Geography and Routing in the Internet

ABDULLAH YASIN NUR, University of Louisiana at Lafayette, USA
MEHMET ENGIN TOZAL, University of Louisiana at Lafayette, USA

The Internet is a network of networks consisting of tens of thousands of Autonomous Systems (ASes). These ASes connect to each other in different forms to enable the "global" Internet communication. In this study, we investigate the geographical characteristics of the visible Internet as well as examine the relation between geography and intra-AS and inter-AS routing policies. We show that the ingress-to-egress subpaths have lower circuitousness compared to the end-to-end paths. Our findings not only demonstrate the efficient backbone infrastructures and routing schemes deployed by ASes but also show the consequences of economical incentives on the adoption of inter-AS paths. We present and examine the existence of a strong correlation between the geographical distance and round trip delay time as well as the lack of a correlation between the geographical distance and hop length in the Internet. We investigate the relation between the geographical distance and intra-AS routing policies by employing cross-AS (X-AS) Internet topology maps. Our results show that more than two thirds of the intra-AS subpaths are congruent with the shortest geographical distance whether or not geographical distance is employed as a custom parameter in routing decisions. Our results provide new insights into the relations between geography and Internet routing which allow the network researchers and practitioners to improve their networking infrastructures, reevaluate their routing policies, deploy geography-aware network overlays and develop more realistic network simulation processes.

CCS Concepts: • **Networks → Network performance analysis**;

Additional Key Words and Phrases: Internet Routing, Geography, End-to-End Paths, Ingress-to-Egress Subpaths

1 INTRODUCTION

The Internet is a network of networks consisting of tens of thousands of Autonomous Systems (ASes). An Autonomous System (AS) in the Internet is a group of networks administered by one or more network operators under a well defined routing policy. The ASes in the Internet are connected to each other in different forms to achieve the "global" Internet communication. Individual users, small businesses and stub ASes located at the edge of the Internet participate in the global infrastructure by means of other ASes called Internet Service Providers (ISPs). Typically, ISPs are business entities providing Internet access service to their customers while getting the same service from one or more upstream ISPs. At the core of the Internet, a small number of ISPs peer with each other through settlement-free interconnections to attain the global communication infrastructure.

An IP (Internet Protocol) packet destined to a geographically distant host in the Internet passes through several routers until it reaches the destination. The routers handling the packet on the way typically belong to different ASes. While the AS accommodating the originating host might be a small, campus or building size AS, the ASes which transit the packet on the way might span over larger regions such as multiple states or countries. The IP packet enters into a transit AS through

Authors' addresses:
Abdullah Yasin Nur, University of Louisiana at Lafayette, Lafayette, LA, 70504, USA, ayasinnur@louisiana.edu.
Mehmet Engin Tozal, University of Louisiana at Lafayette, Lafayette, LA, 70504, USA, metozal@louisiana.edu.

2374-0353/ART $15.00
https://doi.org/10.1145/3239162

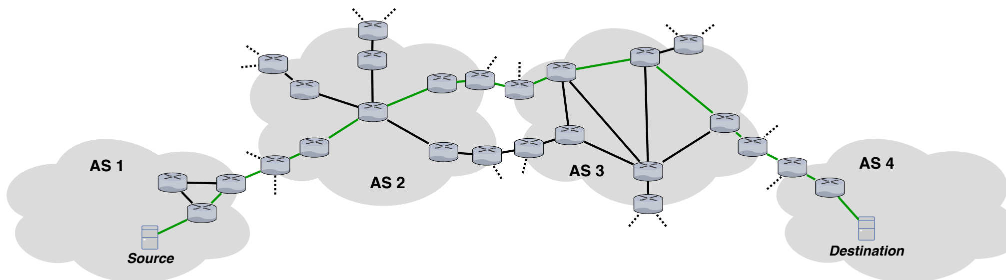


Fig. 1. An example path trace from a source host in $AS1$ to a destination host in $AS4$

an ingress border router, gets routed within the backbone infrastructure of the AS and finally gets relayed to the next AS via an egress border router toward the destination. Figure 1 presents an overly simplified topology and a path trace from a source host in $AS1$ to a destination host in $AS4$. The ASes are illustrated as clouds and the route taken by a packet is shown in green. The transit ASes, $AS2$ and $AS3$, in the figure receive the packet through one of their border routers (ingress), route the packet within their networks and then relay the packet to the next AS through one of their border routers (egress) toward the destination. The ASes providing Internet access service, i.e., ISPs, generally organize their routers by way of Point of Presences (PoPs) and connect these PoPs to each other to build their backbone infrastructures and extend their services geographically [1]. Moreover, ISPs typically exchange traffic and routing information through physical connections at different colocation centers and Internet eXchange Points (IXPs) connecting colocation centers at metropolitan areas. Please note that PoPs, colocation centers and IXPs are not presented in Figure 1 for the sake of brevity.

ASes in the Internet use the Border Gateway Protocol (BGP), a path vector protocol, to exchange *inter-AS* routing and reachability information. On the other hand, they independently design their internal networking infrastructures and employ their own *intra-AS* routing policies. Intermediate System-Intermediate System (IS-IS) and Open Shortest Path First (OSPF) are two widely adopted intra-AS routing protocols in the Internet. Both IS-IS and OSPF protocols are link-state routing protocols allowing routers to exchange intra-AS topology information with their neighbors. The routers use the topology information to compute the “preferred” routes via the shortest path algorithm based on configurable link weights. Several metrics such as bandwidth, latency, traffic load or their combinations can be used as link weights. There are two popular intra-AS routing schemes in the Internet: *hot potato* and *cold potato* routing. In hot potato routing, an AS relays a packet to the next AS as quickly as possible to reduce the cost of transit traffic in its network. On the other hand in cold potato routing, an AS carries the packet in its network to the nearest hand over point toward the destination to ensure quality of service (QoS) requirements.

The overall Internet performance depends on the efficiency of the underlying ASes’ backbone infrastructures; their intra-AS routing decisions as well as their inter-AS routing policies. In this empirical work, we investigate the relations between the geography and routing in the visible Internet. We analyze the geographical properties of the Internet routing in three different but complementary ways: circuitousness of the Internet, the relation between geographical distance and end-to-end path metrics as well as the relation between geographical distance and internal routing decisions. First, we revisit the fundamental work presented in 2002 on path circuitousness [2] and apply it to end-to-end paths as well as ingress-to-egress subpaths covering a large portion of the current, global Internet infrastructure. We define ingress-to-egress subpaths as end-to-end path fragments belonging to the individual ASes.

Subramanian et al. [2] define *distance ratio* as a metric to analyze the relative circuitousness of IP paths in the Internet. Distance ratio is defined as the ratio of the total geographical distances between consecutive IP addresses on a path to the direct geographical distance between the source and the destination IP addresses. Circuitousness simply corresponds to the degree of geographical indirectness in end-to-end paths in the Internet. Different from more recent works, [3] and [4], we characterize the ingress-to-egress subpath circuitousness along with the end-to-end path circuitousness. Our experimental results show that, the end-to-end path distance ratio is exactly one for 9% of the paths in the Internet. Moreover, it changes from one to two in 41% of the paths, two to three in 20% of the paths and three to ten in the remaining 30% of the paths. On the other hand, the distance ratio is exactly one for 72% of the ingress-to-egress subpaths. It changes from one to three for 18% of the subpaths and from three to 7.6 for the remaining 10% of the subpaths. Our findings show that the ingress-to-egress subpaths have lower circuitousness compared to the end-to-end paths. This observation is an indication of efficient backbone infrastructures and routing schemes deployed by ASes. That is, ASes adopt direct or slightly indirect connectivity schemes among the ingress/egress points within their networks. Existence of direct links between ingress and egress points reduces routing delays and improves service quality. However, the end-to-end paths spanning multiple ASes display higher circuitousness, because inter-AS paths prioritize economical incentives rather than performance measures.

Next, we investigate the relations between geography and two end-to-end path metrics reflecting the cumulative effect of multiple routing decisions: round trip time (RTT) delay and hop length. Our results show that there is a linear correlation between geographical distance and RTT delay in end-to-end paths. On the other hand, we do not observe any strong correlation between geographical distance and hop length.

Lastly, we analyze the relations between the geographical distance and intra-AS routing policies. We employed cross-AS (X-AS) Internet topology maps [1] to identify the demarcation points of ASes which correspond to ingress and egress facilities. The X-AS topology map of the Internet captures the ASes in the Internet; multiple (or parallel) connections among the ASes; as well as the abstractions of the endpoints of the connections called cross-border interface (X-BI) knots (or nodes). X-BI knots roughly correspond to AS presences at colocation facilities with some exceptions [1]. We mapped the path traces in our dataset to their respective X-BI knots and marked the ingress and egress X-BI knots where the paths enter into an AS and leave the AS. For each subpath between a particular ingress and egress X-BI knot we computed the geographical distances of the ingress knots to all potential egress knots to investigate if the chosen egress knot is geographically the closest one among all potential egress knots. Our experimental results show that more than two thirds of the intra-AS subpaths are congruent with the shortest geographical distance whether or not the geographical distance is employed as a custom parameter in routing decisions. Moreover, in 13.26% of the cases RTT delay characterizes the intra-AS routing decisions despite that it is not in agreement with the shortest geographical distance. Similarly, in only 3.34% of the cases hop length characterizes the intra-AS routing decisions while being inconsistent with the shortest geographical distance.

Understanding the geographical characteristics of the Internet infrastructure is crucial. Specifically, it allows us to (i) utilize resourceful paths during or after natural disasters [5]; (ii) improve the inter-domain routing processes [6]; (iii) deploy geography-aware network overlays for efficient multimedia communications [7]; (iv) predict path latency for service improvement in the Internet [8]; and (v) develop more realistic Internet topology generation tools [9]. In this study, we investigate the interplay between geography and inter-AS and intra-AS routing in the Internet. Our results provide new insights into the impact of geography on the Internet routing which may contribute the enhancement of the current Internet as well as the design of the future Internet. We show that while ASes have smaller circuitousness in their own networks, the business relations among the ASes cause an increase in the circuitousness of the end-to-end paths. This result may help ISPs to reevaluate their business relations to achieve an improved routing efficiency. We show how ASes route packets within their own networks before relaying them to the next ASes which may help us with improving path predictions in the Internet and developing more representative network simulation processes.

The rest of the paper is organized as follows. Section 2 presents the background. We introduce the details of our approach and the experimental results along with their discussions in Section 3. Section 4 presents the related work. Finally, Section 5 concludes the paper.

2 BACKGROUND

In this section we briefly explain inter-AS and intra-AS routing. In addition, we provide an overview of the complimentary tools and datasets used in this work including traceroute, IP-to-AS mapping and IP-to-Geolocation mapping.

2.1 Inter-AS Routing

Border Gateway Protocol (BGP) is a path vector protocol for exchanging IP prefix reachability information among the ASes in the Internet. The reachability information consists of an IP address prefix, one or more AS paths to reach the prefix and a set of AS path attributes. An AS willing to deliver traffic to an IP address prefix originates a BGP advertisement declaring the prefix and its AS number as the path to the prefix. This advertisement is sent to the neighboring ASes of the originating AS. The neighboring ASes independently decide to employ, drop, aggregate and/or re-advertise the new IP address prefix with or without modifying any AS path attributes. A neighboring AS willing to transit traffic destined to the new IP address prefix re-advertises the prefix to its own neighbors by prepending its AS number to the path. The neighbors of a re-advertising AS repeat the same process. Hereby, multiple AS paths to an IP address prefix gets disseminated in the Internet through neighbor-to-neighbor BGP advertisements while each AS independently selects/employs/re-advertises a path(s) toward the prefix. ASes employ the inter-AS paths according to the business relations, e.g., customer, provider and peer, among them [10]. Typically, an AS re-advertises the reachability information learned from its customers to its other customers, providers and peers; the reachability information learned from its peers and providers to its customers in the routing preference of customer, peer and provider [11, 12]. The reason is that an AS charges its customers; neutralizes its peers and pays to its providers for the traffic exchanged between them.

2.2 Intra-AS Routing

Intermediate System-Intermediate System (IS-IS) protocol distributes routing information within an AS. IS-IS is a link-state routing protocol where the routers exchange topology information throughout the AS network by flooding (or multicast). Each router originates information about its immediate links and the costs of these links [13]. Routers use this information to compute the “preferred” routes via the shortest path algorithm. ASes have full control on their networks. Therefore, they can employ different cost metrics including link bandwidth, average latency and traffic load and compute the paths with respect to those metrics. As a result, network administrators can control and divert the traffic by changing the link attributes.

Open Shortest Path First (OSPF) is another popular intra-AS routing protocol in the Internet. Similar to IS-IS, OSPF is a link-state routing protocol and uses flooding (or multicast) to exchange and distribute topology information in the network. Both protocols use shortest paths algorithm to compute the best path within an AS. While IS-IS natively runs on Layer-2 (Data Link Layer), OSPF runs on Layer 3 (Network Layer). By design, IS-IS easily extends to support new capabilities while the OSPF protocol is comparably more rigid. IS-IS protocol is favored by large ISPs because of its simplicity, extensibility and scalability.

Hot and Cold Potato routing are two popular intra-AS routing schemes employed by ISPs [14]. In Hot-potato routing, an AS delivers a packet to the next AS as early as possible. The behavior is explained by ISPs not willing to keep their networks busy with the transit traffic. Therefore, hot-potato routing aims to deliver packets to the immediate egress point depending on the destination. The term "immediate" can be defined in terms of hop length, geographic distance, link states (such as bandwidth of the links or traffic load on the links), a combination of them, or any other metric defined by ISPs' network administrators. In Cold-potato routing, ISPs deliver packets to the next AS via the nearest egress point toward the destination. When an AS relays a packet to another AS, it loses all the control over the packet. Even if an AS provides the best service, the end user might not experience it, because an intermediate AS on the path may provide a subpar service. Therefore, ISPs use cold-potato routing to carry the traffic inside their networks to the nearest delivery point toward the destination to ensure QoS requirements.

2.3 Complementary Tools and Datasets

Traceroute is a network debugging tool revealing the path trace between two hosts in the Internet. A path trace corresponds to the interfaces (represented by IP addresses) that belong to different routers on a path from a source to a destination. Traceroute is prevalently used for path analysis in the Internet [15, 16]. In this work, we use iPlane [17] and Caida's Ark [18] traceroute datasets collected from multiple vantage points. To achieve a wide coverage, iPlane collects path traces from several vantage points toward all globally routable prefixes in BGP snapshots. Similarly, Caida collects path traces from several vantage points toward all /24 prefixes in routable IPv4 address space. IPlane and Caida datasets contain more than 11 million (11,266,865) and 19 million (19,275,509) path traces, respectively. The datasets include path traces along with the associated round trip times (RTT) and the time-to-live (TTL) values. Altogether, the datasets consist of 46,198 ASes along with 195,362 unique AS pair connections and almost 24 million (23,938,838) unique path traces covering a significant portion of the Internet. Note that we consider only the unique path traces in this study.

IP-to-AS mapping is the process of finding the AS number that originates a given IP address. Different data sources such as Regional Internet registries (RIRs), DNS and BGP routing tables are used for IP-to-AS mapping. In this work, we use Caida's prefix-to-AS mappings dataset [19] which is derived from the BGP advertisements obtained from UO RouteViews project. The dataset contains IP prefixes, prefix lengths and the corresponding AS numbers.

IP-to-Geolocation mapping is used to find the geographical location of a given IP address. Typically, researchers use delay and topology measurements to map IP addresses to latitudes and longitudes. Moreover, some ISPs encode geographic information in their DNS naming conventions. In this work, we use UNDNS tool [16] which maps the DNS information of IP addresses to a location as well as three commercial geolocation databases: DB-IP [20], Maxmind [21] and IP2Location [22]. We improved UNDNS tool key dataset by extending the coverage of the DNS names. Due to the diversity of the naming conventions, UNDNS falls short to map all IP addresses to their geographic locations. We use commercial geolocation databases along with the majority rule to resolve the geographic locations of the unresolved IP addresses.

3 ANALYSES OF GEOGRAPHY AND ROUTING

In this section, we introduce our methodology and present our empirical results. We investigate the geolocation properties of the Internet routing via three different experimental setups. First, we analyze the end-to-end path and ingress-to-egress subpath circuitousness in the Internet. Next,

Algorithm 1 Distance ratio calculation

```
Input: traceList                                  ▷ List of consecutive IP addresses
Output: r                                         ▷ Distance ratio in Equation 1
 1: procedure CALCULATEDR(traceList)
 2:     firstIP = traceList.get(0)
 3:     lastIP = traceList.get(traceList.length − 1)
 4:     distFL = geoDistance(firstIP, lastIP)
 5:     totalDist = 0
 6:     for i = 0 to traceList.length − 1 do
 7:         IP1 = traceList.get(i)
 8:         IP2 = traceList.get(i + 1)
 9:         totalDist = totalDist + geoDistance(IP1, IP2)
10:     end for
11:     if distFL = 0 then
12:         return 1
13:     else
14:         return totalDist / distFL
15:     end if
16: end procedure
```

we examine the relations between geolocation and two end-to-end path metrics reflecting the combined effects of routing decisions at multiple ASes: RTT delay and hop length. Lastly, we investigate the relations between geographical distance and early exit points using cross-AS (X-AS) Internet topology maps.

3.1 Circuitousness

A packet sent from one host to another host typically passes through several routers belonging to different ASes until it reaches its destination. Both inter-AS and intra-AS routing policies affect the path taken by a packet in terms of geography. To illustrate, two interfaces appearing in a path trace in different cities, e.g., A and E, might have a direct connection between their accommodating routers, i.e., $A \rightarrow E$. Alternatively, they might have a circuitous route, e.g. $A \rightarrow C \rightarrow E$, passing through an intermediate city, i.e., C. The route can be even more circuitous by passing through multiple intermediate cities such as $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E$. Subramanian et al. [2] define *distance ratio* as a metric to analyze the relative circuitousness of end-to-end paths in the Internet. Distance ratio is defined as the ratio of the total geographical distances between consecutive IP addresses in a path trace to the direct geographical distance between the source and the destination IP addresses. Circuitousness simply corresponds to the degree of geographical indirectness in end-to-end paths in the Internet. In this study we use circuitousness to examine not only the inter-AS, end-to-end paths but also intra-AS, ingress-to-egress subpaths covering a large portion of the global Internet infrastructure. Formally, let $d(I_i, I_j)$ be the geographical distance between interfaces i and j. Equation 1 presents the distance ratio, r, of a path where n is the number of the interfaces (represented by IP addresses) in the path. Theoretically, the distance ratio assumes a value from one to infinity such that one means no geographical circuitousness.

$$r = \frac{1}{d(I_1, I_n)} \sum_{i=1}^{n-1} d(I_i, I_{i+1}) \tag{1}$$

Algorithm 1 presents the procedure to calculate the distance ratio for a path trace. The algorithm expects a traceList, a list of consecutive IP addresses in a path trace, as input. Lines 2 and 3 get the first and the last IP addresses in the list, respectively. Line 4 calls the method "geoDistance" which maps the input IP addresses to their geographic locations and returns the distance between them.

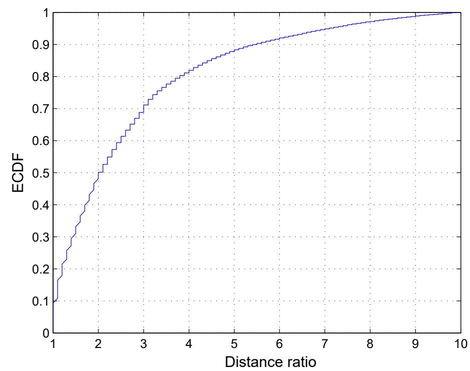


Fig. 2. End-to-end path circuitousness

We use Haversine formula which calculates the *great-circle distance* between two points on the surface of a sphere as suggested in [3]. Haversine formula requires two latitude and longitude pairs to compute the distance between them as shown in Equation 2.

$$\begin{aligned} a &= \sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right) \\ c &= 2arcsin(\sqrt{a}) \\ d &= Rc \end{aligned} \tag{2}$$

In the Equation, ϕ is latitude in radians, λ is longitude in radians, R is Earth's radius and d is the great-circle distance between (ϕ_1, λ_1) and (ϕ_2, λ_2) pairs. The distance corresponds to the shortest distance between two points on the surface of a sphere where the ellipsoidal effects of the earth are ignored. Lines 6 to 10 in Algorithm 1 calculate the total traversed distance between the first and the last IP addresses by summing up the geographic distances between the consecutive IP addresses. Lines 11 and 12 correct the divide-by-zero case when the first and the last IP addresses are located in the same city. Naturally, the circuitousness in this case would be 1. Lines 13 to 15 compute the distance ratio based on Equation 1. Note that we do not consider IP alias resolution, since all interfaces of a router ideally map to the same location.

3.1.1 End-to-End Path Circuitousness. We apply Algorithm 1 to each unique path trace in our traceroute dataset to find the end-to-end path circuitousness. Figure 2 shows the Empirical Cumulative Distribution Function (ECDF) of the distance ratios of 23,938,838 unique end-to-end path traces out of 30,542,374 in total. In addition, Table 1 shows the minimum, first quartile, second quartile (median), third quartile, maximum, mean and standard deviation of the empirical distribution.

Table 1. Summary statistics for end-to-end path distance ratio

Q_0	Q_1	Q_2	Q_3	Q_4	*Mean*	*StdDev*
1	1.3	2	3.3	9.9	2.71	1.89

In Figure 2, the distance ratio equals to one for 9% (2,155,535) of the paths in our dataset. It changes from one to two in 41% (9,815,546) of the paths, two to three in 20% (4,785,151) of the paths

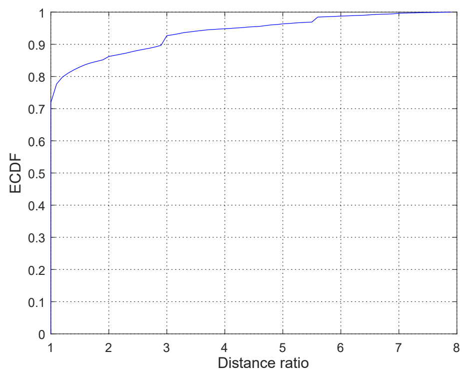


Fig. 3. Ingress-to-egress subpath circuitousness

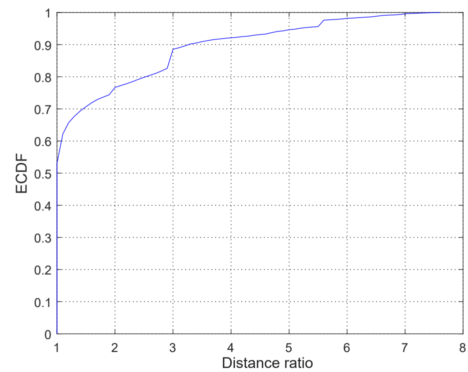


Fig. 4. Ingress-to-egress subpath circuitousness excluding stub ASes

and three to ten in the remaining 30% (7,182,606) of the paths. We analyze the paths having high distance ratios further to understand the reason behind the higher circuitousness. We observe that shorter hop length paths have smaller circuitousness compared to the longer hop length paths for a given distance. Naturally, an increasing hop length contributes the circuitousness of the paths when the geographical distance is fixed. Contrary to the intuition, geographically closer source and destination pairs have higher circuitousness compared to the longer pairs. This observation implies that longer distances have more direct subpaths compared to the sorter ones. In fact, our analysis show that long cables, e.g., continental, trans-Atlantic and trans-Pacific, do not contribute much into circuitousness while crossing very long geographical distances. Finally, the number of ASes on a path increases the circuitousness of the path. This result supports the reported inter-AS path inflation in [23].

After analyzing the end-to-end path circuitousness in the Internet, we analyze the effects of intra-AS routing decisions on circuitousness. In order to study the relation between circuitousness and intra-AS routing, we first parse all path traces in our dataset into subpaths where each subpath is confined to a single AS. For each path trace, we demarcate the subpaths by marking the ingress IP addresses where the path trace enters into an AS and the egress IP address where the path trace leaves the AS. Next, we discard the duplicate ingress-to-egress subpaths to eliminate the potential bias that may be introduced by the frequently observed subpaths. Then, we use Algorithm 1 to compute the distance ratios of all subpaths generated in the previous step. Note that the distance ratio of an end-to-end path is not equal to the sum of the distance ratios of its ingress-to-egress subpaths, because the first term in Equation 1 delineates, r, with respect to the bounding interfaces.

3.1.2 Ingress-to-Egress Subpath Circuitousness. Figure 3 shows the Empirical Cumulative Distribution Function (ECDF) of the distance ratios of 7,988,764 unique ingress-to-egress subpaths in our dataset. In this experiment we excluded 8% (68,788) of the IP addresses because IP-to-AS mapping failed to return their AS numbers. We observed that some ingress-to-egress subpaths frequently appear in end-to-end paths. Table 2 shows the minimum, first quartile, second quartile (median), third quartile, maximum, mean and standard deviation of the distribution.

In Figure 3, the distance ratio equals to one for 72% (5,752,286) of the ingress-to-egress subpaths. It changes between one and three for 18% (1,437,348) of the subpaths and between three and 7.6 for the remaining 10% (799,129) of the subpaths. An immediate comparison between Figure 3 and Figure 2 shows that circuitousness is greatly higher in end-to-end paths compared to the ingress-to-egress

Table 2. Summary statistics for ingress-to-egress subpath distance ratio

Q_0	Q_1	Q_2	Q_3	Q_4	*Mean*	*StdDev*
1	1	1	1.1	7.6	1.44	1.11

subpaths. However, when we analyze the ingress-to-egress subpaths having distance ratio one, we observe that more than half, 53.3% (3,066,087), of these subpaths belong to the stub ASes. Remember that stub ASes are those ASes which do not transit any traffic belonging to other ASes. They are virtually at the edge of the Internet and participate the Internet via upstream ISPs. The subpaths belonging to the stub ASes correspond to the 38.4% of all subpaths in our dataset. In fact, 80.7% (37,281) of the ASes in our dataset have only one ingress/egress point based in the same city. This finding is congruent with the previous studies demonstrating that the majority of the ASes in the Internet are stub ASes spanning a campus or a building [24]. Note that the 80.7% stub ASes contribute 38.4% of all subpaths because stub ASes typically have a single entry/exit point while the transit ASes, typically ISPs, span geographically and have multiple entry/exit points with rich connectivity among them.

When we remove the ingress-to-egress subpaths belonging to the stub ASes, the subpaths having distance ratio one drops to 54.4%. Figure 4 shows the same information in Figure 3 after removing the stub ASes. In addition, Table 3 shows the minimum, first quartile, second quartile (median), third quartile, maximum, mean and standard deviation of the distribution.

Table 3. Summary statistics for ingress-to-egress subpath distance ratio excluding stub ASes

Q_0	Q_1	Q_2	Q_3	Q_4	*Mean*	*StdDev*
1	1	1	1.99	7.6	1.73	1.32

Figure 4 demonstrates that ISPs take direct or linear paths to transit packets within their networks. The figure shows that the ingress-to-egress subpaths have lower circuitousness compared to the end-to-end paths. This observation is an indication of efficient backbone infrastructures and routing schemes deployed by ISPs. That is, ISPs adopt direct or slightly indirect connectivity schemes among the ingress-egress points within their networks. Especially, the existence of direct links between ingress and egress points implies reduced routing delays and improved service quality.

The results in this section can be employed by ISPs to compare their circuitousness against the overall circuitousness distribution in the Internet and improve their backbone infrastructures accordingly. In addition, smaller ISPs may evaluate the circuitousness of their candidate upstream providers against the overall circuitousness distribution in the Internet and use it as a parameter in their decision making processes.

3.2 Geographical Distance Analyses of End-to-End Paths

In our second experiment, we investigate the relation between geographical distance and two end-to-end path parameters: Round Trip Time (RTT) delay and hop length. Both parameters reflect the combined effect of routing decisions that take place at multiple ASes on a path.

3.2.1 Geographical Distance vs. RTT Delay. To analyze the relation between geographical distance and RTT delay we compute the great-circle distance and RTT delay between the first and the last IP addresses of each path trace. Figure 5 shows the conditional RTT delay distribution with respect to the geographical distance as a heat map. In the figure, x-axis shows the geographical distance between source and destination IP addresses on a path. y-axis presents the RTT delay

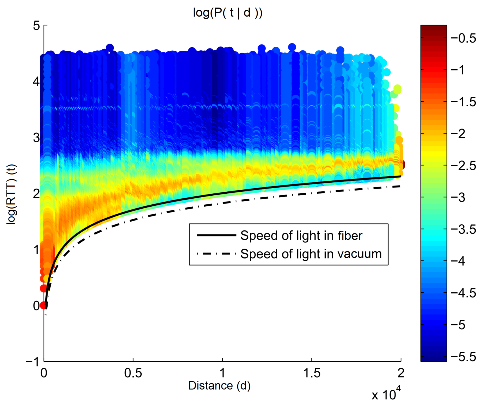


Fig. 5. Geographical distance and RTT delay in end-to-end paths

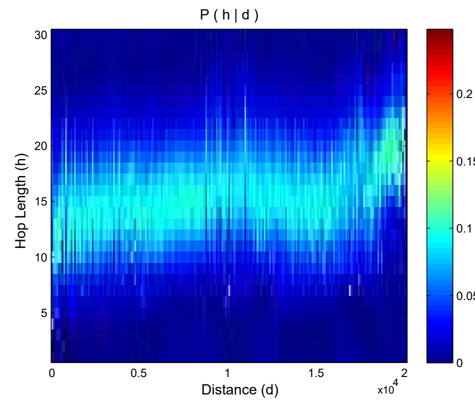
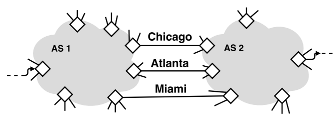


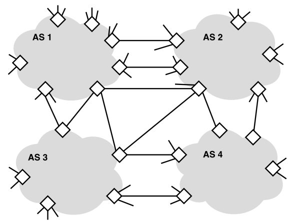
Fig. 6. Geographical distance and hop length in end-to-end paths

between the source and destination IP addresses. Lastly, color encoded z-axis shows the conditional distribution $P(t|d)$ where t is the RTT delay and d is the geographical distance. In the figure, y and z axes are given in logarithmic scales due to the positive skewness toward the larger values in the data. The conditional RTT delay increases from cold colors toward the warm colors. We superimpose the speed of light in vacuum (300 km/ms) and the speed of light in optical fiber (200 km/ms) in the figure. Assuming that the fastest way to transmit data is using fiber optic cables, the speed of light in fiber presents the lower limit of the RTT distribution. We excluded 0.72% (172,359) of the path traces due to the erroneous RTT delays exhibiting below the theoretical limit values. Figure 5 clearly shows that geographical distance and RTT delay have a positive linear correlation shown in logarithmic scale. In fact, the Pearson product correlation coefficient is 0.8997. Moreover, the variability in RTT delay gets smaller as the distance increases. Our analysis show that the change in RTT delay variance is related to the existence or lack of multiple paths or subpaths between geographical locations. We observe that there are multiple paths and subpaths belonging to different ASes in shorter geographical distances. Therefore, the paths demonstrate a greater variability. On the other hand, as the geographical distance of the end-to-end paths increase, the variability in RTT delay decreases. Because, those paths have to take long regional and/or trans-Atlantic/Pacific subpaths where the alternatives are not as many as the shorter distances.

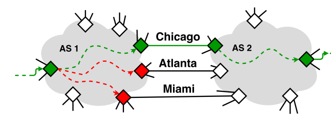
3.2.2 Geographical Distance vs. Interface Hop Length. Huffaker et al. [25] present that the geographical distance has a small effect on the hop distance in Asia-Pacific region. In this experiment, we investigate the same relation using a large scale sample of the Internet consisting of almost 24 million unique path traces. Figure 6 presents the conditional hop length distribution with respect to the geographical distance as a heat map. In the figure, x-axis shows the geographical distance between end-to-end paths. y-axis presents the hop length between source and destination IP addresses. Color encoded z-axis shows the conditional distribution $P(h|d)$ where h is the hop length and d is the geographical distance. Similar to [25], our results do not show a strong correlation between the distance and the hop length for the Internet. In the figure, we observe that the majority of the paths have a hop length changing between 10 to 25 for varying geographical distances. Despite a small decrease between 10,000 and 15,000 km, the hop lengths do not extremely change. Because, long cables, e.g., continental, trans-Atlantic and trans-Pacific, do not contribute much into hop lengths while significantly contributing the geographical distances. Since long distance cabling



(a) X-BI knots connected at multiple cities



(b) Partial X-AS map



(c) Egress X-BI knots of *AS*1 toward *AS*2

Fig. 7. Illustrations depicting partial X-AS Internet topology maps with multiple X-BI knots

does not always introduce additional hops into a path, hop length and geographical distance do not demonstrate a strong correlation. Yet, after roughly 16,000 kilometers we observe a small surge in hop length. Because following the long submarine or regional cables, packets traverse additional routers to reach their destinations. We also note that the longest great-circle distance, 19,991 km, observed in our dataset is from a host in Hamilton, New Zealand (AS681 - University of Waikato) to a host in Córdoba, Spain (AS34977 - Procono SA) with 19 hops.

The results in this section can be employed by the research community to develop more accurate network delay estimation techniques and enhance synthetic network topology generators.

3.3 Geographical Characterization of Intra-AS Routing

In this part, we analyze the relation between geographical distance and intra-AS routing schemes. In hot potato routing, ASes relay packets to the next ASes as quickly as possible to reduce the cost of transit traffic in their networks. That is, when an AS receives a transit packet at an ingress point, it delivers the packet to the immediate egress point toward the destination. Remember that the immediate egress point is not necessarily the geographically closest egress point. In cold potato routing an AS carries a packet in its network until the nearest egress point toward the destination to ensure QoS requirements.

In the following, we investigate the influence of intra-AS routing decisions on the internal geographical distance in terms of the choices of egress points for given ingress points. Note that a transit AS usually has more than one, geographically distributed connection points with a next AS to relay a packet towards its destination. For example, Figure 7a illustrates two ASes having three connections at different colocation centers, in different cities, i.e., Chicago, Atlanta and Miami. In the figure, the transit AS, *AS*1, can relay a packet to the next AS, *AS*2, via any of their joint connection points. The inter-AS routing decisions, i.e., the choice of the next ASes to relay packets toward their destinations, are governed by the business relations among the ASes. On the other hand, the intra-AS routing decisions, i.e., the choice of the next routers (via links) including the egress points to relay packets, are based on performance metrics such as link bandwidth, average latency and traffic load with one caveat. Sometimes, the next AS requests the transit AS to prefer a particular egress point over the others by explicitly setting the BGP Multi-Exit Discriminator (MED) attribute to the lowest value. Please note that in general, ASes do not publicly share their routing parameters and business relations due to security concerns. Considering the illustration given in Figure 7a, *AS*2 may request *AS*1 to use the connection point in Miami over the others for a particular IP prefix by setting the related MED value in its BGP advertisement. In this part, we try to find an answer to the following question: “If there are multiple egress points between a transit AS and the next AS toward a destination and the transit AS chooses one of these egress points, is the choice related to the geographical distance between the ingress and the egress points? If not, is the choice related to any other parameters?"

To answer the research question, we employ cross-AS (X-AS) Internet topology mapping [1] to identify the demarcation points of ASes. The X-AS topology map of the Internet captures the ASes in the Internet as well as multiple (or parallel) connections among the ASes. X-AS topology mapping systematically exploits multiple data sources (BGP advertisements, traceroute paths, geolocation databases and DNS datasets) to construct the X-AS topology map of the Internet. In the implementation a *cross border interface* (X-BI) is defined as an interface (or an IP address) belonging to a border/edge router of an AS. An *X-BI knot* (or *node*) corresponds to a set of X-BIs of a particular AS, located in the same facility regardless of the border router(s) accommodating the X-BIs. Both BGP ("NEXT HOP" field) and traceroute datasets are used to identify X-BIs. IP address to AS mapping tool is used to extract X-BIs that appear in path traces where the paths switch from one AS to another. Next, X-BIs are clustered according to their geolocations to build X-BI knots. X-BI knots represent the demarcation points of ASes and abstract the endpoints of parallel connections between the ASes. Note that ASes have physical presences at colocation centers and the connections between the ASes are established at these centers via cross connects, e.g., optical fiber cable, or switches, e.g., Internet eXchange Point (IXP) access switch, between the edge routers. After X-BI clustering, the topology map consists of a set of isolated, unconnected X-BI knots which roughly correspond to AS presences at colocation facilities with some exceptions [1]. Lastly, traceroute, BGP and IP geolocation datasets are used to discover the cross connections between the X-BI knots. The finished X-AS map, $X = (K, C)$, consists of a set of AS-annotated X-BI knots, K, and a set of X-BI connections, C. Figure 7b illustrates a partial X-AS Internet topology map where the four ASes are connected to each other via multiple X-BI knots at multiple locations. In the figure, clouds represent the ASes, rhombuses show X-BI knots, full links present the connections between the X-BI knots and partial links illustrates connections to not-shown ASes. We refer the interested readers to [1] for the details of X-AS topology mapping and the multigraph abstraction of the Internet.

Next, we map the path traces in our dataset to their respective X-BI knots and mark the ingress X-BI knots and egress X-BI knots where the paths enter into an AS and leave the AS, respectively. Finally, for each subpath between a particular ingress and egress X-BI knot we compute the geographical distances of the ingress X-BI knot to all "potential" egress X-BI knots to examine if the chosen egress X-BI knot is geographically the closest one among all potential egress X-BI knots. Figure 7c illustrates a packet relay scenario from the transit AS, $AS1$, to the next AS, $AS2$, toward its destination. The packet enters into $AS1$ at the left-most ingress X-BI knot and leaves the AS at the egress X-BI knot in Chicago while the potential egress X-BI knots are the ones in Atlanta and Miami. The dashed lines represent the intra-AS paths that may involve multiple routers such that the chosen path is shown in green and the alternatives are shown in red. We compute the geographical distance from the ingress X-BI knot to all plausible egress X-BI knots. In addition, we compute the average RTT delay and average hop length distances between the ingress X-BI knot and all plausible egress X-BI knots to examine the RTT delays and hop lengths as well. In this experiment, we only consider the subpaths having at least two potential egress X-BI knots including the chosen one. Note that the subpaths having only one egress X-BI knot are ignored because there is no alternative egress X-BI knots for comparison.

We analyze 274,195 unique ingress-egress X-BI knot pairs representing the subpaths in this experiment. We observe that in 76.91% (210,896) of the cases, the geographical distance between the ingress X-BI knot and the chosen egress X-BI knot is the shortest among all potential egress X-BI knots. Even though geographical distance is not an explicit decision parameter in intra-AS routing, 76.91% cases are congruent with the geographical distances. When we parse these ingress-egress X-BI knot pairs having the shortest geographical distance further, we observe that 9.22% (25,283) simultaneously satisfy the shortest RTT delay but not the shortest hop length. Moreover, we observe that 12.67% (34,747) satisfy the shortest hop length but not the smallest RTT delay. Lastly, the remaining 55.02% (150,866) satisfy the shortest geographical distance, RTT delay and hop length, altogether.

We observe that in the remaining 23.08% (63,299) of the subpaths, the geographical distances between the ingress X-BI knots and the chosen egress X-BI knots are not the shortest compared to the potential egress X-BI knots. We study these subpaths further to reveal the relation between the ingress X-BI knots and the chosen egress X-BI knots. In 3.34% (9,153) of the cases we do not observe any subpath between the ingress X-BI knots and the geographically closest egress X-BI knots in our dataset. Therefore, we are not able to examine these cases further. The hop lengths between the ingress X-BI knots and the chosen egress X-BI knots are the shortest ones compared to the potential egress X-BI knots in 3.28% (8,992) of the cases where the geographical distances are not necessarily the shortest. On the other hand in 13.26% (36,384) of the cases, the RTT delays are the smallest while the geographical distances are not necessarily the shortest. Lastly, in the remaining 3.2% (8770) of the cases neither geographical distances nor the RTT delays and the hop lengths represent the characteristic of the subpaths between the ingress and the chosen egress X-BI knots. These cases can be attributed to cold potato routing, MED attributes in BGP routing or another metric used in intra-AS routing decisions.

Our experimental results in this section show that more than two thirds of the intra-AS subpaths are congruent with the geographical distance whether or not the geographical distance is employed as a custom parameter in routing decisions. Moreover, in 13.26% of the cases RTT delay characterizes the intra-AS routing decisions despite that it is not in agreement with the shortest geographical distance. Similarly, in only 3.28% of the cases hop length characterizes the intra-AS routing decisions while being inconsistent with the shortest geographical distance.

The results in this section can be employed by the research community in developing better path prediction techniques. Path predictions often play an important role in the analyses of the resilience and the robustness of the Internet.

4 RELATED WORK

Over the past decade, several studies examining the relations between various metrics and routing schemes have been introduced. Wang et al. [26] study the effects of routing on end-to-end Internet path performance. They show that routing policies and iBGP configurations on end-to-end paths are the major causes of performance degradation in the Internet. Tangmunarunkit et al. [27] analyze the hop length inflation due to Internet routing policies. Dongyu et al. [28] focus on the performance analysis of peer-to-peer networks. They analyze the behavior of BitTorrent-like networks and study their scalability, performance and efficiency. Labovitz et al. [29] examine the reliability of the Internet paths in case of path failures. They show that routing changes due to a path failures cause high BGP convergence delays. Teixeira et al. [14] analyze the impact of intra-AS routing on inter-AS routing. They show that geographical distance has an implicit influence on BGP path selection through hot potato routing.

Importance of geolocation in the Internet has also been discussed in several works [30]. Researchers have proposed methods to resolve geolocation of IP addresses by taking advantage of the delay and network topology information [31, 32]. Subramanian et al. [2] use approximately 84 thousand end-to-end paths to analyze the geographic properties of the Internet routing and intra-AS routing in the US. They define *distance ratio* as a metric to analyze the relative circuitousness of IP paths in the Internet. Their results show that paths in the Internet can be highly circuitous. Following studies, [3] and [4], improve the accuracy of their approach by using alternative geolocation techniques and traceroute datasets. In addition, the authors investigate the symmetry of end-to-end paths by analyzing the forward and backward routes between hosts [3]. Landa et al. [33] analyze the effect of geographical distance on round trip delay times in the Internet. They show a strong correlation between geographical distance and RTT delay in the Internet. Spring et al. [23] discusses the effect of inter-AS routing policies on path inflation. They argue that one of the main reason of inter-domain path inflation is the lack of BGP policy controls. Geographically Informed Inter-Domain Routing (GIRO) [6] uses geographic information to assist routing decisions. The authors propose a new IP address allocation technique by integrating the geographic information into the IP address structure. They show that GIRO can significantly reduce the geographic distance of the existing BGP paths.

In this study, we investigate the interplay between geography and routing in the Internet. We differentiate the inter-AS and intra-AS routing schemes by analyzing both end-to-end path and ingress-to-egress subpath circuitousness covering a large portion of the global Internet infrastructure. We demonstrate the relation between RTT delay, hop length and geographical distance by analyzing a large set of end-to-end paths in the Internet. Finally, we exploit X-AS Internet topology maps to demonstrate the impact of intra-AS routing decisions on the internal geographical distances.

5 CONCLUSIONS

The Internet is one of the largest man-made complex systems. Understanding Internet's structural and operational characteristics is essential for network diagnostics, protocol development, architectural improvement and network simulation advancement, among others. In this study, we focus on the geographical characteristics of the visible Internet and its relations with intra-AS and inter-AS routing policies.

We first analyzed the end-to-end path and ingress-to-egress subpath circuitousness in the Internet. Our findings show that the ingress-to-egress subpaths have lower circuitousness compared to the end-to-end paths. This observation is an indication of efficient backbone infrastructures and routing schemes deployed by ASes. That is, ASes adopt direct or slightly indirect connectivity schemes among the ingress-egress points within their networks. However, the end-to-end paths spanning multiple ASes display higher circuitousness, because inter-AS paths prioritize economical incentives rather than performance measures. Next, we examined the relation between geographical distance and RTT delay and hop length. Our experimental results show that there is a linear correlation between geographical distance and RTT delay. The results also demonstrate the lack of a strong correlation between geographical distance and hop length. Finally, we investigated the relation between the geographical distance and intra-AS routing policies by employing X-AS Internet topology maps. Our experimental results show that more than two thirds of the intra-AS subpaths are congruent with the geographical distance whether or not the geographical distance is employed as a custom parameter in routing decisions.

ACKNOWLEDGMENTS

We are grateful to DB-IP for providing their commercial geolocation database. We are also thankful to CAIDA, iPlane, Maxmind and IP2Location projects for publicly sharing their datasets.

REFERENCES

[1] Abdullah Yasin Nur and Mehmet Engin Tozal. 2018. Cross-AS (X-AS) Internet Topology Mapping. *Elsevier, Computer Networks* 132 (2018), 53 – 67.

[2] Lakshminarayanan Subramanian, Venkata N Padmanabhan, and Randy H Katz. 2002. Geographic Properties of Internet Routing. In *USENIX*. Berkeley, CA, USA.

[3] Shiva Prasad Kasiviswanathan, Stephan Eidenbenz, and Guanhua Yan. 2011. Geography-based analysis of the internet infrastructure. In *INFOCOM*. IEEE, Shanghai, China.

[4] Péter Mátray, Péter Hága, Sándor Laki, István Csabai, and Gábor Vattay. 2011. On the network geography of the internet. In *INFOCOM*. IEEE, Shanghai, China.
[5] Jian Wu, Ying Zhang, Z Morley Mao, and Kang G Shin. 2007. Internet routing resilience to failures: analysis and implications. In *ACM CoNEXT*. 25.
[6] Ricardo Oliveira, Mohit Lad, Beichuan Zhang, and Lixia Zhang. 2007. Geographically informed inter-domain routing. In *International Conference on Network Protocols*. IEEE, 103–112.
[7] Ahmed Elmokashfi, Eugene Myakotnykh, Jan Marius Evang, Amund Kvalbein, and Tarik Cicic. 2013. Geography matters: building an efficient transport network for a better video conferencing experience. In *ACM conference on Emerging networking experiments and technologies*. ACM, 369–380.
[8] Harsha V Madhyastha, Thomas Anderson, Arvind Krishnamurthy, Neil Spring, and Arun Venkataramani. 2006. A structural approach to latency prediction. In *IMC*. ACM, 99–104.
[9] Bruno Quoitin, Virginie Van den Schrieck, Pierre François, and Olivier Bonaventure. 2009. IGen: Generation of router-level Internet topologies through network design heuristics. In *International Teletraffic Congress (ITC)*. IEEE.
[10] Lixin Gao. 2001. On inferring autonomous system relationships in the internet. *IEEE/ACM Transactions on Networking* 9, 6 (Dec 2001), 733–745.
[11] Mehmet Engin Tozal. 2016. Enumerating Single Destination, Policy-Preferred Paths in AS-level Internet Topology Maps. In *IEEE Sarnoff Symposium*. Newark, NJ, USA.
[12] Mehmet Engin Tozal. 2018. Policy-Preferred Paths in AS-level Internet Internet Topology Graphs. *Theory and Applications of Graphs* 5, 3 (Mar 2018), 32. Issue 1.
[13] CiscoPress. 2001. Dynamic Routing Protocols. Retrieved June 28, 2018 from http://www.ciscopress.com/articles/article.asp?p=24090&seqNum=4
[14] Renata Teixeira, Aman Shaikh, Tim Griffin, and Jennifer Rexford. 2004. Dynamics of hot-potato routing in IP networks. In *ACM SIGMETRICS Performance Evaluation Review*, Vol. 32. ACM, 307–319.
[15] Yuval Shavitt and Noa Zilberman. 2010. A structural approach for PoP geo-location. In *INFOCOM Conference on Computer Communications Workshops*. IEEE, San Diego, CA, USA.
[16] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson. 2004. Measuring ISP topologies with Rocketfuel. *IEEE/ACM Transactions On Networking* 12, 1 (Feb 2004), 2–16.
[17] iPlane. 2016. Traceroute Dataset. Retrieved August 16, 2016 from http://iplane.cs.washington.edu/
[18] Caida. 2016. Traceroute Dataset. Retrieved August 16, 2016 from http://www.caida.org/data/active/ipv4_routed_24_topology_dataset.xml
[19] Routeviews. 2016. Prefix to AS Dataset. Retrieved August 16, 2016 from http://www.caida.org/data/routing/routeviews-prefix2as.xml
[20] DB-IP. 2016. Geolocation Database. Retrieved August, 2016 from https://www.db-ip.com/
[21] Maxmind. 2016. Geolocation Database. Retrieved August, 2016 from http://www.maxmind.com/
[22] IP2Location. 2016. Lite Geolocation Database. Retrieved August, 2016 from http://lite.ip2location.com/
[23] N. Spring, R. Mahajan, and T. Anderson. 2003. The Causes of Path Inflation. In *SIGCOMM*. Karlsruhe, Germany.
[24] Mehmet Engin Tozal. 2016. The Internet: A System of Interconnected Autonomous Systems. In *IEEE Systems Conference*. Orlando, FL, USA.
[25] Bradley Huffaker, Marina Fomenkov, David Moore, and Evi Nemeth. 2000. Measurements of the Internet Topology in the Asia-Pacific Region. In *Proceedings of INET*.
[26] 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. 375–386.
[27] Hongsuda Tangmunarunkit, Ramesh Govindan, Scott Shenker, and Deborah Estrin. 2001. The impact of routing policy on internet paths. In *INFOCOM*, Vol. 2. IEEE, 736–742.
[28] Dongyu Qiu and Rayadurgam Srikant. 2004. Modeling and performance analysis of BitTorrent-like peer-to-peer networks. In *ACM SIGCOMM computer communication review*, Vol. 34. 367–378.
[29] 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.
[30] Richard Barnes, James Winterbottom, and Martin Dawson. 2011. Internet geolocation and location-based services. *IEEE Communications Magazine* 49, 4 (2011).
[31] Ethan Katz-Bassett, John P John, Arvind Krishnamurthy, David Wetherall, Thomas Anderson, and Yatin Chawathe. 2006. Towards IP geolocation using delay and topology measurements. In *ACM IMC*.
[32] Yong Wang, Daniel Burgener, Marcel Flores, Aleksandar Kuzmanovic, and Cheng Huang. 2011. Towards Street-Level Client-Independent IP Geolocation. In *NSDI*, Vol. 11. 27–27.
[33] Raúl Landa, Joao Taveira Araújo, Richard G Clegg, Eleni Mykoniati, David Griffin, and Miguel Rio. 2013. The large-scale geography of internet round trip times. In *IFIP Networking Conference*. IEEE, Brooklyn, NY, US.