Rooting Out Incorrect RIPE Atlas Probe Geolocations

Katherine Izhikevich[1], Ben Du[2], Manda Tran[2], Sumanth Rao[1], Alisha Ukani[1], Ray Bellis[3], Liz Izhikevich[2]

[1]UC San Diego [2]UC Los Angeles [3]Internet Systems Consortium

Abstract

Geolocation plays a critical role in understanding the Internet. In this work, we analyze and fix operator-misreported geolocation. Using DNS root servers to detect speed of Internet violations, we conservatively infer that at least 2% of vantage points in the largest community-vantage point collection, RIPE Atlas, were not located at their operator-reported geolocation between January 2019 and April 2026. To increase the accuracy of future studies that use operator-reported geolocation, we open source our simple methodology, implement a reporting campaign of violating probes within RIPE, work with operators to fix the misreported geolocations, and release a continually updated dataset of RIPE vantage points that likely misreport geolocation.

1 Introduction

IP address geolocation plays a central role in networking and security: understanding the Internet's topology to improve last mile latency [3, 17], filtering traffic to prevent cyber attacks [42], and identifying malicious activity [27, 37], all rely on geolocation. To geolocate an IP address, methodologies often rely upon a global, human-reported set of vantage points (i.e., probes) to measure from and verify accuracy against. For example, geolocation based on latency measurements often require a network of probes, each with operator-attested geolocation, to be located near every target to accurately estimate every target's geolocation [5, 16, 20, 22, 29, 45].

While necessary, relying on human operators to report accurate probe geolocation can introduce human error. For example, if an operator moves their probe, but fails to update its location, geolocation methodologies relying on the outdated information risk inaccurate conclusions. To account for incorrect geolocation, a few geolocation studies take an additional precaution to filter out probes reporting unlikely locations [11, 18]. However, where misreported probes occur, the magnitude of their error, and whether they are corrected remains poorly understood. As operator-reported geolocations continue to serve as a foundation for many studies [2, 6, 9, 12, 14, 19, 21, 26, 40], it is critical that we understand the successes and pitfalls of our community-contributed datasets and work towards an accurate ground truth.

In this work, we provide the most in-depth analysis to date of operator misreported geolocation, including where violations occur, what causes them, and how operators correct inaccuracies after we report their probes. First, we apply a simple methodology that uses speed-of-Internet calculations between probes and DNS root servers to conservatively test whether operator-reported probe locations are physically possible. We apply this methodology to RIPE Atlas [32], the largest and most geographically distributed community vantage-point collection, using 2.1K anycasted DNS root servers between January 2019 and April 2026. Our methodology infers a lower bound of 731 probes (2.4%) are likely not responding from their reported location, as doing so would exceed the physical limits of the speed of light in Internet fiber. We show that this seemingly small error rate has cascaded into hundreds of incorrect city-level geolocations in a prior study and thousands in scheduled RIPE Atlas measurements.

Errors in misreported geolocation are substantial; roughly 50% of violating probes likely responded at least 1,000 kilometers away from their reported location. We consider how inaccuracies prevail over time, finding that operators are slow to update geolocation and that the number of violating probes is increasing. The majority (52%) of geolocations are updated after the probe violates the speed of Internet for at least 13 weeks. Furthermore, the proportion of violating probes increases over time, highlighting that this problem is not diminishing.

We correct misreported geolocation at the source by working with RIPE Atlas to report physically impossible geolocations, prompting over 90% of responding operators to update their probe locations. To support this effort, we build a reporting pipeline that continually identifies probes that violate speed of Internet and use operator responses to validate our methodology. We open-source our methodology so researchers can identify physically impossible operator-reported geolocations in other settings. Many studies rely on operator-reported geolocation [3, 10, 11, 15, 17, 37, 47], and our approach provides a lightweight method to root out physically impossible geolocations to more accurately measure the Internet.

2 Background and Related Work

Many studies rely on RIPE Atlas to develop geolocation methodologies and understand the Internet, making them dependent on the accuracy of operator-reported locations.

RIPE Atlas. RIPE Atlas is a global measurement platform with 14,405 active vantage points across 183 countries and 4,475 ASes as of April 2026 [32]. It includes volunteer-hosted probes and organization-hosted anchors; anchors are considered more "accurate" [12] because they require RIPE-approved hardware. Unless stated otherwise, we refer to

both as RIPE Atlas probes. Operators report their probe's physical location by placing a pin on a map, and must manually update this location if the probe moves. The reported probe location will be skewed at most by 1 kilometer to protect the probe host's privacy. [1] When a probe becomes disconnected, the RIPE Atlas platform will regularly send the host email notifications to remind them of the disconnection. Critically, if the probe is physically moved, the operator must manually update their reported location. The RIPE platform allows other users to dispute a given probe's geolocation, although prior to our notification campaign, only 32 probe locations were disputed.

Research Using RIPE Atlas. The distributed architecture of the RIPE Atlas platform has proven useful in many research areas, especially in Internet infrastructure geolocation [3, 10, 15, 17, 37, 47]. Scheitle et al. [39] used RIPE Atlas probes to verify geolocation inferred from location hints embedded in reverse DNS records. Weinberg et al. [44] used RIPE Atlas probes to measure the latency to 2K proxy and VPN providers, and found at least 1/3 of VPNs falsified their advertised geolocation. Fanou et al. [16] use speed-of-light checks to identify that 20% of probed Google cache and DNS resolver IPs are wrongly geolocated. Darwich et al. [11] used RIPE Atlas to replicate the results of previous geolocation works. Gharaibeh et al. [18] used RIPE Atlas probes to validate four geolocation databases. RIPE themselves use RIPE Atlas probes to power a public IP geolocation tool, IPmap [13, 33].

Many geolocation studies using RIPE Atlas probes are sensitive to probe-location accuracy because latency-based methods often assume that nearby probes provide evidence for nearby targets. For instance, if RIPE IPmap sees a 0.2 ms RTT to a mislocated probe, it may wrongly assign the target IP to that incorrect location. Probes with incorrect reported locations have a larger impact on results where RIPE Atlas probes are sparsely deployed. Trammell and Kühlewind [41] also found that accuracy depends upon selecting a vantage point close to the target.

Given the importance of accurate geolocation, prior work filters unlikely RIPE Atlas probe locations, but stops short of understanding or correcting misreported geolocation at the source. Gharaibeh et al. [18] identify 24 questionable probes using basic sanity checks, while Darwich et al. [11] use speed-of-Internet violations among RIPE Atlas anchors and probes themselves to filter 9 of 723 anchors and 96 of 10K probes; we further describe their methodologies in Appendix B. We introduce a methodology that uses DNS root servers as the reference points for detecting physically impossible RIPE Atlas probe locations.

[1]We observed the same anonymization for the 6 RIPE Atlas probes we control.

DNS Root Servers. Domain Name System (DNS) root servers provide a globally distributed, operationally managed reference infrastructure for our measurements. The DNS root zone is served by 13 root servers, A through M [24], whose operators deploy thousands of anycasted instances worldwide [1]. Operating a root server instance requires meeting strict technical and administrative criteria that ensure the root server operators retains reliable control, consistent performance, and secure management [8, 25, 35, 36]. These requirements are designed to uphold the service expectations defined by the Root Server System Advisory Committee (RSSAC) [38].

We use DNS root-server instances to validate the reported geolocations of RIPE Atlas probes. Because root servers are anycasted, multiple instances can respond from the same IP address; we therefore use standardized `hostname.bind` responses to identify which instance answered each query [46]. Critically, RIPE Atlas already collects measurements between probes and DNS root servers, allowing us to validate reported probe locations using existing historical data rather than issuing large volumes of new measurements [11].

Precautions When Using DNS Root Servers. Prior work on DNS root servers motivates several precautions in our methodology. Li et al. [31] measured the RTTs from RIPE Atlas probes to C, D, and K roots and found that more than 2/3 queries were directed to a root instance that is not the closest to the Atlas probe. Similarly, Koch et al. [30] found that root servers with larger anycast deployments are more likely to experience inflated latencies. Other studies have analyzed the manipulation of DNS root server responses. In 2016, Jones et al. [28] compared RTTs from RIPE Atlas probes obtained by pings and `hostname.bind` and found only 11 instances of B root server response manipulations and none for L root. They also anecdotally identified one RIPE Atlas probe in New York that was mis-reported to be in Switzerland. In 2020, Wei and Heidemann [43] identified two types of DNS response manipulators: 1) *Overt spoofers* who responds to `hostname.bind` queries with their own identifiers; 2) *Covert delayers* who respond with identifiers that are the same as the root server instances. We minimize the risk of inflated latencies and these manipulators in Section 3.

3 Methodology

In this section we outline our methodology to measure whether probes' self-reported locations are physically impossible. We measure the latency between RIPE Atlas probes and DNS root servers using `hostname.bind` queries. We rely on simple physics to infer whether the measured latency between our measurement source and the destination's reported location exceeds the speed at which light travels in a fiber optic cable.

Obtaining `hostname.bind` Responses. RIPE Atlas operates a suite of built-in measurements that each probe conducts automatically. As part of these measurements, every connected probe issues a `hostname.bind` query to each of the 13 DNS root server identities every 240 seconds [34]. We obtain historical `hostname.bind` responses collected by all connected probes between January 2019 and April 2026 from RIPE Atlas' historical measurements database [23].

Sanitizing `hostname.bind` Responses. Wei and Heidemann [43] identified several types of DNS manipulators that may respond to `hostname.bind` queries in place of the actual root server instances (§2). Critically, they explained intercepting ISPs respond to `hostname.bind` queries with naming conventions that do not follow the naming scheme of A–M roots. To account for this phenomenon, we remove any `hostname.bind` responses that do not match the naming conventions of A – M roots. Covert delayers, who intercept the query and therefore cause inflated latencies, do not affect our results, as we explain in the following section.

Extracting Latency Values. Although `hostname.bind` was originally intended for operational diagnostics, we repurpose these measurements to infer latency between RIPE Atlas probes and individual DNS root server instances. When a RIPE Atlas probe sends a `hostname.bind` query, the responding DNS root server instance returns a hostname identifying itself, and the probe simultaneously measures the round-trip time (RTT) between transmitting the query and receiving the response. After parsing the `hostname.bind` responses, the resulting dataset consists of the initiating probe ID, DNS root server identity (A–M), the hostname of the responding anycast instance, and the RTT measured by the probe.

We sample measurements three times per day (06:00, 12:00, and 18:00 UTC) on the first day of each month, beginning on January 1, 2019. Then, for each `<probe,instance>` pair, we retain the minimum observed RTT to each root letter per day as the representative latency. We use minimum RTTs as the representative RTT because the speed of light in optical fiber or air imposes a hard lower bound. Unlike averaged RTTs, which can be inflated by congestion or queuing effects, the minimum RTT reflects the fastest observed path and cannot be artificially reduced.

Estimating Distance Between Source and Destination. The `hostname.bind` responses encode the location of the responding root server instance by matching an instance listed in `root-servers.org`, which provides the expected latitude and longitude of that instance as reported by the Root Server Operator (RSO). For hostnames not explicitly defined in the `root-servers.org` dataset (7%), we infer their location by matching the site prefix from the dataset. If all listed instances with a matching prefix (e.g., `nnn1-ams[4-6]`) share the same coordinates, we assume an unlisted instance matching the same format (e.g., `nnn1-ams1`) is co-located and assign it those coordinates accordingly. Combining these techniques we are able to geolocate the observed responding hostnames for 94% of our measurements between January 2019 and April 2026, removing the remaining 6% from our study.

To help derive the theoretical minimum latency between a vantage point and probe, we estimate the minimum Haversine distance ("as-the-crow-flies"), hereafter referred to as H_d, between the reported coordinates of every RIPE Atlas probe and every DNS root server instance. By leveraging RIPE Atlas provided geolocation snapshots, we ensure we are using the reported location of the probe from the date of each measurement. Recognizing that root server coordinates are often approximated (e.g., city-center or airport locations), we incorporate a substantial 100 km error radius,[2] so our analysis identifies only significant geolocation discrepancies.

Deriving Speed of Internet Theoretical RTTs. After estimating the minimum distance between each probe and server (H_d), we convert it to a minimum theoretical RTT. We calculate the theoretical *absolute minimum* RTT if the pings were sent at the speed of light in optical fiber, $H_d / \frac{2}{3}c$ [7]. We refer to the speed of Internet as SOI.

Identifying Violating Probes. We identify a RIPE Atlas probe as "violating" (i.e., we believe the probe is not in its operator-reported geolocation) if any RIPE Atlas probe measured latency is lower than the theoretical minimum latency from the self-reported geolocation to at least two DNS root servers operated by distinct root-server operators (a threshold selected to mitigate false positives, as discussed below). For DNS root server historical data, we take note of when within the past 5 years the RTT violation happened, if not presently occurring.

Precautions to Minimize False Positives. Minimizing false positives requires that DNS root servers report accurate locations and that probes can freely communicate with them. We consulted a root-server operator who noted that RSOs are generally highly reliable in tracking where their nodes are installed. Nevertheless, we take several precautions to minimize false positives.

First, we verify that the DNS root servers we rely on do not themselves violate SOI. Specifically, we use CAIDA's Ark [4], a set of 294 globally distributed vantage points whose exact latitude and longitude we confirmed directly with CAIDA operators who themselves deploy the servers. These nodes span 194 autonomous systems, 218 cities, 68 countries, and 6 continents; a map of Ark nodes is available in [4]. Using this independent vantage set, we issue `hostname.bind` queries and measure RTTs to DNS root servers, finding only one I-root instance appears to violate SOI. We flag it for further

[2]The vast majority of cities in the world are significantly less than 100 km in radius.

analysis, finding additional evidence in Section 4.5 that it is likely not in its reported place. For all other root instances, using CAIDA's Ark provides an external sanity check that root-server locations and paths are consistent with physical latency constraints.

Second, we require that a probe violate SOI constraints with at least two DNS root servers operated by distinct root-servers (`a–m`). This mitigates the risk that an error or misconfiguration by a single operator could incorrectly trigger a violation. Finally, we confirm that none of the violating probes originate from ASes known to perform DNS filtering, following Wei and Heidemann [43]. These precautions collectively reduce false positives, albeit at the cost of increasing false negatives.

Third, we remove Starlink probes from analysis. Satellite links can achieve lower latency than terrestrial fiber (e.g., via inter-satellite lasers) [7]. As a result, Starlink probes may legitimately exceed the terrestrially-defined SOI constraint, so we do not classify them as violators. We identify these probes via ASN 14593 and exclude them.[3]

We discuss the implications and limitations of this methodology in Section 4.5.

4 Results

Between January 2019 and April 2026, 731 unique probes (2.4%) violate speed of Internet (SOI). As of April 2026, 155 probes are in violation of speed of Internet. Violating probes are geographically widespread, often responding from thousands of kilometers away from their operator-reported geolocation. Slow operator response times exacerbates the issue: 50% of probes violate for over 13 weeks until operators update. We work with RIPE Atlas to report inaccurate geolocation and find that operators do update their probe locations, in some cases even moving them across continents.

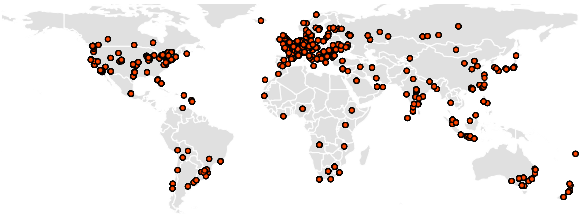

Figure 1: Reported geolocations of probes violating SOI between January 2019 and April 2026—Violating probes are geographically distributed.

4.1 Where and When Probes Violate

Violating probes exist all over the world. In Figure 1, we plot the reported locations all RIPE Atlas probes that violate the speed of Internet. Nearly half of violating probes claim to be in the USA (29%), Germany (11%), or France (9%). The majority of violating probes seem to respond from hundreds to thousands of kilometers from their operator-reported geolocation (Appendix C).

[3]Less than 1% of probes in April 2026 are announced by the Starlink ASN.

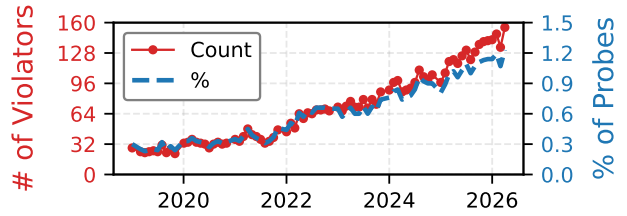

Figure 2: Violations over time—The number of violating probes increased from 22 probes in January 2019 to 155 probes in April 2026. The relative proportion of these probes quintupled during this interval.

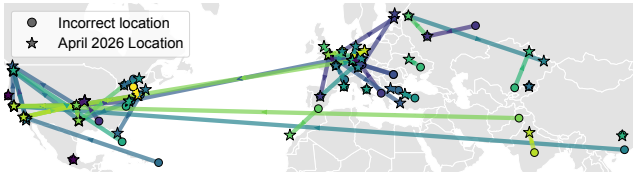

Figure 3: Operators updated their location after our report—Shown are all operator-updated probes above the equator, including substantial cross-country and transcontinental moves.

Across five years, the raw number of probes that violate SOI increases every year. Using historical data from built-in RIPE measurements, Figure 2 illustrates an increase in number and proportion of violating probes, from just 22 (0.23%) in January 2019 to 155 (1.23%) in April 2026. The increasing trend is not driven by older probes accumulating repeated violations: violating probes are significantly younger than other probes.[4] In total, 731 probes (2.4% of all probes) seem to be initiating from a different geolocation at least one point in time. Nearly half of probes (345/731) eventually stop violating SOI, with 35% of those eventually disconnected or abandoned after the SOI is violated.

4.2 Correcting Misreported Geolocation

We work with RIPE Atlas to correct misreported probe geolocation. Between July 2025 and April 2026, we reported to RIPE Atlas 258 probes that were violating SOI constraints at the time of reporting. RIPE Atlas then notified operators

[4]The median probe age is 837 days for violators and 1503 days for non-violators (Mann–Whitney $p < 10^{-200}$, Common-Language Effect Size = 0.38).

that their reported probe geolocation was inconsistent with network measurements. Operators could accept or deny the proposed correction, update the probe location when appropriate, and optionally provide a comment.

As of April 2026, 103 operators responded to our notifications: 93 accepted and updated their operator-reported geolocation, while 10 denied that their location was inaccurate. Notably, 4% of probes whose locations were corrected had violated SOI on only one day, suggesting that even sparse violations can reveal real misreported geolocation. Figure 3 plots operator-corrected locations above the equator, where most violating probes are located, and shows multiple intercontinental updates. This is consistent with our historical analysis: 50% of operators whose probes no longer violate SOI updated their geolocation over 1,000 km, indicating that their probes could have incorrectly reported locations on different continents (Appendix Figure 6).

Some denied reports involve probes hosted on third-party infrastructure, rather than directly operated by the probe owner. Of the ten operators who denied that their location was wrong, four provided comments stating that their probe ran on a virtual private server hosted by a third-party provider. These operators reported that they had confirmed the leased server was in the country listed by the probe. We manually checked all 10 cases and found that they either do not appear to be country-level violations or are located near borders, but still appear largely inconsistent with the reported city-level location. Interestingly, hosting-network probes—probes for which the operator themselves lack ground truth—are disproportionately likely to violate speed of Internet. Using the March 2026 ASdb snapshot [48], we classify probes based on the ASNs announcing their IP addresses. Although most violating probes are ISP-hosted, ISP probes are significantly less likely to violate speed of Internet overall ($p = 6.19 \times 10^{-4}$, risk ratio = 0.56), whereas probes in hosting networks are significantly more likely to do so ($p = 1.81 \times 10^{-6}$, risk ratio = 2.53).[5]

4.3 Understanding Contributing Factors

Through a manual analysis of reported geolocations and round trip times, we discover that tardiness in updating geolocation after a move, and initial misconfiguration of geolocation are the most likely and common contributors to SOI violations.

4.3.1 Tardiness. At times, probes reporting RTTs that violate SOI are simply late to report a move. We see 41% of violating probes appear to stop violating SOI after the operator updates the probe location. For example, Figure 4 shows longitudinal measurements from a probe whose operator reported a location in Germany. From 2019 to 2022, its minimum RTTs to three root letter instances hosted in Germany were compatible with the operator's claimed Germany location (≈19 ms).

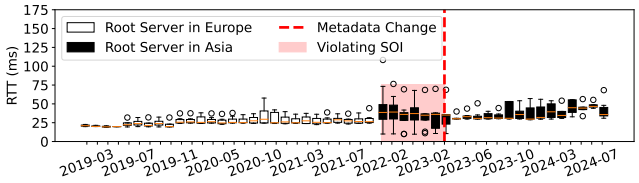

Figure 4: Moving from Germany to Brunei Case Study—The operator's probe reports a SOI violation between February 2022 and April 2023, before updating their location from Germany to Brunei in April 2023.

However, after a brief disconnect in 2022, the root server responses changed drastically. Subsequent measurements show RTTs under 10 ms to two root letters in Bangkok, implying the probe should be within 1000 km of Bangkok, even though its geolocation still shows a location in Germany. Finally, in 2023, the operator updated the probe's geolocation to Brunei, which is consistent with the inferred proximity to Bangkok.

4.3.2 Initial Misreporting. Some long-term violations stem from probes configured with incorrect geolocation at deployment. A quarter (23.3%) of violating probes appear to violate SOI their entire existence (either disconnecting or still violating as of April 2026). An additional 12.72% appear to violate SOI from their initial misconfiguration until the operator updates the location (Section 4.3.1).

Consider Probe 1000011, which violated the speed of Internet from its deployment in October 2019 until June 2023. It reported a location in California until 2023, when the geolocation updated to Virginia. However, its minimum RTT across all 12 root letters remained stable—unexpected for a coast-to-coast move. Moreover, RTTs to root servers located along the United States East Coast (e.g., Miami, Florida; Dulles, Virginia; Washington, D.C.) violated speed of Internet at the California location, but stopped once the location was updated. This suggests the probe was likely near Virginia all along. The location change coincides with an upstream provider switch from Cogent to Lumen, hinting that a network change may have prompted the operator to verify the probe's configuration and update the geolocation. Roughly half (43/93) of responding operators in our RIPE campaign from Section 4.2 appear to have initially misconfigured their probe location.

[5]We use a two-proportion z-test to assess statistical significance. Effect sizes are reported as risk ratios, defined as the ratio of the probability of violation within a given class (e.g., ISPs or hosting providers) to that of probes outside that class.

4.3.3 Lack of Violation-Resolving Update. The remaining 35% of probes (256/731) either never update their reported geolocation (187/256) or update with no violation resolution (69/256), even though their violation patterns change over time. For some probes, their geolocation remains unchanged through April 2026, despite abrupt changes in RTTs to the same root instances or to an entirely different set of instances (see Appendix Figure 9). For others, violations are observed only with respect to a consistent subset of DNS instances; consequently, if those instances do not respond, no violation is observed.

4.4 Impact of Misreported Geolocation

Relying on a small number of incorrectly geolocated probes can subtly cascade into many more incorrect city-level geolocations. We first revisit Du et al.'s evaluation of RIPE IPmap's single-radius engine, which infers city-level target locations from RIPE Atlas latency measurements [13]. Comparing the 2,000 probes used in their study with probes that violated speed of Internet during their measurement period (September 2019–January 2020), we find that 33 probes appear to violate speed of Internet. These probes risk incorrect city-level geolocation for 4.3% of the study's 8,000 candidate IPs. We then analyze the past year of RIPE Atlas measurements scheduled by the RIPE IPmap single-radius engine [23]. During this period, IPmap geolocated 98K candidate IPs using latency to the closest RIPE Atlas probe among 2.8K probes. Of these probes, 28 are likely mislocated, cascading to 1,125 distinct target IPs potentially geolocated to the wrong city.

4.5 Methodological Sensitivity

Our methodology is conservative by design, but SOI-based filtering remains sensitive to the measurement infrastructure used as vantage points. We evaluate our methodology's sensitivity to two parameters: a 100 km location buffer and requiring agreement from at least two independent root letters. On average, removing the buffer adds 26 violations per month, while removing the two-letter requirement adds 43 violations per month. The 100 km buffer is necessary: 9.5% of responding root instances are reported at default IATA coordinates, which may not reflect their true physical locations. We therefore include a buffer to account for the likely discrepancy between the reported and actual instance locations. The two-letter requirement is important in practice: under a one-letter rule, we found 53 probes near Santiago, Chile that violated only a single I-root instance listed in Concepción (netnod-cl-ccp), while nearby root instances did not violate speed of Internet. Notably, in Section 3, we found the same I-root instance violated our sanity check. We reported this case to I-root, who preliminarily agreed that the instance is likely in Santiago, not Concepción; our two-letter rule therefore prevented us from falsely labeling these probes as violators.

A comparison with Darwich et al. [11] further shows that SOI-based filtering is sensitive to which vantage points are used to evaluate probes. Using Darwich et al.'s experiment window (2023-04-24 to 2023-05-03), we identify the same number of violating probes (96), but only 63 overlap. The difference largely stems from the vantage points: Darwich et al. use validated anchors, while we use root DNS servers. For probes identified only by Darwich et al., their anchors are on average 150 km closer than our root DNS servers; conversely, for probes identified only by our method, root DNS servers are on average 234 km closer than their anchors (see Appendix B.1). Nevertheless, using root DNS servers enables analyses without issuing new measurements (e.g., avoiding the millions of RIPE Atlas credits required by [11] for a single snapshot) and allows for retrospective studies.

Our methodology's conservative choices reduce false positives, but also limit what our method can detect. Our focus on RIPE Atlas may not represent all vantage points with operator-reported geolocations, although its scale underscores the importance of improving its accuracy. Our method likely underestimates both prevalence and distance-error magnitude because true violations may remain undetected when paths are slower than ideal, congestion-free links. It may also miss violations in regions where root DNS servers are far from probes or served by few operators, since longer paths loosen RTT bounds and our method requires agreement from at least two distinct root-server operators. Future work should explore new methods to detect SOI violations, including providing tighter—but still accurate—bounds for detecting violations.

5 Conclusion

Operator-reported geolocation remains a critical but weakly validated input to Internet measurement studies. We show that misreported probe locations can persist for months, appear thousands of kilometers from their reported locations, and cascade into downstream geolocation errors. Critically, we take a first step toward correcting misreported geolocation at the source by working with RIPE Atlas to report and fix physically impossible reported geolocations to operators. Further, we open-source our methodology and maintain a continually updated list of currently violating RIPE Atlas probes. A more accurate foundation for Internet measurement begins with rooting out the physically impossible locations hidden in today's community-reported datasets.

References

[1] Root Server Technical Operations Association. 2025. Root Server Instances. https://root-servers.org/.

[2] Vaibhav Bajpai, Steffie Jacob Eravuchira, and Jürgen Schönwälder. 2015. Lessons learned from using the ripe atlas platform for measurement research. *ACM SIGCOMM Computer Communication Review* 45, 3 (2015), 35–42.

[3] Vaibhav Bajpai, Steffie Jacob Eravuchira, and Jürgen Schönwälder. 2017. Dissecting Last-mile Latency Characteristics. *SIGCOMM Comput. Commun. Rev.* 47, 5 (oct 2017), 25–34. https://doi.org/10.1145/3155055.3155059

[4] CAIDA. 2024. Archipelago (Ark) Measurement Infrastructure. https://www.caida.org/projects/ark/.

[5] Matt Calder, Xun Fan, Zi Hu, Ethan Katz-Bassett, John Heidemann, and Ramesh Govindan. 2013. Mapping the Expansion of Google's Serving Infrastructure. In *Proc. of the 2013 IMC*.

[6] Massimo Candela, Enrico Gregori, Valerio Luconi, and Alessio Vecchio. 2019. Using RIPE atlas for geolocating IP infrastructure. *IEEE Access* 7 (2019), 48816–48829.

[7] Aizaz U Chaudhry and Halim Yanikomeroglu. 2022. Optical Wireless Satellite Networks versus Optical Fiber Terrestrial Networks: The Latency Perspective: Invited Chapter. In *30th Biennial Symposium on Communications 2021*. Springer, 225–234.

[8] Internet Systems Consortium. 2025. Hosting an F-Root Node. https://www.isc.org/froot-technical/.

[9] Lorenzo Corneo, Maximilian Eder, Nitinder Mohan, Aleksandr Zavodovski, Suzan Bayhan, Walter Wong, Per Gunningberg, Jussi Kangasharju, and Jörg Ott. 2021. Surrounded by the clouds: A comprehensive cloud reachability study. In *Proceedings of the Web Conference 2021*. 295–304.

[10] Lorenzo Corneo, Maximilian Eder, Nitinder Mohan, Aleksandr Zavodovski, Suzan Bayhan, Walter Wong, Per Gunningberg, Jussi Kangasharju, and Jörg Ott. 2021. Surrounded by the Clouds: A Comprehensive Cloud Reachability Study. In *Proceedings of the Web Conference 2021* (Ljubljana, Slovenia) *(WWW '21)*. Association for Computing Machinery, New York, NY, USA, 295–304. https://doi.org/10.1145/3442381.3449854

[11] Omar Darwich, Hugo Rimlinger, Milo Dreyfus, Matthieu Gouel, and Kevin Vermeulen. 2023. Replication: Towards a Publicly Available Internet scale IP Geolocation Dataset. In *Proceedings of the 2023 ACM on Internet Measurement Conference*. 1–15.

[12] Lily Davisson, Joakim Jakovleski, Nhiem Ngo, Chau Pham, and Joel Sommers. 2021. Reassessing the constancy of end-to-end internet latency. In *IFIP Network Traffic Measurement and Analysis Conference*.

[13] Ben Du, Massimo Candela, Bradley Huffaker, Alex C. Snoeren, and kc claffy. 2020. RIPE IPmap active geolocation: mechanism and performance evaluation. *SIGCOMM Comput. Commun. Rev.* 50, 2 (may 2020), 3–10. https://doi.org/10.1145/3402413.3402415

[14] Rodérick Fanou, Pierre Francois, and Emile Aben. 2015. On the diversity of interdomain routing in africa. In *Passive and Active Measurement: 16th International Conference, PAM 2015, New York, NY, USA, March 19-20, 2015, Proceedings 16*. Springer, 41–54.

[15] Roderick Fanou, Pierre Francois, Emile Aben, Michuki Mwangi, Nishal Goburdhan, and Francisco Valera. 2017. Four years tracking unrevealed topological changes in the African interdomain. *Computer Communications* 106 (2017), 117–135.

[16] Rodérick Fanou, Gareth Tyson, Eder Leao Fernandes, Pierre Francois, Francisco Valera, and Arjuna Sathiaseelan. 2018. Exploring and Analysing the African Web Ecosystem. *ACM TWEB* (2018).

[17] Romain Fontugne, Anant Shah, and Kenjiro Cho. 2020. Persistent Last-mile Congestion: Not so Uncommon. In *Proceedings of the ACM Internet Measurement Conference* (Virtual Event, USA) *(IMC '20)*. Association for Computing Machinery, New York, NY, USA, 420–427. https://doi.org/10.1145/3419394.3423648

[18] Manaf Gharaibeh, Anant Shah, Bradley Huffaker, Han Zhang, Roya Ensafi, and Christos Papadopoulos. 2017. A look at router geolocation in public and commercial databases. In *Proceedings of the 2017 Internet Measurement Conference*. 463–469.

[19] Petros Gigis, Vasileios Kotronis, Emile Aben, Stephen D Strowes, and Xenofontas Dimitropoulos. 2017. Characterizing user-to-user connectivity with RIPE Atlas. In *Proceedings of the 2017 Applied Networking Research Workshop*. 4–6.

[20] Bamba Gueye, Artur Ziviani, Mark Crovella, and Serge Fdida. 2004. Constraint-based geolocation of internet hosts. In *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*. 288–293.

[21] Thomas Holterbach, Cristel Pelsser, Randy Bush, and Laurent Vanbever. 2015. Quantifying interference between measurements on the RIPE Atlas platform. In *Proceedings of the 2015 Internet Measurement Conference*. 437–443.

[22] Zi Hu, John Heidemann, and Yuri Pradkin. 2012. Towards geolocation of millions of IP addresses. In *Proceedings of the 2012 Internet Measurement Conference*. 123–130.

[23] IANA. 2025. RIPE Atlas API. https://atlas.ripe.net/api/v2/measurements/.

[24] IANA. 2025. Root Servers. https://www.iana.org/domains/root/servers.

[25] ICANN. 2025. Hosting IMRS in your network. https://www.dns.icann.org/imrs/host/.

[26] Mattia Iodice, Massimo Candela, and Giuseppe Di Battista. 2019. Periodic path changes in RIPE atlas. *IEEE Access* 7 (2019), 65518–65526.

[27] Katherine Izhikevich, Geoffrey M Voelker, Stefan Savage, and Liz Izhikevich. 2024. Using Honeybuckets to Characterize Cloud Storage Scanning in the Wild. (2024).

[28] Ben Jones, Nick Feamster, Vern Paxson, Nicholas Weaver, and Mark Allman. 2016. Detecting DNS root manipulation. In *International Conference on Passive and Active Network Measurement*. Springer, 276–288.

[29] Ethan Katz-Bassett, John P John, Arvind Krishnamurthy, David Wetherall, Thomas Anderson, and Yatin Chawathe. 2006. Towards IP geolocation using delay and topology measurements. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement*. 71–84.

[30] Thomas Koch, Ethan Katz-Bassett, John Heidemann, Matt Calder, Calvin Ardi, and Ke Li. 2021. Anycast In context: a tale of two systems. In *Proceedings of the 2021 ACM SIGCOMM 2021 Conference* (Virtual Event, USA) *(SIGCOMM '21)*. Association for Computing Machinery, New York, NY, USA, 398–417. https://doi.org/10.1145/3452296.3472891

[31] Zhihao Li, Dave Levin, Neil Spring, and Bobby Bhattacharjee. 2018. Internet anycast: performance, problems, & potential. In *Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication* (Budapest, Hungary) *(SIGCOMM '18)*. Association for Computing Machinery, New York, NY, USA, 59–73. https://doi.org/10.1145/3230543.3230547

[32] RIPE NCC. 2024. RIPE Atlas API. https://atlas.ripe.net/coverage/.

[33] RIPE NCC. 2024. RIPE IPmap. https://ipmap.ripe.net/.

[34] RIPE NCC. 2025. Built-in Measurements. https://atlas.ripe.net/docs/getting-started/built-in-measurements.

[35] RIPE NCC. 2025. K-root FAQs. https://hosted-dns.ripe.net/faq/.

[36] Netnod. 2025. Technical requirements for hosting an I-root node. https://www.netnod.se/technical-requirements-for-hosting-an-i-root-node/.

[37] Audrey Randall, Enze Liu, Ramakrishna Padmanabhan, Gautam Akiwate, Geoffrey M. Voelker, Stefan Savage, and Aaron Schulman.

2021. Home is where the hijacking is: understanding DNS interception by residential routers. In *Proceedings of the 21st ACM Internet Measurement Conference* (Virtual Event) *(IMC '21)*. Association for Computing Machinery, New York, NY, USA, 390–397. https://doi.org/10.1145/3487552.3487817

[38] ICANN Root Server System Advisory Committee (RSSAC). 2025. Service Expectations of Root Servers. https://itp.cdn.icann.org/en/files/root-server-system-advisory-committee-rssac-publications/rssac-001-root-service-expectations-04dec15-en.pdf.

[39] Quirin Scheitle, Oliver Gasser, Patrick Sattler, and Georg Carle. 2017. HLOC: Hints-based geolocation leveraging multiple measurement frameworks. In *2017 Network Traffic Measurement and Analysis Conference (TMA)*. 1–9. https://doi.org/10.23919/TMA.2017.8002903

[40] Andrei M Sukhov and Artem V Onoprienko. 2014. Evaluating the effectiveness of geographic routing based on RIPE Atlas data. In *2014 22nd Telecommunications Forum Telfor (TELFOR)*. IEEE, 107–110.

[41] Brian Trammell and Mirja Kühlewind. 2018. Revisiting the Privacy Implications of Two-Way Internet Latency Data. In *Passive and Active Measurement*, Robert Beverly, Georgios Smaragdakis, and Anja Feldmann (Eds.). Springer International Publishing, Cham, 73–84.

[42] Gerry Wan, Liz Izhikevich, David Adrian, Katsunari Yoshioka, Ralph Holz, Christian Rossow, and Zakir Durumeric. 2020. On the Origin of Scanning: The Impact of Location on Internet-Wide Scans. In *ACM Internet Measurement Conference*.

[43] Lan Wei and John Heidemann. 2020. Whac-A-Mole: Six Years of DNS Spoofing. arXiv:2011.12978 [cs.CR] https://arxiv.org/abs/2011.12978

[44] Zachary Weinberg, Shinyoung Cho, Nicolas Christin, Vyas Sekar, and Phillipa Gill. 2018. How to Catch when Proxies Lie: Verifying the Physical Locations of Network Proxies with Active Geolocation. In *Proceedings of the Internet Measurement Conference 2018* (Boston, MA, USA) *(IMC '18)*. Association for Computing Machinery, New York, NY, USA, 203–217. https://doi.org/10.1145/3278532.3278551

[45] Bernard Wong, Ivan Stoyanov, and Emin Gün Sirer. 2007. Octant: A Comprehensive Framework for the Geolocalization of Internet Hosts.. In *NSDI*, Vol. 7. 23–23.

[46] S. Woolf and D. Conrad. 2007. *Requirements for a Mechanism Identifying a Name Server Instance*. RFC 4892. RFC Editor. http://www.rfc-editor.org/rfc/rfc4892.txt http://www.rfc-editor.org/rfc/rfc4892.txt.

[47] Marco Zennaro, Cristel Pelsser, Franck Albinet, and Pietro Manzoni. 2020. Evaluating the performance of NRENs in deploying IoT in Africa: The case for TTN. In *2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 1–4.

[48] Maya Ziv, Liz Izhikevich, Kimberly Ruth, Katherine Izhikevich, and Zakir Durumeric. 2021. ASdb: a system for classifying owners of autonomous systems. In *Proceedings of the 21st ACM Internet Measurement Conference*. 703–719.

A Appendix

A.1 Ethics

All RIPE data used in this analysis is publicly available at `https://atlas.ripe.net/api/v2/measurements/`. The RIPE location data is anonymized by a few blocks. This work does not raise any ethical issues.

B Prior Work

Given the importance of geolocation, relatively few studies have validated the accuracy of RIPE Atlas probe locations. Gharaibeh et al. [18] were the first to perform a basic "sanity check" in 2017. They considered the reported locations untrustworthy if they had default country coordinates (typically geographic centers) or if two probes were connected to the same router, but reported locations more than 100km apart—thereby identifying 24 questionable probes. In 2020, Darwich et al. [11] applied a recursive method based on speed-of-Internet violations, filtering 9 of 723 anchors and 96 of 10K probes. Their approach compared RTTs exclusively among RIPE Atlas anchors and probes, iteratively removing nodes whose latencies violated physical feasibility constraints. Both methods detect groups with violations, but cannot isolate the specific mislocated probe. Further, no work provides an in-depth analysis of where violations occur or what causes them.

These prior works rely on the same vantage points, the RIPE Atlas probes themselves, to identify other potentially violating RIPE Atlas probes. In contrast, we use a publicly available set of globally distributed nodes that are subject to far stricter operational and management requirements: the DNS root servers.

B.1 Sources of Discrepancies with Darwich et al. [11]

Differences between our method and Darwich et al. [11] arise from both vantage point selection and data availability. Darwich et al. propose a RIPE Atlas sanitization procedure that (1) performs anchor-to-anchor measurements, (2) iteratively removes anchors with SOI violations, and (3) tests probes against the validated anchors. We extract the 96 violating probes reported in their open-source repository and reproduce their time window (2023-04-24 to 2023-05-03), identifying 96 violating probes with 63 overlapping between the two methods. Of the probes reported by Darwich et al., RIPE Atlas contains the necessary geolocation and root DNS measurements for 79 probes, and their repository is missing data for 3 probes. Among the remaining probes, we find three cases where the anchor used in their probe-anchor measurements violates speed of Internet under our analysis, rather than the probe itself.

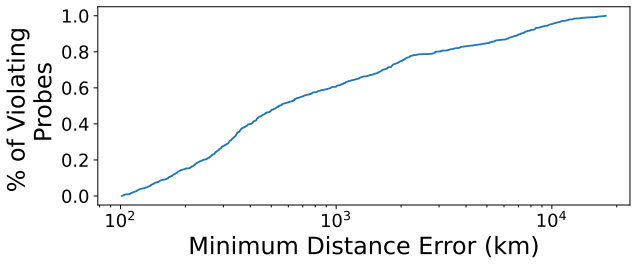

Figure 5: Distance Error—Nearly half of likely misreported probes are likely over 750 kilometers away from their reported locations.

The remaining discrepancies are explained by differences in distances between probes and their respective vantage points. For probes identified only by Darwich et al., the destination anchors are on average 150 km closer to the probes than the root DNS servers used in our analysis, reducing the likelihood of observing violations under our method. Conversely, for probes identified only by our method, root DNS servers (See Figure 10, which plots the geolocations of the instances of each root letter) are on average 234 km closer than the anchors used by Darwich et al., reducing the likelihood of observing violations under theirs. Together, these results show that SOI-based sanitization depends on both the placement and selection of measurement infrastructure.

C Distribution of Probe Distance Errors

The majority of violating probes appear to be hundreds to thousands of kilometers from their operator-reported geolocations. For every <probe, vantage point> pair that violates the SOI constraint, we calculate the minimum error of distance by (1) selecting all VPs that violated SOI, (2) multiplying the minimum observed RTT by SOI, to derive the furthest distance, d, a probe could be if it traveled at exactly SOI, (3) calculate the theoretical distance, d_{theory}, between our VP and the operator-reported probe location, (4) subtract d from d_{theory}, recording the minimum distance error.

We analyze the full distribution of distance errors: nearly 50% exceed 1,000 km (Figure 5). The large distance errors suggest violating probes are likely on different continents, as we discuss in Section 4.3.1 and show in Section 4.2. Figure 6 further shows that 50% of operators whose probes no longer violate SOI updated their geolocation over 1,000 km, indicating that their probes could have incorrectly reported locations on different *continents*.

D Time to Fix Geolocation

The majority of violating probes take at least 13 weeks to update their geolocation. Figure 7 plots the time from when

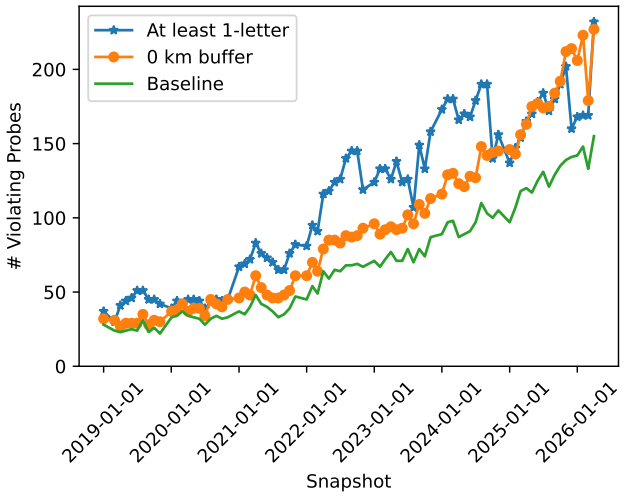

Figure 8: Testing removal of 100 km buffer and independent root letter agreement—On average, removing the buffer adds 26 violations per month, while removing the two-letter requirement adds 43 violations per month.

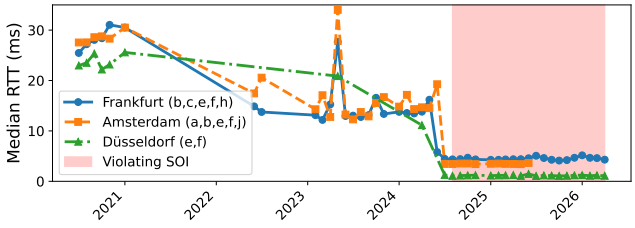

Figure 9: RTTs of a probe that likely moved, but has yet to update its geolocation—This probe shows a significant difference in RTTs across numerous root instances (see the letters in the legend) in the same cities over time.

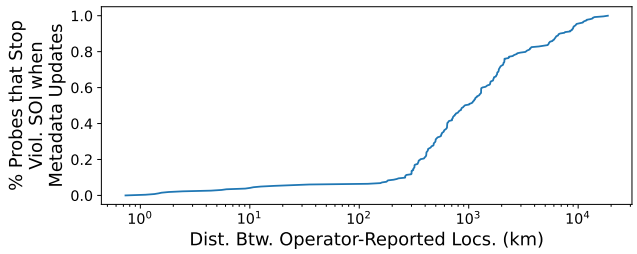

Figure 6: Distance Between Reported Locations Pre- and Post-Violations—Nearly half of operators resolved violations by updating locations over 1,000 km.

a probe violates the SOI until an operator updates its geolocation such that it no longer violates SOI.

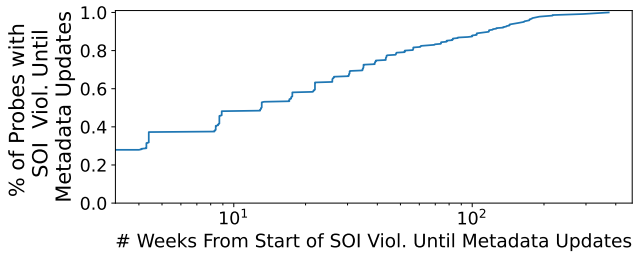

Figure 7: Number of Weeks Between SOI Violation and Geolocation Update—Of the probes that do not disconnect, it takes over 13 weeks for 50% of such probes to update their geolocation. Notably, the SOI violation no longer occurs.

E Analysis Continued

In Figure 8, we show how the number of violating probes follows similar temporal patterns regardless of removing the 100 km buffer or multi-letter agreement criteria.

In Figure 9, we show how a probe could keep the same metadata over time, but experience vastly different RTTs to the same host sites over time. While we cannot say for certain the probe physically moved, the sudden drop in RTTs to Frankfurt, Amsterdam, and Düsseldorf would suggest the probe may have moved closer to these locations.

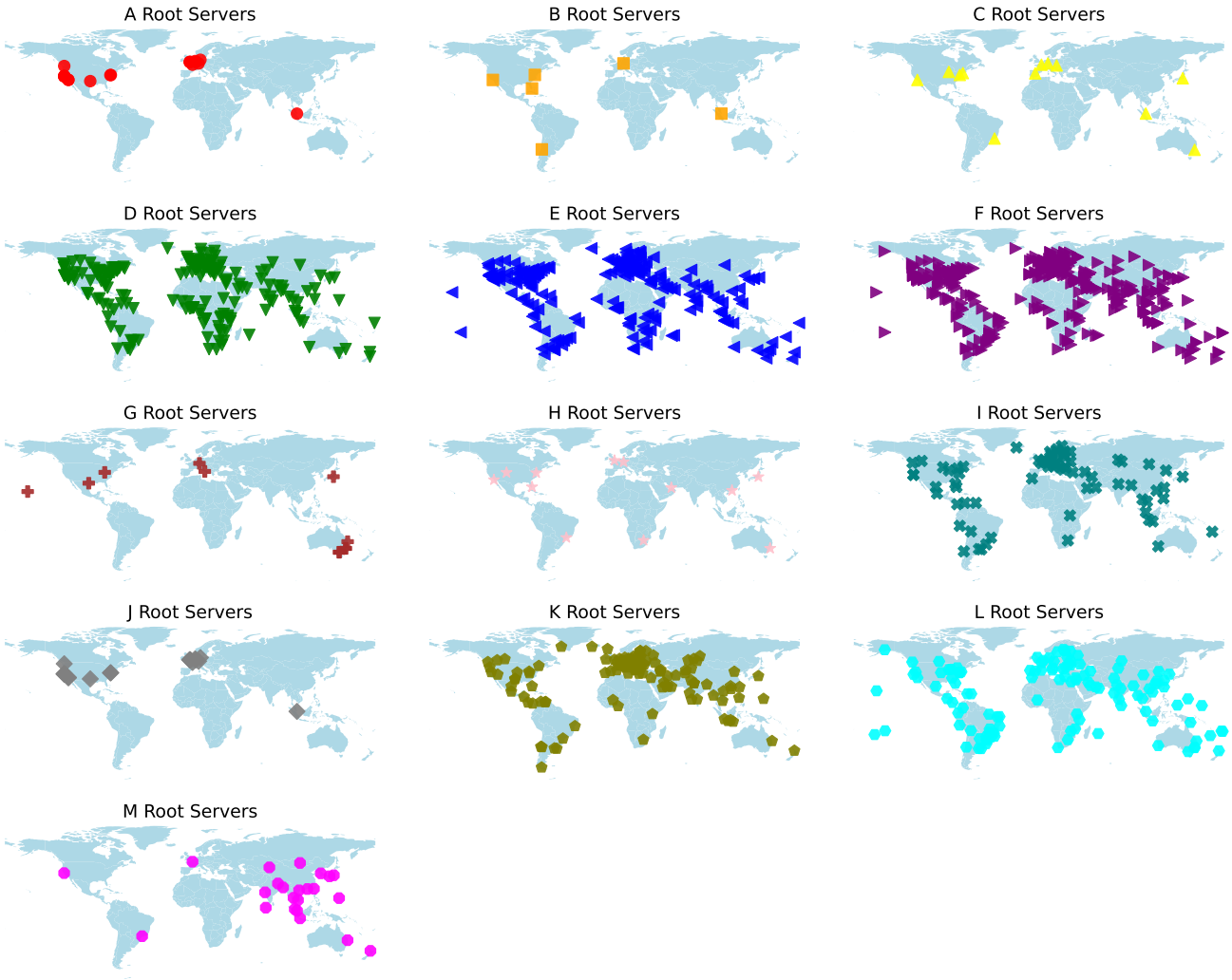

Figure 10: DNS Root Letter Instance Geographic Distribution.