Session 4
IP address geolocation
(subject)

Introduction

IP address geolocation is the process of associating a tuple (longitude; latitude) on an IP address. This information is one of the important metadata that we can retrieve from an IP address.
It is widely used in the industry with examples such as:

- for personalized content delivery based on a client localisation (think about the difference between Netflix's content in France and in Germany).
- for fraud detection , someone trying to connect to your account in a country that is not the usual one
-

But also in research:

- Path latency/performances troubleshooting
- Geopolitical study of the Internet (to find where the traffic is being routed)
- Censorship (where censorship is happening)

Of course, for most applications the geolocation of an IP address is retrieved from the user's GPS position directly. But it might not be always possible:

- the device does not provide GPS information. This is the case for router or CDN Point Of Presence (PoP, i.e. replicas of a CDN).
- Hackers might hide their geographical information

For all of these cases, and simply if you do not have access to user's data, you must find this information by yourself.

Geolocation with latency estimation

Once again here, we will rely on using different VPs to issue ping to a target.

Q1°) What kind of information do you retrieve by pinging an IP address from a single VP?

Q2°) Describe how we can extract a geographical distance from a ping between a VP and an IP address.

Q3°) You made your measurements and issued a ping toward the target from a set of VPs. Which one is giving you the most precise estimation of the target's geolocation?

Q4°) Can you think about a way to use multiple distance information to estimate the geolocation of the IP address? (Hints: think about how GPS work)

Q5°) Within the folder TP4/ of the repository, follow each step and perform geolocation on the set input target IP addresses. You will have to:

- select the VPs from RIPE Atlas to perform geolocation (hints: as all IP addresses to geolocate are located in France, we will only use VPs in France)
- ping each target from each VPs
- retrieve measurement results
- compute the distance between each VPs
- estimate the IP address geolocation using two methods:
    - shortest ping
    - constraint based geolocation (CBG)
- validate the quality of the geolocation