Inferring Internet Topology at Point of Presence Level

Abdullah Yasin Nur
Department of Computer Science
University of New Orleans
New Orleans, LA, USA 70148
Email: ayn@cs.uno.edu

Mehmet Engin Tozal
School of Computing and Informatics
University of Louisiana at Lafayette
Lafayette, LA, USA 70504
Email: metozal@louisiana.edu

***Abstract*—Analyzing point of presence (PoP) level Internet topology offers several advantages for researchers, network operators, and those involved in optimizing and maintaining the Internet infrastructure. It provides insights into how the Internet backbone is organized and how various networks are connected at a fundamental level. Specifically, understanding PoP level Internet topology is essential for optimizing network performance, ensuring reliability, and developing protocols that facilitate seamless communication across the interconnected web of networks on a global scale. PoP maps allow us to go beyond the simple autonomous system (AS) level abstraction by supporting multiple connections among the ASes. In this study, we propose a set of techniques to infer the PoP level topology map of the Internet by utilizing several real-life datasets, including traceroute, DNS, and commercial geolocation databases. Initially, we map the IP addresses collected from the traceroute to their corresponding ASes. Then, we use six geolocation techniques to find the location of IP addresses. Subsequently, we group the IP addresses into PoP nodes based on their geographical locations within ASes and identify the connection between PoP nodes. We validate this approach using various research and commercial networks worldwide. We investigate various features of the map to provide more insight regarding the backbone structure of the Internet.**

***Index terms*—** Point of Presence, PoP, Internet Topology, IP Geolocation, Internet Infrastructure

I. Introduction

The Internet is a global decentralized network of interconnected computer networks that allows computers and other devices all over the world to communicate and share information with each other. Research focused on mapping Internet topology plays a pivotal role in understanding the dynamics of the Internet backbone structure. By gaining insight into the underlying network structure, researchers can develop protocols by either crafting novel protocols and services or refining and optimizing existing ones.

The Internet is also called a network of networks containing more than 60 thousand autonomous systems connected to each other. A group of networks managed by one or more operators under a well-defined routing policy is called an Autonomous System (AS) in the Internet [15]. These systems operate independently, making routing decisions based on the Border Gateway Protocol (BGP) to exchange routing information with other Autonomous Systems. AS numbers are unique identifiers assigned to each autonomous system, facilitating the routing of data across the internet. ASes enable the decentralized and distributed nature of the internet, allowing diverse networks to interconnect and facilitate the exchange of data across the digital landscape. The size of an AS can vary, ranging from a small size, such as that of a campus or building (e.g., University of New Orleans, AS26333), to larger ASes that traverse extensive regions like multiple states (e.g., Cox Communications in the US, AS22773) or even countries (e.g., Lumen Technologies worldwide, AS10753).

ASes typically construct their physical networks in a hierarchical fashion, organizing them into sub-networks known as Point of Presence (PoPs). Each PoP contains several routers and network devices located within the same facility. A PoP is a physical location where different networks, such as Internet Service Providers (ISPs), connect and exchange traffic.

PoP level topology focuses on understanding the layout and interconnections of these PoPs, which serve as crucial nodes in the global network. Analyzing PoP level internet topology involves studying the relationships, traffic exchanges, and data flow between these key points, providing insights into the functioning and efficiency of the broader internet ecosystem. Understanding PoP level internet topology is essential for optimizing network performance, ensuring reliability, and developing protocols that facilitate seamless communication across the interconnected web of networks on a global scale.

Discovering the PoP level topology map is relatively more complex than other levels, including interface, router, and AS level, because IP protocol does not provide geographical information. In our previous work [13], we show that the Internet's geographical properties are crucial for analyzing the packet traversing and routing of the Internet. This work proposes several techniques to discover the PoP level Internet topology map. We use traceroute, IP to AS mapping, and several geolocation techniques, including Vendor location based geolocation, DNS based geolocation, Majority GeoDB based geolocation, Sandwich method geolocation, RTT based geolocation, Singular GeoDB based geolocation. We validate our results across multiple research and commercial networks around the world, and investigate various aspects of the map to offer a deeper understanding of the Internet's backbone structure.

The rest of the paper is organized as follows. Section II presents the related work. We introduce the details of our approach in Section III. Sections IV and V show validation of our approach and present experimental results, respectively. Finally, Section VI concludes the paper.

II. Related Work

The use of Internet topology maps has been pivotal in various areas of networking research and applications [7, 10]. A comprehensive and accurate global topology map of the Internet empowers network researchers to grasp the practical dynamics of the Internet. It provides guidance for network operators in enhancing the reliability and security of their networks, enables network engineers to enhance the efficiency of their systems, and helps developers in creating applications that are attuned to network topology, among other benefits [1].

Discovering the Internet's topology poses challenges due to a scarcity of measurement tools, limited available sources, inadequate support from ASes, and the existing infrastructure of the Internet. Nevertheless, the significance of mapping the Internet's structure has attracted many researchers, which has resulted in extensive studies for creating accurate topology maps in the field [3].

Point of Presence (PoP) is defined as a group of routers that belong to the same AS and are physically located at the same building or campus [9]. PoP level map contains the physical infrastructure information of the backbone networks instead of a simple abstraction. Therefore, it provides precious information to analyze the Internet structure. One of the earlier works, Spring et al. [4] proposed UNDNS, which is a DNS-based geolocation technique to group routers into their geolocations to create a PoP level topology map. In a later work, Madhyastha et al. [6] improved UNDNS's keyset and applied it in their work to improve coverage and accuracy. However, DNS naming is not mandatory, and not all ASes support DNS naming conventions. Therefore, the techniques entirely dependent on DNS geolocation are inadequate to represent the entire Internet. Feldman et al. [8] have suggested a graph-based approach by analyzing patterns and motifs in the traceroute dataset to identify PoPs. They use several geolocation services to discover the coordinates of the PoPs to generate PoP level topology maps. Topology Zoo [11] and the Internet Atlas [12] projects collect network topologies from network providers. Unfortunately, both of the projects are outdated and no longer available.

In this work, we propose a method to discover the PoP level topology map which captures the backbone structure of the Internet. We use several different geolocation techniques to improve our map's accuracy. Compared to the AS level Internet topology maps, PoP level maps capture multiple links among ASes instead of logical relations among them.

III. Methodology

This section presents our techniques for creating accurate PoP level topology maps. We construct PoP level maps in four steps.

1) **IP to AS Mapping:** The initial phase involves mapping the IP addresses collected from the traceroute to their corresponding ASes.
2) **IP Geolocation:** We geolocate the IP addresses extracted from the previous step by applying the following techniques.
    a) Vendor location based geolocation
    b) DNS based geolocation
    c) Majority GeoDB based geolocation
    d) Sandwich method geolocation
    e) RTT based geolocation
    f) Singular GeoDB based geolocation
3) **PoP Clustering:** We group the IP addresses into PoP nodes based on their geographical locations within ASes.
4) **PoP Map Construction:** We identify the connections between the PoP nodes generated in the preceding step using traceroute.

In the first step, we map all unique IP addresses observed in the traceroute to their corresponding ASes. We used the CAIDA IPv4 Prefix-Probing Traceroute Dataset [20] consisting of more than 291 million $(291,776,909)$ path traces. We use the IPv4 Routeviews prefix to AS mappings dataset (pfx2as) obtained from CAIDA [22].

A. IP Geolocation

In the second step, we use several geolocation techniques to infer the coordinates of IP addresses.

1) Vendor Location Based Geolocation: Certain entities offer a worldwide research network that allows researchers to test their applications on a large-scale, geographically distributed platform. In this work, we use RIPE Atlas [23] and Measurement Lab (M-Lab) [24] nodes.

RIPE NCC serves as the Regional Internet Registries (RIRs), tasked with allocating and overseeing Internet number resources in the European and Middle Eastern regions, including IP addresses and autonomous system numbers. In an effort to comprehensively assess the connectivity and real-time reachability measurements of the Internet, they initiated the RIPE Atlas project. This initiative involves volunteers globally deploying RIPE probes or anchors within their networks. RIPE probes, characterized by their compact hardware design, draw power through USB connections and are linked to the Ethernet port on users' routers. For example, AT&T (AS7018) provides a probe with an IP address of 76.236.29.168, which is located in San Francisco. By the time this paper is written, there are $12,050$ probes available. RIPE anchors combine the capabilities of RIPE probes with enhanced measurement functionalities and regional measurement targets within the broader RIPE Atlas network. For example, F5 Networks (AS55002) provides an anchor with an IP address of 107.162.223.5, which is located in Seattle. By the time this paper is written, there are 784 anchors available.

In a similar vein, Measurement Lab (M-Lab) stands as an open and distributed platform, offering researchers, developers, and the broader public a convenient way to measure

and diagnose the performance of their internet connections. Launched in 2009 through a collaborative effort involving the New America Foundation's Open Technology Institute, Google, and academic researchers, M-Lab boasts a network of 191 nodes dispersed across 66 different cities worldwide. As an illustration, Zayo (AS6461) contributes a node featuring the IP address 128.177.119.229, which is located in New York.

Finally, we use the CAIDA Internet eXchange Points (IXPs) dataset [21], which provides details on IXPs' geographical location, facility information, and prefixes. The dataset combines information from several sources, including PeeringDB, Hurricane Electric, and Packet Clearing House. The dataset contains IP addresses allocated by individual member ASes at a specific IXP.

2) DNS Based Geolocation: ASes commonly incorporate geographic details into their Domain Name System (DNS) naming conventions. While the utilization of DNS naming is not mandatory, it remains one of the most valuable sources of information directly accessible from the ASes. Techniques based on DNS for geolocation leverage geographic cues embedded in domain names to deduce locations. For instance, Cox employs the naming convention "ip98-160-200-1.lv.**lv**.cox.net," where the inclusion of "lv" signifies the location as Las Vegas.

We use HOIHO [14], which is a tool designed for extracting geolocation information from DNS names. The project provides an API [26] that returns JSON output containing coordinates of resolved DNS names.

3) Majority GeoDB Based Geolocation: Commercial IP geolocation services employ various techniques and compile databases that associate IP addresses with physical locations, including countries, cities, and/or geographic coordinates. However, the accuracy of these mappings has been questioned in previous research [5]. To address this, we utilize three commercial geolocation databases along with a majority rule approach to resolve the geographic locations of unresolved IP addresses. Basically, we gather geolocation data from all three databases and assign a location to an IP address only if at least two databases reach a consensus on the geolocation. We use the commercial versions of "DB-IP IP address to location" database [16], "IP2Location DB5 Lite" [17], and "IPGeolocation.IO IP to City" [18] databases.

4) Sandwich Method Geolocation: Sandwich Method employs a similar rationale to the mathematical theorem where if $f(x) \leq g(x) \leq h(x)$ holds for all numbers, in the case of $x = y$ where $f(y) = h(y)$, then $g(y)$ must also be equal to them. We apply a similar approach, assuming that it is unlikely for a packet to travel to a city and then backtrack to the same city after traversing an intermediate city. The Sandwich Method locates unresolved IP addresses appearing in path traces by examining the two IP addresses immediately preceding and succeeding a particular IP address. If these two IP addresses are located in the same city, the method assigns the intermediate IP address to the same city. For instance, in a path trace with three IP addresses [A,B,C], if A and C are located in the same city, the method infers that B is also located in that city.

5) RTT Based Geolocation: Our previous work shows that Round-Trip Time (RTT) and geographical distance have a very strong correlation [13]. We employ this correlation to resolve the geolocation of the remaining unresolved IP addresses. Our approach involves gathering resolved IP addresses that appear before or after each IP address in all path traces. We assume that an RTT time difference below a specified threshold suggests that the IP address and the resolved IP address are located in the same city. However, if any other resolved IP address indicates a different location, we refrain from applying this method and retain the IP address as unresolved, recognizing that RTT can be influenced by various factors. We use a conservative threshold value of 3 ms, as suggested in previous work [2], to minimize false negatives.

6) Singular GeoDB based Geolocation: Lastly, we use a single geolocation database, DB-IP database [16], to find the locations of the remaining IP addresses.

B. PoP Map Generation

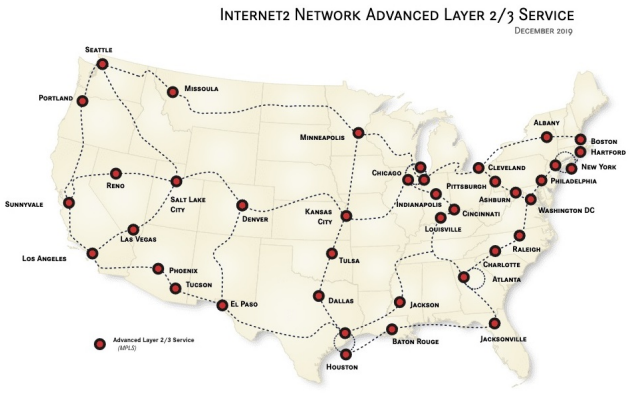
In the third step, we cluster PoPs belonging to individual ASes into PoP nodes based on their geolocations. Our underlying assumption is that each AS has at most one PoP in a given city. Consequently, if multiple PoPs from the same AS are located in the same city, we merge them into a single PoP node. This clustering process is executed by grouping IP addresses according to their associated city information. By the conclusion of this step, we have PoP nodes with the list of IP addresses located in those PoPs.

In the fourth step, we revisit our traceroute dataset to identify connections between PoPs. For each IP address in path traces, we examine whether an IP address immediately preceded or followed another IP address belonging to a different PoP. Upon such identification, we establish a link between the corresponding PoP nodes in our PoP map.

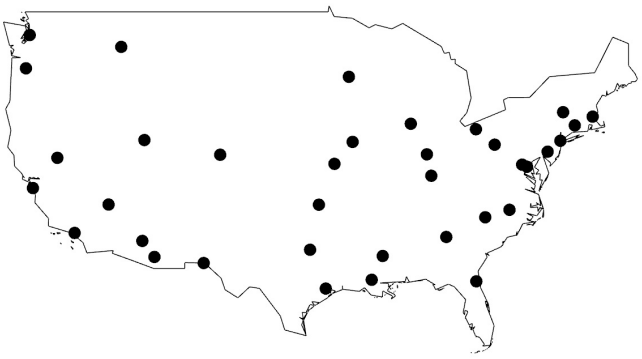
IV. PoP Level Validation

One significant challenge for the network measurement and analysis community is the validation of research outcomes. This challenge arises primarily from the limited availability of comprehensive validation methods in the literature. Existing approaches are often partial or indirect, mainly because ISPs do not share complete details of their networks. This reluctance is driven by concerns related to both security and business confidentiality. The sensitive nature of network infrastructure, coupled with the potential exposure of proprietary information, leads ISPs to restrict the disclosure of intricate details about their networks. Security concerns involve the risk of revealing vulnerabilities that could be exploited by malicious actors. Moreover, business concerns revolve around maintaining a competitive edge and safeguarding proprietary network configurations.

We obtained PoP level topology maps of major research networks and several large scale ISPs from their official websites. These maps visually represent the backbone infrastructures of



(a) Internet2 Layer 2/3 PoPs



(b) Our PoP nodes for Internet2

Fig. 1: Internet2 backbone map comparison

(a) Deutsche Telekom Ethernet PoPs

(b) Our PoP nodes for Deutsche Telekom

Fig. 2: Deutsche Telekom backbone map comparison

the respective networks, offering insights into the organization and connectivity of their PoPs. For research networks, we examine Internet2 (AS11164, ASN11537) and GÉANT (AS20965,AS21320). For commercial networks, we examine Cox Communications (AS22773), Cogent Communications (AS174), GTT Communications (AS3257), Deutsche Telekom (AS3320), and Hurricane Electric (AS6939).

We first examined Internet2, which provides high-speed network for research and education communities in the United States. Figure 1a presents the Internet2 network infrastructure map and Figure 1b presents the output of our method. We used Gephi [25] with Map of Countries and Geo Layout plugins to present coordinates into the map. The figures verify that our method successfully identified Internet2 PoP nodes. When we analyze it closely, we see that Internet2 has three nodes in Chicago, two in Houston, and two in New York. Note that our approach assumes that ASes has only one PoP in each city. Therefore, our method maps these extra nodes into one per city.

In addition to the research network, we compared our map with several a commercial ISP from Europe, Deutsche Telekom. Since Internet2 is based in the United States, we also wanted to see our accuracy in Europe. Figure 2a presents the Ethernet backbone structure of Deutsche Telekom in Europe. Figure 2b presents the output of our method. We successfully identify all of the nodes in Europe. Note that, Deutsche Telekom contains four PoPs in the US, one in Canada, and one in Hong Kong. Our technique successfully identified and captured these PoPs. Due to space limitations, we focused on zooming in on the European location in the map.

We conducted the same experiment for GÉANT (AS20965, AS21320), Cox Communications (AS22773), Cogent Communications (AS174), GTT Communications (AS3257), Deutsche Telekom (AS3320), and Hurricane Electric (AS6939). While the results indicate that our technique effectively captures PoP nodes for both research and commercial networks, we were unable to put them in the paper due to space constraints.

V. Empirical Analysis

A. Datasets Analysis

Traceroute and IP to AS Mapping:

We used the CAIDA IPv4 Prefix-Probing Traceroute Dataset [20] consisting of more than 291 million $(291,776,909)$ path traces. We found $1,070,949$ unique IP addresses appeared in the dataset. We used the CAIDA IPv4 Routeviews prefix to AS mappings dataset (pfx2as) [22] to map these IP addresses to their corresponding ASes. We were able to map $937,670$ IP addresses to their corresponding ASes. We removed the remaining 12% $(133,279)$ IP addresses since they do not have a valid AS number. Moreover, we observed $58,335$ unique ASes appearing in the traceroute dataset.

Geolocation Methods

Vendor Location Based Geolocation:

We gathered $26,133$ IP addresses from RIPE probes, 1190 IP addresses from RIPE anchors, and 1719 IP addresses from M-Lab. Notably, 1160 IP addresses were observed in both the RIPE anchors and probes sets. In summary, we collected information on $27,872$ unique IPv4 addresses, including their respective locations. Additionally, from the IXP dataset, we obtained details on $34,620$ unique IPv4 addresses along with their geographic locations. Interestingly, out of a total of $62,502$ IP addresses, only 1648 IP addresses were present in the traceroute dataset.

DNS Based Geolocation:

We use CAIDA's *"DNS Names for IPv4 Routed /24 Topology"* dataset, which provides DNS names for every routed /24 in the IPv4 address space [19]. The dataset contains $45,850,783$ unique DNS names, each associated with its respective IP addresses. We use HOIHO [26] to extract location information from DNS. HOIHO was able to obtain a valid geolocation for $5,715,300$ DNS entries. We were able to resolve the geolocations of $39,789$ IP addresses.

Majority GeoDB based geolocation:

In our study, we used three different commercial geolocation databases. To resolve the remaining unresolved IP addresses, we implemented a majority voting approach. In case two geolocation databases agree on the location of an IP address, we move that IP address to our certain list. Majority voting successfully resolved $642,665$ IP addresses, which corresponds to $\%68.53$ of all IP addresses.

Sandwich Method Geolocation:

The sandwich method locates unresolved IP addresses appearing in path traces and examines the two IP addresses that immediately appear before and after a given IP address. If these two IP addresses are at the same location, the sandwich method assigns the intermediate IP address to that location. Sandwich method successfully resolved $20,381$ IP addresses.

RTT based Geolocation:

Our assumption is that if the round-trip time difference is less than a 3ms threshold between a resolved IP address and an IP address that appears in a path trace immediately before or after the resolved IP address, then these addresses are located in the same location. This technique successfully resolved $127,135$ IP addresses.

Singular GeoDB based Geolocation:

Lastly, we used the DB-IP geolocation database to resolve the remaining $106,052$ unresolved IP addresses. At the end, all IP addresses were successfully mapped to a geolocation.

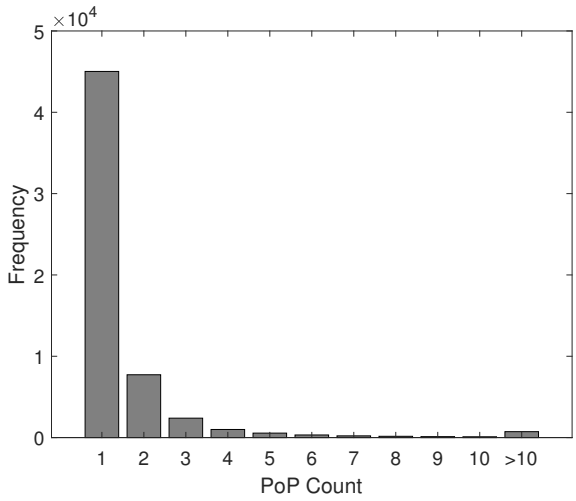


Fig. 3: PoP Node Count Distribution

TABLE I: Summary statistics for PoP Node Count Distribution

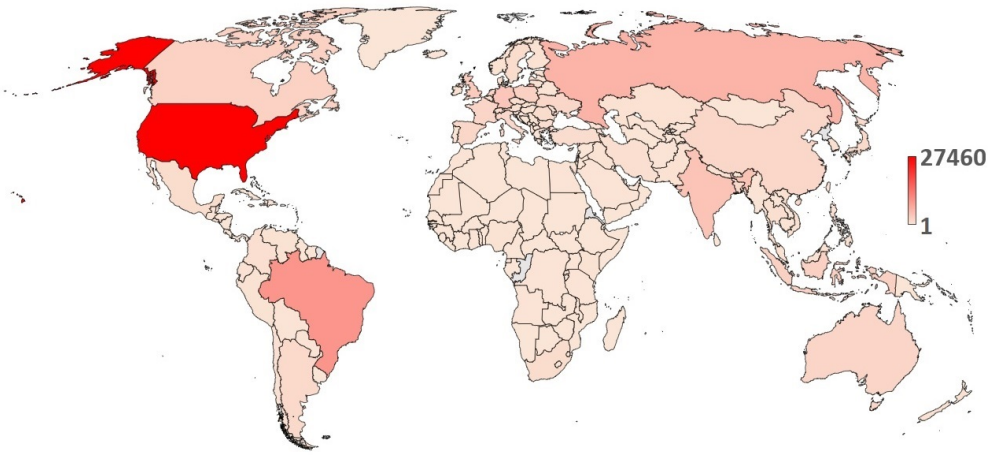
$Level$	Q_0	Q_1	Q_2	Q_3	Q_4	$Mean$	$StdDev$
PoP Node	1	1	1	1	451	1.7799	5.6242

B. PoP Level Topology Map Analysis

In this part, we analyze the PoP level topology map to provide more insight about the backbone structure of the Internet.

We found $103,830$ PoP nodes distributed over $58,335$ ASes. Figure 3 presents the node distribution of the Internet at the PoP level. The x-axis shows the number of the PoP nodes per AS and the y-axis shows their frequencies. In addition, Table I shows the summary statistics for the distribution. We observe that 77.18% of the ASes $(45,023)$ contain only 1 PoP. It is evident that a significant portion of ASes primarily resides at the edge of the Internet, not offering internet access to other ASes. In the conventional tier classification, these ASes are referred to as stub-ASes. The number of AS that contains more PoP significantly drops when the PoP count increases. Our observations show that only 100 ASes contain more than 100 Pops. The maximum PoP count belongs to Cogent Communications (AS174) with 451 PoPs distributed around the world.

Figure 4 shows the PoP node count by countries in the world as a heatmap and a table to present the top 10 list. Most of the PoP nodes belong to the United States with $27,460$, which corresponds to 26.45% of all PoPs that we observe. The nearest competitor is Brazil with $9,560$ nodes. One of the observation is that China is ranked at 15 with $1,679$ nodes, lower than several smaller countries.



Top 10 Countries

Country	PoP Node Count
United States	27460
Brazil	9530
Russia	5944
India	3942
Germany	3550
United Kingdom	3042
Indonesia	2402
Poland	2353
Canada	2269
Netherlands	1861

Fig. 4: PoP Node Count by Country in the World

TABLE II: Top 10 States in terms of Node Count

States	PoP Node Count
California	3464
Texas	2297
New York	2072
Florida	1409
Illinois	1371
Virginia	1329
Georgia	890
Pennsylvania	880
Colorado	854
Washington	848

Table II shows the top 10 states which contain the majority of PoPs in the United States. Especially California gives an interesting insight since it contains $3,464$ PoP nodes. If California were treated as a separate country, it would hold the sixth position globally in terms of ranking.

Next, we check the node degree distribution of the Internet at the PoP level. We define the degree of an AS as the number of connections it has to other ASes. Since some ASes have connections from several PoPs, the degree of an AS is defined as the total connection of all of its PoPs. Figure 5 shows the histogram for degree distribution at the PoP level. We observe that 80.33% of the ASes have less than five connections, whereas only 9.18% of the ASes have more than ten connections. This observation is consistent with the above observation where most ASes primarily reside at the edge of the Internet, not offering internet access to other ASes. They get service from one or two ASes from their single location PoP. Table III shows the summary statistics for the distribution. Only 47 ASes have more than 1000 connections, whereas only 6 ASes have more than $10,000$ connections. The maximum connection count belongs to Cogent Communications (AS174) with $30,225$.

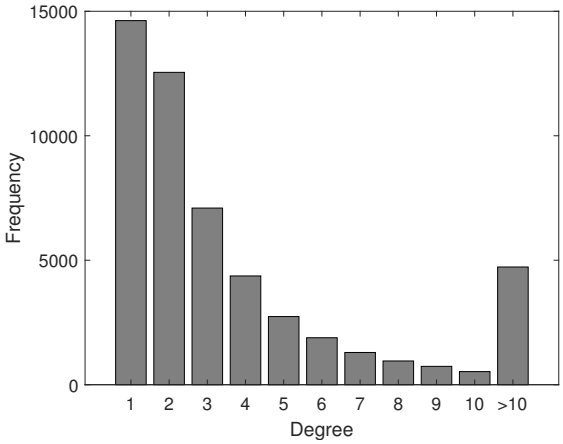


Fig. 5: AS Degree Distribution at PoP Level

TABLE III: Summary statistics for AS Degree Distribution at PoP Level

Level	Q_0	Q_1	Q_2	Q_3	Q_4	*Mean*	*StdDev*
AS Degree	1	1	2	4	30225	11.21	241.67

VI. Conclusions

Point of Presence (PoP) level Internet topology refers to the detailed structure and arrangement of PoPs in the Internet's infrastructure. A PoP is a strategic location where different ASes interconnect to facilitate the exchange of data and traffic. PoP level topology focuses on understanding the layout and interconnections of these PoPs, which serve as crucial nodes in the global network. In this work, we propose a set of techniques to infer the PoP level topology map of the Internet by utilizing several real-life datasets, including traceroute, DNS, and commercial geolocation databases. We validate our results across multiple research and commercial networks around the world and investigate various aspects of the map to offer a deeper understanding of the Internet's backbone structure.

ACKNOWLEDGEMENTS

We are grateful to DB-IP [16], IP2Location [17], and IP-Geolocation.IO [18] for sharing their commercial geolocation databases. We also thank CAIDA, Route Views, and HOIHO projects for publicly sharing their datasets and the ISPs that post their backbone maps on their websites.

REFERENCES

[1] W. Willinger and M. Roughan, "Internet Topology Research Redux", ACM SIGCOMM eBook: Recent Advances in Networking, 2013

[2] A. Y. Nur and M. E. Tozal, "Cross-AS (X-AS) Internet Topology Mapping", Computer Networks, vol. 132, pp. 53-67, 2018

[3] R. Motamedi, R. Rejaie, and W. Willinger, "A Survey of Techniques for Internet Topology Discovery", IEEE Communications Surveys & Tutorials, vol. 17, no 2, pp. 1044–1065, 2015

[4] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with Rocketfuel", IEEE/ACM Transactions On Networking, vol. 12, no 1, pp. 2-16, 2004

[5] A. Y. Nur, "Accuracy and Coverage Analysis of IP Geolocation Databases", IEEE International Balkan Conference on Communications and Networking (BalkanCom), 2023

[6] H. V. Madhyastha, T. Isdal, M. Piatek, C. Dixon, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "iPlane: An Information Plane for Distributed Services", OSDI, 2006

[7] M. E. Tozal, "Autonomous System Ranking by Topological Characteristics: A Comparative Study", IEEE Systems Conference, 2017

[8] D. Feldman, Y. Shavitt, and N. Zilberman, "A Structural Approach for PoP Geo-location", Computer Networks, vol. 56 no. 3, 2012

[9] Y. Shavitt and N. Zilberman, "Improving IP Geolocation by Crawling the Internet PoP Level Graph", IEEE IFIP Networking Conference, 2013

[10] A. Y. Nur and M. E. Tozal, "Identifying Critical Autonomous Systems in the Internet", Springer Journal of Supercomputing, SI: Cyber Threats against Critical Infrastructure and their Countermeasures, vol. 74, no 10, 2018

[11] S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, "The Internet Topology Zoo", Journal on Selected Areas in Communications, vol. 29, no 9, pp. 1765-1775, 2011

[12] R. Durairajan, S. Ghosh, X. Tang, P. Barford, and B. Eriksson, "Internet Atlas: a Geographic Database of the Internet", ACM Workshop on HotPlanet, 2013

[13] A. Y. Nur and M. E. Tozal, "Geography and Routing in the Internet", ACM Transactions on Spatial Algorithms and Systems, vol. 4, no. 4, pp. 1-16, 2018

[14] M. Luckie, B. Huffaker, A. Marder, Z. Bischof, M. Fletcher, and Kc Claffy, "Learning to Extract Geographic Information from Internet Router Hostnames", ACM SIGCOMM Conference on emerging Networking EXperiments and Technologies (CoNEXT), December 2021

[15] M. E. Tozal, "The Internet: A System of Interconnected Autonomous Systems", IEEE Systems Conference, 2016

[16] DB-IP Geolocation Database - October 2023 - https://www.db-ip.com/

[17] LITE Geolocation Database - October 2023 - http://lite.ip2location.com/

[18] IPGeolocation.IO Geolocation Database - October 2023 - https://ipgeolocation.io/

[19] The CAIDA UCSD IPv4 Routed /24 DNS Names Dataset - October 2023 - https://www.caida.org/catalog/datasets/ipv4_dnsnames_dataset/

[20] The CAIDA UCSD IPv4 Prefix-Probing Dataset - October 2023 - https://www.caida.org/catalog/datasets/ipv4_prefix_probing_dataset/

[21] The CAIDA UCSD IXPs Dataset, 2023-10 - https://www.caida.org/catalog/datasets/ixps/

[22] The CAIDA UCSD Routeviews Prefix to AS mappings Dataset (pfx2as) for IPv4 and IPv6, 2023-10 - https://www.caida.org/catalog/datasets/routeviews-prefix2as

[23] RIPE Atlas - October 2023 - https://atlas.ripe.net/

[24] Measurement Lab (M-Lab) - October 2023 - https://www.measurementlab.net/

[25] Gephi - https://gephi.org/

[26] CAIDA Hoiho (Holistic Orthography of Internet Hostname Observations) API - https://api.hoiho.caida.org/