Geolocation system

NETMET Lab project

Practical information

- The drop-off of the project is before the last lab of the year: **01/26/2021 (26 of January, 2021) at 08:00AM**.
- You must deliver a **report** that explains your solution in detail **with your code attached**. You must also give a justification of why you chose your solution, with the advantages and drawbacks.
- Finally you must deliver a **JSON file with a map between the IP addresses and their geolocation**.
- The project will be evaluated via the relevance of the solution, its justification and the accuracy of the solution.
- You can do the project alone or in binomial (e.g. *max. 2 people*).
- This mini-project will count for **10%** of your final NETMET grade.

Prelude

Is able to geolocate IP addresses is an active field of research and is still an open question. It's also an economic challenge that a lot of private industries try to tackle.

Description

The project is guessing as correctly as possible a list of IP addresses. You can use any technique or *combination of techniques* you can think of as long as the solution is free and open sourced. **You are not allowed to use the data of a geolocation database** (MaxMind, GeoIP, …).
Of course you can use the techniques / technologies we have seen and you will see during the labs (multilateration, topology, DNS, ...) but you can also come up with new ones you imagined or found in a paper. Please cite the paper(s) that influenced you in your report.

Infrastructure

You can use the PlanetLab Europe nodes in the same way you already used in the previous labs. Due to the number of IP addresses to tackle (~100), you will have to build a system that automatically handles the task for you which means you will have to write some code. Also, you would have to demonstrate in class your solution so it has to be easy enough to use.

API

It can be tricky to evaluate if your solution goes to the good direction without having any reference to compare.
Beside some of dataset that you could find on the Internet, you can directly test your solution on the actual IP address list via this API : http://ares.planet-lab.eu:8000/ (Note that it uses the distant port 8000, the port 80 is taken for an other project).You can also check the interactive documentation here http://ares.planet-lab.eu:8000/ui/.
This API is composed of a single POST route. You provide as a body ip addresses mapped to their coordinates. As a response, the API gives you a score. The perfect score is *0.0* but I don't expect you to have it. See this score as a relative number that allows you to check if your solution is taking the right way.
You can provide a partial list of IP to the API and the score will be computed on this partial list. This way you can for instance provide only one IP address mapped with the geolocation and compare your solution on this particular case.

References

Gueye, B., Ziviani, A., Crovella, M., & Fdida, S. (2006). Constraint-based geolocation of internet hosts. *IEEE/ACM Transactions on Networking (TON)*, *14*(6), 1219-1232.

Katz-Bassett, E., John, J. P., Krishnamurthy, A., Wetherall, D., Anderson, T., & Chawathe, Y. (2006, October). Towards IP geolocation using delay and topology measurements. In *Proceedings of the 6th ACM SIGCOMM conference on Internet measurement* (pp. 71-84). ACM.

Wang, Y., Burgener, D., Flores, M., Kuzmanovic, A., & Huang, C. (2011, March). Towards Street-Level Client-Independent IP Geolocation. In *NSDI* (Vol. 11, pp. 27-27).

Hu, Z., Heidemann, J., & Pradkin, Y. (2012, November). Towards geolocation of millions of IP addresses. In *Proceedings of the 2012 Internet Measurement Conference* (pp. 123-130). ACM.

Cicalese, D., Joumblatt, D. Z., Rossi, D., Buob, M. O., Augé, J., & Friedman, T. (2016). Latency-based anycast geolocation: Algorithms, software, and data sets. *IEEE Journal on Selected Areas in Communications*, *34*(6), 1889-1903.