Does IP geolocation answer the right questions? †

Marwan Fayed
Cloudflare, Inc.

IP geolocation has its origins in the 1990s, at a time when IP addresses were long-lasting and often statically located in networks, two attributes that are no longer reliably true of IP addresses today. It also predates large-scale NAT devices, ubiquitous use of VPNs and other IP-sharing mechanisms, alongside mobile devices and their networks. Underlying geolocation mechanisms have been refined and improved, but are otherwise unchanged. In 2025 and beyond, it behooves us to ask what question IP geolocation truly answers, then identify the set of actual questions that need answers — and the extent to which they are answered using IP geolocation information.

Consider a strawperson example: One of the earliest documented use-cases of IP geolocation was to restrict the posting and sale of Nazi paraphernalia to users in France. Can we say that geolocation today would feasibly be effective for the same purpose? A positive outcome conservatively requires that (i) IP geolocation is accurate to the level of country, and (ii) users in a country are bound to IP addresses physically located in the country. Today, however, neither clause is guaranteed to be true. In addition, emitting metadata at finer granularities such as postcodes conveys confidence that creates unrealistic human expectations.

Fundamentally we use “where is a logical IP address in the physical world?” as a proxy for “where is a user located?” Furthermore, a chain of proxy questions may also exist, for example, using “where is a user located?” for “is this user permitted to access a given resource?” The second chain is subtle, but important: Presumably an expatriate of country A located in country B should not be restricted from governmental or equivalent websites intended only for residents and citizens of A. In an Internet enforced strictly by IP geolocation, the expat has few options, and none that are intuitive.

In this short paper, we contextualize IP geolocation in the Internet 25 years after its introduction, and bring attention to alternatives. We show why even a perfectly accurate IP geolocation and its metadata is, in many cases, a non-viable proxy for any question related to users — and suggest that newer application-layer capabilities and features are better alternatives.

For purposes of exposition, the examples that follow assume perfect IP geolocation precision.

IP Geolocation in 2025 and an increasingly mobile Internet

Mobile devices generate more than 50% of HTTP requests, worldwide.

1. Border-level Precision. In the context of mobile devices along national borders, it may still fail because mobile towers only know that a device is in range. For example, consider Germany and its border with nine other countries. In the EU, mobile devices are permitted to

† Submitted to the IAB Workshop on IP Address Geolocation (ip-geo)

roam freely on networks outside of a subscriber's country. As a result, German subscribers in Germany may attach to non-German cell towers, and vice versa.

As a result, a person in one country might well appear from behind an IP address geolocated to a neighbouring country. In this context the device's own knowledge is moot. For example, devices can and do reveal country-registration information to the mobile operator, but a cell tower knows only that a device is in range and not if the device is on the 'correct' side of the border. Alongside, GPS is application-layer information inaccessible from the network-layer.

One outcome is that German policies enforced using geolocation will unintentionally not apply to some Germans, and also affect non-Germans attached to towers in Germany.

In addition, non-neighbouring borders can also be problematic, for example with eSIM roaming services, which violates users' expectation and manifests as "the pizza problem."

2. The "Pizza Problem" relates to user expectations in context of dual-hop oblivious proxy services supported by the MASQUE working group (and that enable Apple's Private Relay service). Consider a user of the service who searches for "pizza places near me." In this model, the second-hop oblivious proxy has no knowledge of the client's IP address, and no geolocation by extension. Accurate IP geolocation is misleading, and the user would be shown pizza near to the second hop that, for most users, is nowhere near to the user.

The problem today is solved in two parts. First, the first hop encodes and transmits location information in the HTTP header to the second hop. (The first hop either infers the device location using IP geolocation, or knows the location because it runs a service on the device directly.) The second hop then selects and egresses from a manually geofeeded address corresponding to the location reported for the device. If the operator deploys the second hop at many locations, then either each location must maintain a unique set of IP addresses geofeeded to all all possible locations (for unicast egress); or alternatively, design, implement, deploy, and maintain mechanisms to route return traffic back to the intended server (anycast egress). Each option incurs cost and complexity that hinder deployability.

Beyond 2025 in an IPv6 Internet, or in Space!

Separate from questions about IPv6 as an answer is the tenability of IP geolocation data as the Internet evolves. The timeliness and feasibility of IP geolocation data faces challenges at IPv6 scales. In addition, the definition and value of a location in space is unclear.

3. An IPv6 Internet presents two challenges. First, every major operating system implements RFC 4941, meaning portions of *client* addresses change at most every 24 hours by default. Second, the number of IPv6 addresses may render complete geolocation infeasible in a timely fashion, because completeness likely requires scanning all addresses in some form. Incomplete geolocation may be acceptable in some applications or when answering some questions, but completeness is required for any service that relies on IP geolocation for compliance or to enforce copyright restrictions (e.g. in streaming services).

4. IP in space is the basis of the tiptop working group's charter. In this context IP-location is likely a poor descriptor when any single location is violated by celestial motion. In space, an

object's location may be less useful than its relative location to other objects. Also, any object's geolocation would change drastically while it is being geolocated. Similarly, geofeed specifications would have to change from labels consisting of string representations of known locations, to some combination of celestial location with timestamps with formulas describing motion and trajectory. Each server, for every new connection it receives, would have to compute the current location of a source IP, only for the location to be invalidated by the time the connection is established.

Rethinking the questions, and reasonable answers

We should recognize two attributes that were true when IP geolocation emerged, but are no longer true today: (1) IP geolocation was the only available mechanism to associate an IP address with a physical location; and (2) it became the de facto proxy answer to "what should this user see?" based on the understanding that users were directly attached to their IP.

The same was true of telephone numbers, although we accept that phone numbers no longer indicate location. For example, +1 country code numbers are structured with area codes ("NPA") that were originally intended to indicate sub- or metro-regions, and subsequent digits ("NXX") that point to the nearest central office or central exchange. These attributes were true and reliable in an age of landlines, and were used for call routing. The same was also true for mobile phones — but has since lost any meaning because of local number portability. The Internet is experiencing a similar transition away from number-to-location reliability, which raises questions about its meaning, and its use.

Supplemental, even alternative, sources of user-location are increasingly available at the application-layer. The HTTP API, for example, is far more accurate and in the control of the user at the location. On the server-side, HTTP's geolocations directive controls cross-origin visibility of the location. These are mechanisms that directly answer, "where is a user located?" Moreover, emerging secure transport mechanisms such as Privacy Pass present opportunities to safely transmit *unlinkable* signals to servers that directly answer, "should this user have access?".

Geolocation at layer-3 has served the Internet for more than 25 years. Now is the time to ask if it is sufficient for the years to come, and acknowledge emerging technologies that are closer to users and that have the potential to answer finer-grained questions.