RESEARCH ARTICLE OPEN ACCESS

Language Geography Based Communication Medium Using Geographic Information System

Sangavi VP*, Kanchana Devi V*, Kavya Alluru***
*(School of Computing Science and Engineering, VIT University Chennai Campus)
**(School of Computing Science and Engineering, VIT University Chennai Campus)
*** (School of Computing Science and Engineering, VIT University Chennai Campus)

----------**********----------

Abstract:

We travel across various places for personal or professional reasons and encounter many challenges. One of the major challenges would be to communicate with the local people at the destination city. The traveler may not know the cultural background of the area or the language is spoken in that area. The aim of the paper is to track the location of the user and identify the majorly spoken languages at that specific region/location and aid to communicate through translation by means of text-to-speech and back in speech-to-text in a specific language of the target area. Using Geographic Information System (GIS) services, track the location and get the majorly spoken languages. The user gets to choose one of the target languages and communicate. To communicate with the native people, the user types the text in English and it gets translated to native chosen language, the result is in the form of text/ speech. The user successfully communicates with native people in their own language without prior knowledge. This can be used by frequent travelers and during emergency situations to communicate across various countries or states.

Keywords **—Language Geography, Geography Information System (GIS), Document Object Model(DOM), Reverse Geocoding, Translation, Global Positioning System(GPS).**

----------**********----------

1. INTRODUCTION

When we travel across various places for personal or professional reasons we encounter many challenges. One of the major challenges would be to communicate with the people at that specific region. A traveler may not know the cultural background of the area or the language is spoken in that area. Hence the traveler is inarticulate. The language spoken in a region is usually well defined based on its cultural background and inhabitants. The project was aimed to develop a system wherein a person could travel without any prior knowledge of the language spoken in that region. The related concepts are described below.

I.1. LANGUAGE GEOGRAPHY

Language geography is the branch that studies the geographic distribution of language(s). Saroj Kausik *et al* in their research on Location Based Learning based prototype [1] wherein the learners use E-learning systems based on their location. The system contained location based content learning about cultural, geographical and architectural specialties of ancient places. The research is similar to the concepts of language geography used in the existing project. The two streams of language geography are as follows:

a) Geography of languages:

It is the subject concerned with the distribution of language with history and space. It also helps in analysis of the language dispersal across time and space.

b) Geolinguistics:

It is the subject concerned with various factors concerned with the distribution of the language and its dialects over geographic areas. The factors considered to analyze are cultural effect/ cultural transition, political rule over the area, economic conditions prevailing in the geographical region. It is also due to the interjection of factors leading to the language distribution.

I.2. TRACKING LOCATION

To obtain the user location and plot it in the application rapidly, both GPS and the Network Location Provider are used. The research work done by Kusum Tara *et al* [2] related to the work done in this project on Google Maps. It dealt Google Maps and Arduino based vehicle tracking with GPS and GSM. The app needs to find user's way while traveling; it tracks to get the geographical location (i.e. latitude and longitude) of the phone at regular intervals (in current system it is set to 5 minutes).

Tasmia Reza *et al* [3] were motivated to develop a system wherein they described about plotting the location on a map backdrop. This is analogous to one of the intents (as in Figure 4) of this project. This information collected while the user is travelling, is stored in the Location object that is retrieved through the fused location provider API. To receive the location updates at regular intervals the app has to be connected location services and make a location request. The location updates are received at regular intervals through the *requestLocationUpdates()* method. Once the user is ready, the *onConnected()* callback provided by Google API Client get forwards us to the location updates. To store parameters for requests to the fused location provider, a *LocationRequest* is created.

When the app wants to stop the location updates when the activity is no longer in focus, such as when the user switches to another app or to a different activity in the same app. The app doesn't collect information.

I.3. FETCHING DATA USING DOM

The Document Object Model (DOM), this structure is used to fetch the data from various web page sources such as Wikipedia. It defines the logical structure of documents and the way a document is accessed and manipulated. Resembling the work done in this project Yuancheng Li *et al* [4] in their project extract blocks of information from web pages and filtering unrelated information related to the user's interest. They have used HTML parser to construct a DOM tree.

In the current project, the DOM specification, the term "document" is used in the broad sense - increasingly, XML is being used as a way of representing many different kinds of information that may be stored in diverse systems, and much of this would traditionally be seen as data rather than as documents. Nevertheless, XML presents this data as documents, and the DOM may be used to manage this data. With the Document Object Model, programmers can create and build documents, navigate their structure, and add, modify, or delete elements and content. Anything found in an HTML or XML document can be accessed, changed, deleted, or added using the Document Object Model, with a few exceptions.

To get data from the website based on DOM structure Open source parser jsoup is a Java library for working with real-world HTML. It provides a very convenient API for extracting and manipulating data, using the best of DOM, CSS, and jquery-like methods. jsoup implements the WHATWG HTML5 specification and parses HTML to the same DOM as modern browsers do. It performs functions such as; scrape and parse HTML from a URL, file, or string find and extract data, using DOM traversal or CSS selectors manipulate the HTML elements, attributes, and text clean user-submitted content against a safe white-list, to prevent XSS attacks output tidy HTML .

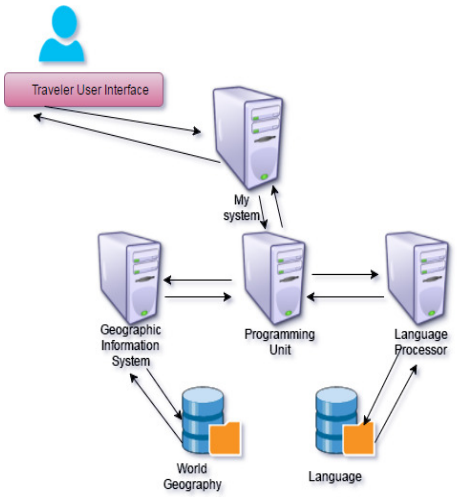
The work proposed by Seungwoo Lee *et al* [5] dealing with noise elimination based on machine learning model. Resembling to this, the language retrieval is done using parsing object. Jsoup deals with all varieties of HTML; jsoup creates a parse trees.

I.4. REVERSE GEOCODING

Reverse geocoding is the process of reverse coding of a latitude, longitude pair to a comprehensible address or place name. This helps in the identification of the neighborhoods. The research work done by Mengyu Ma *et al* [6] retrieve the location using PostGIS to determine the spatial relations between coordinates and spatial objects. They have also mentioned the use of Geohash method .Upon combining reverse geocoding with routing services and geocoding ; it becomes a vital component of mobile system. This is matching with work done by Jorg Roth [7] dealing with the approach to generate meaning text information that represents the current location.

When a GPS coordinate is entered the street address is interpolated from a range assigned to the road segment in a reference dataset that the point is nearest to. This approach to reverse geocoding does not return actual addresses, only estimates of what should be there based on the predetermined range. Alternatively, coordinates for reverse geocoding can also be selected on an interactive map, or extracted from static maps by georeferencing them in GIS with predefined spatial layers to determine the coordinates of a displayed point. Many of the same limitations of geocoding are similar with reverse geocoding.

II. METHODOLOGY



First let us begin with the flow of the project (as shown in Figure 2). It can be explained with an example. Let us assume that the user is in Vandalur at 11:10 am; first, the application would get its coordinates i.e. 12.8932° N, 80.0816° E. Next the coordinates obtained would be reverse geocoded and the region is obtained and plotted on map as Vandalur, Tamil Nadu, India with the time stamp with language as Tamil. Upon clicking the language the user gets a text box where he gets to type the message. The message is translated and given to the

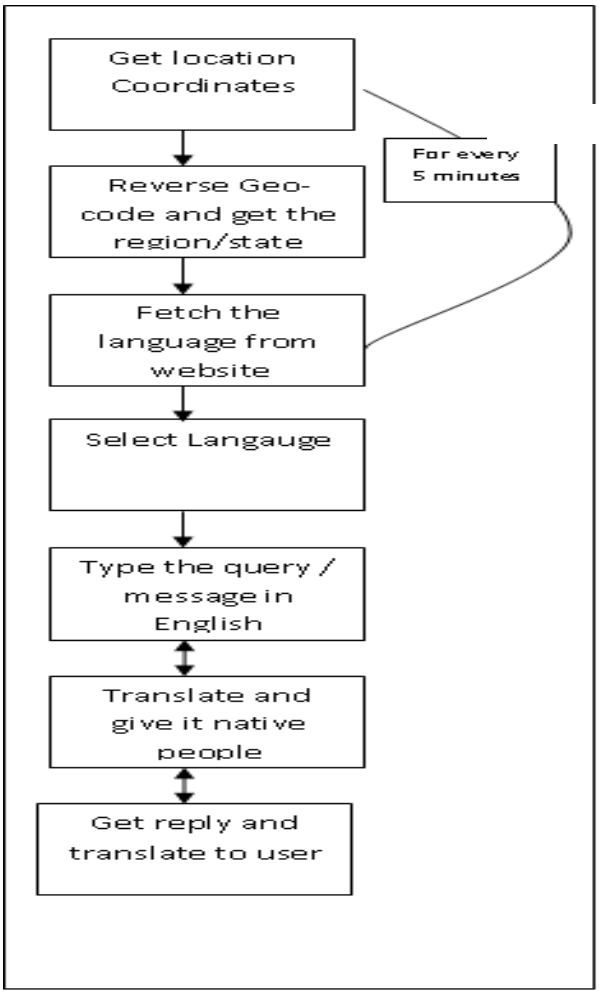


Fig 1 : Flow Chart

native people and later their reply is captured, translated back to English and displayed to the user. The user can communicate as long as he wants and the location is tracked for every five minutes and added to map.Hence the presented project aims at developing a system were the person could travel without any prior knowledge of the language spoken at that region. He need not confine the aim of exploring new places with the thought of not being able to communicate. The project also aims at being handy to solve the whole issue. The initial phase of the system begins with capturing the current coordinates of the user with the time stamp. This is done by using GPS in the phone. This fetches geographic coordinates i.e. the latitude and longitude of the location. The coordinates are reverse-geocoded to the address of the location. The address is processed and region is obtained. The region is used to get the language. The name of the region is queried on the website to obtain the

Fig 2 : Architecture Diagram

language. If language for the region is not found then the state of the location is obtained from the address and used to query the language. If more than one language is found then all the languages come as a list where the user gets to choose one of the languages to communicate.

The location with the region, the time stamp and language options are plotted on the map. The location is captured and the language is fetched for every five minutes and added on the existing map. The user gets to choose one of the languages from the list and starts communicating. User types his query or question on the message box in English. This text gets translated to the language selected by the user. The translated message is fetched back and used it with the native people to communicate. The native replies back and it is translated back to the traveler/ user in English. The whole process is performed for languages and libraries which are supported by android system.

Figure 1 represents the overall architecture of the project. The location of the traveler user obtained. The app reverse geocodes it and obtains the language geography. Later based on the language obtained language processor translates it to the native language.

II. MODULAR DESCRIPTION

Module 1: Identifying and choosing Language Geography

The location of the user is tracked with network providers and gets the coordinates i.e. latitude and longitude. The coordinates are reverse geocoded to get the locality, state and country of the location along with the time stamp. The language/s is/are fetched based on the user's location from Wikipedia using jsoup DOM parser. The user gets to choose one of the languages. If specific location's language geography is unavailable the general language geography is specified for the area and displayed for the user to choose. This process is repeated at regular intervals, with fastest interval 1 minute and interval 5 minutes.

Module 2: Communication with the natives

The chosen language is used to communicate with the native people. The language is mapped based on the information available about the location on Wikipedia. The target language is set in the Google Translate App. The user types in English and it is translated to native language through Google Translate App. Text to speech and speech to text translation is performed.

ALGORITHM

1. Through app tracking the *traveler_location* across route
   1.1 if languages for *traveler_location* is found using jsoup
      1.1.1 *language_choices* are fetched and *Choosen_language* from *language_choices* by *User*
   1.2 Otherwise *getCapitalCity* of State of the *traveler_location*
      1.2.1 Repeat 1.1.1
2. *User* types the *text_message* in English on Google Translate App
3. Google Translate App translates it to *Choosen_language* and sends it to user
4. The *user* receives it and replies in speech to the Google Translate App
5. It is translated and replied back to *user* in English *text_message*
6. Process repeated until the communication is terminated

III. RESULTS

The app was deployed on a phone with Android 5.1, 1 GB unused internal memory, 1 GB RAM with GPS on and mobile network on. The app was tested at various locations to check its tracking efficiency. The location was tracked at regular intervals and corresponding languages spoken at that region was displayed to the user for selection. The selected language was translated from English to the chosen language for communication with the natives.

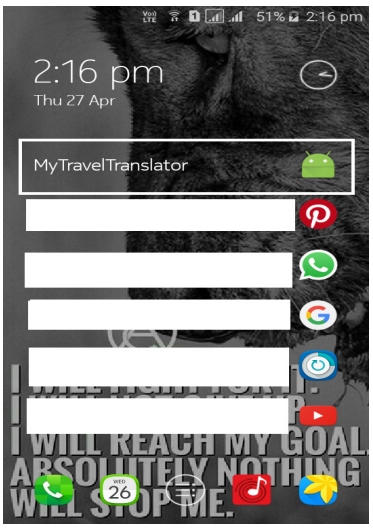


Fig 3: App on Device

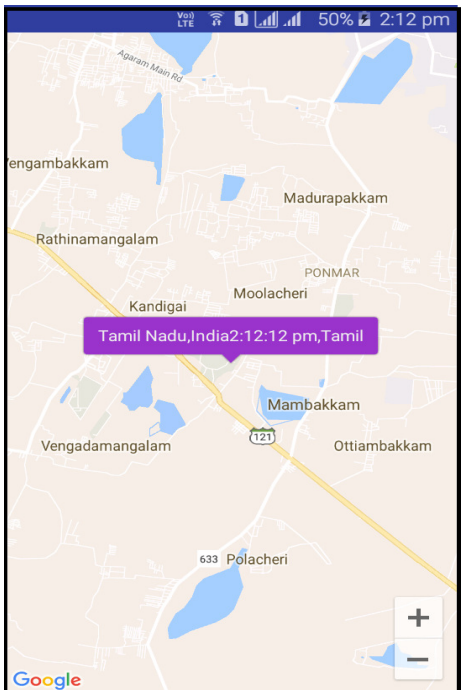


Fig 4: Tracking and Language Retrieval

The Figure 3 shows the look on the app when on deploying it on the phone. It comes in the name of *MyTravelTranslator*.When the app is clicked with GPS and Network on the App tracks the location and gets the coordinates. The coordinates are reverse geocoded to get the locality, state, country, time stamp, and the language spoken and plots it on the map. In the Figure 4 a purple color marker with state, country, time stamp and language. The locality is not displayed as the related webpage was not found to get the language.

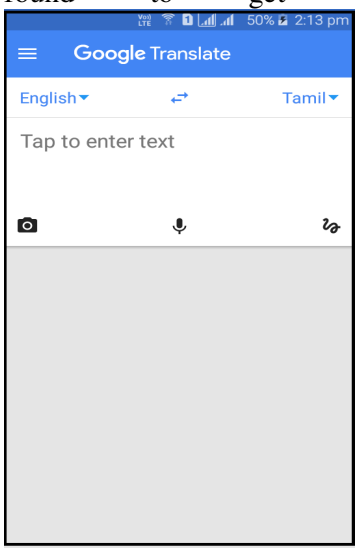


Fig 5 : Input Text

Fig 6 : Translated Output

The Figure 5 depicts the textbox that appears as a result of clicking the marker. The user types the text/query which he wants to know in native language to communicate with the people.The final intent of the app is in Figure 6, which contains the query in English and the translated text in Tamil i.e the native language of the location.The whole communication process

continues until the user gets information, thathe/she wants to know.

IV. DISCUSSION

The Figure 7 shows the Memory Usage, CPU Usage, Network Usage when the app executes on the phone across various intents. It can be observed the memory allocated is 11.73 MB upon using the app. The CPU used is less than 40% and network required is less than 256.00 kB/s.

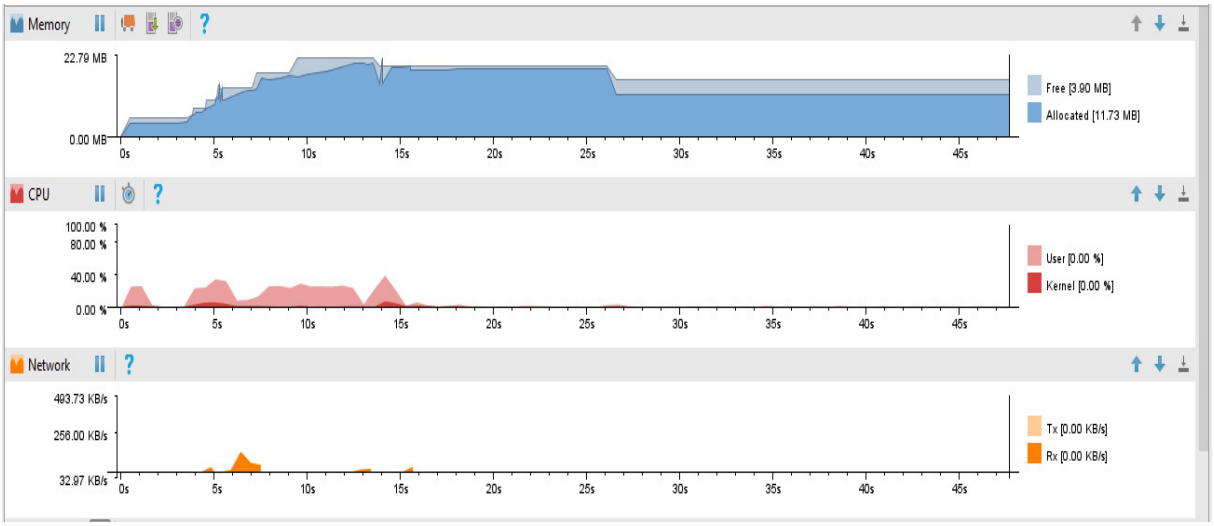


Fig 7: Performance of the App

V. CONCLUSION

The system with GPS on and a proper internet connection can be used to communicate anywhere in the world. It can be used by travelers or tourism department. The location is dynamically marked on the map with a time stamp at regular intervals with the language options. Upon clicking a specific language, the translator is initiated with language pair set as English and specific language. The user can able communicate with the people in their own native language with neither any prior knowledge of region nor about the language is spoken in that region.

Through future enhancement, the user could be giving the message in the language that the user knows and translating it to the user selected language. With improvements, it can be integrated with major phone systems, web callers, and at various linguistic gatherings. It can also be configured to capture various dialects and translate it efficiently.

REFERENCES

[1] Saroj Kaushik, Shivendra Tiwari. Fusion of navigation technology and elearning systems for on-the-spot learning. Wireless Communications and Applications (ICWCA 2012), IET, 2012.

[2] Kusum Tara, Md. Marufi Rahman, Jannatul Robaiat Mou. Real time google map and arduino based vehicle tracking system. Electrical, Computer Telecommunication Engineering (ICECTE), IEEE, 2016.

[3] Tasmia Reza,Sunjida Sultana, Mahdia Tahsin. An innovative implementation of indoor positioning system using gps. Electrical Engineering and Information Communication Technology (ICEEICT), IEEE, 2016.

[4] Yuancheng Li, Jie Yang. A novel method to extract informative blocks from web pages. Artificial Intelligence, 2009. JCAI '09, IEEE, 2009.

[5] Seungwoo Lee, Yeongsu Kim. Svm-based web content mining with leaf classification unit from dom-tree. Knowledge and Smart Technology (KST), IEEE, 2017.

[6] Mengyu Ma, Zhinong Zhong, Ning Guo. An efficient reverse geocoding method based
on global subdivision model. Geoinformatics,IEEE, 2016.

[7] Jorg Roth. Generating meaningful location descriptions. Innovations for Community Services (I4CS), IEEE, 2015.

[8] Goodchild, M.F., 2009. Geographic information system. In *Encyclopedia of Database Systems* (pp. 1231-1236). Springer US.

[9] Yang, C., Kafatos, M., Wong, D.W., Wolf, H.D. and Yang, R., Mason George Intellectual Properties Inc, 2010. *Geographic information system*. U.S. Patent 7,725,529.

[10] Jenson, S.K. and Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric engineering and remote sensing*, *54*(11), pp.1593-1600.

[11] Liu, D. and Liu, R.Y., 2002. Geographic information system.

[12] Iwamura, K., Mine, R. and Kazama, Y., Hitachi Ltd, 2013. *Geographic information system*. U.S. Patent 8,359,324.

[13] Chuvieco, E., Aguado, I., Yebra, M., Nieto, H., Salas, J., Martín, M.P., Vilar, L., Martínez, J., Martín, S., Ibarra, P. and De La Riva, J., 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. *Ecological Modelling*, *221*(1), pp.46-58.

[14] Gabrysch, S., Cousens, S., Cox, J. and Campbell, O.M., 2011. The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. *PLoS medicine*, *8*(1), p.e1000394.

[15] Graham, R.L., English, B.C. and Noon, C.E., 2000. A geographic information system-based modeling system for evaluating the cost of delivered energy crop feedstock. *Biomass and bioenergy*, *18*(4), pp.309-329.