ping-pong between a pair of ASes as detailed in Section 4.2.

4 PATH TRACE ANOMALIES

While tracing toward a destination, probe packets may be dropped on the path, ignored by the routers, or responded with an IP address belonging to a neighboring router on the path instead of the actual router due to the misconfiguration of routers and middleboxes. In this section, we first identify IP addresses that have repetitive patterns at the end of a path trace and filter them.

Then, we analyze loops observed within a trace at both AS level and router level. Finally, we identify a new type of anomaly, named *bounce-back*, where traces backtrack over the same set of subnetworks they have traversed.

4.1 Trailing Repetitions

Middleboxes such as firewalls and NAT devices can cause the repetition of IP addresses at the end of a trace. Some path traces have a mixed set of IP addresses repeating at the tail caused by a border router, a NAT box, a firewall, or a honeypot that replies with a set of IP addresses. Note that Atlas Anchors are expected to be beyond any middlebox that alters traffic indicating that these traces have deviated from the path to its intended destination. In the dataset, there were path traces with 100 hops without reaching its destination. We analyzed such long path traces for anomalies and identified possible combinations of repetitions. In addition to having "$\ldots$-IP_1 - IP_1" as a regular repetition at the end, we observed multiple IPs repeat with varying patterns such as "$\ldots$-IP_1 - IP_3 - IP_2-IP_1" and "$\ldots$-IP_2 - IP_1 - IP_2 - IP_1." In the majority of the cases, we observed one or multiple of the last three IP addresses as repeating within the trace. Hence, we analyzed the last three IPs of path traces to identify repetitions of any one of them within the trace. We filtered such anomalies if there were at least two repetitions of any combination of the same IP or different IPs.

We observed that 10,483,414 out of 433,155,232 traces (i.e., 2.4%) had a combination of at least two of the last three IP addresses repeating within the traces in the 2 months data. When we analyzed the ASes of the IP addresses involved in the trailing repetitions, we observed that 335 ASes had traces with different lengths of trailing repetitions. We filtered such traces from the first occurrence of the repeating IP address before any of the analysis. These random patterns are not informative of the underlying path but most likely a result of a middlebox or routing misconfiguration.

4.2 AS Level Loops

We analyzed AS level loops by converting all IP addresses to their AS number and identifying any loop where a trace traverses the same AS domain after leaving into another AS. We analyzed the length of the AS loops by considering the number of intermediate ASes that are traversed by the trace before re-entry to the same AS. For instance, if a path trace traverses "AS_A-AS_B-AS_C-AS_A," we considered this as a length 3 loop since AS_A is observed again after two other ASes. Figure 6 shows the distribution of the loop lengths. Considering 2 months data, we observe that 56% of traces have no AS level loop while 32% of path traces have a loop with only one other AS in between and 6.4% have a loop with two ASes in between. We also observe AS loops that have 15 to 20 ASes in between indicating packets re-enter the same AS after being exchanged by many ASes. Overall, we obtain similar probabilities for observing a particular AS loop length at different time scales.

Figure 7 further explores the number of times the same AS is traversed after a path trace departs to other ASes. Similar to AS loop lengths, the probability distribution function shows the consistency of AS re-entries at different time scales. We observe that 44% of path traces have an AS that is re-visited only once and 6.5% have an AS re-visited twice in the 2 months data. While a higher number of re-entries are possible with a smaller probability, we observe some ASes are re-visited 7 to 11 times after the trace encounters other AS domains on its path.

Further analyzing path traces with such a high number of revisits, we observed a **ping-pong** behavior between two ASes in most of them. That is, path traces would leave an AS_A into AS_B and then return back to the AS_A and further into AS_B multiple times. Overall, 17,675,940 path traces (i.e., 4.1% of traces) had a ping-pong between two ASes and there were 1,225 unique AS pairs that would bounce packets between them. Among traces with such ping-pong between ASes, only 75,427 (i.e., 0.43% of ping-pongs) were able to reach their destination while the significant majority

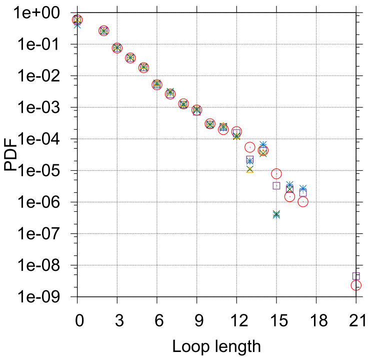

Fig. 6. AS level loop length (y:log scale).

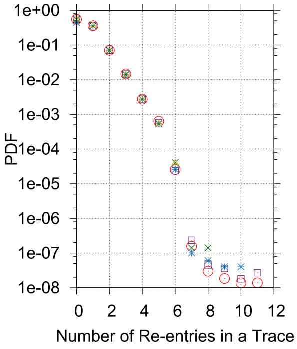

Fig. 7. AS re-entries (y:log scale).

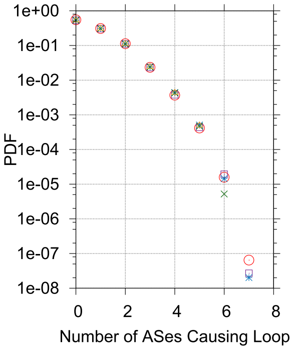

Fig. 8. Multiple AS loops in a trace (y:log scale).

did not reach their destination. This behavior might be due to the hot potato routing where ASes have advertised the destination network to each other.

Finally, Figure 8 presents the number of different ASes looping in the same path trace. We observe that 37% of path traces have only one AS looping back and 14% of traces have two ASes looping in the same trace. While the majority of traces with an AS loop have re-visits of only one AS, there are traces that revisited five to seven different ASes on its path in a week's data.

Overall, we observe about half of the traces contain an AS level loop where the same AS is revisited on the path toward its destination. As we had merged sister ASes based on [66], the existence of such revisits of an AS indicates BGP level errors. Moreover, in a non-negligible number of paths, multiple such AS re-entries of the same or different ASes occur, indicating considerable misconfiguration of BGP routers. We believe this behavior is due to hot-potato routing where a network gets rid of a packet as soon as possible. The AS level loops could hinder the end-to-end traffic flows as they might be redundantly transmitting packets over additional AS networks.

4.3 Router Level Loops

After analyzing AS level loops, we investigated router level loops by identifying traces that cross the same router multiple times. Note that we have performed IP alias resolution using both the probe based [28] and analytical [27] approaches and identify loops due to alias IPs in addition to the traditional IP repetitions. In a considerable number of path traces (i.e., 12% of all traces), we observe that the same IP address is repeated twice as in "**aa**," which might not be an actual routing loop but a proceeding or succeeding router responding with the TTL exceeded message instead of the actual router on the path. For instance, on some of the submarine backbone links, we observed the first router would not check for the TTL and the downstream router's IP would be repeated in the path trace. Hence, we did not include such cases in the following analysis.

Analyzing segments of traces belonging to each AS, we observed that 367 out of 3,380 ASes contained a path with a routing loop while the others (i.e., 89% of ASes) did not have any path segment with a routing loop *indicating that their routers were configured such that routing loops were not manifested*. Figure 9 shows the ranks of ASes by the number of traces with loops and Figure 10 shows the ranks of ASes by the number of routers that are involved in loops. Note that each figure is independently ranked at each time frame. For instance, in Figure 9 we observe that in the first hour, the first day, the first week, the first month, and 2 months data there are 73, 94, 119, 140, and 147 ASes, respectively, that have only one router involved in routing loops. As we extend the analyzed time frame from 1 hour to 2 months, we see that the number of routers involved

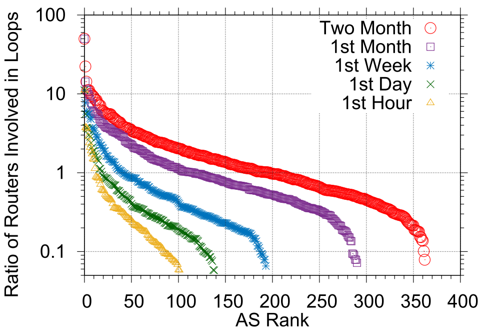

Fig. 9. AS rank by the number of path traces with a loop (y:log scale).

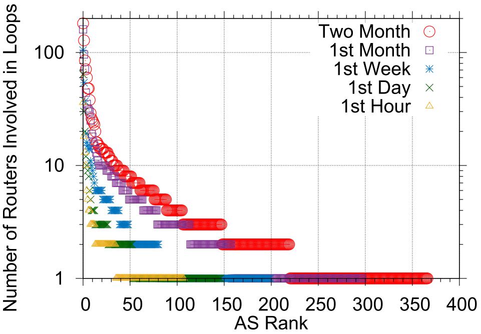

Fig. 10. AS rank by the number of routers involved in loops (y:log scale).

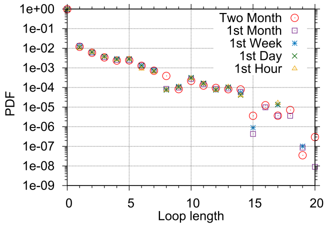

Fig. 11. Router level loop length (y:log scale).

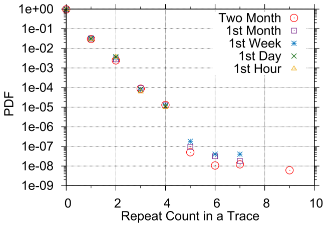

Fig. 12. Router re-entries (y:log scale).

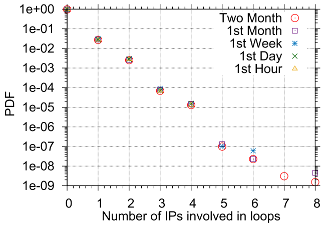

Fig. 13. Multiple loops in a trace (y:log scale).

in routing anomalies increase. For the 2 months data, we observe that while there are ASes with hundreds of routers that are involved in loops, the significant majority of ASes with routing loops have a couple of routers which are involved in the loop. In the majority of the ASes, a couple of routers are observed in the loop indicating that even a handful of misconfigured routers could lead to a considerable number of redundantly longer paths at the router level.

Figure 11 presents the PDF of a trace with a given loop length. Note that a loop length of 1 indicates a router's IP address(es) is being observed after a different IP address in the trace. We observe that 95.6% of traces have no routing loop in them indicating only 4.4% of end-to-end path traces have a routing loop, after filtering trailing repetitions (as described in Section 4.1). We observe longer loop lengths with smaller probabilities. In 30% of the routing loops an IP address is visited after another IP address, and in 25% a router is visited after two other IP addresses.

Figure 12 further explores the number of times a router is observed in the same end-to-end path trace. We observe that 62% of routing loops happen with a router being repeated only once in 2 months data and 5.6% of loops have the same router being repeated a second time (i.e., the router's IP address(es) are observed three times in the trace). We observe the same router appearing four times in the same path trace with a probability of less than 10^{-4}. Note that trailing repetitions where a set of IP addresses were observed multiple times due to a middlebox are filtered (see Section 4.1) and occurrence of the same router multiple times indicates misconfiguration of routers or a temporal behavior.

Finally, Figure 13 presents the number of different routers being looped back in the same trace. We observe that in 91% of the traces that have loops, there is one router that repeats, 8.5% have two routers looping back, and 0.23% have three routers looping in the same trace. In the extreme, we

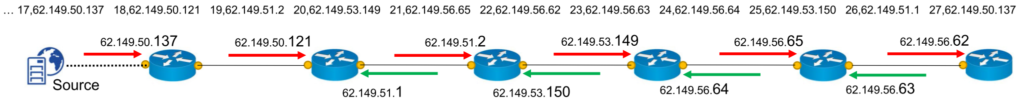

Fig. 14. Bounce-back sample.

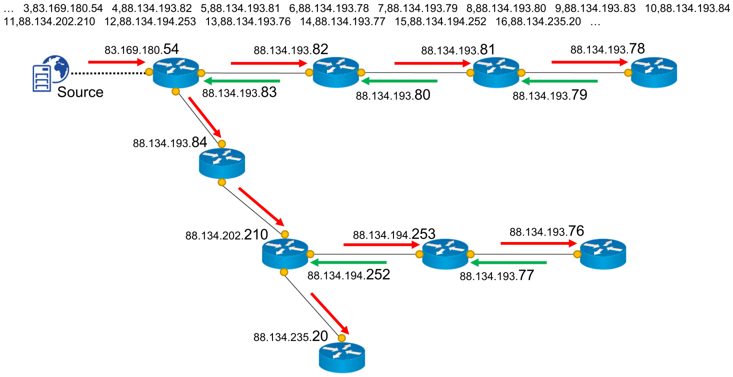

Fig. 15. Multiple bounce-backs sample.

see four different routers being re-visited in the same path trace in a day's data and eight routers in a month's data.

Overall, we observe that end-to-end path traces have similar loop length distributions, number of re-entries, and number of multiple loops at the router level and the AS level. However, while 4.4% of end-to-end path traces have a routing loop, 44% of traces have an AS level loop. This indicates end-to-end paths have much higher likelihood of AS level loops caused by BGP misconfigurations (i.e., more than 10 times) than router level loops caused by inconsistencies of routing tables.

4.4 Bounce-Back

Analyzing path traces, we observed that some routers are being traversed in both forward and backward directions. For instance, Figure 14 presents a trace where the packets enter the routers from the eastbound interface (shown with the red arrow) and then return from the westbound interface (shown with the green arrow) of the same set of routers. Note that the subnetworks of these routers' IPs indicate a point-to-point link with a /30 or /31 subnet mask [58, 67]. Similarly, Figure 15 presents a trace that is bounced back twice, first by three routers and then by two routers.

We analyzed path traces for any point-to-point subnet where IP addresses differ by one. We also included alias IP addresses and unresponsive routers, if they were inside a bounce-back and not the outermost bounce. Figure 16 presents the length distribution of bounce-backs. For instance, bounce-back in Figure 14 has a length of four as four subnetworks are traversed in both directions. Note that we excluded bounce-backs that involved one subnetwork as they could be due to ICMP replies from non-incoming interfaces. We observe that 574,237 of path traces (i.e., 0.13% of all) contain a bounce-back while the remaining significant majority of path traces do not contain a bounce-back. Further analysis of the trace segments with bounce-backs revealed that only seven ASes were involved in these bounce-backs and in some instances, a couple of ASes were part of the bounce-back segment.

Finally, Figure 17 analyzes the number of multiple bounce-backs in a particular path trace. For instance, Figure 15 has two bounce-backs. We observe that only a very small fraction of path traces

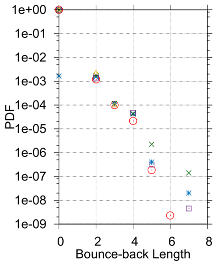

Fig. 16. Bounce-back lengths (y:log scale).

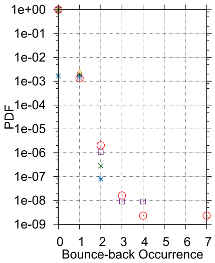

Fig. 17. Multiple bounce-backs (y:log scale).

contain multiple bounce-backs but in the extreme, there was a path trace with seven bounce-backs of different lengths.

Overall, the reported bounce-backs are an underestimation of the actual bounce-backs in path traces as we only identified point-to-point subnetworks (i.e., a subnet mask of /30 or /31). Detection of the underlying multi-access subnetworks depends on the similarity of the distance of the subnetwork IP addresses to the vantage point [58, 67], but traces with bounce-backs would invalidate the subnetwork as the distances of subnetwork IP addresses differ by more than one hop. Hence, a new subnetwork detection approach is needed with traces that include bounce-backs. Moreover, to our knowledge, this is the first study to identify bounce-back behavior in trace datasets. While the number of ASes where the bounce-back behavior is detected is very small, their existence indicates a misconfiguration of routing tables that wastes network bandwidth as packets traverse the same routers twice without getting closer to the destination.

5 PATH DYNAMICS WITHIN AUTONOMOUS SYSTEMS

In this section, we analyze the stability of path traces within each AS considering *path prevalence*, i.e., the probability that a particular path is observed, and *path persistence*, i.e., the probability that a route remains unchanged for a particular duration [12]. In our analysis, we focused on the *path frequencies*, i.e., how often a particular path is observed between the ingress/source and egress/destination pair (called as *ingress-egress* for brevity), and the *path distance* with respect to the shortest observed path among 2 months of trace data. We detect ingress and egress of an AS as routers preceding and succeeding the AS in the path trace, respectively. If the source or destination is within the AS, they are considered as the start or end of the trace segment, respectively. We present details of sample ASes and then average analysis for all 3,380 ASes observed in the dataset.

5.1 Path Prevalence Analysis

We first analyze the prevalence of a particular router level path between an ingress-egress pair. Figure 18 presents a 15-minute snapshot of path traces belonging to the highest transit degree AS, as ranked by the customer cone size [66]. Each point indicates the presence of a trace sample between the ingress-egress pair colored by the observation *frequency* on the left and by the path *length* on the right. The y-axis indicates a specific ingress-egress pair through which the path trace is collected, and the x-axis indicates the timestamp of the trace starting from May 1, 2017 midnight. Note that these ingress-egress pairs are ranked from the most prevalent (at the bottom) to the least considering the 2 months data and only present a 15-minute snapshot of the most observed 50 paths. Even though end-to-end paths are traced every 15 minutes by the RIPE Atlas, path segments of intermediate ASes are observed with a greater frequency as they are transit on

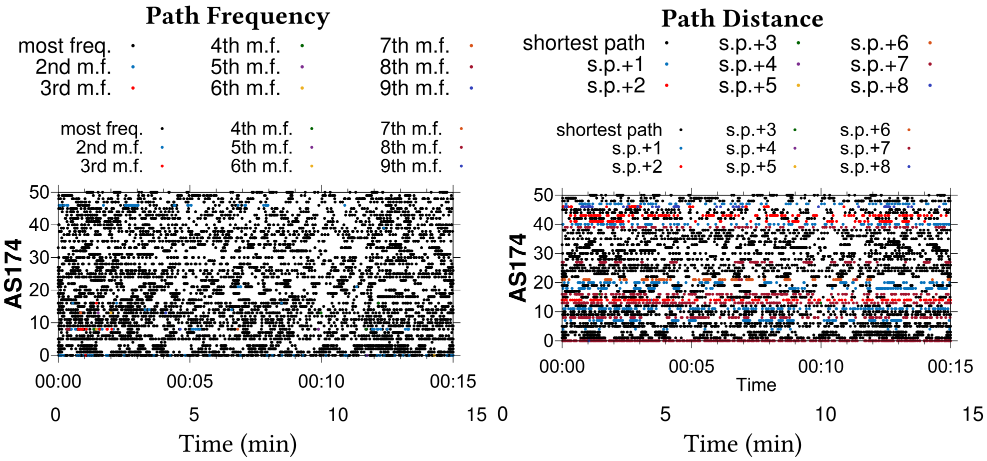

Fig. 18. Path observation visualization of frequency and path length of a backbone AS.

the path toward the destination. Hence, we obtain multiple snapshots of the ingress-egress pair of backbone ASes in the 15-minute interval.

In the *path frequency* graphs on the left, the most frequently observed paths between the ingress-egress pair are colored "black," the second most frequent as "blue," the third most frequent as "red," and so on. In the *path distance* graph on the right, the shortest paths observed between the ingress-egress pair are colored as "black," the paths one hop longer than the shortest paths are colored as "blue," the two hop longer paths as "red," and so on. Note that the frequency is determined based on the 2 months data and the shortest path is the minimal hop path observed between the ingress-egress pair anytime between May 1 and June 30, 2017.

The first observation is that, comparing the observation *frequency* and path *length* based ranking of the paths, we observe that *length* based ranking seems to be alternating more while the *frequency* based ranking seems to be more stable. This indicates that these *ASes are handling their traffic based on their preferred paths that are not necessarily the shortest paths.* We observe that load balancers swap flows between multiple paths as indicated by different colors such as the pairs #1, 8, 13, 16, 22, and 46 in the *path frequency* graphs on the left. Even though frequency based figures are dominated by black dots indicating observance of the most frequent path, there are ingress-egress pairs with nine different paths over the 15-minute sample.

When considering the *path length* based graphs on the right, we observe that some of the paths have considerably different path lengths. There are paths dominated by non-black dots for some ingress-egress pairs indicating that packets often cross the AS through non-shortest paths while shorter paths are observed for the pair. In the presented 15 minute sample, we observe 18 pairs (out of 50) that were often the non-shortest path with some being eight hops longer than the shortest observed path.

Figure 19 presents the prevalence of paths ranked by the *frequency* on the left and path *length* on the right between the ingress-egress for two top-ranked backbone ASes as ranked by the customer cone size [66]. We observe that the ASes 174 and 3,257 trace the same path for 67% and 97% of ingress-egress pairs, respectively. Other top-ranked ASes had distributions similar to the AS 174. Note that the ranking in the x-axis is based on the path observation ratio of the *most frequent path* for the ingress-egress pair on the left and the *shortest length path* on the right. If there are multiple paths for the same ingress-egress pair, they appear at the same x-value with different markers indicating the usage ratio of that path. Note that paths that are almost equally observed

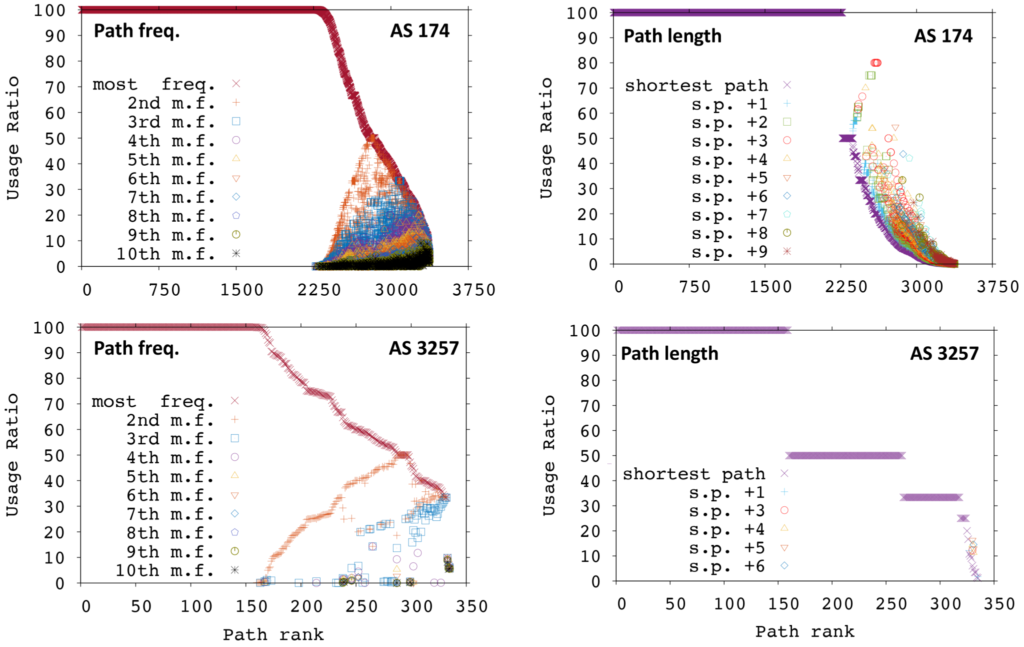

Fig. 19. Path usage ratio for sample backbone ASes by path frequency (*left*) and path length (*right*).

at frequency graph overlap at 50% for red, 33% for orange, 25% for blue, 20% for purple, and 17% for yellow indicating that there are 2, 3, 4, 5, and 6, respectively, alternative paths chosen equally likely. Among the backbone ASes, AS 3,257 seems to have the majority of its paths taking one path and hence almost all paths seem to be at the same shortest path length. Additionally, in most of the cases with the presence of non-shortest paths between an ingress-egress path pair, the shortest paths are used less often than the longer ones. This indicates *ASes often prefer to use longer paths, probably to reduce congestion at the core of their network.*

Paths can have a fewer number of path lengths compared to the path frequencies as some of the differing paths would have the same hop distance, and hence they are combined under the same path length in the plots. For instance, in AS 3,257, we observe 10 different paths when ranked by *path frequency*, but there are 6 different path lengths in the *path length* plots. Also, we do not observe a path that is two hops longer than the shortest path (i.e., s.p.+2). In the *frequency* based ranking of AS 174, we observe more than 2,200 ingress-egress pairs with only one path in between while there are about 1,100 pairs with 10 different paths between them. Toward the end of the distribution, we observe the most frequently utilized path is observed in about 10% of the traces between an ingress-egress pair. In the *path length* based ranking, we observe that the shortest path is often the least utilized path when there are alternative paths at different lengths as reflected with purple markers at the bottom of the distribution between the 2,250th to 3,400th ingress-egress pair. In a couple of ingress-egress pairs, paths that are more than five hops longer than the shortest path are utilized in 70–80% of the path traces.

Figure 20 presents the prevalence average of all paths of the top six ASes by *frequency* (i.e., the most frequent, the second most frequent, the third most frequent, and so on) on the left and by *length* (i.e., the shortest path, one hop longer than the shortest, two hops longer, and so on) on the right. We observe that on average, ASes 3,356, 1,299, 174, 2,914, 3,257, and 6,762 utilize the most frequent path for any ingress-egress pair on 85%, 82%, 86%, 89%, 81%, and 89% of all paths,

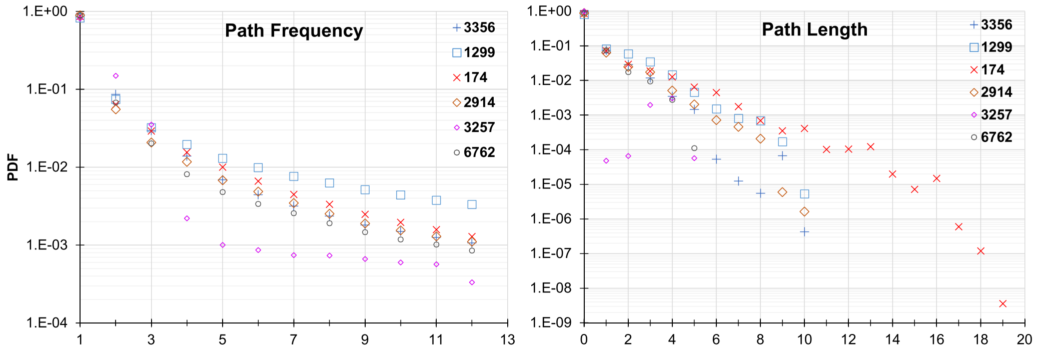

Fig. 20. PDF of top ranked ASes by path frequency (*left*) and path length (*right*) (y:log scale).

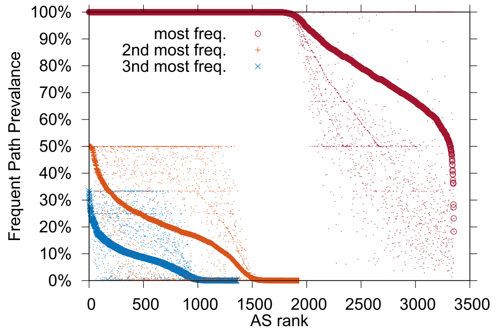

Fig. 21. Path observation averages of frequent paths for all ASes ranked independently.

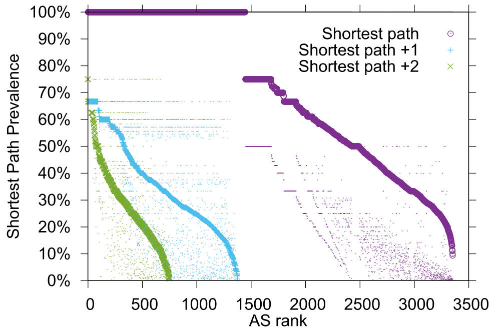

Fig. 22. Path observation averages of shortest paths for all ASes ranked independently.

respectively. Similarly, these ASes utilize the shortest path on 88%, 81%, 85%, 89%, 99%, and 90% of all paths, respectively. On average, for the top six backbone ASes, we observe the most frequent path in 85% of all traces, the second most frequent path in 8.2%, the third most frequent path in 2.8%, and the fourth most frequent path in 1.2% of all traces. Similarly, we observe the shortest path in 89%, one hop longer than the shortest path in 6.1%, two hops longer paths in 2.6%, and three hops longer paths in 1.6% of all traces.

We observe AS 174 has paths that are 12 to 19 hops longer than the shortest path while it has up to 12 alternative paths for an ingress-egress pair. A reason for observing much longer paths could be due to the occasional use of MPLS tunnels that are not reflected in some path traces (i.e., TTL propagation is disabled within the MPLS core) or actual variation between hop count of alternative paths utilized for traffic engineering.

Figure 21 presents the average prevalence of the top three *most frequent paths* of each ingress-egress pair along with the min and max observation of the frequent path for each AS. The small dots indicate the min and max utilization of the frequent path observed for any of the AS' ingress-egress pairs. We observe that 1,317 ASes (i.e., 40% of analyzed ASes) use a single path between all of their ingress-egress pairs. There are very few ASes that utilize the most frequent path of less than half of the ingress-egress pairs. The usage ratio of the most frequent paths varies widely as shown with smaller dots for the min and max of any path in the AS. Additionally, there are 327 ASes (i.e., 9.4%) with exactly two alternative paths between any of its ingress-egress pairs. Note that some of the large ASes had a couple of hundreds of alternative paths between an ingress-egress pair (not shown). This could be due to multiple random load balancers that randomly send

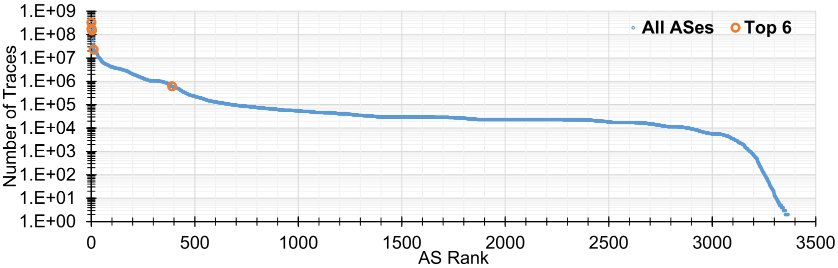

Fig. 23. ASes ranked by number of traces observed within their domains (top ranked six ASes are marked) (y:log scale).

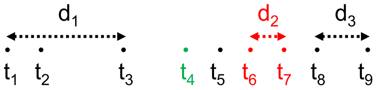

Fig. 24. Path duration calculation.

packets downstream. Traffic flows through such ingress-egress pairs would have a larger variation in the delay and jitter.

Similarly, Figure 22 shows the average prevalence of the three *shortest paths* of each ingress-egress pair along with the min and max observation ratio of such paths for each AS. We observe that 1,319 ASes (i.e., 43% of analyzed ASes) use the shortest path between all of their ingress-egress pairs. This indicates two ASes use the same length path even with different paths for all of their path pairs. Additionally, there are 624 ASes (i.e., 19%) with exactly the shortest path or one hop longer path between any of its ingress-egress pairs. We observe that for many ASes the shortest path or one hop longer paths are utilized less often than the two hops longer paths. Compared to the most frequent paths in Figure 21, we observe that only 155 AS (4.6% of all) with alternating paths have the same length paths and the vast majority of the frequent paths are longer than the shortest paths (which was also observed in the top-ranked ASes in Figure 19).

5.2 Path Persistence Analysis

In this section, we focused on the persistence of the paths, i.e., for how long a path was stably utilized over the course of 2 months. To this end, we first looked at the number of traces passing over each AS. Figure 23 shows the ASes ranked by the number of traces that pass through their domain. We marked the top six ASes analyzed in the previous section. Overall, the most traced AS was observed in 76% of all path traces. Also, 0.15% of the ASes have more than 100 million traces traversing their network domain, 9.4% of the ASes have more than a million traces, 85% of the traces have more than 10,000 traces, and 6.4% of the ASes had less than 1,000 traces through their network. Since we are focused on the duration of the paths in the course of 2 months, we filtered out the path pairs which do not have at least three traces for each hour. As a result, in this section, we analyzed 2,543 ASes out of the 3,380 ASes.

In order to find the duration of a particular path, we sorted paths by their timestamps similar to Figure 18. For each set of the consecutive path, we identified its duration from the first observation to the last one as shown in Figure 24. We ignored pairs that are separated by more than an hour to eliminate any bias that would be caused by infrequent tracing of a path. Note that as RIPE traces a destination from a source approximately every 15 minutes, we skip an ingress-egress pair only if none of the five or six observations are through this ingress-egress pair consecutively. If samples are separated by long durations, they could be incorrectly interpreted as long stability.

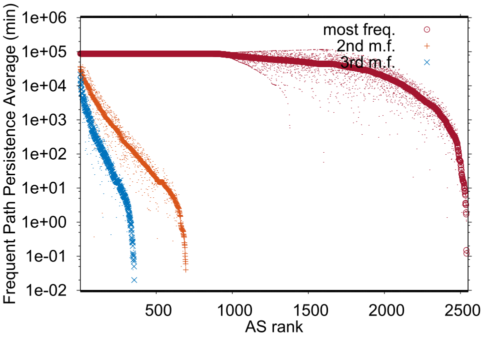

Fig. 25. Average observation duration of frequent paths for all ASes ranked independently.

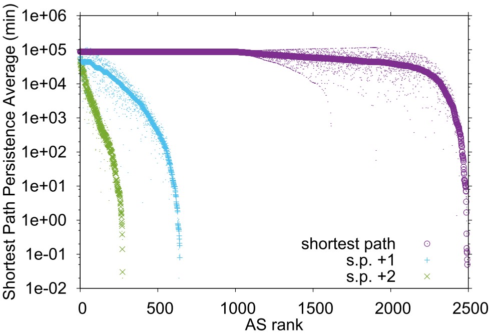

Fig. 26. Average observation duration of shortest paths for all ASes ranked independently.

The arrows indicate the time span for a particular path with multiple observations. For instance, we observe a path three times at t_1, t_2, and t_3. This path's duration is calculated as $t_3 - t_1$. Since there is no continuous interval for paths observed at t_4 or t_5, they are not considered in the time duration calculation. Hence, some of the ASes are not present in this analysis as neither of their ingress-egress pairs were traced over the same path consecutively within an hour (either because of router-level fluctuations or AS level fluctuations, which was observed with a higher frequency). Note that we may not observe path fluctuations that happen faster than the sampling period.

Figure 25 presents the persistence of the top three *most frequent* paths along with the 95% confidence interval (shown as smaller dots) for each AS, ranked by the average duration independently. The most observed paths stay the same for 87,840 minutes, i.e., the 61-day data collection period. Note that we have only 2,541 ASes in the plots as the remainder (i.e., two ASes) did not have any consecutive observations of the frequent paths within an hour period. We observe that the average duration of the most frequent paths ranges from 61 days for all paths of 865 ASes (i.e., 34% of the analyzed ones) to less than a minute for ASes at the tail. We even observe ASes with the second and the third most frequent paths being observed for a couple of seconds on average. This indicates the AS is *fluctuating flows between alternative paths and hence we do not observe the same path consecutively.* While load balancing between multiple paths is beneficial for traffic engineering, end users would benefit from consistent paths for the same flow.

Overall, we observe that only 27% of analyzed ASes have a second alternative path, indicating the rest always used the same path even if they were not periodically observed between the ingress-egress pair (i.e., 39% of the ASes). Note that the difference is because we ignore time durations for which there is no sample for an hour. Also, 14% of ASes have a third alternative path observed between any of its ingress-egress pairs.

Similarly, Figure 26 shows the persistence of the three *shortest paths* for each ingress-egress pair along with the 95% confidence interval (shown as smaller dots) for each AS. Similar to the frequent paths, we observe high variation between duration of the shortest paths. In this case, we have 934 ASes (34% of analyzed) whose paths are always the shortest path even if they changed (i.e., 8 ASes had alternative paths of equal length). Note that, in these figures, we have 2,447 out of 2,543 ASes presented as remaining ASes did not have stable observations of the three shortest paths within an hour period. Compared to frequency based results of Figure 25, this indicates 94 ASes had three most frequent paths appearing consecutively in an hour but neither of their three shortest paths appeared consecutively even as a single pair in an hour during the 2 months of the data collection. Overall, we observe that *only 29% of the analyzed ASes have a path that is*

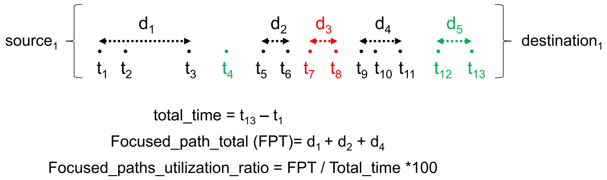
source$_1$
d_1
d_2
d_3
d_4
d_5
destination$_1$
t_1 t_2 t_3 t_4 t_5 t_6 t_7 t_8 t_9 t_{10} t_{11} t_{12} t_{13}
total_time = $t_{13} - t_1$
Focused_path_total (FPT)= $d_1 + d_2 + d_4$
Focused_paths_utilization_ratio = FPT / Total_time *100

Fig. 27. Path utilization ratio calculation.

one hop longer than the shortest and 12% of the ASes have a path that is two hops longer than the shortest observed path between any of its ingress-egress pairs.

5.3 Utilization of Path Analysis

In this section, we focus on the total utilization of a path over the possible alternative paths between each ingress-egress pair of all ASes. In order to find the total utilization time ratio of a particular path, we sorted paths by their timestamps. For each set of consecutive paths, we identified its duration from the first observation to the last one as shown in Figure 27. The arrows indicate the time span for a particular path with multiple observations. In the previous section, we ignored the observations which have a gap longer than an hour in two timestamps. For this analysis, we included all observations into the duration calculation regardless of how distinct the timestamps were. For instance, we observe a path three times at t_1, t_2, and t_3. This specific observation duration d_1 is calculated as $t_3 - t_1$ irrespective of the duration between instances. Since there is no continuous interval for paths observed at t_4, it is not considered in the time duration calculation.

Later, we calculated the total time frame (i.e., the observation time) t_{total} for this specific pair by $t_{total} = t_{13} - t_1$. As we rank paths based on their *frequency* and their *length*, we sum up all durations of the currently focused path by $F_{path} = d_1 + d_2 + d_3 + d_4$. Next, we calculate the ratio of the focused path over the total time frame by $FPR = F_{path}/t_{total}$ for the black-colored path. For each AS, we calculate FPR for every pair within the AS and then calculate the average of the FPR for all path pairs.

In Figure 28, *lower-left*, we ranked 2,541 ASes based on the average of their total time ratios with respect to the frequency. Also, in Figure 28, *upper-left*, we present the alternative path averages for each pair within the specific AS presented in Figure 28, *lower-left*. For instance, in Figure 28, *lower-left*, 1,316 ASes (i.e., 52% of the analyzed) have utilized a single path all the time for all of their ingress-egress pairs. However, 1,227 ASes have more than one route in between their ingress-egress pairs.

Similarly, in Figure 28, *lower-right*, we ranked the ASes based on the average of their total time ratios with respect to the *length*. In Figure 28, *upper-right*, we present the alternative path average for each pair within the specific AS shown in Figure 28, *lower-right*. For instance, in Figure 28, *lower-right*, 1,532 ASes (i.e., 60% of the analyzed) have utilized the shortest paths in all of their paths. Different from the frequency figures in Figure 28, *upper-left*, in Figure 28, *upper-right*, we see that even though the shortest path is utilized all the time, these ASes have alternative paths between their ingress-egress pairs. Furthermore, 1,011 ASes have utilized different length paths.

In both *upper* figures showing the average number of alternative paths, we occasionally see a large number of routes between the path pairs of an AS. For instance, AS 32,934 was ranked as the 2,518th based on the total utilization and had 39 ingress-egress pairs in its domain. There were a total of 1,448 different path traces over these 39 ingress-egress pairs, as presented with an average of 37 alternative paths in the *upper* figure. Similarly, in the AS 14,420, the shortest path was utilized 77% of the time, whereas one hop longer path was utilized 15% of the time. There were

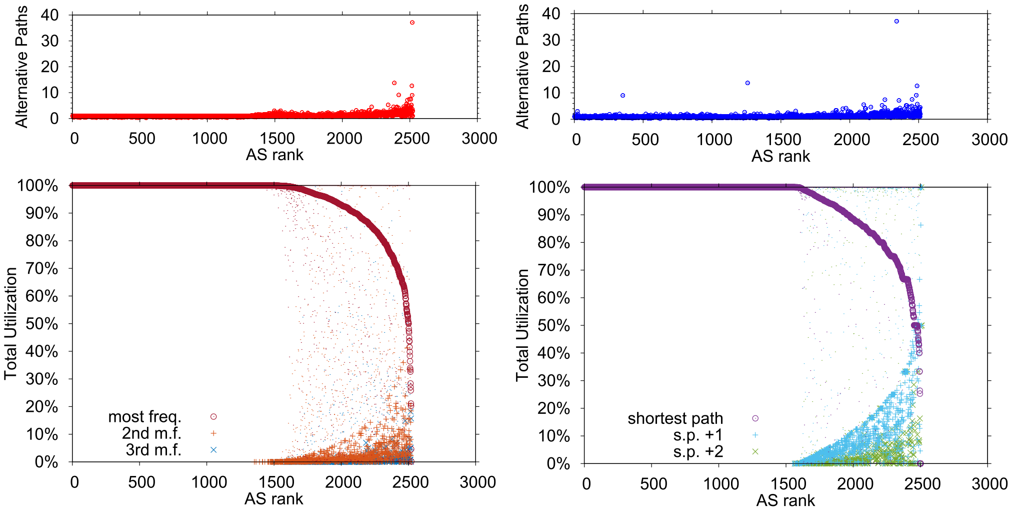

Fig. 28. Total path utilization ratio for all ASes ranked by the path frequency (*bottom-left*) and shortest path (*bottom-right*), along with path alternatives between source-destination pairs for all ASes by the path frequency (*top-left*) and shortest path (*top-right*).

19 ingress-egress pairs, with a total of 141 different paths between these pairs, resulting in an average of 7.4 alternative paths for the AS. Overall, we observe that more than half of the analyzed AS networks always use the same ingress-egress path, indicating they do not employ any load balancing. Also, 60% of the analyzed ASes utilize the same length paths while some of the paths differ.

6 RELATED WORK

While there has been extensive analysis of the end-to-end routing over the last two decades [12, 68], it is still focus of measurement studies as the Internet evolves with new technologies such as Multiprotocol Label Switching (MPLS) [69], Multi-Path TCP (MPTCP) [70], Software-defined networking (SDN) [71], cloud computing [72], and edge computing [73]. Prior Internet path measurements mainly focused on the packet-loss, packet corruption, out-of-order packet delivery, routing convergence, and bottleneck bandwidth detection [12–14, 17, 74].

Overall, our findings confirm the prior works that report end-to-end paths observed between any given source-destination pair has great diversity [12–14]. However, none of the earlier studies focused on the inra-AS path characteristics. While some of the ingress-egress paths only use a single path, others alternate traffic over tens of different paths. We observe that only 40% of the ASes use a single path between any of their ingress-egress routers while the rest of them transmit their traffic over different paths. Interestingly, some of the largest ASes had a couple of hundred alternative paths between their ingress-egress routers indicating a very high level of load-balancing by their routers. This is aligned with prior work that indicated tier-1 ASes have higher use of MPLS tunneling [33]. Researchers also showed that there is considerable dynamism of BGP announcements [18, 74, 75]. Our results also indicate that while AS level paths are stable for short durations of an hour or a day, there is considerable change over the longer duration as end-to-end paths traverse through different ASes.

Routing anomalies have also been reported previously [22, 41, 76–78]. We similarly identify anomalies in four categories, namely, trailing repetitions, AS level loops, router level loops, and bounce-backs. While earlier work has pointed to the first three anomalies, this study is the first to show the presence of the bounce-backs across a network domain.

7 CONCLUSION

In this article, we investigated the characteristics of Internet paths. Utilizing 2 months of traces that are collected every 15 minutes between 16,577 source nodes and 183 destinations, we performed a detailed analysis of path dynamics of end-to-end path paths along with intra-AS paths. Overall, we observed significantly more trace anomalies and path dynamism at the AS level than at the router level.

Analyzing end-to-end paths for routing anomalies, we observe that 18% of ASes contain routing loops within their network indicating misconfiguration of routers. Some of the ASes had over 100 routers causing loops in the path traces through their networks. We observe a much higher rate of anomalies in the AS level, with 44% of path traces containing an AS loop, while 4.4% of end-to-end traces contained a router level loop. Additionally, we discovered that few of the ASes bounce-back packets, where some traces through their network traverse routers in both forward and backward directions.

About half the ASes had a single path between all of its ingress-egress pairs while others had altering paths. Considering end-to-end paths, we observe that only 22% cross the same ASes and hence would result in differing end-to-end paths. These results indicate that while individual ASes have more stable paths internally, end-to-end paths are much more dynamic. While earlier studies focused on the end-to-end paths and found highly dynamic paths (such as [8] reporting that over 60% of the end-to-end source-destination pairs to traverse more than 10 different paths and 50% of the pairs never have the same path), our results reveal that the main cause of this behavior is due to the BGP level dynamism rather than the router level dynamism.

Analyzing the prevalence and the persistence of path segments, we see that frequent or shortest paths' average observation duration varies significantly from 2 months of measurement period to less than a minute. Results indicate that about half of the ASes utilize the same path for all of its ingress-egress pairs while others distribute the traffic over multiple paths. The alternate paths, however, are often a couple of paths with only a few ASes having tens of alternative paths on average. Finally, the majority of path segments through a network domain are not the shortest paths. While utilizing non-shortest paths could help avoid congestion, there are paths that are much longer than the observed shortest path.

REFERENCES

[1] Turgay Korkmaz and Marwan Krunz. 2001. Multi-constrained optimal path selection. In *Proceedings of the 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'01). IEEE*, Vol. 2. IEEE, 834–843.

[2] Mehmet H. Gunes, Sevcan Bilir, Kamil Sarac, and Turgay Korkmaz. 2007. A measurement study on overhead distribution of value-added Internet services. *Computer Networks* 51, 14 (2007), 4153–4173. DOI: http://dx.doi.org/10.1016/j.comnet.2007.05.001

[3] Bingdong Li, Jeff Springer, George Bebis, and Mehmet Hadi Gunes. 2013. A survey of network flow applications. *Journal of Network and Computer Applications* 36, 2 (2013), 567–581. DOI: http://dx.doi.org/10.1016/j.jnca.2012.12.020

[4] Sandeep Kumar Singh, Tamal Das, and Admela Jukan. 2015. A survey on Internet multipath routing and provisioning. *IEEE Communications Surveys and Tutorials* 17, 4 (2015), 2157–2175.

[5] Jiayue He and Jennifer Rexford. 2008. Toward Internet-wide multipath routing. *IEEE Network* 22, 2 (2008).

[6] Hasan T. Karaoglu, Murat Yuksel, and Mehmet H. Gunes. 2011. On the scalability of path exploration using opportunistic path-vector routing. In *Proceedings of the 2011 IEEE International Conference on Communications (ICC'11).* IEEE, 1–5.

[7] Hasan T. Karaoglu, Mehmet Burak Akgun, Mehmet Hadi Gunes, and Murat Yuksel. 2012. Multi path considerations for anonymized routing: Challenges and opportunities. In *Proceedings of the 2012 5th International Conference on New Technologies, Mobility and Security (NTMS'12).* IEEE, 1–5.

[8] Zakaria Al-Qudah, Mohammad Alsarayreh, Ibrahim Jomhawy, and Michael Rabinovich. 2016. Internet path stability: Exploring the impact of MPLS deployment. In *Proceedings of the Global Communications Conference (GLOBECOM'16)*. IEEE, 1–7.

[9] Ahmet Soran, Murat Yuksel, and Mehmet Hadi Gunes. 2017. Multiple graph abstractions for parallel routing over virtual topologies. In *Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. IEEE, 904–909.

[10] Anthony Olufemi Adeyemi-Ejeye, Mohammed Alreshoodi, Laith Al-Jobouri, Martin Fleury, and John Woods. 2017. Packet loss visibility across SD, HD, 3D, and UHD video streams. *Journal of Visual Communication and Image Representation* 45 (2017), 95–106.

[11] Asif Ali Laghari, Hui He, and Muhammad Ibrahim Channa. 2018. Measuring effect of packet reordering on quality of experience (QoE) in video streaming. *3D Research* 9, 3 (2018), 30.

[12] Vern Paxson. 1997. End-to-end routing behavior in the Internet. *IEEE/ACM Transactions on Networking* 5, 5 (1997), 601–615.

[13] Stefan Savage, Andy Collins, Eric Hoffman, John Snell, and Thomas Anderson. 1999. The end-to-end effects of Internet path selection. In *ACM SIGCOMM Computer Communication Review*, Vol. 29. ACM, 289–299.

[14] Yaron Schwartz, Yuval Shavitt, and Udi Weinsberg. 2010. On the diversity, stability and symmetry of end-to-end Internet routes. In *Proceedings of the INFOCOM IEEE Conference on Computer Communications Workshops, 2010*. IEEE, 1–6.

[15] Dan Pei, Xiaoliang Zhao, Daniel Massey, and Lixia Zhang. 2004. A study of BGP path vector route looping behavior. In *Proceedings of the 24th International Conference on Distributed Computing Systems*. IEEE, 720–729.

[16] Feng Wang, Lixin Gao, Jia Wang, and Jian Qiu. 2005. On understanding of transient interdomain routing failures. In *Proceedings of the 13th IEEE International Conference on Network Protocols (ICNP'05)*. IEEE, 10 pp.

[17] Craig Labovitz, Abha Ahuja, Abhijit Bose, and Farnam Jahanian. 2001. Delayed Internet routing convergence. *IEEE/ACM Transactions on Networking (TON)* 9, 3 (2001), 293–306.

[18] Sharad Agarwal, Chen-Nee Chuah, Supratik Bhattacharyya, and Christophe Diot. 2004. The impact of BGP dynamics on intra-domain traffic. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 32. ACM, 319–330.

[19] Feng Wang, Zhuoqing Morley Mao, Jia Wang, Lixin Gao, and Randy Bush. 2006. A measurement study on the impact of routing events on end-to-end Internet path performance. In *ACM SIGCOMM Computer Communication Review*, Vol. 36. ACM, 375–386.

[20] Forough Golkar, Thomas Dreibholz, and Amund Kvalbein. 2014. Measuring and comparing Internet path stability in IPv4 and IPv6. In *Proceedings of the International Conference and Workshop on the Network of the Future (NOF'14)*. IEEE, 1–5.

[21] Nick Feamster, David G. Andersen, Hari Balakrishnan, and M. Frans Kaashoek. 2003. Measuring the effects of Internet path faults on reactive routing. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 31. ACM, 126–137.

[22] Matthew Roughan, Tim Griffin, Morley Mao, Albert Greenberg, and Brian Freeman. 2004. Combining routing and traffic data for detection of IP forwarding anomalies. *ACM SIGMETRICS Performance Evaluation Review* 32, 1 (2004), 416–417.

[23] Athina Markopoulou, Gianluca Iannaccone, Supratik Bhattacharyya, Chen-Nee Chuah, and Christophe Diot. 2004. Characterization of failures in an IP backbone. In *Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'04)*, Vol. 4. IEEE, 2307–2317.

[24] M. A. Canbaz, J. Thom, and M. H. Gunes. 2017. Comparative analysis of Internet topology data sets. In *Proceedings of the 2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS'17)*. 635–640. DOI: http://dx.doi.org/10.1109/INFCOMW.2017.8116451

[25] M. Abdullah Canbaz, Khalid Bakhshaliyev, and Mehmet Hadi Gunes. 2017. Analysis of path stability within autonomous systems. In *Proceedings of the 2017 IEEE International Workshop on Measurement and Networking (M&N'17)*. IEEE, 1–6.

[26] Ripe NCC. RIPE Atlas. Retrieved on June 23, 2019 from https://atlas.ripe.net/anchors/map/.

[27] Mehmet H. Gunes and Kamil Sarac. 2009. Resolving IP aliases in building traceroute-based Internet maps. *IEEE/ACM Transactions on Networking* 17, 6 (Dec. 2009), 1738–1751. DOI: http://dx.doi.org/10.1109/TNET.2009.2014227

[28] Ken Keys, Young Hyun, Matthew Luckie, and Kim Claffy. 2013. Internet-scale IPv4 alias resolution with MIDAR. *IEEE/ACM Transactions on Networking* 21, 2 (2013), 383–399.

[29] M. Gunes and K. Sarac. 2007. Importance of IP alias resolution in sampling Internet topologies. In *IEEE Global Internet (GI)*.

[30] Edeline Korian and Donnet Benoit. 2017. A first look at the prevalence and persistence of middleboxes in the wild. In *ITC29* (2017-09-04). http://orbi.ulg.ac.be/bitstream/2268/212334/3/PID4875431.pdfhttp://hdl.handle.net/2268/212334.

[31] Joel Sommers, Paul Barford, and Brian Eriksson. 2011. On the prevalence and characteristics of MPLS deployments in the open Internet. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*. ACM, 445–462.

[32] CAIDA. Archipelago (Ark) Measurement Infrastructure. http://www.caida.org/projects/ark/.

[33] Benoit Donnet, Matthew Luckie, Pascal Mérindol, and Jean-Jacques Pansiot. 2012. Revealing MPLS tunnels obscured from traceroute. *ACM SIGCOMM Computer Communication Review* 42, 2 (2012), 87–93.

[34] ANT Censuses of the Internet Address Space. https://ant.isi.edu/address/.

[35] Constantine Dovrolis, Krishna Gummadi, Aleksandar Kuzmanovic, and Sascha D Meinrath. 2010. Measurement lab: Overview and an invitation to the research community. *ACM SIGCOMM Computer Communication Review* 40, 3 (2010), 53–56.

[36] Muhammed Canbaz, Khalid Bakhshaliyev, Ahmet Aksoy, and Mehmet Gunes. IM - Internet Measurement Platform. https://im.cse.unr.edu/.

[37] Chiara Orsini, Alistair King, Danilo Giordano, Vasileios Giotsas, and Alberto Dainotti. 2016. BGPStream: A software framework for live and historical BGP data analysis. In *Proceedings of the 2016 Internet Measurement Conference*. ACM, 429–444.

[38] CIDR Reports. http://www.cidr-report.org/as2.0/.

[39] UCLA Internet Research Lab. http://irl.cs.ucla.edu/index.html.

[40] Team Cymru. IP to ASN mapping. http://www.team-cymru.com/IP-ASN-mapping.html.

[41] Brice Augustin, Xavier Cuvellier, Benjamin Orgogozo, Fabien Viger, Timur Friedman, Matthieu Latapy, Clémence Magnien, and Renata Teixeira. 2006. Avoiding traceroute anomalies with Paris traceroute. In *Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement*. ACM, 153–158.

[42] A. Lakhina, J. W. Byers, M. Crovella, and P. Xie. 2003. Sampling biases in IP topology measurements. In *Proceedings of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE INFOCOM'03) (IEEE Cat. No. 03CH37428)*, Vol. 1. 332–341 vol.1. DOI: http://dx.doi.org/10.1109/INFCOM.2003.1208685

[43] Srikanth Kandula and Ratul Mahajan. 2009. Sampling biases in network path measurements and what to do about it. In *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference (IMC'09)*. ACM, New York, NY, 156–169. DOI: http://dx.doi.org/10.1145/1644893.1644912

[44] Hakan Kardes, Mehmet Gunes, and Talha Oz. 2012. Cheleby: A subnet-level Internet topology mapping system. In *Proceedings of the 2012 4th International Conference on Communication Systems and Networks (COMSNETS'12)*. IEEE, 1–10.

[45] P. Barford, A. Bestavros, J. Byers, and M. Crovella. 2001. On the marginal utility of network topology measurements. In *ACM Internet Measurements Workshop*.

[46] Yuval Shavitt and Eran Shir. 2005. DIMES: Let the Internet measure itself. *SIGCOMM Computer Communication Review* 35, 5 (2005), 71–74. DOI: http://dx.doi.org/10.1145/1096536.1096546

[47] J. L. Guillaume and M. Latapy. 2005. Relevance of massively distributed explorations of the Internet topology: Simulation results. In *Proceedings of IEEE INFOCOM*.

[48] Yuval Shavitt and Udi Weinsberg. 2009. Quantifying the importance of vantage points distribution in Internet topology measurements. In *Proceedings of IEEE INFOCOM 2009*. IEEE, 792–800.

[49] Q. Chen, H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. 2002. The origin of power laws in Internet topologies revisited. In *Proceedings of IEEE INFOCOM*.

[50] Ka-Cheong Leung, Victor O. K. Li, and Daiqin Yang. 2007. An overview of packet reordering in transmission control protocol (TCP): Problems, solutions, and challenges. *IEEE Transactions on Parallel and Distributed Systems* 18, 4 (2007), 522–535.

[51] Bernard Fortz, Jennifer Rexford, and Mikkel Thorup. 2002. Traffic engineering with traditional IP routing protocols. *IEEE Communications Magazine* 40, 10 (2002), 118–124.

[52] How Does Load Balancing Work? Retrieved April 21, 2019 from http://www.cisco.com/warp/public/105/46.html. ([n.d.]).

[53] Configuring Load-Balance Per-Packet Action. Retrieved April 21, 2019 from https://www.juniper.net/documentation/en_US/junos/topics/usage-guidelines/policy-configuring-per-packet-load-balancing.html. JUNOS Policy Framework Configuration Guideline.

[54] Brice Augustin, Timur Friedman, and Renata Teixeira. 2011. Measuring multipath routing in the Internet. *IEEE/ACM Transactions on Networking (TON)* 19, 3 (2011), 830–840.

[55] R. Govindan and H. Tangmunarunkit. 2000. Heuristics for Internet map discovery. In *Proceedings of the 19th Annual Joint Conference of the IEEE Computer and Communications Societies Conference on Computer Communications (Cat. No. 00CH37064) (IEEE INFOCOM 2000)*, Vol. 3. 1371–1380. DOI: http://dx.doi.org/10.1109/INFCOM.2000.832534

[56] N. Spring, M. Dontcheva, M. Rodrig, and D. Wetherall. 2004. *How to Resolve IP Aliases*. Technical Report. University of Washington.

[57] Ken Keys. 2010. Internet-scale IP alias resolution techniques. *ACM SIGCOMM Computer Communication Review* 40, 1 (2010), 50–55.

[58] Mehmet H. Gunes and Kamil Sarac. 2007. Inferring subnets in router-level topology collection studies. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*. ACM, 203–208.

[59] M. B. Akgun and M. H. Gunes. 2011. Link-level network topology generation. In *Proceedings of the 2011 31st International Conference on Distributed Computing Systems Workshops*. 140–145. DOI: http://dx.doi.org/10.1109/ICDCSW.2011.25

[60] M. Tozal and Kamil Sarac. 2010. TraceNET: An Internet topology data collector. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*. ACM, 356–368.

[61] Jean-Francois Grailet, Fabien Tarissan, and Benoit Donnet. 2016. TreeNET: Discovering and connecting subnets. In *Proceedings of the 8th International Workshop on Traffic Monitoring and Analysis (TMA'16)*.

[62] Mehmet H. Gunes and Kamil Sarac. 2009. Analyzing router responsiveness to active measurement probes. In *International Conference on Passive and Active Network Measurement*. Springer, 23–32.

[63] B. Yao, R. Viswanathan, F. Chang, and D. Waddington. 2003. Topology inference in the presence of anonymous routers. In *IEEE INFOCOM*.

[64] Mehmet Hadi Gunes and Kamil Sarac. 2008. Resolving anonymous routers in Internet topology measurement studies. In *Proceedings of the 27th Conference on Computer Communications (INFOCOM 2008)*. IEEE, 1076–1084.

[65] Hakan Kardes, Mehmet Hadi Gunes, and Kamil Sarac. 2015. Graph based induction of unresponsive routers in Internet topologies. *Computer Networks* 81 (2015), 178–200.

[66] CAIDA. The Caida UCSD AS to organization mapping dataset. http://www.caida.org/data/as_organizations.xml.

[67] M. Engin Tozal and Kamil Sarac. 2011. Subnet level network topology mapping. In *Proceedings of the IEEE 30th International Performance Computing and Communications Conference (IPCCC'11)*. IEEE, 1–8.

[68] Romualdo Pastor-Satorras and Alessandro Vespignani. 2007. *Evolution and Structure of the Internet: A Statistical Physics Approach*. Cambridge University Press.

[69] Daniel Awduche, Joe Malcolm, Johnson Agogbua, Mike O'Dell, and Jim McManus. 1999. *RFC 2702: Requirements for Traffic Engineering Over MPLS*. Technical Report.

[70] Michael Scharf and Alan Ford. 2013. *RFC 6897: Multipath TCP (MPTCP) Application Interface Considerations*. Technical Report.

[71] Hyojoon Kim and Nick Feamster. 2013. Improving network management with software defined networking. *IEEE Communications Magazine* 51, 2 (2013), 114–119.

[72] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia. 2010. A view of cloud computing. *Communications of the ACM* 53, 4 (2010), 50–58.

[73] Mahadev Satyanarayanan. 2017. The emergence of edge computing. *Computer* 50, 1 (2017), 30–39.

[74] Craig Labovitz, G. Robert Malan, and Farnam Jahanian. 1998. Internet routing instability. *IEEE/ACM Transactions on Networking* 5 (1998), 515–528.

[75] Matthias Wübbeling, Till Elsner, and Michael Meier. 2014. Inter-AS routing anomalies: Improved detection and classification. In *Proceedings of the 2014 6th International Conference on Cyber Conflict (CyCon'14)*. IEEE, 223–238.

[76] Rahul Hiran, Niklas Carlsson, and Phillipa Gill. 2013. Characterizing large-scale routing anomalies: A case study of the China telecom incident. In *Proceedings of the International Conference on Passive and Active Network Measurement*. Springer, 229–238.

[77] Kevin Vermeulen, Stephen D. Strowes, Olivier Fourmaux, and Timur Friedman. 2018. Multilevel MDA-Lite Paris traceroute. In *Proceedings of the Internet Measurement Conference 2018*. ACM, 29–42.

[78] Urs Hengartner, Sue Moon, Richard Mortier, and Christophe Diot. 2002. Detection and analysis of routing loops in packet traces. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet Measurement*. ACM, 107–112.

Received September 2018; revised April 2019; accepted May 2019