# Internet PoP Level Maps

Yuval Shavitt and Noa Zilberman

School of Electrical Engineering, Tel-Aviv University, Israel
{shavitt,noa}@eng.tau.ac.il

**Abstract.** Inferring the Internet Point of Presence (PoP) level maps is gaining interest due to its importance to many areas, e.g., for tracking and studying properties of the Internet. In this chapter we survey research towards the generation of PoP level maps. The chapter introduces different approaches to automatically classify IP addresses to PoPs and discusses their strengths and weaknesses. Special attention is devoted to the challenge of validating the generated PoP maps in the absence of ground truth. The chapter next describes general IP geolocation techniques, points out weaknesses in geolocation databases, as well as, in constraint-based approaches, and concentrates on PoPs geolocation techniques, discussing validation and lack of ground truth availability. The third part of the chapter describes how to generate maps with PoP-to-PoP connectivity and analyzes some of their properties. At the end of the chapter, some applications of PoP level maps, such as Internet distance maps, evolution models and homeland security are introduced and discussed.

## 1 Introduction

The study of the Internet topology attracted a great deal of work over the years. A good survey of these efforts can be found in the "Internet topology discovery" chapter of this book, as well as in an earlier survey by Donnet and Friedman [9]. Internet topology maps are used for a vast number of applications, such as building models of the Internet [37], studying the robustness of the network [10], network management [47] and improving routing protocols design [40]. There are several levels Internet maps are presented at, each level of abstraction is suitable for studying different aspects of the network. The most detailed level is the IP level, which represents separately each and every entity connected to the network. Many projects map the Internet at the IP level, such as Skitter [23], RIPE NCC's Test Traffic Measurement [14], iPlane [31], DIMES [8], Ark [24], and more. This level is far too detailed to suit practical purposes, and the large number of entities makes it very hard to handle. One level above the IP level is the router level, aggregating multiple IP interface addresses to a router, using alias resolution, as done by projects such as Mercator [15], MIDAR [27], Ally [49], and RadarGun [4]. While being less detailed than the IP level, this level of aggregation is still highly detailed and difficult to handle. The most coarse level is the Autonomous System (AS) level. It is most commonly used

E. Biersack et al. (Eds.): Data Traffic Monitoring and Analysis, LNCS 7754, pp. 82–103, 2013.

to draw Internet maps, as it is relatively small (tens of thousands of ASes) and therefore relatively easy to handle: there is only one node for every AS, and may have only one edge between every pair of ASes. There are different methods to discover the Internet's AS-level topology, from using traceroutes, as done in Ark, iPlane and DIMES, through BGP announcements, as done by Routeviews [52] to Internet Routing Registries (IRR) [34]. One limitation of using AS information for Internet mapping is that AS sizes may differ by orders of magnitude. While a large AS can span an entire continent, and a small one can serve a small community, yet both seem identical at the AS level map.

An interim level between the AS and the router graphs is the PoP level. Service providers tend to place multiple routers in a single location called a Point of Presence (PoP), which serves a certain geographical area. A PoP is defined as a group of routers which belong to the same AS and are physically located at the same building or campus.

Figure 1 demonstrates the Internet aggregation levels. The figure presents for clarity only the AS, PoP, and router levels. Every AS, marked by a large circle, is made of a network of routers, marked by small light gray circles. The routers may be part of a PoP (colored dark gray), or reside outside of a PoP. A router which is not part of a PoP will still be connected to other routers, eventually connecting to a PoP. The points of presence are connected to other PoPs within the same AS as well as to PoPs outside their AS, thus creating AS level connectivity.

The technological nature of PoPs varies between service providers as well as within the same network. Some PoPs operate entirely on the IP level, while other PoPs employ MPLS and VPLS switching. In many cases, service provider mix switching and routing within the same PoP, combining both MPLS and IP. In more rare cases, in Optical Transport Networks, the PoP may only serve as a channel based cross connect. A good example of this mix is shown in CenturyLink's network [5]: In some cities, such as Atlanta, Los Angeles, and New York City, both IP and MPLS/VPLS are used. In other cities, such as Sacramento, Duluth, and Cambridge, MA, there is an IP PoP, while in cities such as New Orleans, San Antonio, and San Diego only MPLS/VPLS is used. Additional examples can be found in the TeliaSonera network map [51] and XO network map [54] [1]. Service provider also tend to distinguish between different types of PoPs, often referring to the hierarchy in the network, e.g., access or backbone PoP [5] or to the area it covers, e.g., a metro PoP [54]. A declining trend is to refer to PoPs by their capacity, such as GigaPoP [25] or TeraPoP[2] [39].

When studying the entire network, and not only specific ISPs, PoP maps give a better level of aggregation than router level maps with a minimal loss of information. PoP level graphs provide the ability to examine the size of each AS network by the number of physical co-locations and their connectivity instead of by the number of its routers and IP links. Points of presence can be annotated with geographical location, as well as information about the size of the PoP.

[1] This information was also confirmed with a large networking equipment provider.
[2] As called by Qwest, before Qwest was acquired by CenturyLink.

PoP maps can also preserve routing information by annotating links connecting Pops that belong to different ASes with the type of relationship (ToR). Thus, using PoP level graphs it is possible to detect important nodes of the network and understand network dynamics as well as many more applications.

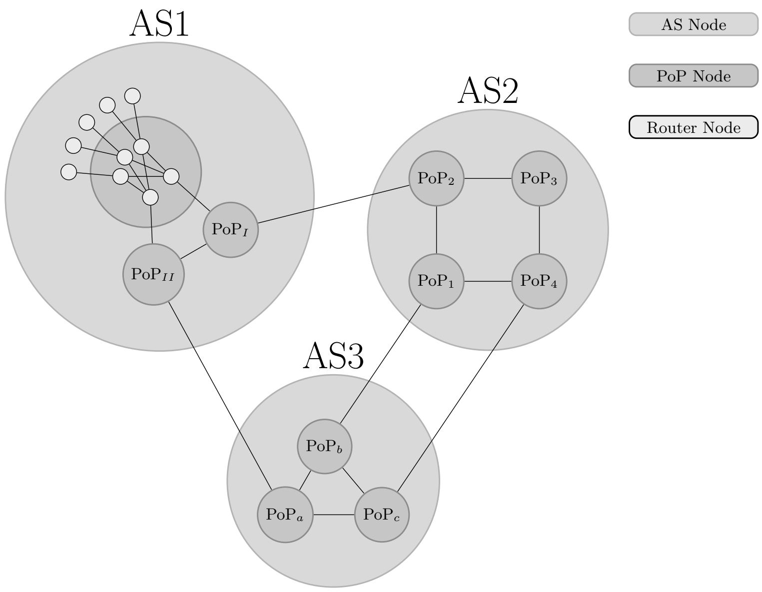


**Fig. 1.** The Internet's Levels of Aggregation

This chapter surveys the study of Internet PoP level maps, proving an overview of all works done so far in this field. The chapter is organized as follows: Section 2 discusses classification of IP addresses into PoPs and surveys existing works in this field. Section 3 describes some methods for assigning a location to points of presence. The generation of PoP-level connectivity maps is presented in Section 4 and some analysis of the maps is provided. The validation efforts of PoP level maps is surveyed in Section 5. In Section 6 we discuss applications of the PoP level graphs by various disciplines. Last, section 7 concludes this chapter.

## 2 IP Classification into PoPs

The first attempts to explore the PoP level graph were done by Andersen *et al.* [3] and Spring *et al.* [49]. Spring *et al.* [49] tried to infer ISP topologies both on the router and the PoP level. The focus of their contribution was in alias resolution and router identification based on in-order IP identifiers and introducing Rocketfuel, their mapping engine. The PoP resolution was entirely DNS based. To this end, they inferred ISP naming convention. For example, s1-bb11-nyc-3.0.sprintlink.net is indicated to be a Sprint backbone (bb11) router in

New York City (nyc). The naming convention was deduced from the list of router names they gathered during the alias resolution and router identification stage with some city names taken from [36]. For routers with no DNS names or where the names lacked location information, the locations of neighbor routers were used. The generated PoP map did not distinguish between backbone network nodes, data centers, or private peering points.

Ten ISPs were tested by Spring *et al.* and the number of PoPs discovered per ISP ranged from 11 (AS4755, VSNL India) to 122 (AS2914, Verio US). The PoPs' analysis showed that the designs of PoPs were relatively similar: generic PoPs are built from a few routers connected in a mesh while in large PoPs the access routers connect one or more routers from a neighboring domain and to two backbone routers for redundancy. The backbone routers connect both to routers within the same PoP as well as to routers in other PoPs that connect to the ISP's backbone. The result showed that small PoPs had for redundancy two backbone PoPs, but in large PoPs with 20 routers or more, the number of backbone routers varied significantly, from two to thirty.

Andersen *et al.* [3] used passive monitoring of BGP messages to cluster IP prefixes into PoPs: In the preprocessing stage, BGP messages are grouped into time intervals of $I$ seconds and massive updates due to session resets are filtered. The clustering stage is based on a distance metric, which is a function that determines how closely two items are. The distance metric used is the correlation coefficient between every pair of BGP update vectors. $u_p^{(t)}$ denotes the update vector for each prefix p:

$$u_p^{(t)} = \begin{cases} 1 \; if \; p \; updated \; during \; interval \; t \\ 0 \qquad\qquad otherwise \end{cases} \tag{1}$$

$C(p1, p2)$ is the correlation coefficient between two prefixes, with $\overline{u_p}$ being the average of $u_p^{(t)}$ and $\sigma_p$ its variance.

$$C(p1, p2) = \frac{\frac{1}{n}\sum_{t=1}^{n}(u_{p1}{}^{(t)} - \overline{u_{p1}})(u_{p2}{}^{(t)} - \overline{u_{p2}})}{\sqrt{\sigma_{p1}^2}\sqrt{\sigma_{p2}^2}} \tag{2}$$

A Single-linkage clustering algorithm [57] is applied for grouping prefixes. Using the distance metric presented by Equation 2, each pairwise distance between two prefixes is computed and prefixes with time window of 30 seconds are grouped.

Andersen *et al.* used BGP updates from two upstream feeds: a commercial feed via Genuity (AS 1), and an Internet2 feed via the Northeast Exchange (AS 10578). Due to their configuration, only the best route to every prefix was recorded, thus some paths were omitted from their dataset. The clustering was conducted on 2338 prefixes announced by UUNET (AS 701) and 1310 prefixes announced by AT&T (AS7018) and ended up with 6 clusters in UUNET and 5 clusters in AT&T, with the number of clusters strongly dependent on the number of pairwise comparisons during the clustering phase. The analysis observed the effect of the number of matches on the number of clusters and their accuracy.

The validation was conducted based on three methods: IP address similarity (the number of IP addresses that separate two prefixes), Ratio of shared to unshared traceroute path length (in hops) and DNS-based PoP comparison. The last means that they extracted a router location from the ISP's naming convention, managing to assign 97% of UUNET hops and somewhat less for AT&T. Their results showed that correlation-based clustering grouped the UUNET prefixes into about 1200 clusters while with over 95% PoP-level accuracy as well as 900 clusters in AT&T with 97% accuracy. The accuracy is defined as a match between the naming conventions. The concluding observation is that clusters that are announced and withdrawn together tend to be located at the same PoP.

The iPlane project [32,31] generates PoP level maps based on the Rocketfuel's approach, with several improvements: First, they determine the DNS names assigned to network interfaces, using two data sources: Rocketfuel's undns utility [49] and data from the Sarangworld project [1]. DNS alone is not enough, as some interfaces have no DNS names, others have no rules to infer their DNS name and for some interfaces may be misnamed, thus incorrect locations can be inferred [56]. For the last, interfaced are probed using ICMP ECHO packets and interfaces where the RTT is smaller than expected are filtered. The main new contribution in this work is an algorithm that clusters router interfaces based on their responses when probed from a large number of vantage points. iPlane estimates the number of hops on the reverse path back from a router to the vantage point, guessing the initial TTL value used by the router. The assumption is that routers in the same AS and geographically co-located take the same reverse path back to the vantage point from which they were probed, while routers that are not co-located will not display similar reverse path. iPlane detects about 135K PoPs, about 56K of them in singleton clusters, meaning a single router in a cluster. We discuss further the iPlane project in Section 4.

The PoP extraction algorithm proposed by DIMES [12] is based on the fact that in most cases [16,42] the PoP consists of two or more backbone/core routers and a number of client/access routers. The client/access routers are connected redundantly to more than one core router, while core routers are connected to the core network of the ISP. The algorithm takes a structural approach and looks for bi-partite subgraphs[3] with certain weight constraints in the IP interface graph of an AS; no aliasing to routers is needed. The bi-partites serve as cores of the PoPs and are extended with other nearby interfaces.

The algorithm works on the Interface graph of each ISP separatly. It starts by removing all edges with delay higher than $PD_{max\_th}$, PoP maximal diameter threshold, and edges with number of measurements below $PM_{min\_th}$, the PoP measurements threshold. As a result the ISP interface graph is partitioned to several components, each is a candidate to become one or more PoPs. Next, the algorithm looks at the bi-partites in each component and uses the rich connectivity there between the sources (parents) and destinations (children) to check for node colocation based on link delays between the groups. If *parent* and *child*

[3] A bipartite graph is a graph whose vertices can be divided into two disjoint sets $U$ and $V$ such that every edge connects a vertex in $U$ to one in $V$.

groups are connected, then the weighted distance between the groups is calculated (If they are connected, by definition more than one edge connects the two groups); if it is smaller than a certain threshold the pair of groups is declared as part of the same PoP. Last, a unification of loosely connected parts of the PoP is conducted. For this end, the algorithm looks for connected components (PoP candidates) that are connected by links whose median distance is very short (below $PD_{max\_th}$).

The number of PoPs discovered by DIMES is in the range of 4000 to 5000.

Yoshida *et al.* [55] mapped PoP-paths in Japan using thirteen dedicated measurement nodes and measuring the delay between these nodes. They tried to map the core network delay, derived from the end-to-end delays and access delays and their corresponding PoP level paths, using a set of delay equations:

$$delay(src, dst) = ad_{src} + ad_{dst} + \sum_{p,q \epsilon N} x_{p,q} \times cd_{p,q} + E_{src,dst} \quad (3)$$

In the equation, $N$ denotes a set of candidate PoP locations of a measured ISP; $p$ and $q$ satisfy $p, q \epsilon N$; $ad_{src}$ and $ad_{dst}$ denote the access delay at the source and the destination; $cd_{p,q}$ denotes a core delay between $p$ and $q$; $E_{src,dst}$ is the measurement error of the delay; $x_{p,q} = 1$ if a direct path between $p$ and $q$ exists and the path is used to connect between src and dst, otherwise $x_{p,q} = 0$. $delay(src, dst)$, $ad_{src}$ and $ad_{dst}$ are measurable through end-to-end measurements and $cd_{p,q}$ can be derived leveraging the distance between $p$ and $q$. To solve the equation, several restrictions are applied. One of the assumptions used is that the network connections are deployed along other infrastructure services, such as railroads and expressways.

The work distinguished between five types of networks, differing by the way the backbone routers are structured and by the way layer two is used. For example, is layer three being used in every location in the network, or are layer three routers being used only in highly populated cities.

A different approach to PoP level maps is presented by Rasti *et al.* [41]. They term an eyeball AS as an individual Autonomous Systems that directly provides service to end-users and use the eyeball ASes to estimate the PoP-level footprint. The basic assumption is that each AS must have a PoP in areas it has a high concentration of customers. Therefore, the AS eyeball offers a view of that AS's PoP-level infrastructure, referred to as PoP-level footprint. The algorithm begins by gathering a large number of end-user IP addresses, collected by crawling P2P applications. The users are then mapped to cities using geolocation services (discussed in Section 3) and are grouped to ASes based on Routeview's BGP tables [52]. Given the locations of the users, the geographical regions where the AS offers service to end-users is inferred using KDE (Kernel Density Estimation). To extract the PoP footprint, local maxima D(i) are detected in the density function, with the highest peak denoted by $D_{max}$. PoPs are indicated by any peak D(i) that is within a given range from $D_{max}$, meaning $D(i) > \alpha \times D_{max}$, with $\alpha$ set to 0.01. The work focused on 672 ASes and found an average of 13.6 PoPs per AS when using 40km range as the kernel function bandwidth.

To conclude, there are several different approaches to the classification of IP addresses into PoPs. Yet, grouping the IP addresses into PoPs is just the first stage of generating PoP level maps, as we discuss in the following sections.

## 3 Geolocation of PoPs

An important feature of PoP level maps is the ability to assign a geographical location to PoPs. The assignment is done using geolocation mapping services, providing longitude and latitude or a city and a country per IP address. Geolocation mapping services can be divided to several groups. For mobile devices, GPS is the most common approach to locate a device. A second group of geolocation mapping services is geolocation databases, holding a table mapping every IP address to its geographical location. Geolocation databases range from free services to services that cost tens of thousands of dollars a year. The most basic services use DNS resolution as the basis for the database [49], while others use proprietary means such as random forest classifier rules, hand-labeled hostnames [2], user's information provided by partners [7], and more.

Another group of geolocation mapping services is based on network measurements. IP2Geo [36] was one of the first to suggest a measurement-based approach to approximate the geographical distance of network hosts. A more mature approach is constraint based geolocation [19], using several delay constraints to infer the location of a network host by a triangulation-like method. Later works, such as Octant [53] used a geometric approach to localize nodes within a 22 miles radius. Katz-Bassett *et al.* [26] suggested topology based geolocation using link delay to improve the location of nodes. Yoshida *et al.* [55] used end-to-end communication delay measurements to infer PoP level topology between thirteen cities in Japan. Eriksson *et al.* [11] applied a learning based approach to improve geolocation. They reduced IP geolocation to a machine learning classification problem and used Naive Bayes framework to increase geolocation accuracy.

One online geolocation service that allows querying specific IP addresses is Spotter, which is based on a work by Laki *et al.* [29]. Spotter uses a probabilistic geolocation approach, which is based on a statistical analysis of the relationship between network delay and geographic distance. To approximate the location of a target, Spotter measures propagation delays from landmarks to the target, and then converts the delays into geographic distances based on a delay-distance model. The resulting set of distance constraints is used to determine the targets estimated location with a triangulation-like method.

Not many works have focused on the accuracy of geolocation databases, but those who did showed them to be inaccurate: In 2008, Siwpersad *et al.* [48] examined the accuracy of Maxmind [33] and IP2Location [20]. They assessed their resolution and confidence area and concluded that their resolution is too coarse and that active measurements provide a more accurate alternative. Gueye *et al.* [17] investigated the imprecision of relying on the location of blocks of IP addresses to locate Internet hosts and concluded that geolocation information coming from exhaustive tabulation may contain an implicit imprecision. Muir and Oorschot [35] conducted a survey of geolocation techniques used by geolocation databases and examined means for evasion/circumvention from a security standpoint. Shavitt and Zilberman [45] studied extensively seven geolocation services both on IP and PoP level. Their results show that the information in the databases may be largely biased at the ISP level: Using a small ground truth database provided by CAIDA of 25K addresses and described in [22], they found that some of the databases place all the IP addresses of a certain ISP in a single location, typically the ISP headquarters' city; while for other ISPs correct location is provided. Additionally, correlation was found between databases: some databases, such as Maxmind [33] and IP2Location [20] have an extremely small median distance between an IP address' geolocations, below 10km, while for other databases, such as GeoBytes [13] and HostIP.info [21] the median distance may be above 500km. The differences between databases may be in the range of countries: in one example case, a 10-nodes PoP was located by some databases in Singapore, by others in Australia, while two more databases pointed to China and Afghanistan, as shown in Figure 2. Constraint based approaches are many times no better than geolocation databases. They inherently have an inaccuracy in the range of tens to hundreds of kilometers [45], and strongly depend on the location of the vantage points. A non optimal location of vantage points may lead to an error in the range of hundreds of kilometers, and more. Poese *et al.* [38] studied five databases and showed that while on the country-level they are rather accurate, the databases are highly biased towards a few popular countries. Using ground truth information from one large European ISP and using DNS names as clues for two large other major ISPs, Poese *et al.* showed that the evaluated databases performed poorly on those ISPs.


**Fig. 2.** Mismatch Between Databases - An Example

Most of the PoP extraction algorithms described in Section 2 use a crude method of geolocation as the basis for their geolocation: DNS names. This is an easy to use method, leveraging the fact that the router's location is often written in the router's name used by the ISP. However, DNS suffers from several problems: many interfaces do not have a DNS name assigned to them, and incorrect locations are inferred when interfaces are misnamed [56]. In addition, rules for inferring the locations of all DNS names do not exist, and require some manual adjustments. DIMES' maps take an approach based on geolocation databases: it uses the geographic location of each of the IPs included in a PoP, as denoted by at least three geolocation databases (typically more) and take the median location. A range of error, indicating the radius within 50% of the IP location votes reside, is assigned per PoP and the location of a PoP is further refined based on these locations alone. As all the PoP IP addresses should be located within the same campus, the location confidence of a PoP is significantly higher than the confidence that can be gained from locating each of its IP addresses separately. An example of a PoP level map generated by this method is given in Figure 3. Note that PoPs that appear in mid ocean are actually located on islands. Rasti *et al.* [41] use two geolocation databases: MaxMind GeoIP City [33] as a main source and IP2Location [20] to corroborate MaxMind's information, discarding IP addresses missing city location in one of the databases or where there is more than 100km deviation between the two sources. This approach is somewhat inaccurate, as Shavitt and Zilberman [45] have shown that these two databases are the most similar out of seven tested databases, with a median of five kilometer distance between IP addresses locations within them. This means that an error is Maxmind database is highly likely to appear in IP2Location database as well.

One way to improve the location provided by geolocation databases is to use the PoP level graph itself, starting at PoPs with a known location (such as universities) or a location with high level of confidence and crawling the


**Fig. 3.** DIMES PoP Level Map, 2010

graph to improve the location of neighboring nodes: The algorithm starts by identifying and marking the PoPs whose location is certain. The algorithm then discovers PoPs that are located in the same place as the marked PoPs, based on the PoP-level link delay. Following, it attempts to find the optimal location of non-marked PoPs based on the ratio of PoP-link delay to PoP distance from marked PoPs. First, all the possible locations for all the PoP IP addresses from all databases are examined to find the one that has the best ratio. If the best location is not satisfying, multilateration is used. The algorithm then iterates and tries to improve the location of unmarked PoPs using the location of newly marked PoPs.

The algorithm can then be extended for IP geolocation. It was shown that 80% of the IP addresses where within 1mS and two hops from a PoP and for the rest of the IPs multilateration can be applied from the nearest PoPs. The algorithm was corroborated by a large ISP where geolocation databases placed all of its IP addresses in a single location. The algorithm correctly distributed the ISP's IP addresses to near PoP locations around the globe.

Assigning a geographical location to PoPs is therefore a difficult task which is hard to validate without ground truth information.

## 4 PoP-Level Connectivity

The connectivity between PoPs is an important part of PoP level maps. DIMES generates PoPs connectivity graph using unidirectional links. They define a link $L_{SD}$ as the aggregation of all unidirectional edges originating from an IP address included in a PoP $S$ and arriving at an IP address included in a PoP $D$. Each of the IP level links has an estimate of the median delay measured along it, with the median calculated on the minimal delay of up to four consecutive measurements. Every such four measurements comprise a basic DIMES operation. All measured values are roundtrip delays [12]. A Link has the following properties:

- Source and Destination PoP nodes.
- The number of edges aggregated within the link.
- Minimal and Maximal median delays of all IP edges that are part of the PoP level link.
- Mean and standard deviation of all edges median delays.
- Weighted delay of all edges median delays. The edge's weight is the number of times it was measured.
- The geographical distance between source and destination PoP, calculated based on the PoPs geolocation.

A weighted delay of a link is used to mitigate the effect of an edge with a single measurement on the overall link delay estimation, where a link is otherwise measured tens of times through other edges. iPlane uses the inter-cluster connectivity to generate PoP level connectivity, with a similar definition of PoP level links, only using bidirectional links. The delay measured on links is not very different than in DIMES': For every inter-PoP link, iPlane considers all the corresponding inter-IP links that are measured in traceroutes. From every traceroute in which any such inter-IP link is observed, obtain one latency sample for the link as the difference in RTTs to the either end of the link and drop all latency samples that are below 0. Compute the latency for the inter-PoP link as the median of all the remaining samples for it. If there are no samples left after ignoring all the negative latency samples, the latency of the link is indicated as -9999 (about 6% of the links).

Using a dataset comprised of 478 million traceroutes conducted in weeks 42 and 43 (late October) of 2010, measured by 1308 DIMES agents and 242 iPlane vantage nodes and applying DIMES' algorithm to it result in a PoP level map that contains 4750 PoPs, 82722 IP addresses within the PoPs and 102620 PoP level links [46]. The links are an aggregation of 1.98M IP level edges. All the PoPs have outgoing links, with only 2 PoPs having only incoming links and one PoP with no PoP level links (only IP-level). As a full PoP level map is too detailed to display, a partial map is shown in Figure 4. The figure demonstrates the connectivity between randomly selected 430 ASes on the PoP level.

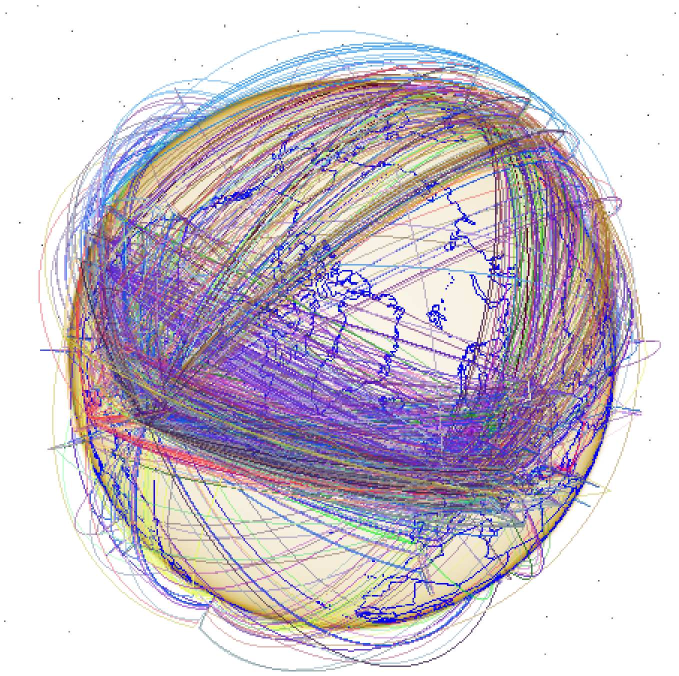

**Fig. 4.** An Internet PoP Level Connectivity Map - A Partial DIMES Map of Week 42, 2010

Most of the IP edges that are aggregated into links are unidirectional: 96.6%. This is a characteristic of active measurements: vantage points are limited in number and location, thus most of the edges can be measured only one way. However, at the PoP level, 18.8% of the links are bi-directional: six times more than the bi-directional edges. This demonstrates one of the advantages of using a PoP level map, as it provides a more comprehensive view of the networks' connectivity without additional resources. The average number of edges within a unidirectional link is 6.9, and the average number of edges within a bidirectional link is 72.9. This is not surprising, as it is likely that most of the bidirectional links will connect major PoPs, within the Internet's core and thus be easily detected.

An additional view of edges aggregation into links is given by Figure 5. The X-axis shows the number of edges aggregated into a link, while the Y-Axis is the number of PoP-level links. The graph shows a Zipf's law relation between the two, as 81.5% of the links aggregate ten edges or less, and less than 2.5% aggregate 100 edges or more. The large number of edges per link is explained by the fact that a measured edge is not a point-to-point physical connection: Take two routers, A & B, connected by a single fiber; If one of the routers has 48 ports, and one measures through each and every port, he will detect 48 edges between the two routers (incoming port i on router A and the single connected incoming port of router B).

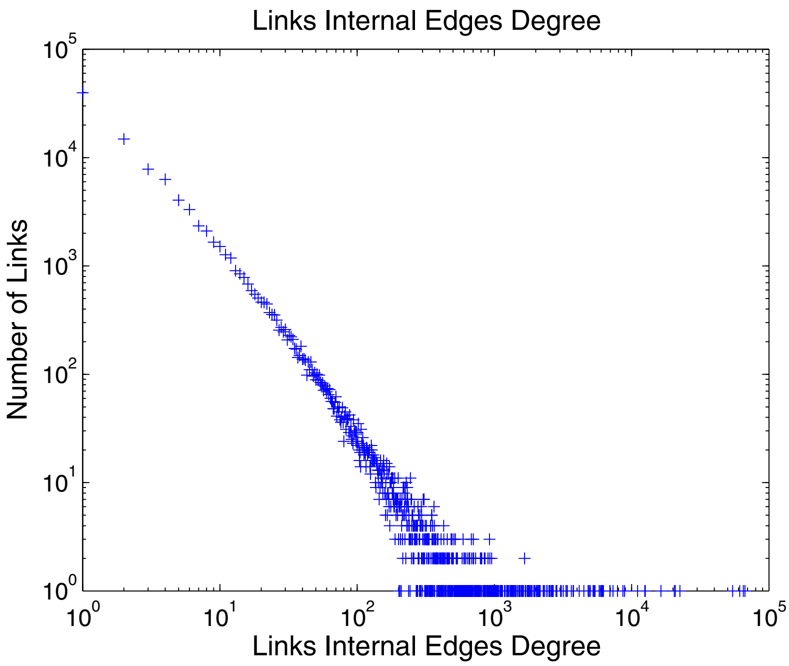


**Fig. 5.** Number of Edges within a Link vs. Number of PoP Level Links in DIMES Dataset

The number of links per PoP also behaves according to Zipf's law, as shown in Figure 6. The figure shows the total number of links per PoP, the number of outgoing links (source PoP) and the number of incoming links (Destination port). The connectivity between PoPs is very rich: only twenty two PoPs have one or two links to other PoPs, while 70% of the PoPs have ten or more links to other PoPs.

Many of the links are between PoPs that are co-located, which we define as links with a minimal delay of 1mS or less, and over 90% of the PoPs have such links. Almost all the PoPs (over 97%) are connected to PoPs outside their AS.

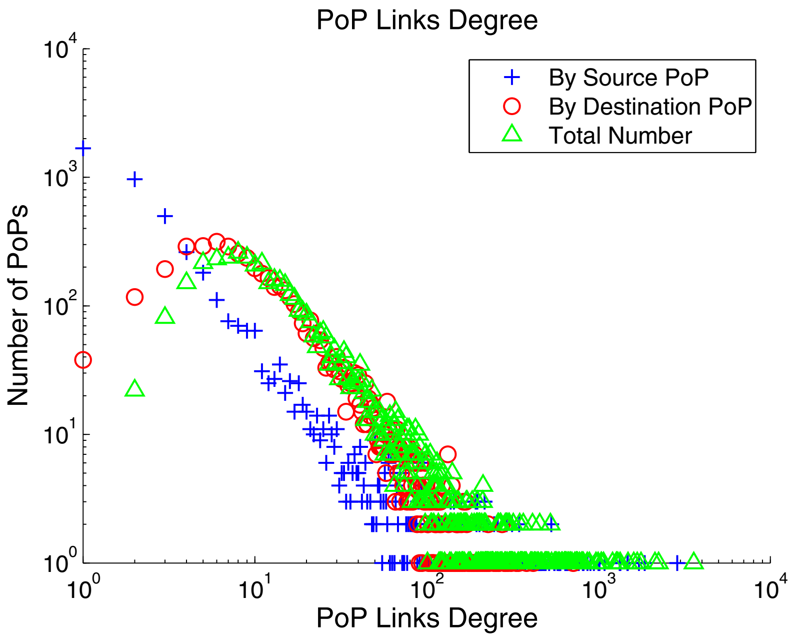


**Fig. 6.** Number of Links per PoP vs. Number of PoPs in DIMES Dataset

Figure 7 shows the minimum, weighted average and maximal delay per link, plotted on a log-log scale with the delay (X-scale) measured in milliseconds. The solid black line shows the cumulative number of measurements up to a given link delay. We omit from this plot links that include only a single edge, which distort the picture as their minimal, weighted and maximal delay are identical. An interesting attribute of this plot is that all three plotted delay parameters behave similarly and are closely grouped. As all the links are an aggregation of multiple edges, this indicates the similarity in the delay measured on different edges. One can also see that most of the measurements represent a delay of 200ms or less, and that the extreme cases are rare (see the cumulative measurement line). In almost all the cases where a minimal delay of 1sec or more are measured, this is a link that is made of a single edge. The same logic applies also for links with a small maximal delay, meaning the maximal delay was defined by only one or two measured edges. Here, however, a small maximal delay may also indicate co-located PoPs.

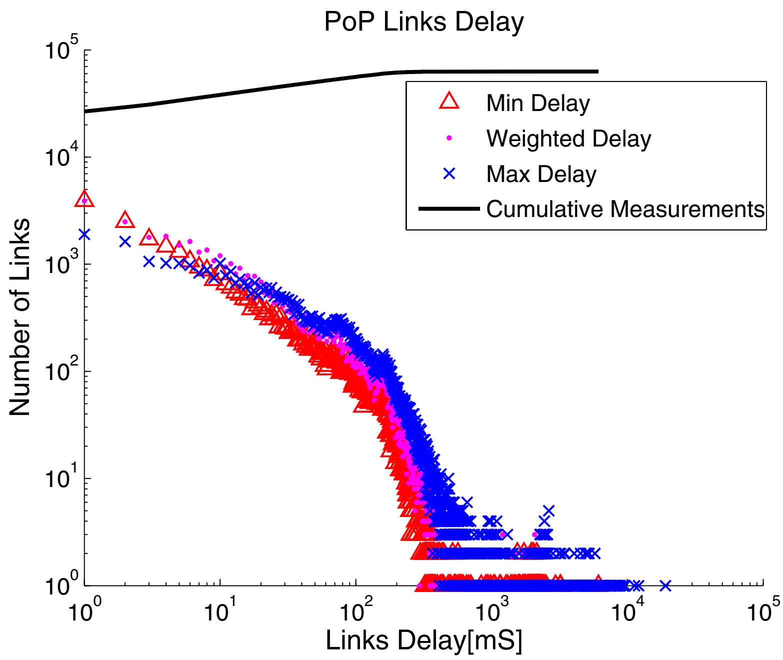


**Fig. 7.** Links Delay vs. Number of Links in DIMES Dataset

Traceroute measurements are known to introduce delay errors [18,44]. The errors tend to be of an additive nature, though sometimes a measured single-edge delay may be lower than its physical delay, due to an additive delay of the previous edge in the measured traceroute. This phenomenon is demonstrated by Figure 8: The X-Axis of the figure shows the estimated minimal link delay (in milliseconds), and the Y-Axis shows the spread of edge delay measurements. The figure focuses on the interesting range of delays, up to 500mS link delay and one second edge delay. A few measurements exist outside these boundaries, but their contribution to this discussion is small. Figure 8 clearly demonstrates the effect of a single edge measurement error: some links have a minimum delay of zero yet some of their measurements reach one second. Thus the aggregation of multiple edges into PoP level links significantly cleans noise from the collected data.

The Internet Topology Zoo [28] is a project that stores a large number of PoP level maps obtained from service providers' websites. The project provides PoP connectivity maps in GML format, converted from the original image file. The maps are annotated, when possible, with the following information:

- link types or speeds.
- longitudes and latitudes of nodes obtained through geocoding of PoP locations.
- a URL showing where the data was obtained.

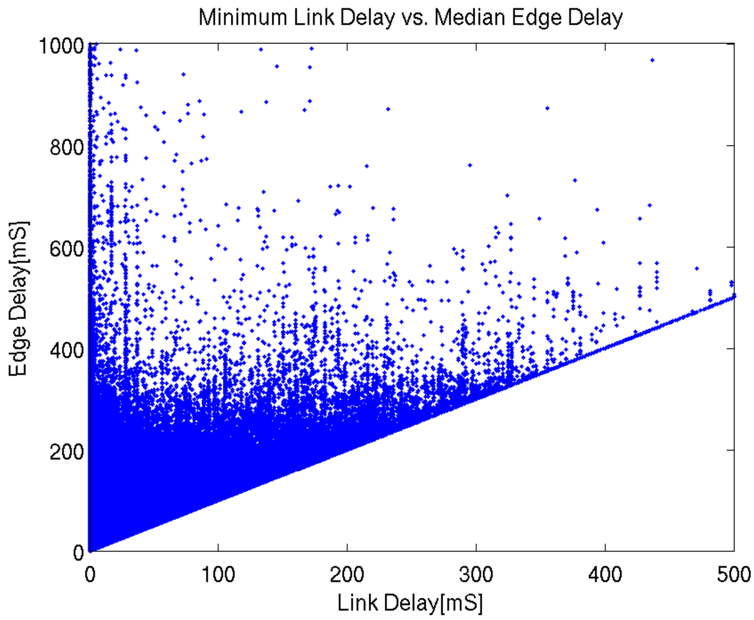


**Fig. 8.** Link Delay vs. Edge Delay in DIMES Dataset

- the date-of-record, i.e., the date that the map was representative of the network.
- the date the network map was obtained.
- a classification of the type of network.

In the Topology Zoo analysis, 59 research and education PoP-level networks and 82 commercial networks were studied. Most of the networks under study were backbone networks, and operated on country level or lower. Half of the networks had more than 21 PoPs, and 10% had 51 PoPs or more. The node degree observed was very low, ranging in different networks from 1.66 to 4.5, however this was only Intra-AS connectivity and was also biased as many of the maps were partial. For example, Sprint's network appeared to contain only 11 PoPs, and in AT&T only the MPLS network was present.

## 5 PoP Maps Validation

An important question when examining PoP level maps is how the map was validated. Accuracy is the most important validation evaluation aspect, and it entails multiple facets.

- How accurate is the classification of IP addresses to PoPs?
- How accurate is the assignment of PoPs to geographic locations?
- How accurate is the inference of PoP level links and their delay?

In addition, one may also want to evaluate the coverage of PoPs, meaning how many of the actual ISP's PoPs are covered by the extracted PoPs map, and how many IPs of a PoP are assigned to it. The effort required to validate PoP level maps is thus considerably high.

Spring *et al.* [49] verified completeness with the help of three ISPs. The ISPs verified that no PoPs or inter-PoP links were missing. However, in two of the cases there were spurious links. In addition, some access PoPs were missing. Further validation was conducted on the router level, both for completeness, impact of measurement reductions and alias resolution. The alias resolution, used for PoPs detection, failed for about 10% of the IP addresses, and in Sprint network, 63 out of 303 routers were resolved incorrectly.

Andersen *et al.* [3] did not focus on the validation of their PoP maps results, rather they presented the impact of different aspects of their clustering algorithm on the results. The PoP level maps were in fact used to validate the clustering results.

The iPlane PoP level maps [31], which are mostly based on the Rocketfuel's approach, focus their validation efforts on the inter-PoP connectivity. The validation uses measurements taken from 37 Planet Lab nodes to destinations in 100 random prefix groups. The first step in the validation is end-to-end latency error estimation. Next, the two path based and latency based delay estimations are compared to the results of Vivaldi [6]. They find that 73% of their predictions obtained using the path composition approach are within 20 ms of the actual latency.

DIMES [12] validation efforts are mainly divided between two aspects of the PoP maps: the PoP extraction and its geolocation. On the extraction level, the stability of the algorithm is evaluated, as well as the best time period for maps generation. Two weeks period is found to be both with a high level of network coverage and yet flexible enough to reflect changes in the network. As the extraction is measurement based, the effect of repeatedly targeting specific networks and adding iPlane's PoP IP addresses to the measurements target was evaluated and shown to improve PoPs detection in small networks, but have a very small effect in large ASes. The assigned location of the PoPs was confirmed by several commercial ISPs, one of them a large global provider. Further comparison of maps was manually conducted with maps published on major ISPs websites, such as Sprint, QWest, AT&T, British Telecom and more. The location of research facilities' PoPs, which is known, was also validation for 50 such institutes. 49 out of the 50 were correctly located within 10km from the actual location, and the last one failed due to an error in two of the three geolocation databases used.

AS eyeballs [41] was validated by comparing the AS eyeballs results with public PoP maps information published by 45 ASes. The scope of the averaging done using the KDE method is controlled by the bandwidth of a kernel function. The validation showed that when kernel bandwidth was 40km, for 60% of ASes only 20% of the PoP locations matched the service provider's map. However, for the top 10% ASes the locations match was over 50%. On the average, 41% of the PoP locations matched the location on the reference ISP's map. Increasing the kernel bandwidth to 80km increased the match to 60%, but decreased the number of PoPs found. Rasti *et al.* found that two causes for inaccuracy in their approach were the existence of multiple PoPs within a short distance and the placement of some PoPs away from major end users concentrations. They also compared their map with DIMES' map and found that for 80% of the eyeball ASes, the identified PoPs were a superset of DIMES'.

The Internet Topology Zoo [28] maps originate at the network operators, and are thus considered reliable. While an ISP may present a somewhat simplified network map, this aspect can be considered negligible. A possible concern is the accuracy of maps' translation into transcripts: The maps are manually annotated by the project's team, with one researcher doing the annotation and another reviewing his work, however both works are manual. The project also omits large networks with graphic links that are tangled or hard to follow.

For all the cases presented above, the validation of the generated PoPs was a very hard task: While service providers provide graphic maps of their PoPs, the PoP's actual details and the address range used within the PoP's routers are being kept confidential. PoP maps are therefore best validated when checked by the ISP, yet this is not possible on a large scale map.

## 6 Discussion

As we have shown in Section 2, extracting PoP level maps was originally considered as a simple task. As years passed, our understanding of the difficulties in PoP extraction grew. The approach of PoP classification using DNS, used by the Rocketfuel project and Andersen *et al.*, was shown to fail as times goes by, and less routers respond to name queries or alternatively contain inaccurate information [56]. In addition, incorrect DNS resolution leads to discovery of inter-PoP links that do not actually exist [50]. Furthermore, these works were limited in span and covered only a small portion of the network.

Of all the works presented before, only iPlane and DIMES generate PoP level maps on a periodic basis and on a large scale: iPlane update their map on a bi-monthly basis, while DIMES generate bi-weekly maps. Both mapping efforts attempt to cover the entire Internet and not only a specific ISP or a region. The two works present two different extremes of the accuracy-coverage tradeoff: iPlane tries to cluster as many routers as possible into PoPs, and may include some non-PoP IPs, whereas DIMES extracts less PoPs but with a very high level of certainty that a discovered PoP includes only IP addresses belonging to PoPs. Singleton (IP addresses with a single low delay link to a PoP) which used to be part of the DIMES PoP level maps, are omitted in their newer maps to avoid assigning end-user IPs to PoPs. Consequently, the iPlane maps are considerably larger than DIMES ones, but the accuracy of mapping IP to PoP is lower. Each map can therefore be used for different types of analysis, depending on the research question.

PoP level maps can be used for a variety of applications. Understanding network topology and dynamics is one clear usage, as was done by Spring *et al.* [49]. Teixeira *et al.* [50] used PoP level topologies to study path diversity in ISP networks. The PoP level maps can also be used to evaluate and validate results of other properties of the networks, as done by Andersen *et al.* [3] who used them to check their clustering algorithm. Several works have considered the PoP level topology for delay estimation and path prediction [30,31].

A new look at the Internet's topology is through dual AS/PoP maps: maps of the Internet that combine both the AS and the PoP level graph views, leveraging the advantages of each level of aggregation. One application of dual AS/PoP maps is the study of types of relationships between ASes. Using the geographical location of PoPs, one can explore not only the connectivity between ASes on the PoP level, but also how the relations between service providers change based on the location of the PoPs. Some work in this field was done by Rasti *et al.* [41], who looked at AS connectivity at the "Edge" in AS1267 (Infostrada) and AS8234 (RAI). They found that actual peering is significantly more complex than expected, e.g., a single PoP may use five peering PoPs in different ASes for upstream. Another application is distance estimation: instead of using router-level path stitching, one can find the shortest path between every two nodes on the dual map. The shortest path can then be used to find the distance between the two nodes. PoP level maps reduce the number of edges used for the path stitching, as multiple routers are aggregated into a single PoP, and the delay-based distance estimation is more accurate as the delay estimation of a PoP level link is better than that of a single IP-level edge. Last, the PoP location can be used to improve geolocation of each node and thus the distance estimation between the pair of nodes.

PoP level maps may also be useful for research related to homeland security. Schneider *et al.* [43] used DIMES' PoP level maps to study the mitigation of malicious attacks on networks. They considered attacks on Internet infrastructure and found that cutting the power to 12% of the PoPs and 10% of power stations will affect 90% of the networks integrity. Following, they suggested ways to improve the robustness of the network by using link changes.

Annotating the PoP level maps with geographic, economic and demographic information, one can achieve an understanding of the dynamics of the Internet's structure at short and medium time scales, in order to identify the constitutive laws of Internet evolution. These can be used to develop a realistic topology generator and a reliable forecast framework that can be used to predict the size and growth of the Internet as economies grow, demographics change, and as-yet unattached parts of the world connect.

## 7 Conclusion

In this chapter we presented Internet PoP level maps and surveyed related works. While PoP level maps provide a good view of the network, annotated with geographic location, only a few works focused on the generation of such maps, and currently only two projects provide large scale PoP level maps on periodic basis. As it is hard to corroborate the generated maps, we presented different approaches that are taken: some prefer extending the size of the map with the possibility of including non PoP IP addresses while others prefer smaller maps with a higher level of accuracy. The geographic location of a PoP is taken from geolocation databases or using measurement based tools. An error in either one can significantly affect the location annotation, thus different approaches are taken to mitigate this effect. We discussed the connectivity of generated PoP level maps and some of their characteristics. The PoP level maps have a high level of connectivity and the effect of delay measurements' errors is mitigated by the aggregation of IP level edges to PoP level links. PoP level maps have many applications in a vast range of research areas, and can be leveraged to study unexplored aspects of the network as well as its evolution.

**Acknowledgments.** We would like to thank Lior Neudorfer for contributing Figure 1. This work was partly supported by the Kabarnit Cyber Consortium (2012-2014) under Magnet program, funded by the chief scientist in the Israeli ministry of Industry, Trade and Labor.

## References

1. Sarangworld project, `http://www.sarangworld.com/TRACEROUTE/`
2. Quova (2010), `http://www.quova.com`
3. Andersen, D.G., Feamster, N., Bauer, S., Balakrishnan, H.: Topology inference from BGP routing dynamics. In: Internet Measurement Workshop, pp. 243–248 (2002)
4. Bender, A., Sherwood, R., Spring, N.: Fixing ally's growing pains with velocity modeling. In: Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement (IMC 2008), pp. 337–342 (2008)
5. CenturyLink Business. CenturyLink network maps, `http://www.centurylink-business.com/demos/network-maps.html` (accessed: October 10, 2012)
6. Dabek, F., Cox, R., Kaashoek, F., Morris, R.: Vivaldi: a decentralized network coordinate system. In: Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM 2004, pp. 15–26 (2004)
7. Digital Envoy. NetAcuity Edge (2010), `http://www.digital-element.com/our_technology/edge.html`
8. DIMES. Distributed Internet Measurements and Simulations, `http://www.netdimes.org/`
9. Donnet, B., Friedman, T.: Internet topology discovery: a survey. IEEE Communications Surveys Tutorials 9(4), 56–69 (2007)
10. Doyle, J.C., Alderson, D.L., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R., Willinger, W.: The robust yet fragile nature of the internet. Proceedings of the National Academy of Sciences of the United States of America 102(41), 14497–14502 (2005)

11. Eriksson, B., Barford, P., Sommers, J., Nowak, R.: A Learning-Based Approach for IP Geolocation. In: Krishnamurthy, A., Plattner, B. (eds.) PAM 2010. LNCS, vol. 6032, pp. 171–180. Springer, Heidelberg (2010)
12. Feldman, D., Shavitt, Y., Zilberman, N.: A structural approach for PoP geolocation. Computer Networks 56(3), 1029–1040 (2012)
13. Geobytes. GeoNetMap (2010), `http://www.geobytes.com/`
14. Georgatos, F., Gruber, F., Karrenberg, D., Santcroos, M., Uijterwaal, H., Wilhelm, R.: Providing active measurements as a regular service for ISPs. In: Proceedings of the Passive and Active Measurements Workshop, PAM 2001 (2001)
15. Govindan, R., Tangmunarunkit, H.: Heuristics for internet map discovery. In: Proceedings of IEEE INFOCOM, Tel Aviv, Israel (2000)
16. Greene, B.R., Smith, P.: Cisco ISP Essentials. Cisco Press (2002)
17. Gueye, B., Uhlig, S., Fdida, S.: Investigating the Imprecision of IP Block-Based Geolocation. In: Uhlig, S., Papagiannaki, K., Bonaventure, O. (eds.) PAM 2007. LNCS, vol. 4427, pp. 237–240. Springer, Heidelberg (2007)
18. Gueye, B., Uhlig, S., Ziviani, A., Fdida, S.: Leveraging Buffering Delay Estimation for Geolocation of Internet Hosts. In: Boavida, F., Plagemann, T., Stiller, B., Westphal, C., Monteiro, E. (eds.) NETWORKING 2006. LNCS, vol. 3976, pp. 319–330. Springer, Heidelberg (2006)
19. Gueye, B., Ziviani, A., Crovella, M., Fdida, S.: Constraint-based geolocation of Internet hosts. IEEE/ACM Transactions on Networking 14(6) (2006)
20. Hexsoft Development. IP2Location (2010), `http://www.ip2location.com`
21. hostip.info, hostip.info (2010), `http://www.hostip.info`
22. Huffaker, B., Dhamdhere, A., Fomenkov, M., Claffy, K.: Toward Topology Dualism: Improving the Accuracy of AS Annotations for Routers. In: Krishnamurthy, A., Plattner, B. (eds.) PAM 2010. LNCS, vol. 6032, pp. 101–110. Springer, Heidelberg (2010)
23. Huffaker, B., Plummer, D., Moore, D., Claffy, K.: Topology discovery by active probing. In: Symposium on Applications and the Internet (SAINT), pp. 90–96. SAINT (2002)
24. Hyun, Y.: Archipelago measurement infrastructure, `http://www.caida.org/projects/ark/`
25. Internet2, Internet2 gigapop list, `http://eng.internet2.edu/gigapoplist.html` (accessed: October 10, 2012)
26. Katz-Bassett, E., John, J.P., Krishnamurthy, A., Wetherall, D., Anderson, T., Chawathe, Y.: Towards IP geolocation using delay and topology measurements. In: Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement (IMC 2006), pp. 71–84 (2006)
27. Keys, K., Hyun, Y., Luckie, M., Claffy, K.: Internet-Scale IPv4 Alias Resolution with MIDAR. IEEE/ACM Transactions on Networking (2012)
28. Knight, S., Nguyen, H., Falkner, N., Bowden, R., Roughan, M.: The internet topology zoo. IEEE Journal on Selected Areas in Communications 29(9), 1765–1775 (2011)
29. Laki, S., Mátray, P., Hága, P., Sebök, T., Csabai, I., Vattay, G.: Spotter: A model based active geolocation service. In: Proceedings of IEEE INFOCOM, Shanghai, China (2011)
30. Lee, D., Jang, K., Lee, C., Iannaccone, G., Moon, S.: Scalable and systematic internet-wide path and delay estimation from existing measurements. Computer Networks 55(3) (2011)
31. Madhyastha, H.V.: An information plane for internet applications. Thesis, University of Washington (2008)

32. Madhyastha, H.V., Anderson, T., Krishnamurthy, A., Spring, N., Venkataramani, A.: A structural approach to latency prediction. In: Proceedings of the 6th ACM SIGCOMM Conference on Internet Measurement (IMC 2006), pp. 99–104 (2006)
33. MaxMind LLC. GeoIP (2010), http://www.maxmind.com
34. Merit Network. Internet routing registries, http://www.irr.net/
35. Muir, J.A., Oorschot, P.C.V.: Internet geolocation: Evasion and counterevasion. ACM Computing Surveys 42(1), 1–23 (2009)
36. Padmanabhan, V.N., Subramanian, L.: An investigation of geographic mapping techniques for Internet hosts. In: Proceedings of the 2001 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM 2001), pp. 173–185 (2001)
37. Pastor-Satorras, R., Vespignani, A.: Evolution and Structure of the Internet: A Statistical Physics Approach. Cambridge University Press (2004)
38. Poese, I., Uhlig, S., Kaafar, M.A., Donnet, B., Gueye, B.: IP geolocation databases: unreliable? ACM SIGCOMM Computer Communication Review 41(2) (2011)
39. QWest Business. Network maps, http://shop.centurylink.com/largebusiness/enterprisesolutions/networkMaps/ (accessed: October 10, 2012)
40. Radoslavov, P., Tangmunarunkit, H., Yu, H., Govindan, R., Shenker, S., Estrin, D.: On characterizing network topologies and analyzing their impact on protocol design. Technical report, Computer Science Department, University of Southern California (2000)
41. Rasti, A.H., Magharei, N., Rejaie, R., Willinger, W.: Eyeball ASes: from geography to connectivity. In: Proceedings of the 10th Annual Conference on Internet Measurement (IMC 2010), pp. 192–198 (2010)
42. Sardella, A.: Building next-gen points of presence, cost-effective PoP consolidation with juniper routers. White paper, Juniper Networks (2006)
43. Schneider, C.M., Moreira, A.A., Andrade, J.S., Havlin, S., Herrmann, H.J.: Mitigation of malicious attacks on networks. Proceedings of the National Academy of Sciences 108(10), 3838–3841 (2011)
44. Schwartz, Y., Shavitt, Y., Weinsberg, U.: A Measurement Study of the Origins of End-to-End Delay Variations. In: Krishnamurthy, A., Plattner, B. (eds.) PAM 2010. LNCS, vol. 6032, pp. 21–30. Springer, Heidelberg (2010)
45. Shavitt, Y., Zilberman, N.: A geolocation databases study. IEEE Journal on Selected Areas in Communications 29(9) (2011)
46. Shavitt, Y., Zilberman, N.: Geographical Internet PoP Level Maps. In: Pescapè, A., Salgarelli, L., Dimitropoulos, X. (eds.) TMA 2012. LNCS, vol. 7189, pp. 121–124. Springer, Heidelberg (2012)
47. Siamwalla, R., Sharma, R., Keshav, S.: Discovering Internet Topology. Technical report, Cornell University (1998)
48. Siwpersad, S.S., Gueye, B., Uhlig, S.: Assessing the Geographic Resolution of Exhaustive Tabulation for Geolocating Internet Hosts. In: Claypool, M., Uhlig, S. (eds.) PAM 2008. LNCS, vol. 4979, pp. 11–20. Springer, Heidelberg (2008)
49. Spring, N.T., Mahajan, R., Wetherall, D., Anderson, T.E.: Measuring ISP topologies with Rocketfuel. IEEE/ACM Transactions on Networking 12(1), 2–16 (2004)
50. Teixeira, R., Marzullo, K., Savage, S., Voelker, G.M.: In search of path diversity in ISP networks. In: Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement (IMC 2003), pp. 313–318 (2003)
51. TeliaSonera International Carrier. Network map, http://www.teliasoneraic.com/NetworkFlash/index.htm (accessed: October 10, 2012)

52. University of Oregon Advanced Network Technology Center. Route views project, http://www.routeviews.org/
53. Wong, B., Stoyanov, I., Sirer, E.G.: Octant: A comprehensive framework for the geolocalization of Internet hosts. In: Proceedings of the 4th USENIX Symposium on Networked Systems Design and Implementation, NSDI 2007 (2007)
54. XO. Complete network assets, http://www.xo.com/SiteCollectionImages/about-xo/xo-network/maps/map_complete_1600.gif (accessed: October 10, 2012)
55. Yoshida, K., Kikuchi, Y., Yamamoto, M., Fujii, Y., Nagami, K., Nakagawa, I., Esaki, H.: Inferring POP-Level ISP Topology through End-to-End Delay Measurement. In: Moon, S.B., Teixeira, R., Uhlig, S. (eds.) PAM 2009. LNCS, vol. 5448, pp. 35–44. Springer, Heidelberg (2009)
56. Zhang, M., Ruan, Y., Pai, V., Rexford, J.: How DNS misnaming distorts Internet topology mapping. In: Proceedings of the Annual Conference on USENIX 2006 Annual Technical Conference (ATEC 2006), pp. 34–34 (2006)
57. Zupan, J.: Clustering of large data sets. Chemometrics Research Studies Series. Research Studies Press (1982)