International Journal of Research in Engineering, Science and Management
Volume-2, Issue-3, March-2019
www.ijresm.com | ISSN (Online): 2581-5792

Tracking Geolocation using IP Address

Vikas Bhowate[1], Mansi Pakhale[2], Mansi Patil[3], Ashwini Dharmik[4], Kalyani Katekar[5]
[1]*Assistant Professor, Dept. of Information Tech., St. Vincent Pallotti College of Engg. & Tech., Nagpur, India*
[2,3,4,5]*Student, Dept. of Information Tech., St. Vincent Pallotti College of Engg. & Tech., Nagpur, India*

***Abstract*: The geographical location of Internet IP addresses is important for academic research and commercial usage as well as to track customer behavior on a particular website. Thus, both commercial and academic databases and tools are available for mapping IP addresses to geographic locations. Evaluation of the accuracy of these mapping services becomes a complex function since obtaining diverse large scale ground truth is laborious. This project is divided into two basic tasks, the former being the one that check for the IP addresses from the user and accordingly provides with his location details mentioning the latitude and longitudinal values, latter part of the project is the study of D3.js ; the types of charts and visualization that can be used to make a dashboard which can effectively display the information collected during the IP address tracking . The project uses the free predefined database provided by Geolite2, to track the IP addresses offline. This way we are able to check the customer action on a website; his current locational details and the number of attempts he makes to hit the site . The work would provide an insight into the strength and weaknesses of IP geolocation databases, and discuss their accuracy and encountered anomalies**

***Keywords*: Geographic information system, IP address, Web and internet service, Internet topology, D3.js**

1. Introduction

Lately, geolocation services have become a necessity in many fields and for many applications. The fields like banking, cybersecurity, disaster management, and many more demand for the location values of a user to uncover him from different parts of the globe. Although the end user is usually not aware of it, many websites visited everyday use geolocation information for targeted localized advertising, localized content (such as local news and weather), and compliance with local law. Despite the fact that IP addresses do not reveal the exact address of a user, it can definitely track the general area or city from where the user is accessing the data around the world in an instance. Geolocation services have access to a wide range of databases which provide them the information needed to locate someone online through their IP address. The database that is used for locating the user in this work is from the Geolite2 API specifying the locational details such as the geoname_id, country_iso_code, latitude and longitude values, accuracy_radius and postal_code along with the major city, continent_code and country_code. Since the databases considered from Geolite2 API are opensource databases, the accuracy rate of the geolocation of the user tends to be less than that of the databases used for commercial purpose at Maxmind-digital mapping company .Geolocation not only helps to locate geographical coordinates of a customer bur can also assist by improving visualization, thus simplifies largescale data analysis. Once the latitude and longitude values of a client are noted, the relevant information captured following the code can be effectively displayed by the means of visualization; heat maps and charts using the Javascript library-D3.js. D3 allows us to bind arbitrary data to a Document Object Model (DOM), and then apply data-driven transformations to the document. It can be used to generate an HTML table from an array of numbers, or, use the same data to create an interactive SVG bar chart with smooth transitions and interaction. The information gathered can then be presented on a dash board for the analysis process. To calculate the attempts made by a particular customer on a specific website along with his geolocation to analyze for which zone is the client request coming from.

2. Proposed system

The visualization of analyzed data is the project which aims to create a dashboard for the users who would analyze their visitors and other activities performed by user in a graphical format. Representation of any statistics with the help of diagrams makes it easier for the user for analysis.

The model of this project will try to demonstrate the client server architecture. The application once launched will provide options to view the actions performed weekly, monthly and years.

A. Block diagram

The block diagram for the proposed system has following main components as shown in fig 3.1 that explains block level representation of the project. It shows the interaction between the modules and the flow of the project. The system design has following main component.

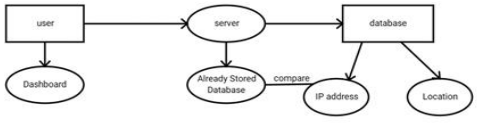

Fig. 1. Block diagram

- *Server:* a server is a computer program or a device that provides functionality for other programs or devices, called "clients".

- *User:* In this, user is a client which is a piece of computer hardware or software that accesses a service made available by a server. The server is often on another computer system, in which case the client accesses the service by way of a network.
- *Database:* A database is an organized collection of data, generally stored and accessed electronically from a computer system.
- *Dashboard:* The dashboard will be a Web analytics service that will provides statistics and basic web traffic details for marketing purposes.

B. Working of the proposed system

The following steps shows the working of proposed system:

- The proposed system will first request for visitor's IP address.
- The system will fetch the visitor's IP address by using a php code.
- A database is already created which includes the range of IP addresses and its corresponding parameters such as latitude and longitude.
- After getting the IP the system compares the obtained IP with the data from the table which is being stored in postgres database.
- If the comparison of the digits of the acquired IP and the list of IP in database matches, then it will proceed further.
- After getting the required row of IP, everything corresponding to the requested data will be displayed in the form of table.
- Since the visitor hits the page we can track the day, date and time of the hit.
- This data can be used in analyzing and getting complete statistics of web traffic on our page.

3. Technical challenges

In this section a set of specific technical challenges are presented, motivated by the complexities that occur during the IP address tracking of a user. While there are techniques that can help providers build their databases which can specify the locational values, ensuring that the resulting data is as accurate as possible is a non-trivial challenge. A number of factors contribute to making it hard to arrive at accurate results. There is the general challenge of lack of ground truth. There is unfortunately no definitive, comprehensive, publicly available, and most importantly, trustable datasets that ties IP address to physical locations. The data sets present with the providers contains mostly approximate data values and not the accurate information. Moreover, the complexity faced during the work of this project was, the dataset used comes from the open source site of Geolite2 API which has got just the approximate values of the geographical location of the user and thus the work lacks the efficiency that the commercial database of the Maxmind DB provide. Other challenges faced during the address tracking is the unvalidated registry data.Some IP geolocation solution providers take the data that they ingest from registries at face value – assuming that the entries are correct, and not validating the information contained within them. However, the information may be wrong, whether intentionally or accidentally.

4. Conclusion

This paper describes how the information of client's location is obtained with the help of the free predefined database provided by Geolite2, to track the IP addresses offline. Today many websites use this data which has been collected from the geolocation to understand their customer and better serve their clients with relevant information like localized news and promotions. The reason behind this analysis is to gather statistical information needed by the server to analyze its clients. This data may prove to be useful during some emergency such as occurrence of natural calamities, management of traffic, security, content localization and geo advertising. This data can be further used to create the analyzed data in visual context to make it easy for the user to analyze his own user statistics diagrammatically.

References

[1] MaxMind LLC, GeoIP, http://www.maxmind.com, 2010
[2] https://www.maxmind.com/en/open-source-data-and-api-for-ip-geolocation
[3] https://dyn.com/blog/finding-yourself-the-challenges-of-accurate-ip-geolocation/
[4] https://d3js.org/