Accuracy and Coverage Analysis of IP Geolocation Databases

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

***Abstract*—Identifying the geographical location of Internet hosts is crucial for researchers, governments, and commercial entities. While public and commercial geolocation services are commonly employed for this task, their accuracy in locating Internet hosts remains questionable. This paper studies the accuracy and coverage of four popular geolocation databases; MaxMind, DBIP, IP2Location, and IPGeolocationIO. We assess the consistency and comprehensiveness of these services by analyzing the entire IPv4 space. Furthermore, we investigate the issue at the prefix level since geolocation databases typically provide location data at that level. Finally, we create a ground truth dataset by employing a DNS-based approach and publicly available vendor locations to evaluate the accuracy of the databases. Our findings indicate that these databases provide comprehensive coverage, whereas their accuracy is far from satisfactory. Therefore, it is essential to use the information obtained from these databases cautiously and verify its accuracy before making any decisions based on it.**

Index terms— IP Geolocation; Geolocation Databases; Geographic Information Systems; Accuracy; Reliability

I. Introduction

The Internet is one of the largest human-engineered, decentralized network of networks serving billions of people worldwide. It is the primary communication medium for critical infrastructures such as electricity, finance, and transportation. With the emergence of innovative applications and technologies, such as cloud computing and the Internet of Things, it is evident that this trend will only continue to expand in the future.

Internet Protocol (IP) to geolocation refers to the process of mapping an IP address to a physical location of the device using that address. It is a challenging task because IP protocol does not provide any geolocation information. Researchers suggest delay-based geolocation [3, 4] and topology-based geolocation [5]. Delay-based algorithms typically use latency metrics collected from known geographically distributed locations to locate target hosts. Topology-based algorithms extend the delay-based techniques by considering the topology with an assumption that the topologically close addresses are also physically close. Moreover, several commercial geolocation databases in the market combine several methods to increase their accuracy and sell their databases [24, 25].

Identifying the geographical location of Internet hosts is crucial for researchers, governments, and commercial entities. Specifically, geolocation is used by many applications, including content personalization [10], advertising [14], e-commerce [15], content delivery networks [12], credit card fraud protection [13], and law enforcement [11]. Additionally, understanding the geographical characteristics of the Internet infrastructure allows us to utilize resourceful paths during or after natural disasters [6]; improve the inter-domain routing processes [7]; deploy geography-aware network overlays for efficient multimedia communications [18]; predict path latency for service improvement in the Internet [8]; develop more realistic Internet topology generation tools [19].

IP geolocation services map IP addresses to their physical locations such as a country, city, and/or geographic coordinates. In addition to the challenges of maintaining and updating them, the accuracy of geolocation databases is highly questionable [16, 17], particularly due to the absence of information about the techniques used to construct them. In numerous instances, the geolocation service providers are the only source of information regarding the accuracy of their databases. Some vendors declare accuracy metrics without disclosing the methods used to obtain them.

In this paper, we study the accuracy and coverage of four popular geolocation databases; MaxMind [25], DBIP [24], IP2Location [26], and IPGeolocationIO [27]. Our study reveals that the databases cover the majority of the IPv4 space (more than 85%). However, they exhibit numerous inconsistencies when we compare their results pairwise. We observe that they have a 620 km distance discrepancy on average. The primary challenge in conducting such research is the scarcity of ground truth reference information, specifically a comprehensive and diverse collection of IP addresses with established geographical locations to compare with geolocation databases. We create a ground truth dataset using a DNS-based approach [2] and publicly available vendor locations. The ground truth dataset contains $6,345,323$ unique IP addresses with their locations. We use the ground truth dataset to evaluate the accuracy of the databases. Our results show that the four databases' average distance discrepancy mean is 376 km.

The rest of the paper is organized as follows. Section II presents the related work. We introduce the details of our ground truth dataset creation in Section III. Section IV demonstrates our experimental results and comparisons. Finally, Section V concludes the paper.

II. Related Work

Mapping IP addresses to their physical location is a significant task for several reasons, including location-based content delivery, advertising and marketing, fraud prevention, and assistance to law enforcement. Content providers can deliver location-specific content to users based on their location using IP geolocation [10]. For instance, news platforms can provide users with news tailored to their specific location based on the users' IP addresses. Additionally, location information helps some companies to deal with copyright and licensing agreements that limit the availability of certain titles based on geographical region. For example, Netflix users in the United States see a different selection of content than those in another country. Businesses can utilize IP geolocation for targeted advertising [14]. Businesses may send localized marketing and promotions to their clients by knowing where they are. For instance, users located in New Orleans may see advertisements for local dining establishments, shops, or tourist attractions. On the other hand, if the users live in a different city, they would likely come across advertisements specific to that area. IP geolocation can be used for detecting and preventing fraudulent activities, such as credit card fraud [13]. Specifically, credit card vendors can utilize geolocation information to detect anomalies and determine whether a transaction is legitimate. Finally, IP geolocation can play a role in law enforcement by providing information about the location of a device accessing the Internet [11]. This information can be used to track down individuals engaging in illegal activities online.

Researchers suggested several methods for IP geolocation by utilizing network measurement and Internet data mining approaches. Network measurement methods use delay and network topology information, whereas Internet data mining approaches use diverse information mined from the Internet, including WHOIS databases, reverse DNS, and public vendor locations. Padmanabhan and Subramanian [3] provided a technique that involves sending ICMP packets from landmark servers across different geographic locations to the target IP address, where the location of the target IP is then estimated based on the proximity of the closest landmark server in terms of latency. Gueye et al. [4] propose constraint-based geolocation technique that estimates a position using a sufficient number of distances to some fixed points. Katz-Bassett et al. [5] propose topology-based geolocation by leveraging network topology along with network delay measurements, using traceroute queries from landmark servers to the IP target.

DNS-based geolocation methods use geographic hints encoded in domain names to infer locations. UNDNS [2] is one of the most popular DNS decoders, which is a database of regular expressions that have been manually compiled to extract geographical hints and other relevant details from hostnames. DRoP [20] determines the geographic location of hostnames by utilizing rules that are automatically generated by identifying patterns across all the hostname terms associated with a given domain. Dan et al. [21] provide a machine learning approach to IP geolocation using reverse DNS hostnames. Their method divides the hostname into individual terms, compares them to a geolocation dictionary to create a set of characteristics, and subsequently employs a binary classifier to analyze the hostname and features obtained.

Several research studies have indicated that public and commercial databases offer low-resolution geolocation and are not dependable regarding city-level accuracy. Hufaker et al. [22] utilized a majority vote system across all participating databases to determine the location of an IP address block and evaluated the databases based on the resultant location. Shavitt and Zilberman [23] assessed the consistency of databases by using a ground truth dataset of IP addresses with verified Points of Presences. Gharaibeh et al. [17] analyzed the reliability of router geolocation by using 1.6 million router interface IP addresses and a ground-truth dataset of 16,586 router interface IP addresses. Their findings indicate that the databases' accuracy at the country and city levels need improvement because they are inadequate for geolocating routers correctly.

This work analyzes the accuracy and coverage of four major commercial geolocation databases. We examine their consistency and coverage using the entire IPv4 space (more than 4 billion IP addresses). Additionally, we analyze the problem at the prefix level since geolocation databases provide locations at the prefix level. Finally, we create $6,345,323$ IP addresses as a ground truth and analyze the accuracy of these databases.

III. Ground-Truth Location Dataset

A. Vendor location based Geolocation

Some organizations provide a global research network that allows researchers to develop, deploy, and test new network services and applications on a large-scale, geographically distributed platform. Volunteer organizations worldwide join these types of networks and make the network globally distributed. In this work, we use RIPE Atlas [29] and Measurement Lab (M-Lab) [30] nodes.

RIPE NCC is the regional Internet registry (RIR) for Europe, the Middle East, and parts of Central Asia. They created RIPE Atlas to provide a worldwide collection of probes that gauge the connectivity and reachability measurements of the Internet in real-time. Volunteers around the world deploy RIPE probes or RIPE anchors to their own networks. RIPE probes are compact hardware devices powered by USB that users connect to the Ethernet port on their router. By the time this paper is written, there are $11,981$ probes available. RIPE anchors are a combination of RIPE probes with increased measurement capabilities, and regional measurement targets that are part of the more extensive RIPE Atlas network. By the time this paper is written, there are 802 anchors available. For example, DigitalOcean (AS number 14061) provides an anchor with an IP address of 104.131.160.184, which is located in New York City.

Similarly, Measurement Lab (M-Lab) is an open, distributed platform that provides researchers, developers, and the general public with an easy way to measure and diagnose the performance of their internet connections. They provide 195 nodes

TABLE I: General Characteristics of the Databases

Database	Prefix Count	Country	City	Coordinates
IPGeolocationIO	4,786,915	249	66,442	891,840
DBIP	3,304,193	243	118,599	364,360
IP2Location	3,123,918	243	72,026	96,879
MaxMind	3,422,806	244	97,735	131,613

from 66 different cities around the world. For example, TATA Communications (AS number 6453) provides a node with an IP address of 63.243.240.78, which is located in Los Angeles.

We collect $24,810$ IP addresses from RIPE probes, 1176 IP addresses from RIPE anchors, and 1746 IP addresses from M-Lab. We observe 1159 IP addresses in both RIPE anchors and RIPE probes set. In total, we collect $26,573$ unique IPv4 addresses with their location.

B. DNS-based Geolocation

Autonomous Systems (ASes) typically encode geographic information in their DNS naming conventions. Although DNS naming usage is not mandatory, it is still one of the most valuable sources of information directly from the ASes. DNS-based geolocation methods use geographic hints encoded in domain names to infer locations. To illustrate, Comcast uses the naming convention *"te-0-0-0-5-sur03.chicago302.**il.chicago**.comcast.net"* which denotes the location of Chicago, Illinois.

We use UNDNS, which is a tool for extracting geolocation information from DNS names [2]. In our previous work [1], we updated and improved their key dataset to extend the coverage and accuracy of the DNS names. Note that UNDNS provides city and country names and does not provide coordinates. In order to receive coordinates from cities and countries, we use Google Geocoding API [31]

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 [28]. In the dataset, we have $39,228,837$ unique DNS names. UNDNS was able to obtain a valid geolocation for $6,318,932$ DNS entries. In the DNS geolocation list, we have 182 IP addresses with locations that are the same as the vendor list. Therefore, we obtain a total of $6,345,323$ unique IP addresses with their locations in our final ground truth dataset.

IV. Experimental Results

A. Database Overview

In this work, we use one commercial database (IPGeolocationIO) and three freely available databases (DBIP, IP2Location, and MaxMind). Unfortunately, none of the databases share their creation process. Their database contains entries with an IP address block *(e.g., 8.21.216.0, 8.21.216.255, or 8.21.216.0/24 prefix)*, several useful information associated with the block, such as country code, city, latitude, and longitude. For example, DBIP has the following entry in their database: *"8.21.216.0, 8.21.216.255, NA, US, Louisiana, New Orleans, 29.9511, -90.0715"*.

TABLE II: Database Coverage

Database	Missing Prefix Count	Missing IP Location Count	Missing IP Location Percentage
IPGeolocationIO	33	320,025,007	7.451%
DBIP	53	592,718,656	13.8%
IP2Location	3511	604,528,896	14.075%
MaxMind	3356	606,552,311	14.122%

Table I shows the general characteristics of the given databases. All four databases use ISO 3166-1 Alpha-2 country code for representing countries. There are 249 countries in Alpha-2 representation. IPGeolocationIO contains at least one IP block for each country, whereas MaxMind does not have an entry for five, and DBIP and IP2Location do not have an entry for six countries. For counting the unique coordinates, we rounded the latitude and longitude values to 3 decimal places, which gives 0.1 km accuracy. Our observations show that IPGeolocationIO contains the most unique coordinates, whereas IP2Location has the minimum values. Note that these values do not present accuracy.

B. Coverage Analysis

In this subsection, we analyze the coverage of four databases regardless of their correctness. Table II presents the missing prefixes and missing IP location counts. Our observations show that IPGeolocationIO misses only 33 prefixes with 7.451% of the IPv4 geolocation, whereas MaxMind misses 3356 prefixes with 14.122% of the IPv4 geolocation. Note that we check all possible IP addresses in IPv4 space, which is 2^{32} corresponding to $4,294,967,296$ unique IP addresses. Nearly 18 million IP addresses are reserved for private networks, and these IP blocks do not have valid geolocation in the databases. Even though it does not negatively affect their coverage, we put them in the missing IP location class.

C. Pairwise Analysis for entire IPv4

In this part, we compare each database pair's consistency with respect to distance discrepancy. We check each IPv4 addresses location in both databases and calculate the distance between their location. Note that the location of the IP address might be incorrect in both databases, correct in one of the databases, or correct in both. For this comparison, our focus is to check the consistency between databases, and the correctness of the location is not considered.

Distance between two locations are calculated by using Haversine formula which calculates the *great-circle distance* between two points on the surface of a sphere as suggested in [9]. Haversine formula requires two latitude and longitude pairs to compute the distance between them as shown in Equation 1.

$$\begin{aligned} a &= \sin^2\left(\frac{\phi_2-\phi_1}{2}\right)+\cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right) \\ c &= 2arcsin(\sqrt{a}) \\ d &= Rc \end{aligned} \tag{1}$$

where 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)

TABLE III: Distance Discrepancy Between Database Pairs (All IPv4 Address Space)

Database Pair	NA	[0-50]	(50-100]	(100-500]	(500-1000]	(1000-5000]	(5000-10000]	(10000-20000)
DBIP - IP2Location	604,538,688	2,365,689,331	133,830,516	534,390,632	151,668,019	341,294,973	148,673,068	14,882,069
	14.08%	55.08%	3.12%	12.44%	3.53%	7.95%	3.46%	0.35%
DBIP - IPGeolocationIO	592,735,790	2,324,329,277	114,562,202	569,818,688	150,193,737	367,815,423	154,934,574	20,577,605
	13.80%	54.12%	2.67%	13.27%	3.50%	8.56%	3.61%	0.48%
DBIP - MaxMind	606,562,103	1,218,031,227	144,166,629	641,207,622	525,602,246	988,410,849	154,080,532	16,906,088
	14.12%	28.36%	3.36%	14.93%	12.24%	23.01%	3.59%	0.39%
IP2Location - IPGeolocationIO	604,545,793	2,558,782,455	129,227,732	527,565,715	150,778,376	284,573,277	25,043,536	14,450,412
	14.08%	59.58%	3.01%	12.28%	3.51%	6.63%	0.58%	0.34%
IP2Location - MaxMind	609,290,959	1,287,448,723	152,153,011	597,749,849	469,627,305	1,128,616,275	34,922,087	15,159,087
	14.19%	29.98%	3.54%	13.92%	10.93%	26.28%	0.81%	0.35%
IPGeolocationIO - MaxMind	606,569,208	1,122,022,035	146,886,889	702,221,727	467,161,664	1,194,610,344	38,839,202	16,656,227
	14.12%	26.12%	3.42%	16.35%	10.88%	27.81%	0.90%	0.39%

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.

Table III presents the distance discrepancy between each database pair. In case one of the databases could not locate an IP address, we put it in the NA (not available) class. The table shows that DBIP, IP2Location, and IPGeolocationIO agree with each other more than they agree with MaxMind. When we check a distance between 0 to 50 km, these three databases have around 55.26% on average. However, their pairwise comparison with MaxMind gives 28.15% on average for the same distance range.

DBIP-IP2Location: We observe that at least one database could not locate around 604 million IP addresses corresponding to 14.08% in the entire IPv4 space. Moreover, around 2.3 billion IP addresses are located within a 50 km distance, corresponding to 55.08% in the entire IPv4 space. Interestingly, 15.29% of the IP addresses are located more than 500 km distance. The maximum distance discrepancy is $19,727$ km for prefix 66.198.44.0/24. DBIP located the prefix in Quito, Ecuador (-0.202, -78.494), whereas IP2Location located it in Singapore (1.289, 103.850).

DBIP-IPGeolocationIO: Our observations show that around 2.3 billion IP addresses are located within a 50 km distance, corresponding to 54.12% in the entire IPv4 space. Comparing the DBIP-IPGeolocationIO pair with DBIP-IP2Location, DBIP-IPGeolocationIO has more than 6 million IP addresses in the 10000-20000 km discrepancy range. The maximum distance discrepancy is $19,910$ km for prefix 167.114.26.40/29. DBIP located the prefix in Jakarta, Indonesia (-6.176, 106.857), whereas IP2Location located it in Santander, Colombia (7.124,-73.109).

DBIP-MaxMind: Comparing DBIP with the other two databases, DBIP has the most discrepancy with MaxMind. Only 1.2 billion of the IP addresses are located within a 50 km distance, whereas the overall average discrepancy is 864 km. The maximum distance discrepancy is $19,750$ km for prefix 45.138.10.232/30. DBIP located the prefix in Shire of Cocos, West Island (-12.145,96.821), whereas MaxMind located it in San Jose, Costa Rica (9.933,-84.084).

IP2Location-IPGeolocationIO: Our observations show that these two pairs had the most agreement, where 2.5 billion locations are within 50 km. The overall average distance discrepancy is 296 km. The maximum distance discrepancy is $19,732$ km for the 168.205.92.34 IP address. IP2Location located the IP address in Buenos Aires, Argentina (-34.603, -58.381), whereas IPGeolocationIO located it in Nantong, China (32.078, 121.260).

IP2Location-MaxMind:

As we stated above, database comparisons with MaxMind have the most disagreement in locations with an overall 760 km average distance. Between MaxMind pairs, this pair has the lowest average distance discrepancy with 684 km. The maximum distance discrepancy is $19,665$ km for prefix 161.123.66.0/24. IP2Location located the prefix in Auckland, New Zealand (-36.866, 174.766), whereas MaxMind located it in Rabat, Morocco (34.012, -6.848).

IPGeolocationIO-MaxMind: This pair has the lowest agreement within the 50 km range with only 26.12%. Additionally, 16.35% disagreement in the 100-500 km range and 27.81% agreement in the 1000-5000 km range are the highest numbers compared to other pairs. The maximum distance discrepancy is $19,899$ km for prefix 77.81.118.64/30. IPGeolocationIO located the prefix in Hamilton, New Zealand (-37.763,175.246), whereas MaxMind located it in Seville, Spain (37.384,-5.970).

D. Pairwise Analysis for Prefixes

Geolocation databases provide geolocation information for IP blocks. For example, MaxMind provides geolocation information of the prefix $1.0.0.0/24$ as $[-37.8333, 145.2375]$ latitude and longitude. In this part, we check each database pair's distance discrepancy regarding the prefixes. Figure 1 shows the cumulative distribution function (CDF) of the database pairs with respect to distance discrepancy. In addition, Table IV shows the minimum, first quartile, second quartile (median), third quartile, maximum, mean, and standard deviation of the distribution. Note that -1 represents not available locations. It is evident that the database pairs have a significant discrepancy

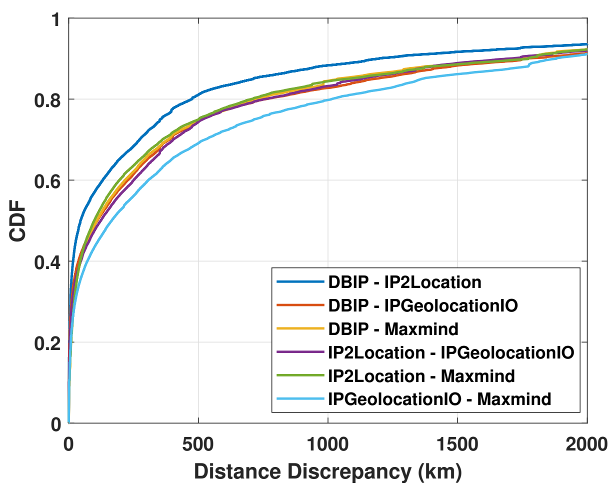

Fig. 1: Pairwise Databases Distance discrepancy CDF (Zoom in to 2000 km)

TABLE IV: Summary Statistics for Distance Discrepancy in km Between Database Pairs (Prefixes)

Database Pair	Q0	Q1	Q2	Q3	Q4	Mean	StdDev
DBIP - IP2Location	-1	4	45	353	19728	511.31	1408.89
DBIP - IPGeolocationIO	-1	6	110	508	19910	642.63	1526.06
DBIP - MaxMind	-1	13	108	513	19750	603.69	1466.04
IP2Location - IPGeolocationIO	-1	7	124	516	19732	596.92	1385.16
IP2Location - MaxMind	-1	15	98	501	19666	569.07	1400.56
IPGeolocationIO - MaxMind	-1	14	166	715	19900	670.53	1445.13

with a mean between 511 to 670 km. Additionally, the third quartile (75%) shows us that the discrepancy is between 353 to 715, with an average of 517 km.

E. Database Coverage and Accuracy over the Ground Truth

This subsection discusses the accuracy of four databases with respect to the ground truth dataset that contains $6,345,323$ unique IP addresses with their locations. In order to assess the accuracy, we use distance discrepancy as our metric. For each IP address in the ground truth database, we check the location provided by databases, then find the distance between the ground truth location and database location. Figure 2 shows the histogram of the distance discrepancy. We observe that all four databases cover the majority of the IP addresses in the ground truth. DBIP and IPGeolocationIO could not locate 2 IP addresses, IP2Location 14 and MaxMind 19 IP addresses. IPGeolocationIO located 71.05% of the IP addresses (4,508,464)

TABLE V: Summary Statistics for Distance Discrepancy in km Between Database and Ground Truth

Database	Q0	Q1	Q2	Q3	Q4	Mean	StdDev
DBIP	-1	0.9	16.53	349.36	19821.64	316.89	777.86
IP2Location	-1	1.9	83.97	514.64	19820.02	423.43	833.45
IPGeolocationIO	-1	2.21	4.98	205.63	19820.56	334.46	1025.41
MaxMind	-1	15.98	146.21	617.88	19834.57	432.20	817.31

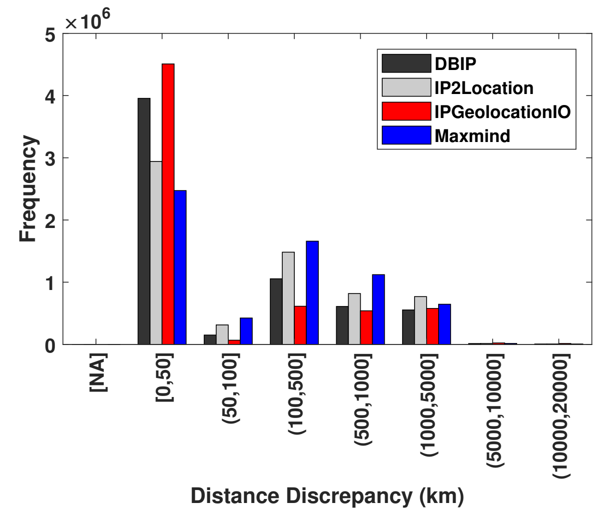

Fig. 2: Databases vs Ground Truth Distance Discrepancy

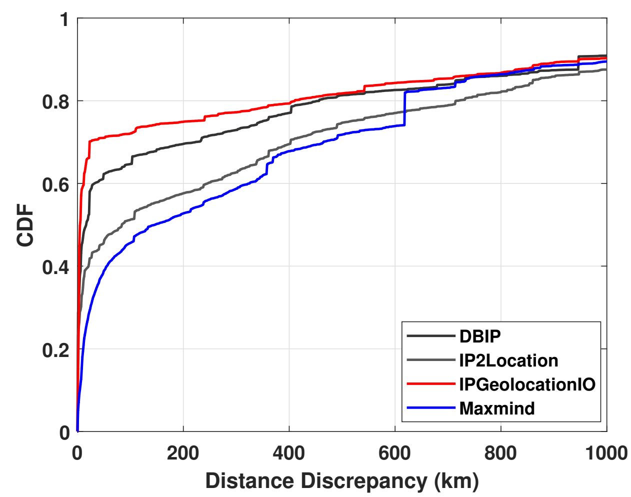

Fig. 3: Databases vs Ground Truth Distance Discrepancy CDF (Zoom in to 1000 km)

within a 50 km distance. On the other hand, MaxMind has the lowest accuracy within the 50 km range, where MaxMind located 38.97% IP addresses (2,472,911). All four databases locate around 25,000 IP addresses with more than 5000 km discrepancy from the ground truth location.

Figure 3 shows the cumulative distribution function (CDF) of the database pairs with respect to distance discrepancy. In addition, Table V shows the minimum, first quartile, second quartile (median), third quartile, maximum, mean, and standard deviation of the distribution. DBIP has the lowest mean at 316.89 km, and MaxMind has the highest mean at 432.2 km. Even though IPGeolocationIO located most IP addresses within a 50 km distance, their distance discrepancy mean is 423.43 km. The reason is that they located 23,215 IP addresses in 5000-10000 km interval and 14,719 IP addresses in 10000-20000 km interval. These numbers are the highest within all four databases.

V. Conclusions

In this paper, we evaluate the coverage and accuracy of four widely used IP geolocation databases. We examine their consistency and coverage using the entire IPv4 space and prefix level. Our pairwise comparisons show that databases have significant disagreement providing locations with a 620 km overall average distance discrepancy. Additionally, we create $6,345,323$ IP addresses as ground truth and analyze the accuracy of these databases. Our results show that most databases provide comprehensive coverage over IPv4 space. However, our findings indicate that the accuracy of these databases is questionable. Therefore, it is essential to use the information obtained from these databases with caution and verify its accuracy before making any decisions based on it.

Acknowledgements

We are grateful to IPGeolocationIO for providing their commercial geolocation database. We are also thankful to CAIDA, MaxMind, DBIP, and IP2Location projects for publicly sharing their datasets.

References

[1] A. Y. Nur and M. E. Tozal, "Cross-AS (X-AS) Internet Topology Mapping", Computer Networks, Vol. 132, 2018

[2] N. Spring, R. Mahajan, and D. Wetherall, "Measuring ISP Topologies with Rocketfuel", ACM SIGCOMM, 2002

[3] V. Padmanabhan and L. Subramanian, "An Investigation of Geographic Mapping Techniques for Internet Hosts", ACM SIGCOMM, 2001

[4] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, "Constraint-based Geolocation of Internet Hosts", IEEE/ACM Transactions on Networking, 2006

[5] E. Katz-Bassett, J. P. John, A. Krishnamurthy, D. Wetherall, T. Anderson, and Y. Chawathe, "Towards IP Geolocation Using Delay and Topology Measurements", Internet Measurement Conference (IMC), 2006

[6] J. Wu, Y. Zhang, Z. M. Mao, and K. G. Shin, "Internet Routing Resilience to Failures: Analysis and Implications", ACM CoNEXT, 2007

[7] R. Oliveira, M. Lad, B. Zhang, and L. Zhang, "Geographically Informed Inter-Domain Routing", IEEE International Conference on Network Protocols, 2007

[8] H. V. Madhyastha, T. Anderson, A. Krishnamurthy, N. Spring, and A. Venkataramani, "A Structural Approach to Latency Prediction", ACM IMC 2006

[9] A. Y. Nur and M. E. Tozal, "Geography and Routing in the Internet", ACM Transactions on Spatial Algorithms and Systems (TSAS), vol. 4.4, 2018

[10] A. Hannak, P. Sapiezynski, A. M. Kakhki, B. Krishnamurthy, D. Lazer, A. Mislove, and C. Wilson, "Measuring Personalization of Web Search", ACM International Conference on World Wide Web, 2013

[11] C. A. Shue, N. Paul, and C. R. Taylor, "From an IP Address to a Street Address: Using Wireless Signals to Locate a Target", Proceedings of the Workshop on Offensive Technologies (WOOT), USENIX, 2013

[12] C. Huang, D. A. Maltz, J. Li, and A. Greenberg, "Public DNS System and Global Traffic Management", IEEE INFOCOM , 2011

[13] J. Akhilomen, "Data Mining Application for Cyber Credit-card Fraud Detection System", Springer Proceedings of the Industrial Conference on Data Mining, 2013

[14] G. C. Bruner and A. Kumar, "Attitude Toward Location-based Advertising", Journal of Interactive Advertising, vol. 7.2, 2007

[15] D. J. B. Svantesson, "E-commerce Tax: How the Taxman Brought Geography to the'borderless' Internet", Revenue Law Journal, vol. 17, 2007

[16] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye, "IP Geolocation Databases: Unreliable?", SIGCOMM Computer Communication Review, vol. 41.2, 2011

[17] M. Gharaibeh, A. Shah, B. Huffaker, H. Zhang, R. Ensafi, and C. Papadopoulos, "A Look at Router Geolocation in Public and Commercial Databases", ACM IMC, 2017

[18] A. Elmokashfi, E. Myakotnykh, J. M. Evang, A. Kvalbein, and T. Cicic, "Geography Matters: Building an Efficient Transport Network for a Better Video Conferencing Experience", ACM Conference on Emerging Networking Experiments and Technologies, 2013

[19] B. Quoitin, V. Van den Schrieck, P. François, and O. Bonaventure, "IGen: Generation of Router-Level Internet Topologies through Network Design Heuristics", IEEE International Teletraffic Congress, 2009

[20] B. Huffaker, M. Fomenkov, and K. C. Claffy, "DRoP: DNS-based Router Positioning", ACM SIGCOMM Computer Communication Review, vol. 44.3, 2014

[21] O. Dan, V. Parikh, and B. D. Davison, "IP Geolocation through Reverse DNS", ACM Transactions on Internet Technology (TOIT), vol. 22.1, 2021

[22] B. Hufaker, M. Fomenkov, and KC clafy, "Geocompare: a Comparison of Public and Commercial Geolocation Databases", Technical Report, CAIDA, 2011

[23] Y. Shavitt and N. Zilberman, "A Geolocation Databases Study", IEEE Journal on Selected Areas in Communications, vol. 29.10, 2011

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

[25] GeoLite2 Geolocation Database - January 2023 - https://www.maxmind.com

[26] LITE Geolocation Database - January 2023 - https://lite.ip2location.com

[27] Ip Geolocation IO Geolocation Database - January 2023 - https://ipgeolocation.io/

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

[29] RIPE Atlas - January 2023 - https://atlas.ripe.net/

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

[31] Geocoding API - https://developers.google.com/maps/documentation/geocoding