IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received March 19, 2019, accepted April 1, 2019, date of publication April 11, 2019, date of current version April 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2909691

Using RIPE Atlas for Geolocating IP Infrastructure

MASSIMO CANDELA[1], ENRICO GREGORI[2], VALERIO LUCONI[2], AND ALESSIO VECCHIO[3], (Member, IEEE)
[1]RIPE NCC, 1012AB Amsterdam, The Netherlands
[2]IIT-CNR, 56124 Pisa, Italy
[3]University of Pisa, 56122 Pisa, Italy

Corresponding author: Massimo Candela (massimo.candela@ing.unipi.it)

This work was supported in part by the University of Pisa, under Grant PRA_2017_37, and in part by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence).

ABSTRACT The vast majority of studies on IP geolocation focuses on localizing the end-users, and little attention has been devoted to localizing the elements of the Internet infrastructure, i.e., the routers and servers that make the Internet work. In this paper, we study the maximum theoretical accuracy that can be achieved by a geolocation approach aimed at geolocating the Internet infrastructure. In particular, we study the effects on localization accuracy produced by the position of landmarks and by the strategy followed for their enrollment. We compare two main approaches: the first is more centralized and controlled, and uses well-connected machines belonging to the infrastructure as landmarks; the second is more distributed and scalable and is based on landmarks positioned at the edge of the network. The study is based on an extensive set of measurements collected using the RIPE Atlas platform. The results show that the uniform and widespread diffusion of landmarks can be as important as their measurement accuracy. The study is carried out at both the worldwide and regional scale, including regions that were scarcely observed in the past. The results highlight that the geographical characteristics of the Internet paths are dependent on the considered region, thus suggesting the use of specifically calibrated models. Finally, the study shows that geolocating IP infrastructure with active measurements is feasible in terms of precision and scalability of the overall system.

INDEX TERMS Internet, IP geolocation, network topology, wide area networks.

I. INTRODUCTION

Estimating the geographic position of a device connected to the Internet is important in many situations: automatic language selection, offering of specific digital content, tax and regulation enforcement are just a few of the numerous possible examples. Thus, in the last years, IP geolocation systems have been adopted for finding the position of end-users starting from their IP addresses.

More recently, researchers, operators, and law enforcement became interested in finding the position of the machines that compose the infrastructure of the Internet. In this context, geographical information can be fundamental for troubleshooting network malfunctions, for studying the impact of changes in network configurations, for understanding the relations between geography and Autonomous Systems (ASes), or for evaluating the amount of local traffic that passes through another country. For instance, the relations between geography and intra-AS and inter-AS routing policies was studied using position information of internal and border routers [1]. Results show that a significant fraction of ingress-to-egress subpaths approximately follow the geographical-shortest path, thus indicating efficiency of infrastructure design and routing policies within Internet Service Providers (ISPs). Geolocation of infrastructural nodes was also used to verify the location of the servers of some large proxy service providers [2], demonstrating that one third of commercial proxies are located in a country that does not correspond to the advertised one. In [3] and [4], geolocation of infrastructural nodes was used to infer whether the traffic between end-points in the same country remains within its borders or flows through other countries. Similarly, the servers running web tracking and advertisement systems were geolocated to quantify the amount of traffic that crosses the data protection borders [5]. Thanks to an infrastructure-specific geolocation system [6], it was possible to conclude that 90% of tracking of European citizens is carried out inside the EU28 borders.

The associate editor coordinating the review of this manuscript and approving it for publication was Ting Wang.

Various commercial services offer APIs that can be used to retrieve the location associated to a given IP address. These services generally rely on data offered by Regional Internet Registries (RIRs). Registries, in turn, are based on administrative information about the company that owns the IP address [7]. As a consequence, the reported position can be rather far from the actual one, especially for large organizations.

An alternative approach is followed by active IP geolocation techniques, which use latency measurements to determine the position of hosts. A set of machines with known location, called *landmarks*, are used to measure the network delay between themselves and the *target* host. Measured delays are converted into physical distances and then used to estimate the position of the target.

In this paper we do not describe yet another geolocation system. Instead, we study and discuss the maximum theoretical accuracy that can be achieved by any geolocation approach based on a world-scale measurement platform, such as RIPE Atlas, and aimed at geolocating the Internet infrastructure. With Internet infrastructure we mean the set of routers and servers not belonging to end-user networks. In this broad definition are included, for instance, the routers of ISPs and IXPs, DNS servers, and content providers. Differently from the existing literature our study is based on the following main elements:

- We evaluate the effects on localization accuracy produced by the position of landmarks and by the strategy adopted for their enrollment. In detail, we evaluate two radically different approaches. In the first, geolocation takes place by means of landmarks that, like targets, belong to the Internet infrastructure. This strategy corresponds to a more centralized and controlled approach, where landmarks are well-connected and well-positioned machines but limited in number. In the second, targets are geolocated by means of nodes at the fringes of the Internet, mostly hosted in domestic environments. This corresponds to a community-based approach, where landmarks are machines with reduced requirements, accessing the Internet via residential connections, but available in great quantity. The two approaches also differ in terms of expected scalability and precision of measurements. The first approach is expected to be more precise but less scalable, and the opposite applies to the second.
- We use the RIPE Atlas system, which is one of the largest open Internet measurement platforms available nowadays [8]. The study thus relies on a set of measurements that is significantly larger than the ones used in previous literature, both in terms of amount of data and spatial distribution.
- The study is carried out across different geographical scales. We first focus on a worldwide scale, then, thanks to the pervasiveness of RIPE Atlas, we are able to deepen the analysis for different regions. We considered seven regions: Europe, North America, South America, Asia, Oceania, Africa, and Middle East. This is a significant advance with respect to existing literature, where the focus was generally on just a single region.

Our analysis of the maximum accuracy is first conducted by assuming a uniform distribution of the infrastructure IP addresses around the globe (and in each region). We then refine our study by assuming a more realistic distribution, with infrastructural IP machines placed at the major urban areas of the world.

The rest of the paper is structured as follows: Section II summarizes existing literature on IP geolocation; Section III describes the adopted methodology; in Section IV, the data collection phase is described; Section V illustrates general results, whereas in Section VI results for the different regions are presented; Section VII presents results obtained when the distribution of infrastructural elements is not uniform; finally, Section VIII concludes the paper.

II. RELATED WORK

A considerable amount of work has been done about IP geolocation and the estimation of physical distances via latency measurements during the last years. The basic approach to IP geolocation relies on a set of active landmarks (hosts with known position) for collecting end-to-end delays towards the target IP address. Delay values are then used to compute the position of the target according to a variety of techniques. One of the simplest methods co-locates the target with the landmark reporting the lowest latency value, in this case the set of possible position estimates is discrete and restricted to the positions of the landmarks [9]. More sophisticated techniques are multilateration-based and in some cases rely on the definition of constraints that can be used to restrict the position of the target to a specific area [10]–[13]. Other interesting works exploit alternative delay-distance conversion techniques [14]–[16], or try to define delay-distance models that are tailored to not-richly connected regions [17]. Similarly, statistical approaches can be used to define the probability density of the position of the target [18]–[22]. This can be useful also for understanding the quality of the estimated position.

More complex measurement methods have also been subject to investigation. In particular, the use of traceroute-based measurements has been considered to include the number of intermediate hops in the geolocation process [20], [23]–[26]. Combinations and improvements of the above methods can be found in [11], [27], [28]. Some other works also considered the possibility of introducing nodes with known location as passive landmarks (e.g. web servers) to improve localization accuracy [12], [23], [29]–[31]. In [32], the operator's Points of Presence (PoP) available in various cities are detected with path exploration and collected in a database. Hence, end-users can be geolocated at city-level based on the PoP traversed by active measurements. Another technique, based on two-tier neural networks, first tries to find an area where the target should be located, and then improves the accuracy by exploiting just the landmarks in that area [33].

The amount of landmarks used in previous studies is in general in the order of few dozens, with some exceptions in the order of a few hundreds. The most used platform is PlanetLab [34], as it provides a convenient platform for running distributed experiments.

Various geolocation databases have been evaluated, including commercial ones, and reported as not accurate [7], [35]–[38]. In [39], authors use geolocation information provided by a popular broadband performance tool to populate a database of IP locations in South Korea, showing an increase of precision with respect to common commercial databases.

Among the above mentioned works, only [24], [28], [40], [41] focus on the geolocation of IP infrastructure instead of end-users. In particular, the approaches described in [40] and [41] use the information obtained by reverse DNS queries for the geolocation of routers, as location codes are sometimes used to generate parts of their names. However, this method can be used only when a reverse DNS is set, and it suffers from the lack of uniform naming rules. As a consequence, it is not clear to which extent such technique can be applied to different parts of the world, or across autonomous systems that are managed according to different internal naming rules. In [24], one-way delay measurements collected by means of synchronized, high-resolution clocks, are used to evaluate the location of routers along the paths between landmarks. This approach, however, is not applicable with common machines, as it requires special purpose hardware. The HLOC framework, presented in [28], is the only available tool which uses RIPE Atlas as a primary method of geolocating routers. In particular, HLOC uses the reverse DNS approach described in [40] together with information extracted from public databases to produce a set of hints about the location of the target. Afterwards, the locations are validated or disproved with active measurements. This approach allows to reduce to reduce the number of active measurements to be performed. Results show that the positioning constraints derived by latency measurements frequently disprove the information originated by the reverse DNS approach. None of the reported papers provides information specific for all the different world's regions and neither about the effects of different topological and performance properties of the adopted landmarks.

III. METHOD

In this paper, we focus on computing the maximum accuracy that can be achieved by an IP geolocation method based on a world-scale measurement platform. First, we evaluate the characteristics of the error model associated to the distance estimation process on the base of a large set of delay measurements collected using RIPE Atlas. Then error characteristics, derived for a number of different measurement approaches (global vs regional, infrastructure- vs edge-based), are used to compute the maximum accuracy that can be achieved. This last step is based on the Cramér–Rao Lower Bound (CRLB), that provides a lower bound on the variance of any unbiased estimator [42].

A. PRINCIPLE OF OPERATION

The procedure generally used for geolocating an IP address can be summarized as follows: first, the distances between a set of landmarks and the target are estimated; then, a geometric technique (e.g. multilateration) is used to compute the position of the target on a global reference system.

Distances between target and landmarks are usually estimated by measuring the network delay. In particular, we assume that each landmark measures the Round Trip Time (RTT) towards the target, i.e. the amount of time needed for a probe packet (e.g. ICMP Echo Request) to reach the target and for the reply (ICMP Echo Reply) to reach back the source. The observed RTT is then used to compute the One-Way Delay (OWD), calculated as $RTT/2$, under the assumption that the same path is traveled by both probes and replies. Finally, OWDs are used to estimate distances according to the following relationship:

$$\hat{d}_i = p \cdot o_i \tag{1}$$

where $\hat{d}_i$ is the estimated distance between the ith landmark and the target, o_i is the OWD observed by the ith landmark, and p is a conversion coefficient. The relation is assumed to be linear, as the only component of o_i that depends on the real distance is the propagation delay, which increases linearly with distance.

The above model and assumptions are coherent with the vast majority of the currently available methods about IP geolocation.

B. ERROR MODEL

IP geolocation methods are intrinsically affected by inaccuracies. Delay measurements are influenced by: *i*) the presence of intermediate nodes, whose number differs from path to path, *ii*) the presence of cross-traffic on the probed path, *iii*) reply generation times, which depend on actual hardware and software configurations, and *iv*) the communication technology, both in terms of bandwidth and latency. In addition, Internet paths from landmarks to targets are not straight lines: they are composed of multiple links between adjacent routers, assuming the final form of polylines. As a consequence, each path is characterized by a different level of circuitousness.

Some of these inaccuracies can be mitigated by collecting multiple OWD values and then using the smallest one in the conversion from delay to distance, as the presence of queues and traffic load can temporarily increase the observed delay. This approach, which has been adopted in almost all studies about IP geolocation, has been included also in the presented method according to the details illustrated in Section IV. Other sources of errors, in particular the fraction of OWD due to the circuitousness of the considered path, are more difficult to be excluded from the localization process: simple delay measurement methods based on the ping tool are unable to provide information about the number and the position of intermediate routers. Using more sophisticated (and more heavyweight) probing methods, such as traceroute, may provide information only on the number of hops, but not on the total amount of circuitousness of the considered path. The inaccuracy introduced by circuitousness, together with the variable technological and runtime factors, contributes to the overall estimation error that affects the distance estimation process.

The relationship between the real and the estimated distance can thus be modeled as

$$\hat{d}_i = d_i + e_i \tag{2}$$

where d_i is the real distance between the ith target, and e_i is the overall error affecting the measurement.

We collected two large datasets to characterize the error in Equation 2 when the target is part of the Internet infrastructure. In particular, a first dataset defines the distance estimation error when a target is probed by a landmark positioned in the infrastructure as well; the other dataset defines the distance estimation error when the target is probed by a landmark positioned at the fringes of the Internet.

C. MAXIMUM ACCURACY

Errors in estimated distances obviously influence the accuracy of the geolocation process. In this paper, we focus on the maximum accuracy that can be achieved by an IP geolocation method based on such perturbed data. To quantify the maximum accuracy of a geolocation method we use the technique described in [43], which is based on computing the CRLB of IP geolocation. The CRLB is the theoretical lower bound of the variance of an unbiased estimator of a deterministic (yet unknown) parameter. In our case the unknown parameter is the position of the target. The CRLB is often used for computing the maximum accuracy of positioning systems [44], [45]. The CRLB defines a lower bound to the mean square error (MSE) of a position estimate. Localization accuracy can be thus defined as the square root of the minimum mean square error (RMSE). Better accuracy corresponds to smaller values of RMSE and vice-versa.

IV. DATA COLLECTION

RIPE Atlas is a large-scale system aimed at measuring Internet connectivity and reachability. Its final goal is to provide a detailed understanding of the state of the Internet on a global scale. RIPE Atlas is composed of devices belonging to two categories: *probes* and *anchors*. Probes are characterized by low-end hardware and they are mostly hosted by end-users in home or office networks. Anchors are more powerful machines with dedicated bandwidth. Anchors are hosted in professional environments, such as IXPs, ISPs, and datacenters. Hosting of probes and anchors is always based on a voluntary basis, but with relevant differences. Probes are freely given to whoever may be interested, and there are virtually no requirements to become a host. Anchors are instead assigned according to very strict requirements, as they should be always available and well-connected. In addition, the volunteers should incur the cost of purchasing and maintaining the required equipment.

In practice, probes collect measurements from the edge of the Internet, according to an approach that resembles crowdsourcing-based ones (i.e. because of the heterogeneous hosting environments) [46]. Anchors, on the contrary, collect measurements from the infrastructure of the Internet, in more controlled and stable conditions.

At the time of writing, there are more than 11 000 probes and 331 anchors involved in Internet measurements. While RIPE Atlas has a bias towards Europe, where many devices are located, the great number of vantage points allows in any case an excellent world-wide coverage.

A. INFRASTRUCTURE AND EDGE DATASETS

Our goal is to evaluate how much geolocation accuracy is affected by the position of landmarks, in terms of distance from the "core" of the Internet, and by their connection quality. Thus, we collected two datasets: the *Infrastructure dataset*, and the *Edge dataset*. The two datasets include measurements run from landmarks located in the infrastructure of the Internet and at the fringes of it, respectively. The set of targets is the same for both datasets and it is composed of machines representing the Internet infrastructure. Note that to build a dataset, the real location of both landmarks and targets must be known, as their positions are used to characterize the error model and in the end to evaluate the accuracy of IP geolocation (real positions are used as ground truth).

1) INFRASTRUCTURE DATASET

This dataset includes measurements collected using all RIPE Atlas anchors as landmarks. The set of probed targets is composed by the same RIPE Atlas anchors plus the nodes belonging to the NLNOG Ring [47]. The number of NLNOG Ring nodes involved in the data collection was 494. The total amount of landmarks and targets in the dataset is 331 and 825, respectively. These nodes are a representative sample of infrastructural machines since they are all hosted in ISPs, IXPs, and datacenters, without specific geographical requirements. In detail, the 331 anchors belong to 317 different autonomous systems, and are located in 260 cities spread over 90 countries; the NLNOG Ring nodes belong to 402 different autonomous systems, and are located in 245 cities spread over 45 countries. Overall, the set of all targets spans 648 different autonomous systems, 425 cities, and 97 countries.

2) EDGE DATASET

This dataset includes measurements collected using RIPE Atlas probes as landmarks. The set of targets is again composed by RIPE Atlas anchors plus NLNOG Ring nodes (as we are interested in localizing the infrastructural Internet). The distribution of RIPE Atlas probes is not uniform on the globe, with some regions much more covered (e.g. Europe) than others (e.g. South America). Thus, to reduce the differences in terms of landmark density for the various regions, we limited the number of probes in the experiments to 500 per region.

We believe that this limitation does not affect the obtained results, as a real geolocation system that uses more than 500 landmarks for locating a single target is unpractical (and probably also unfeasible, as the high volume of traffic towards the target could be easily mistaken for a denial of service attack). The probes inside each region were selected as evenly as possible among the various countries: the number of probes belonging to each country is approximately equal to M divided by the number of countries, where M is the minimum between 500 and the total number of probes in the region. The number of involved probes is 500 in Europe, 500 in North America, 210 in South America, 484 in Asia, 260 in Oceania, 242 in Africa, and 238 in Middle East. The total amount of landmarks and targets in the Edge dataset is 2434 and 825, respectively.

B. CALIBRATION AND COLLECTION DETAILS

The total amount of landmark-target pairs for the two datasets is $\sim$ 2.3 million. To not overload the RIPE Atlas platform we had to limit the number of measurements. We thus ran a preliminary calibration campaign using a subset of the landmark-target pairs where, for each pair, we collected 100 OWD values. We observed that the minimum value computed on the first 10 measurements was very close to the minimum computed on the whole set of 100 measurements. In particular, the minimum obtained using only 10 measurements is, on average, 2.5% larger than the one obtained with 100 measurements for anchors, and 1.5% for probes. Since the introduced distortion was small, we decided to collect 10 measurements for each landmark-target pair in the two datasets.

For each pair, collection of OWD samples was carried out with a one-hour interval between ICMP echo requests, to reduce cross measurement interferences and to reduce the possibility of triggering ICMP rate limiting on the traversed routers [48]. In the end, the two datasets complessively include a total of $\sim$ 23 million measurements.

C. REMOVAL OF OUTLIERS

An initial check was carried out to detect landmark-target pairs characterized by inconsistent measurements. We found that some pairs were reporting communication speeds faster than the speed of light. This was due to an erroneous configuration of the geographical coordinates of some landmarks. For the Edge dataset 18 sources were reporting inconsistent values. They have been removed from the dataset and reported to the RIPE Atlas team. No inconsistent values were detected in the Infrastructure dataset. This was somehow expected since anchors are managed by operators and not end-users.

V. GLOBAL RESULTS

The two datasets were used to characterize the delay-distance relationship and the error model. From these intermediate results we computed the maximum accuracy that is possible to achieve using the two approaches (Edge- and Infrastructure-based) on a global scale.

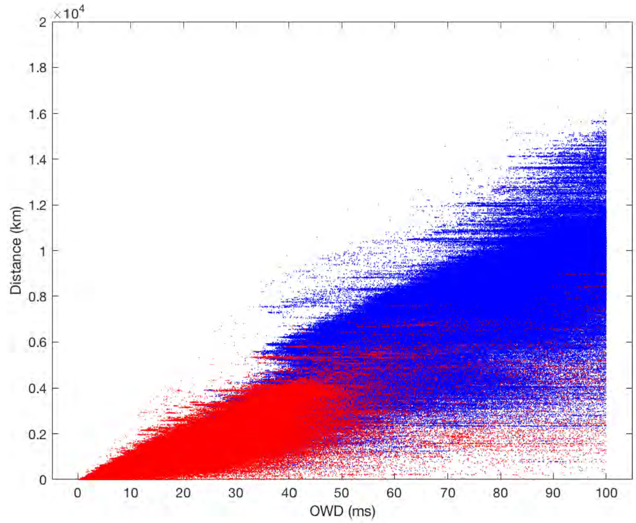

FIGURE 1. The real distance between landmarks and targets is plotted against the observed minimum OWD (both Infrastructure and Edge dataset). Measurements are depicted in red when the endpoints are located in the same region, in blue when in different regions.

A. FROM DELAYS TO DISTANCES

Fig. 1 shows the scatter plot of the minimum observed OWD versus the real distance for each landmark-target pair involved in the measurement phase in both datasets. Two main clusters of measurements can be identified, separated at a value of OWD approximately equal to 30 ms. The two clusters mostly represent measurements where landmark and target belong to the same region and measurements where the endpoints belong to different regions. This behavior has been discussed also in [46]. It is known that geolocation with landmarks placed at large distance from the target (e.g. in another continent), and so with high OWD values, is not effective [38].

Above all, the presence of two clusters suggests that two different models should be used for the delay-distance relationship depending on the presence/absence of long-haul links on the path. Since we are interested in a single and homogeneous model characterized by reduced error, we removed from the datasets all measurements with OWD larger than 30 ms. From another perspective, we preferred the possibility of defining a model that is consistent, uniform, and faithful to the data, albeit being limited to the selected threshold.

Filtering using the 30 ms threshold removed more than 95% of the measurements collected by landmarks placed in a different region with respect to the target. It is fair to assume these measurements would have produced poor results if fed in a geolocation method. On the contrary, approximately 80% of the measurements between landmarks and targets belonging to the same region fall inside the 30 ms threshold, confirming the validity of the latter for the separating the two clusters.

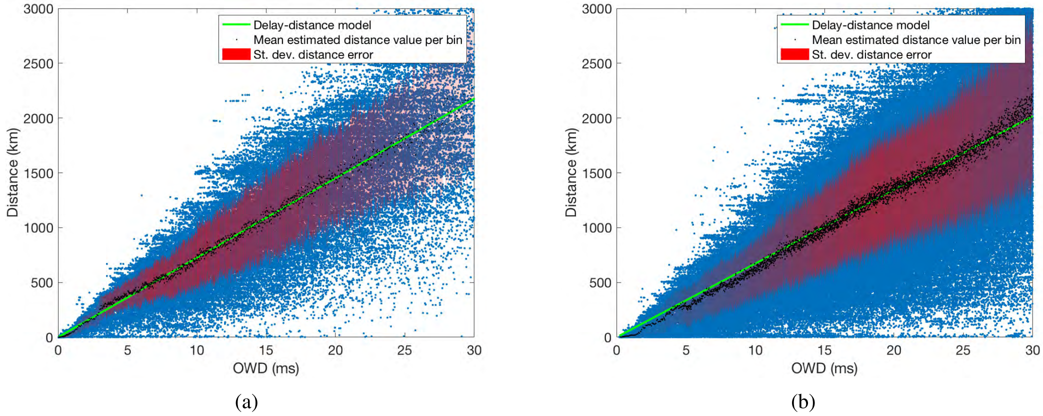

FIGURE 2. Delay-distance scatterplot for the Infrastructure and Edge datasets. The delay-distance linear conversion is represented by the green line. Red vertical bands represent the standard deviation of the distance estimation error. (a) Infrastructure dataset. (b) Edge dataset.

Coherently with the above choice, we will assume that a landmark contributes to localizing a target only if their distance falls within the validity interval of the model. This, in turn, limits the geographical area where the maximum accuracy derived from the model can be computed. The term *coverage* will be used to indicate the fraction of land where the model can be applied.

We then quantified the delay-distance coefficient (term p in Eq. 1) for both the Infrastructure and the Edge dataset. Fig. 2a shows a scatter plot of the real distance versus the observed minimum OWD, for the measurements in the Infrastructure dataset. The conversion factor from delays to distances p, computed with linear regression, is represented with a green line in the plot. The obtained value is 71.89 km/ms. Fig. 2b shows the same scatterplot for the Edge dataset. In this case the value of p is 67.11 km/ms, approximately 7% slower. To better understand the reasons behind this difference, we analyzed $\sim$ 200 million traceroute results (successfully reaching the target address) produced by RIPE Atlas during a week of its ordinary monitoring activities. This data collection has been carried out for a single week because the average number of hops of Internet paths is known to change at a slow rate [49]. Moreover, we were interested in a snapshot of the status of the network at a specific point in time, namely when the measurements used for the experiment above were performed. Results show that the average traceroute length of the path to the target is 10.6 hops for anchors and 13.2 hops for probes. In general, anchors are able to reach another host, and thus also the target, with a reduced number of hops, compared to probes, thanks to their better topological position (closer to the "core" of the Internet). This explains the difference in the observed p values, as the larger number of intermediate routers for probes introduces additional delay in communication latency.

B. ERROR AND ACCURACY CHARACTERIZATION

We verified that the distribution of the error in RIPE Atlas measurements (term e in Eq. 2) is compatible with a Gaussian distribution. Despite the existence of previous literature (such as [22], [43]) suggesting a Gaussian distribution model, the confirmation obtained using RIPE Atlas is important for two reasons. First, the amount of data we used is particularly large, thus increasing confidence in obtained results. Second, measurements are in part influenced by the adopted platform, thus a confirmation obtained using RIPE Atlas is a step forward in terms of generality, as existing literature is based on data collected with PlanetLab and PingER.

To this purpose, each dataset has been divided using bins of variable width. The size of each bin, in terms of observed OWD, has been selected to contain exactly 100 measurements. For all measurements, we computed the error e_i as the difference between the real distance d_i and the estimated distance $\hat{d}_i$. The black dots in Fig. 2 correspond to the mean value of $\hat{d}$ for each single bin. The red band has an amplitude that is equal to two times the standard deviation of e values, and it is vertically centered on the mean value of $\hat{d}$. We used the Kolmogorov-Smirnov test to verify if the distribution of values of e within a single bin is compatible with a normal distribution. The null hypothesis (normality) is not rejected in 91.7% of bins for anchors, and 98.2% of bins for probes. These results confirm that the maximum accuracy can be computed by means of a CRLB-based technique where the error is Gaussian.

For both datasets, it can be observed that the standard deviation of e increases proportionally to the measured delay (the amplitude of the red bands). In other words, the higher the value of OWD the higher the error on the estimated distance. The relationship between the standard deviation of the error and OWD is approximately linear. We thus evaluated the ratio between the standard deviation of the error with respect to OWD. The obtained values are 22.5 km/ms and 23.8 km/ms for anchors and probes respectively. In other words, each additional millisecond in the OWD increases the standard deviation of the error affecting the estimated distance by approximately 22 km for anchors and 24 km for probes.

The CRLB defines a lower bound on the mean square error of a position estimate:

$$MSE \geq trace\{I^{-1}\} \tag{3}$$

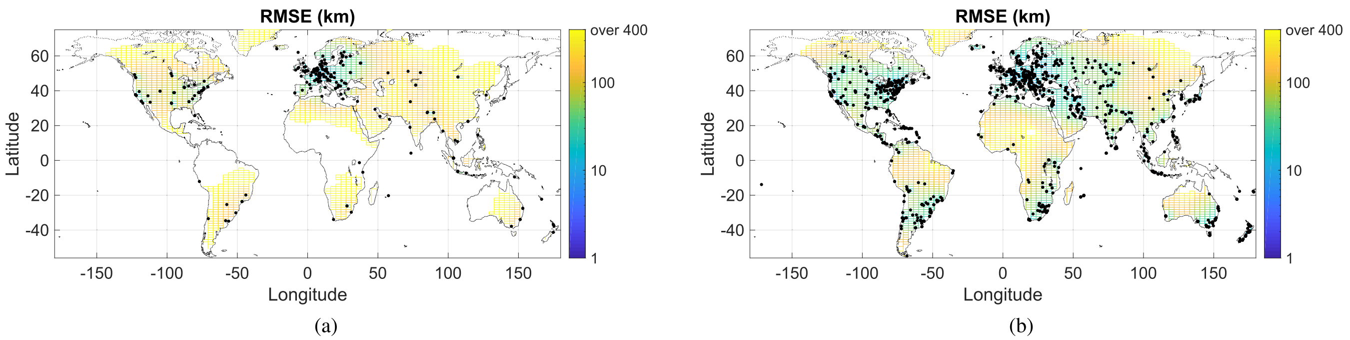

FIGURE 3. Global accuracy when using all RIPE Atlas anchors and probes. (a) Anchors. (b) Probes.

TABLE 1. RMSE for the Edge- and Infrastructure-based geolocation when using a given number of landmarks or all the available ones.

Infrastructure				Edge			
Landmarks	Avg. RMSE (km)	Median RMSE (km)	Coverage	Landmarks	Avg. RMSE (km)	Median RMSE (km)	Coverage
10	781	615	1.3%	10	827	651	1.3%
20	660	501	13.6%	20	682	512	11.5%
50	548	390	40.1%	50	596	399	35.6%
100	513	314	60.3%	100	482	320	51.9%
200	481	256	68.2%	200	406	249	69.3%
313	486	254	71.0%	2166	148	78	91.1%

where I is the Fisher Information Matrix (FIM). According to [43], the elements of the FIM for IP geolocation can be defined as follows:

$$I_{1,1} = \sum_{i=1}^{N} \frac{(1+2k^2)\cos^2(\alpha_i)}{\sigma_i^2}$$

$$I_{2,2} = \sum_{i=1}^{N} \frac{(1+2k^2)\sin^2(\alpha_i)}{\sigma_i^2}$$

$$I_{1,2} = I_{2,1} = \sum_{i=1}^{N} \frac{(1+2k^2)\sin(\alpha_i)\cos(\alpha_i)}{\sigma_i^2}$$

where N is the number of landmarks, α_i is the angle between the target and the ith landmark, σ_i is the standard deviation of the error, and k is the proportionality factor between σ and distance:

$$\sigma_i = k \cdot d_i. \tag{4}$$

Accuracy, defined as the square root of the minimum MSE, has been calculated by using the previously determined error characteristics in the above equations. According to collected measurements k is equal to 0.313 for anchors and to 0.355 for probes, i.e. distance estimation is more reliable when using anchors with respect to probes.

Figures 3a and 3b show the accuracy values on the world's surface for Infrastructure- and Edge-based geolocation when using all RIPE Atlas anchors and probes. The position of landmarks is depicted as black dots.

White areas correspond to those parts of the world that, with current RIPE Atlas diffusion, are outside the validity interval of the model. It is important to highlight that if an area is not covered, this does not mean that a host cannot be localized. It just means that the maximum localization accuracy cannot be computed according to the adopted error model. More in detail, roughly speaking, the vertical dispersion of measurements on the right-hand side of Fig. 1, lets reasonably affirm that the distance estimation error associated to delays above 30 ms is going to be equal or worse than the one below the threshold (always considering its proportionality to the delay). Thus, in areas not covered by the proposed model, localization accuracy will be definitely disappointing.

For both anchors and probes, Europe and North America register the best accuracy levels. This is due to the relatively high density of landmarks in the two regions.

When all the landmarks are used, the median accuracy values that can be achieved are equal to 254 km for Infrastructure-based localization, and 78 km for Edge-based localization (Table 1). These values are obtained when using 313 and 2166 landmarks respectively, which are the total number of anchors and probes in the dataset after having removed measurements with OWD above 30 ms. The covered world's surface is 71% and 91.1% for Infrastructure and Edge-based localization, respectively.

The better performance obtained by Edge-based localization is due to the strong difference in the number of landmarks, which contributes more positively than the topological and technological advantages of the infrastructural landmarks. To clearly understand how much the number of landmarks and the technological/topological factors contribute to the accuracy of IP geolocation, we repeated the analysis using the same number of landmarks for the two approaches. Table 1 shows the accuracy obtained when using a fixed number of landmarks, randomly chosen among the set of available ones. In particular, the reported values of accuracy are the average and median on the world surface obtained in 50 repetitions of the experiment, each repetition with a different random selection of landmarks.

TABLE 2. Main characteristics at regional level.

Infrastructure	Asia	Europe	North Am.
Landmarks	37	191	55
p (km/ms)	106.51	68.33	95.63
k	0.39	0.30	0.25
Avg. number of hops	10.5	7.7	8.0
Avg. geographical distance (km)	1561	891	1447
Avg. hop length (km)	149	116	181

Edge	Asia	Europe	North Am.	South Am.	Oceania	Africa	Middle Ea.
Landmarks	484	500	500	210	260	242	238
p (km/ms)	97.13	62.60	73.16	61.40	96.67	80.13	21.12
k	0.36	0.33	0.39	0.42	0.33	0.41	1.03
Avg. number of hops	13.1	10.2	12.4	13.9	9.5	11.2	17.0
Avg. geographical distance (km)	1559	1132	1301	970	1520	1103	193
Avg. hop length (km)	119	111	105	70	160	98	11

When the number of landmarks used is small, Infrastructure-based geolocation is able to provide slightly better accuracy and coverage than Edge-based geolocation. This is due to the more favorable properties of the delay-distance model of anchors with respect to probes, both in terms of error and conversion factor. When the number of landmarks used is large, the Edge-based approach becomes the better one. Mostly because of the more uniform distribution of the landmarks on the globe. Anchors are particularly dense in Europe and, as the number of landmarks used gets larger, finding a set that uniformly covers the other regions becomes more difficult, as confirmed by the coverage. This acts as a balancing factor with respect to the slightly better error characteristics of anchors. When all anchors and probes are used the much larger number of probes makes Edge-based geolocation definitely more favorable. It is also important to notice that the increase of accuracy is not proportional to the amount of involved landmarks. For instance, in Edge-based geolocation, when the number of landmarks is increased from 200 to more than two thousand the accuracy does not improve by a factor of ten.

In summary, the total number of landmarks used in the geolocation process and their distribution on the world surface appear to be more important factors than their technological and topological properties.

VI. REGIONAL RESULTS

The vast majority of existing IP geolocation techniques use a single geolocation model for the entire world, not taking into account the technological and topological differences across regions. Also the uneven availability of landmarks is in general not considered.

We carried out a detailed analysis of IP geolocation accuracy on a regional scale, following the approach illustrated in Section V. As mentioned, the considered regions are: Europe, North America, South America, Asia, Oceania, Africa, and Middle East. Regional datasets have been generated by extracting from the global datasets the measurements collected when landmark and target are in the same region.

For each region, the delay-distance conversion factor and the error characteristics have been re-computed. In detail, for Edge-based geolocation this procedure has been carried out for all the seven regions, whereas for Infrastructure-based geolocation this has been done only for Europe, North America, and Asia. For the remaining regions the amount of available data was too small to draw statistically sound conclusions because of the limited number of anchors. Table 2 shows a summary of the numerical results of the analysis.

For Europe, North America, and Asia the delay-distance conversion factor p is always higher for the Infrastructure-based dataset than for the Edge-based one. This is expected as anchors are generally better connected with respect to probes. Limiting the analysis to the Edge-based dataset, the value of p does not seem to be correlated to the technological level of the different regions (in terms of Internet infrastructure). To better study this phenomenon we performed an additional data collection campaign to evaluate the number of intermediate routers on the involved paths. In detail, we used traceroute to collect the number of intermediate routers for each landmark-target pair appearing in the Infrastructure and Edge datasets. Table 2 reports for each region the average geographical distance between landmarks and targets, the average number of hops, and the average geographical distance traveled at each hop.[1] A linear correlation between the regional value of p and the average hop length can be observed. The Pearson correlation coefficient between p and the average hop length, for the Edge dataset, is 0.92 [50]. In the end, the delay-distance conversion factor seems to be heavily affected by the average hop length, which generally increases in regions with a large surface (e.g. Africa and Asia with respect to Europe). In Middle East the value of p is particularly small. This is probably due to the much smaller (with respect to the other regions) average geographical distance between landmark-target pairs, and the conversely larger average number of hops. In this case, using the global value of p would introduce significant distortions in the geolocation process.

We also computed the proportionality coefficient between the distance and the standard deviation of the error

[1] More precisely, this value is computed by dividing the distance between the two endpoints by the number of hops. Thus, it corresponds to the average distance traveled "towards" the target at each hop, which is equal to the real distance associated to an hop only when the path is not circuitous.

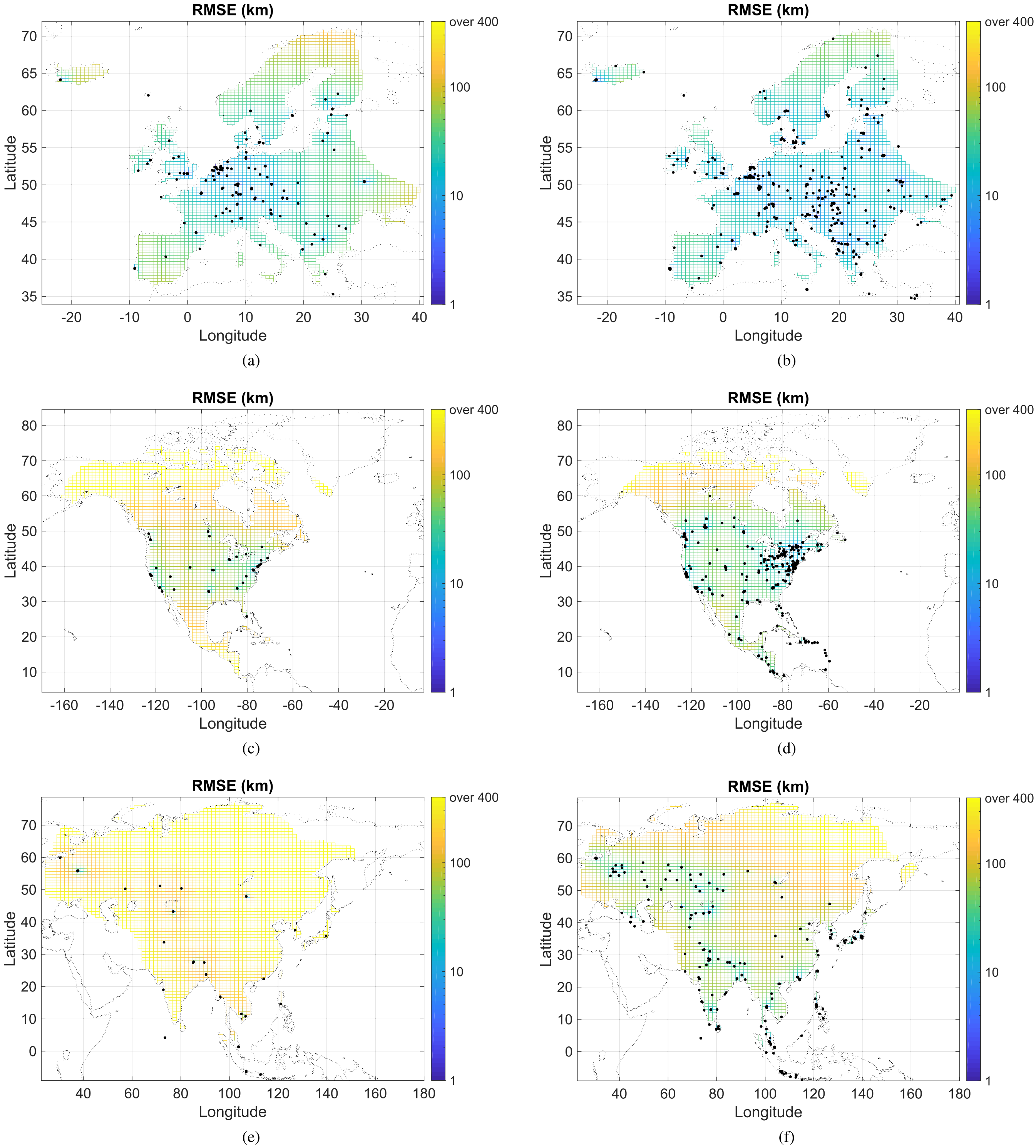

FIGURE 4. **Maximum accuracy in Europe, North America, and Asia when using RIPE Atlas anchors or probes. (a) Europe, anchors. (b) Europe, probes. (c) North America, anchors. (d) North America, probes. (e) Asia, anchors. (f) Asia, probes.**

(k introduced in Section 4) for the different regions, the values are shown in Table 2. Relatively small values are observed in North America and Europe when using Infrastructure-based geolocation (0.25 and 0.30 respectively). This can be explained by the connection quality of anchors, which reduces run-time distortions, and by the density of network elements in such regions, which increases the probability of finding direct paths to the targets. The worst situation, on the contrary, is found in Middle East, where the standard deviation of the distance estimation error is approximately

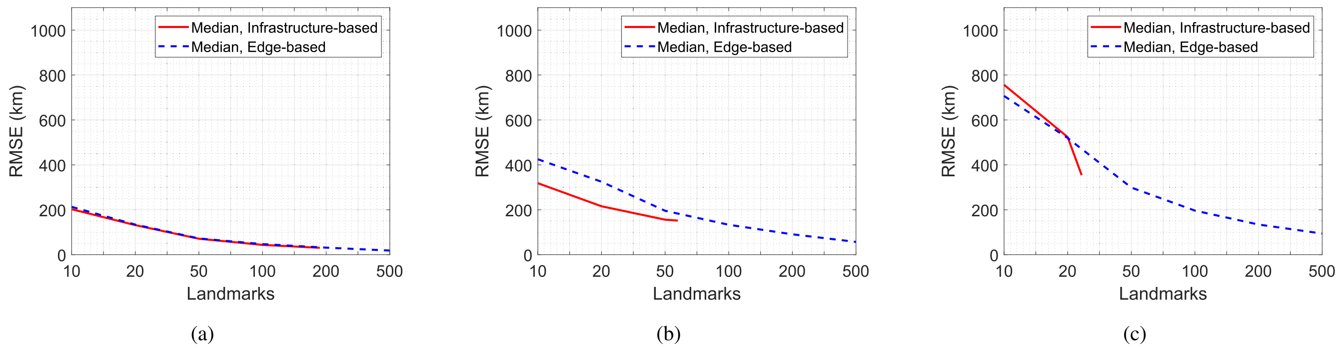

FIGURE 5. **RMSE of Infrastructure- and Edge-based geolocation when using a variable number of landmarks. (a) Europe. (b) North America. (c) Asia.**

equal to the distance between the two endpoints (k is equal to 1.03). This could be due to increased levels of circuitousness (i.e., the polyline formed by path segments may deviate significantly from the great circle distance).

Fig. 4 shows the maximum accuracy that can be achieved when using all RIPE Atlas anchors or probes in Europe, North America, and Asia (for probes the number is capped at 500). Europe and North America register the maximum accuracy (lowest RMSE values). In Europe the error is relatively small in the central part of the region, while it increases in peripheral countries. This phenomenon is less evident for probes, which are more uniformly spread. In North America localization accuracy is better along the coastal areas, where landmarks are more abundant [51]. In Asia, the reduced number of landmarks has a strong impact on the level of accuracy for Infrastructure-based geolocation. Coverage is complete in Europe, and almost complete in North America and Asia, with the exception of Greenland and the land in proximity of the Bering Strait (other areas are white just because they are outside the considered region, e.g. Middle East and Oceania in Fig. 4e). For South America, Oceania, Middle-East, and Africa a comparison between the Infrastructure-based and the Edge-based approach is not possible because of the reduced number of anchors in such regions. Accuracy achieved with probes is available in the Appendix.

Results shown in Fig. 4 have been obtained using the number of landmarks reported in Table 2. We then evaluated the maximum accuracy that can be achieved with the two approaches when using the same number of landmarks. The curves in Fig. 5 represent the median RMSE obtained in Europe, North America, and Asia across 50 repetitions. Also in this case, each repetition is characterized by a randomly chosen set of landmarks among the available ones. Both Infrastructure- and Edge-based geolocation perform better in Europe with respect to North America and Asia. For instance, when 20 landmarks are used, RMSE is below 150 km in Europe, in the order of 200-330 km in North America, and $\sim$ 500 km in Asia. The reason behind these differences can again be found in the smaller surface of Europe with respect to North America and Asia.

In Europe and North America, Infrastructure-based geolocation performs similarly or better than Edge-based geolocation, when using the same number of landmarks. In particular, in North America, the same accuracy that is achieved with 20 anchors is obtained with approximately 40 probes. In Asia, the Infrastructure-based geolocation starts performing better than the Edge-based one when the landmarks used are at least 20, but the relatively small number of anchors does not allow to draw sound conclusions.

Similar figures for the other regions are included in the Appendix but only for Edge-based geolocation (because the number of anchors in such regions is too small).

VII. NON-UNIFORM DISTRIBUTION OF TARGETS

The previously presented results, both global and regional ones, assume a uniform distribution of targets on the land surface of the world or on the surface of the considered region. Such results thus provide an indication of the maximum accuracy in the absence of prior information about the spatial distribution of infrastructural nodes. In this section we evaluated the maximum accuracy of IP geolocation under the assumption that targets are not uniformly distributed on the land surface (or on the considered region), and are more concentrated where RIPE Atlas anchors and NLNOG Ring nodes are located. The rationale behind this choice is that both RIPE Atlas anchors and NLNOG Ring nodes are hosted in the infrastructural part of the Internet. This is a realistic assumption, as the Internet infrastructure is concentrated in a few highly interconnected locations.

In detail we evaluated the maximum accuracy at the location of target nodes, without using the same node as a landmark when the target corresponds to an anchor. Global results, over all target positions, are presented in Table 3. The accuracy is significantly better, if compared to the one obtained assuming a uniform distribution of targets previously reported in Table 1. The reason can be found in the presence of landmarks that are, on average, closer to the target. From another perspective, if we assume that no infrastructural nodes can be located in large empty regions (such as the central part of Australia), the accuracy of IP geolocation gets better. The definition of coverage provided in Section V-A cannot be applied to the non-uniform case, as the set of positions is discrete. Thus, we tried to preserve the same concept by defining coverage, for the non-uniform case,

TABLE 3. RMSE and coverage for the Edge- and Infrastructure-based geolocation at the position of targets (over all targets positions).

Infrastructure				Edge			
Landmarks	Avg. RMSE (km)	Median RMSE (km)	Coverage	Landmarks	Avg. RMSE (km)	Median RMSE (km)	Coverage
10	695	532	9.7%	10	676	612	5.3%
20	397	324	68.3%	20	626	491	52.7%
50	192	150	91.1%	50	289	249	86.7%
100	91	61	97.5%	100	170	141	94.7%
200	45	18	97.9%	200	102	83	98.5%
313	36	6	97.9%	2166	10	3	100.00%

TABLE 4. RMSE and coverage for Edge-based geolocation at the position of targets (over all targets positions) when using 100 landmarks, compared to assuming uniform distribution.

	Uniform			Non uniform		
Region	Avg. RMSE (km)	Median RMSE (km)	Coverage	Avg. RMSE (km)	Median RMSE (km)	Coverage
Africa	889	266	88.0%	10	2	94.3%
Asia	405	196	89.5%	61	16	100.0%
Europe	81	47	100.0%	29	28	100.0%
Middle-East	313	118	52.7%	17	3	99.6%
North Am.	417	133	65.7%	47	30	100.0%
Oceania	375	141	88.5%	3	1	100.0%
South Am.	271	130	80.8%	28	2	100.0%

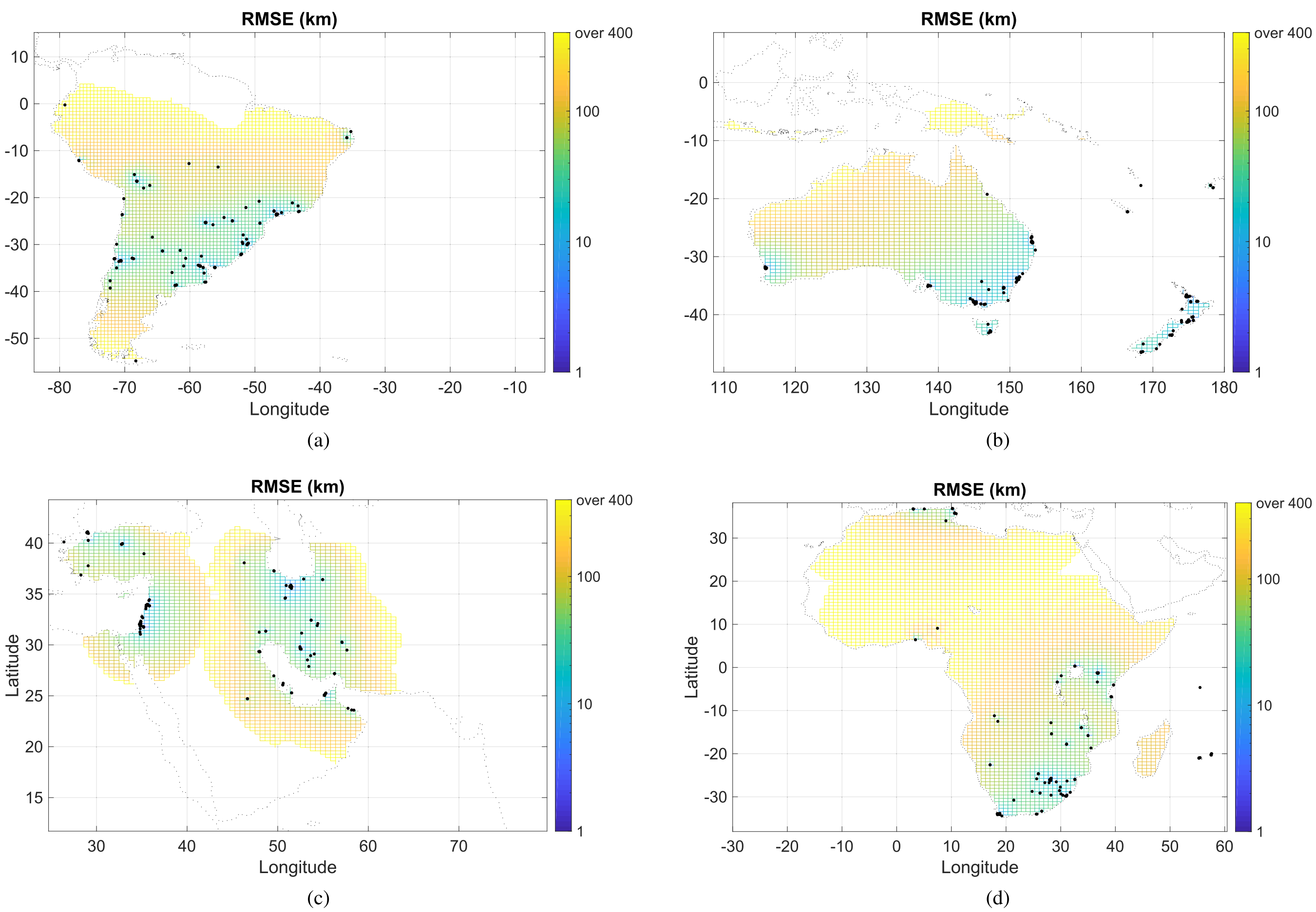

FIGURE 6. Maximum accuracy when using RIPE Atlas probes. (a) South America. (b) Oceania. (c) Middle East. (d) Africa.

as the fraction of targets that can be successfully localized, i.e. are within the validity domain of the model.

Table 4 reports the accuracy for Edge-based geolocation when using 100 landmarks, assuming both a uniform distribution of targets (for comparison) and a non-uniform one. The average improvement factor is in general more than 10x. In some cases, the improvement with respect to uniform distribution can be particularly large. For Infrastructure-based geolocation results are available only for Europe (as for the other regions the number of RIPE anchors is below 100):

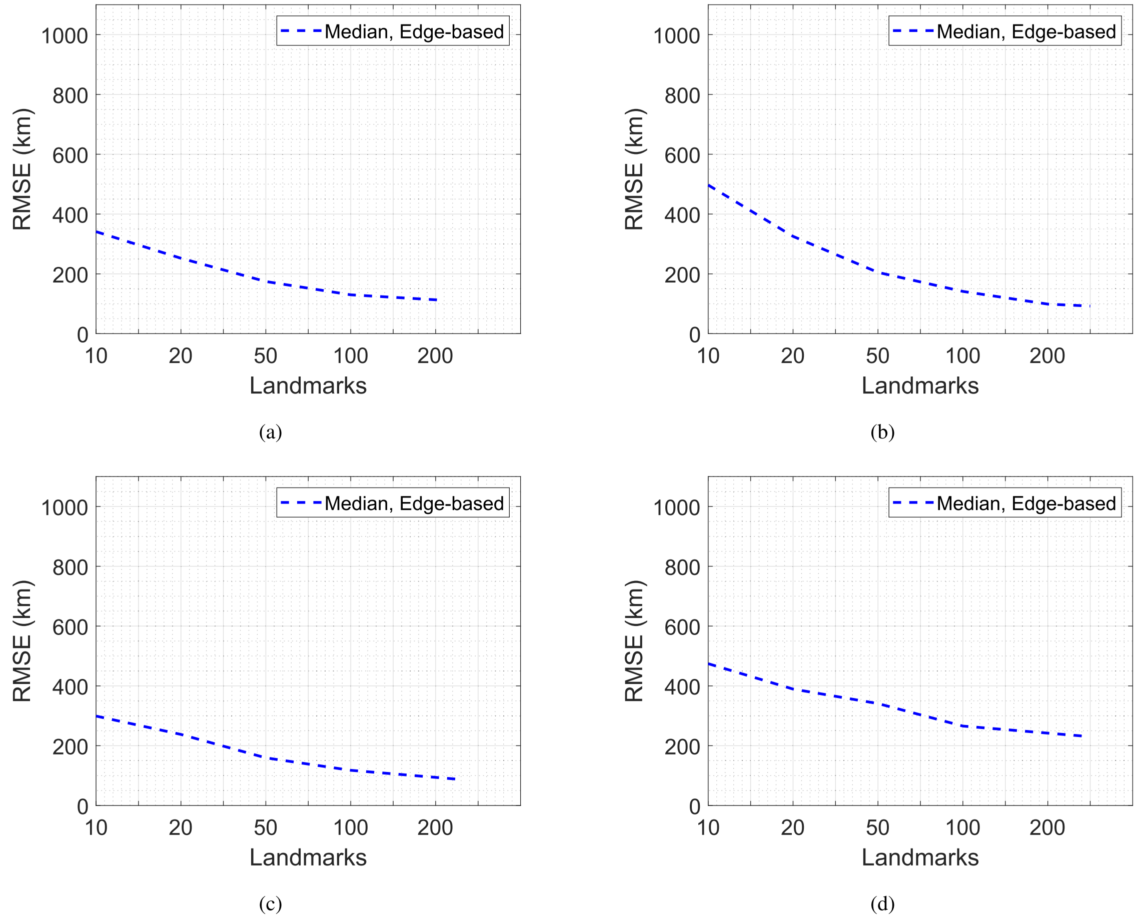

FIGURE 7. RMSE of Edge-based geolocation when using a variable number of landmarks. (a) South America. (b) Oceania. (c) Middle East. (d) Africa.

considering a non-uniform distribution of targets, the median accuracy is 9 km (avg. 17 km), against 71 km (avg. 75 km) obtained when the distribution is uniform. For both distributions of targets, a 100% coverage has been obtained.

These last results are extremely important as they clearly indicate that active geolocation is a suitable tool for geolocating the Internet infrastructure with acceptable precision.

VIII. CONCLUSION

In this paper we studied the maximum accuracy achievable by an IP geolocation system for localizing infrastructural Internet nodes in different scenarios. In particular, the CRLB defines a fundamental limit on the performance of any unbiased estimator, and thus represents a baseline against which localization methods can be compared. In the presented results we adopted the landmarks offered by the RIPE Atlas measurement platform. We compared a scenario composed by well-connected landmarks belonging to the infrastructure of the Internet with another scenario composed by best-effort landmarks hosted by end-users in domestic environments (i.e., at the fringes of the Internet). The maximum accuracy was characterized at both global and regional levels, to highlight differences, and to show that using a single geolocation model for the entire world is not a suitable approach. Thanks to the diffusion of the RIPE Atlas platform, the study presented in this paper was conducted at an unprecedented large scale with respect to previous literature, for what concerns both the number of used landmarks and the number of collected measurements.

We showed that infrastructural landmarks are slightly more precise than edge landmarks in estimating distances, as highlighted by the smaller error in the distance estimation process. This is mainly due to their more favorable position in the Internet, which allows them to reach other nodes, on average, with a smaller amount of hops. However, when the number of landmarks involved in geolocation increases, the advantage of using infrastructural nodes is less evident. This suggests that the number of landmarks and their geographical distribution can be more decisive factors than their technological and topological properties. It must be however noted that the number of landmarks for performing latency measurements is an extremely relevant parameter when a geolocation system must be implemented in a real context, as it directly affects the feasibility and scalability of the whole system. Hence, the use of more precise landmarks as the infrastructural ones is not to be discarded a priori. We also showed that speed, intended as the amount of km travelled per ms, is directly correlated to the average hop length.

Finally, we showed that when considering a more realistic distribution of targets (a non-uniform one, with targets located where the infrastructural nodes actually are), the accuracy achieved is significantly higher than the one obtained with a uniform distribution. This happens because landmarks are on average much closer to the target in the non-uniform case. Again, accuracy still depends on the number of landmarks used, but, for the same number, infrastructural landmarks seem to show better results than edge landmarks.

Building a worldwide-scale geolocation system implies several strategic decisions that affect the feasibility and precision of the system. This paper, gives a quantitative evaluation of the main factors influencing the system.

Reproducibility: The data collected by RIPE Atlas is publicly available at [52]. In particular, the datasets used for the analysis presented in this paper are marked with the tag "mcgl".

APPENDIX

Localization accuracy that can be achieved when using RIPE Atlas probes in South America, Oceania, Middle East, and Africa. Fig. 6 shows the location of landmarks (probes) and the variation of accuracy in the regions. In all the four regions, the presence of large areas without landmarks can be observed. Fig. 7 shows how accuracy changes when increasing the number of landmarks for Edge-based geolocation.

REFERENCES

[1] A. Y. Nur and M. E. Tozal, "Geography and routing in the Internet," *ACM Trans. Spatial Algorithms Syst. (TSAS)*, vol. 4, no. 4, pp. 11:1–11:16, Oct. 2018.

[2] Z. Weinberg, S. Cho, N. Christin, V. Sekar, and P. Gill, "How to catch when proxies lie: Verifying the physical locations of network proxies with active geolocation," in *Proc. Internet Meas. Conf. (IMC)*, 2018, pp. 203–217.

[3] P. Gigis, V. Kotronis, E. Aben, S. D. Strowes, and X. Dimitropoulos, "Characterizing user-to-user connectivity with RIPE atlas," in *Proc. Appl. Netw. Res. Workshop*, 2017, pp. 4–6.

[4] RIPE NCC. *IXP Country Jedi*. Accessed: Jan. 1, 2019. [Online]. Available: https://www.ripe.net/analyse/internet-measurements/ixp-country-jedi

[5] C. Iordanou, G. Smaragdakis, I. Poese, and N. Laoutaris, "Tracing cross border Web tracking," in *Proc. Internet Meas. Conf. (IMC)*, 2018, pp. 329–342.

[6] RIPE NCC. *RIPE IPmap*. Accessed: Jan. 1, 2019. [Online]. Available: https://ipmap.ripe.net

[7] Y. Shavitt and N. Zilberman, "A geolocation databases study," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 10, pp. 2044–2056, Dec. 2011.

[8] R. N. Staff, "RIPE atlas: A global Internet measurement network," *Internet Protocol J.*, vol. 18, no. 3, pp. 2–26, 2015.

[9] V. N. Padmanabhan and L. Subramanian, "An investigation of geographic mapping techniques for Internet hosts," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 31, no. 4, pp. 173–185, 2001.

[10] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, "Constraint-based geolocation of Internet hosts," *IEEE/ACM Trans. Netw.*, vol. 14, no. 6, pp. 1219–1232, 2006.

[11] B. Wong, I. Stoyanov, and E. G. Sirer, "Octant: A comprehensive framework for the geolocalization of Internet hosts," in *Proc. USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, vol. 7, 2007, p. 313.

[12] Y. Wang, D. Burgener, M. Flores, A. Kuzmanovic, and C. Huang, "Towards street-level client-independent IP geolocation," in *Proc. USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, vol. 11, 2011, pp. 365–379.

[13] D. Cicalese, D. Z. Joumblatt, D. Rossi, M.-O. Buob, J. Augé, and T. Friedman, "Latency-based anycast geolocation: Algorithms, software, and data sets," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 6, pp. 1889–1903, Jun. 2016.

[14] P. Hillmann, L. Stiemert, G. D. Rodosek, and O. Rose, "Dragoon: Advanced modelling of IP geolocation by use of latency measurements," in *Proc. 10th Int. Conf. Internet Technol. Secured Trans. (ICITST)*, 2015, pp. 438–445.

[15] P. Hillmann, L. Stiemert, G. Dreo, and O. Rose, "On the path to high precise IP geolocation: A self-optimizing model," *Int. J. Intell. Comput. Res.*, vol. 7, no. 1, pp. 682–693, Mar. 2016.

[16] Z. Dong, R. D. W. Perera, R. Chandramouli, and K. Subbalakshmi, "Network measurement based modeling and optimization for IP geolocation," *Comput. Netw.*, vol. 56, no. 1, pp. 85–98, 2012.

[17] S. Ding, X. Luo, D. Ye, and F. Liu, "Delay-distance correlation study for IP geolocation," *Wuhan Univ. J. Natural Sci.*, vol. 22, no. 2, pp. 157–164, Apr. 2017.

[18] I. Youn, B. L. Mark, and D. Richards, "Statistical geolocation of Internet hosts," in *Proc. 18th Int. Conf. Comput. Commun. Netw. (ICCCN)*, 2009, pp. 1–6.

[19] B. Eriksson, P. Barford, J. Sommers, and R. Nowak, "A learning-based approach for IP geolocation," in *Proc. Int. Conf. Passive Act. Netw. Meas. (PAM)*. Berlin, Germany: Springer, 2010, pp. 171–180.

[20] B. Eriksson, P. Barford, B. Maggs, and R. Nowak, "Posit: A lightweight approach for IP geolocation," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 40, no. 2, pp. 2–11, 2012.

[21] S. Van Langen, X. Zhou, and P. Van Mieghem, "On the estimation of Internet distances using landmarks," in *Proc. Int. Conf. Next Gener. Teletraffic Wired/Wireless Adv. Netw. (NEW2AN)*. Berlin, Germany: Springer, 2004.

[22] S. Laki, P. Mátray, P. Hága, T. Sebõk, I. Csabai, and G. Vattay, "Spotter: A model based active geolocation service," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2011, pp. 3173–3181.

[23] E. Katz-Bassett, J. P. John, A. Krishnamurthy, D. Wetherall, T. Anderson, and Y. Chawathe, "Towards IP geolocation using delay and topology measurements," in *Proc. Internet Meas. Conf. (IMC)*, 2006, pp. 71–84.

[24] S. Laki, P. Mátray, P. Hága, I. Csabai, and G. Vattay, "A model based approach for improving router geolocation," *Comput. Netw.*, vol. 54, no. 9, pp. 1490–1501, 2010.

[25] S. Laki, P. Matray, P. Haga, I. Csabai, and G. Vattay, "A detailed path-latency model for router geolocation," in *Proc. Int. Testbeds Res. Infrastruct. Develop. Netw. Communities Workshops (TridentCom)*, 2009, pp. 1–6.

[26] M. Jia, Y. Zhu, and H. Yang, "An IP geolocation algorithm using bandwidth measurement and distance estimation," in *Proc. Int. Conf. Elect. Inf. Technol. Rail Transp.*, Y. Qin, L. Jia, J. Feng, M. An, and L. Diao, Eds. Berlin, Germany: Springer, 2016, pp. 519–527.

[27] A. Ziviani, S. Fdida, J. F. de Rezende, and O. C. M. B. Duarte, "Improving the accuracy of measurement-based geographic location of Internet hosts," *Comput. Netw.*, vol. 47, no. 4, pp. 503–523, 2005.

[28] Q. Scheitle, O. Gasser, P. Sattler, and G. Carle, "HLOC: Hints-based geolocation leveraging multiple measurement frameworks," in *Proc. Netw. Traffic Meas. Anal. Conf. (TMA)*, 2017, pp. 1–9.

[29] D. Li *et al.*, "IP-geolocation mapping for involving moderately-connected Internet regions," Microsoft Res., Tech. Rep., 2009.

[30] J. Chen, F. Liu, X. Luo, F. Zhao, and G. Zhu, "A landmark calibration-based IP geolocation approach," *EURASIP J. Inf. Secur.*, vol. 2016, no. 1, p. 4, 2016.

[31] C. Guo, Y. Liu, W. Shen, H. J. Wang, Q. Yu, and Y. Zhang, "Mining the Web and the Internet for accurate IP address geolocations," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2009, pp. 2841–2845.

[32] S. Zu, X. Luo, S. Liu, Y. Liu, and F. Liu, "City-level IP geolocation algorithm based on PoP network topology," *IEEE Access*, vol. 6, pp. 64867–64875, 2018.

[33] H. Jiang, Y. Liu, and J. N. Matthews, "IP geolocation estimation using neural networks with stable landmarks," in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, Apr. 2016, pp. 170–175.

[34] B. Chun *et al.*, "PlanetLab: An overlay testbed for broad-coverage services," *ACM SIGCOMM Comput. Commun. Rev. (CCR)*, vol. 33, no. 3, pp. 3–12, 2003.

[35] B. Gueye, S. Uhlig, and S. Fdida, "Investigating the imprecision of IP block-based geolocation," in *Proc. Int. Conf. Passive Active Netw. Meas. (PAM)*. Berlin, Germany: Springer, 2007, pp. 237–240.

[36] M. Gharaibeh, A. Shah, B. Huffaker, H. Zhang, R. Ensafi, and C. Papadopoulos, "A look at router geolocation in public and commercial databases," in *Proc. Internet Meas. Conf. (IMC)*, 2017, pp. 463–469.
[37] I. Poese, S. Uhlig, M. A. Kaafar, B. Donnet, and B. Gueye, "IP geolocation databases: Unreliable?" *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 2, pp. 53–56, Apr. 2011.
[38] B. Trammell and M. Kühlewind, "Revisiting the privacy implications of two-way Internet latency data," in *Proc. Int. Conf. Passive Act. Netw. Meas. (PAM)*. Cham, Switzerland: Springer, 2018, pp. 73–84.
[39] Y. Lee, H. Park, and Y. Lee, "IP Geolocation with a Crowd-sourcing Broadband Performance tool," *ACM SIGCOMM Comput. Commun. Rev. (CCR)*, vol. 46, no. 1, pp. 12–20, 2016.
[40] B. Huffaker and M. Fomenkov, "Drop: DNS-based router positioning," *ACM SIGCOMM Comput. Commun. Rev. (CCR)*, vol. 44, no. 3, pp. 5–13, 2014.
[41] J. Chabarek and P. Barford, "What's in a name?: Decoding router interface names," in *Proc. 5th ACM Workshop HotPlanet*, 2013, pp. 3–8.
[42] S. M. Kay, *Fundamentals Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
[43] G. Ciavarrini, M. S. Greco, and A. Vecchio, "Geolocation of Internet hosts: Accuracy limits through Cramér–Rao lower bound," *Comput. Netw.*, vol. 135, pp. 70–80, Apr. 2018.
[44] M. Angjelichinoski, D. Denkovski, V. Atanasovski, and L. Gavrilovska, "Cramér–Rao Lower bounds of RSS-based localization with anchor position uncertainty," *IEEE Trans. Inf. Theory*, vol. 61, no. 5, pp. 2807–2834, May 2015.
[45] Y. Qi and H. Kobayashi, "Cramér–Rao Lower bound for geolocation in non-line-of-sight environment," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 3, May 2002, pp. III-2473–III-2476.
[46] G. Ciavarrini, V. Luconi, and A. Vecchio, "Smartphone-based geolocation of Internet hosts," *Comput. Netw.*, vol. 116, pp. 22–32, Apr. 2017.
[47] NLNOG. *NLNOG Ring*. Accessed: Jan. 5, 2018. [Online]. Available: https://ring.nlnog.net/
[48] R. Ravaioli, G. Urvoy-Keller, and C. Barakat, "Characterizing ICMP rate limitation on routers," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 6043–6049.
[49] T. Böttger *et al.* (2018). "The elusive Internet flattening: 10 years of IXP growth." [Online]. Available: https://arxiv.org/abs/1810.10963
[50] B. Ratner, "The correlation coefficient: Its values range between $+1/-1$, or do they?" *J. Targeting, Meas. Anal. Marketing*, vol. 17, no. 2, pp. 139–142, Jun. 2009.
[51] M. Balch. *World Internet Map*. Accessed: Aug. 5, 2018. [Online]. Available: http://wnm.ixty.net/ranks/world
[52] RIPE NCC. *RIPE Atlas*. Accessed: Jan. 1, 2019. [Online]. Available: https://atlas.ripe.net

MASSIMO CANDELA received the master's degree in computer engineering from Roma Tre University, Italy. Since 2013, he has been a Software Engineer with the Research and Development Department, RIPE Network Coordination Centre, focusing on network measurements and network data representation. He is currently leading the development of IPmap, a tool which geolocates components of the Internet infrastructure.

ENRICO GREGORI received the Laurea degree in electronic engineering from the University of Pisa, in 1980. He has contributed to several national and international projects on computer networking. He has authored more than 100 papers in the area of computer networks. His current research interests include Internet measurements and data analysis, ad hoc networks, sensor networks, wireless LANs, and quality of service in packet switching networks.

VALERIO LUCONI received the master's and Ph.D. degrees in computer engineering from the University of Pisa, in 2012 and 2016, respectively. He is currently a Postdoctoral Researcher with IIT-CNR, Pisa. His research interests include the Internet measurements, the Internet topology, IP geolocation, network neutrality, and network monitoring.

ALESSIO VECCHIO is currently an Associate Professor with the Department of Information Engineering, University of Pisa, Italy. His research interests include computer networks, mobile, and pervasive computing. He is currently serving as an Associate Editor for the IEEE ACCESS and an Area Editor in the *Pervasive and Mobile Computing*.