Rethinking Geolocalization on the Internet

Augustin Laouar
École Normale Supérieur de Lyon

Loïc Desgeorges
Université Claude Bernard Lyon 1

Paul Schmitt
California Polytechnic State University

Francesco Bronzino
École Normale Supérieur de Lyon

Abstract

Location underpins critical Internet services, yet our primary mechanism for Internet localization, IP-based geolocation, fails to meet the needs of all stakeholders. User location is conflated with network location, leading to a fundamental mismatch between the goals of content providers, infrastructure operators, and regulators. As users increasingly adopt privacy-preserving technologies that obscure their network identity, this mismatch becomes more pronounced, making localization even more challenging. We argue that the problem cannot be solved by simply improving the accuracy of incumbent mechanisms that are inappropriately applied today to solve multiple, unrelated problems. Instead, we require a new approach for localization on the Internet.

1 Context

The modern Internet relies on IP geolocation as the predominant method to determine device and user positions, using their *public IP address* as the key identifier. Commercial providers (*e.g.*, IPInfo, MaxMind) infer locations from *network-level evidence* tied to these addresses, combining static data (RIR allocations, WHOIS, routing) with dynamic signals (reverse DNS, telemetry, latency) to probabilistically map each IP.

Unfortunately, IP geolocation has been long known to be an unreliable, often greatly inaccurate. Prior work has sought to improve accuracy through various techniques [9], but all remain limited by their reliance on the public IP address as the primary identifier. Unfortunately, this dependency becomes increasingly problematic as the core assumption, that network identifiers can reliably map to physical geography, further breaks down.

Privacy-preserving tools such as VPNs and Tor have grown in adoption, helping users protect privacy and bypass censorship or geo-restrictions. These tools complicate accurate IP geolocation by hiding real IP addresses. While traditionally limited by complexity and overhead, newer solutions now integrate anonymization directly into operating systems and browsers, easing mainstream adoption. Apple's iCloud Private Relay (PR) [1], Google's IP Protection [4], and Microsoft's Edge Secure Network [6] are examples of privacy preserving architectures that route traffic through multi-hop tunnels built on the MASQUE protocol, forming a performant privacy-preserving overlay network [8].

If-and-when mainstream systems and browsers enable relays by default, a large share of Web traffic will traverse them, further complicating geolocation. Relay providers try to mitigate this by publishing egress IPs and their logical locations. Apple, for example, discloses IP prefixes mapped to cities or regions [2] to guide IP geolocation services in locating their users. However, this approach introduces a fundamental inconsistency: advertised egress IP prefix locations are often significantly disconnected from their actual infrastructure locations.

This evolving landscape leads to two key effects. First, the concept of geolocation is becoming confusing: services face the task of mapping IP addresses to both infrastructure and user locations, merging different objectives into a single API. For example, with relay services, IP-geolocation databases must decide whether to map the relay egress node or the users behind them, who may be located hundreds of kilometers away. This mismatch produces discrepancies that will grow with adoption and threaten services requiring accuracy. Second,

most current patches depend on commercial entities whose motivations are self-interest rather than the integrity of Internet infrastructure[1]. This creates an ecosystem subject to the decisions of private companies.

These factors raise a fundamental question: should we continue patching the IP geolocation ecosystem, or does the shift toward privacy-preserving browsing call for a clean slate approach? We argue that it is time for the networking community to *reimagine geolocation* for an era where IP addresses are no longer identifiers of users.

2 Discussion

However, IP geolocation remains highly effective for its intended purpose: *locating infrastructure*. CDNs combine it with traceroute, latency probes, BGP inspection, and real-user monitoring to optimize content delivery, while operators and researchers use it to detect anomalies, study attacks, and analyze Internet topology. The real issue arises when infrastructure and user localization are conflated, underscoring the need for a dedicated service for localizing *users*.

2.1 User Localization: a Wishlist

Rethinking user localization requires meeting several fundamental properties while addressing their trade-offs.

Accuracy. Must provide reliable, quantifiable distance error for a user's real location, not infrastructure, though this conflicts with privacy and verification requirements that follow.

Verifiability. Services must trust the reported position, using lightweight checks (e.g., latency, BGP, attestation). This verification process must balance trust with user privacy

Privacy-conscious. Users should control the granularity of shared location (country to city/coordinates). This granularity control creates inherent trade-offs with accuracy and verifiability, as coarser location data may be less useful for services while being harder to verify independently.

Scalable. The system should support Internet-scale usage with low overhead. This might be challenging if verification mechanisms are implemented for all localization requests.

Frictionless. As user experience is a key factor on today's Internet, users should not be oversolicited by location requests or verification procedures.

Open. Prior works provide partial building blocks toward a privacy-preserving geolocation system [5, 7], but do not offer a complete solution. To ensure transparency, trust, and broad interoperability, such a system should be open, publicly specified through standardization bodies, and built from the ground up for independent implementation and verification.

2.2 Rethinking User Geolocalization

The properties listed in the previous Section involve inherent trade-offs, and designing a system that satisfies all of them remains largely unresolved. However, we believe a practical solution is possible. We sketch a high-level design called Geo-Certification Authorities (Geo-CAs), where a trusted third party attests both the user's position (furnished via reliable signals) and the minimum spatial granularity required by each service. Trust among the third party, user, and location-based service (LBS) is anchored in a certificate chain, analogous to X.509. To ensure scalability and minimize latency, the third party operates offline, issuing long-lived certificates that define each LBS's authorization scope and short-lived tokens attesting user positions, without being involved in subsequent connections. One possible implementation could exchange certificates and tokens during the TLS handshake, thereby embedding localization proofs directly into the secure channel setup.

The Geo-CA workflow (Figure 1) unfolds in four phases: **(i) LBS registration**—the service obtains a certificate attesting its authorized spatial granularity; **(ii) User registration**—the client receives a bundle of signed geo-tokens for multiple granularity levels (*e.g.*, exact point, city, country); **(iii) Service authorization**

[1] While we assume the companies involved do not wish to degrade Internet localization, their design choices may not generalize.

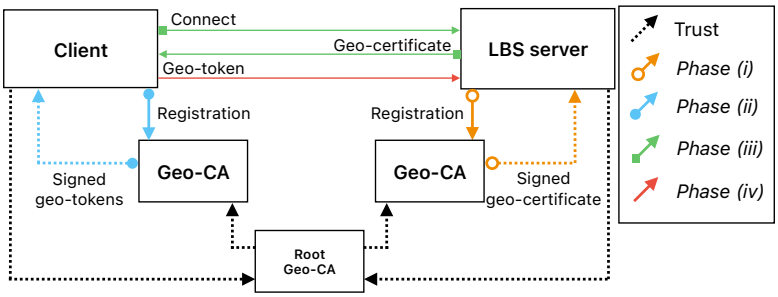

Figure 1: Geo-CA workflow.

attestation—the service presents its Geo-CA certificate, which the client verifies against its trusted root Geo-CAs; and **(iv) Client attestation**— the client sends a geo-token of the requested granularity, which the LBS verifies for validity.

This design surfaces several open challenges: preventing token replay (possibly via adapted DPoP [3]), protecting user coordinates through privacy-preserving issuance (e.g., zero-knowledge proofs, blind signatures) while balancing privacy and verifiability, deciding update frequency to trade off freshness against privacy and overhead, avoiding centralization risks seen in Web PKI, and establishing standards for expressing and verifying least-privilege spatial granularity.

References

[1] Apple Inc. *iCloud Private Relay — Overview*. `https://www.apple.com/icloud/docs/iCloud_Private_Relay_Overview_Dec2021.pdf`. White paper, accessed 16 Jun 2025. 2021.

[2] Apple Inc. *Private Relay Egress IPs*. `https://mask-api.icloud.com/egress-ip-ranges.csv`. CSV file, accessed 20 Jun 2025. 2025.

[3] Daniel Benjamin et al. *OAuth 2.0 Demonstrating Proof-of-Possession at the Application Layer (DPoP)*. Tech. rep. RFC 9449. Internet Engineering Task Force, July 2023. URL: `https://www.rfc-editor.org/info/rfc9449`.

[4] Google Chrome Team. *IP Protection: Privacy Sandbox Explainer*. `https://developer.chrome.com/docs/privacy-sandbox/ip-protection/`. Accessed 16 Jun. 2025. 2025.

[5] Scott Hendrickson et al. *Privacy Pass Geolocation Hint Extension*. Work in Progress. July 2023. URL: `https://www.ietf.org/archive/id/draft-hendrickson-privacypass-geo-extension-00.html`.

[6] Brandon Maslen. *Introducing Microsoft Edge Secure Network*. `https://techcommunity.microsoft.com/discussions/edgeinsiderannouncements/introducing-microsoft-edge-secure-network/3367243`. Accessed: 2025-10-03. 2022.

[7] Markus Pauly. *An HTTP Geolocation Hint for IP Addresses*. Work in Progress. July 2023. URL: `https://datatracker.ietf.org/doc/html/draft-pauly-httpbis-geoip-hint`.

[8] Tommy Pauly et al. *Proxying IP in HTTP*. RFC 9484, Internet Engineering Task Force. Apr. 2023. URL: `https://www.rfc-editor.org/rfc/rfc9484`.

[9] Aviram Zilberman et al. "A Survey on Geolocation on the Internet". In: *IEEE Communications Surveys & Tutorials* (2024). DOI: `10.1109/COMST.2024.3518398`.