Introduction

Human overpopulation is among the most pressing environmental issues. Due to continuous exponential growth in population, the problems related to the future are significantly rising. One of the problem is to maintain the connectivity within the peoples. The connectivity here is could be any mode of communication or mode of transportation. Now, the world is in IT Age thus communication is no problem but transportation-related problems still persists. The need of transportation is not only limited to the availability of modes of transportation but it is also dependant on medium of transportation. To satisfy this need humans invented subways and flyovers. But the ever going rise in demand of transportation is still increasing. So we have to develop more new ways to tackle problems related to transportation. Due to significant development in the field of GNSS the location-based researches are improved. Now, most of the smartphones have one or more GNSS embedded in them. These system allows to track their daily routine, provide location-based services, location-based recommendation, etc. So we can use these system to identify the need of an individual or general public with the help of analysis of their data and these technique is classified as Activity Recognition. Many of the users uses activity recognition in daily basis to track their fitness routine, to enjoy automated assistance in their work, etc. Thus we are going to use the same technology to determine the mode of transportation of an individual in transit. This process can become the preliminary for the other processes like transit time calculation, optimal route suggestion, and it also helps researchers in pre-processing the trajectorial data because different mode of transportation affects the nature of data and hence may requires a different approach to process.

The project starts from manually retrieving the GPS data through smartphones. Then the data is fed in the pre-processing stage. This stage resolves the problems regarding the raw data. Like handling semantic errors, removing outliers etc. Then the extracted data is processed through the segmenting stage to decompose the whole data into small homogenous datasets. This segmentation would be motivated from Time Series segmentation. Then the data enters the final stage, which is to identify the mode of transportation in each segment through analysis. Our Aim is to automate the process as much possible.

The challenge is to identify the mode of transportation because the modes of transportation don’t follow the same characteristics everywhere, every time. Example the speed of the vehicle is much more in the general time as compared to prime time. Here prime time is the time when traffic is most like morning time 8:00 to 10:00 and evening time 5:00 to 7:00 when the general employee’s office time starts or ends. And this time varies for different places. Also, the average speed in India is less than the western countries.

Technology Used

For the data extraction, we are using the GPS embedded in our smartphones to log the position. Then to visualise, we are using the QGIS and Google Maps to get more insights into the extracted data. We are going to use the Python programming language to develop the application for our project. This application will use many of the Python’s modules among which most basic are Numpy, Pandas, Matplotlib etc.

Motivation

The implementation is motivated by the concept of Time Series segmentation. To implement this project we are using multiple implementations which follows different approaches to meet the same goals. Thus this project also provides a comparison between different approaches with the help of Precision, Recall and F1 score. This comparison will help us to choose the optimal approach to fulfil our requirement.

Field of Application

This project comes under the field of Spatio-Temporal Trajectorial data analysis. It could be used by the researchers in the further analysis of Trajectorial data. This can also be used by Location-based services to refine their output.

Technical terms

Trajectory

The trajectory is the path followed by an object (here, a person) in motion. Here, a trajectory is the set of points that refer to the position of the person at any instant of time.

Trajectory τ is the set of sequence of points P i , such that

τ = [ P i : P i = { x i , y i , <z i >, Ti} ]

where T i < T i+1

x i and y i are the coordinates of the object in Geographics coordinate system at any instant T i

Segmentation

The process of decomposing a large dataset into relatively smaller homogenous chunks. Segmentation is used in Businesses as Market segmentation or Customer segmentation, Image Segmentation etc. Segmentation can be either done on the basis of similarity of one or more features or the similarity in the geometry of the data. The former one is known as Feature-based segmentation and the latter one is Geometric segmentation.

We are using the Feature-based segmentation.

GNSS

GNSS stands for Global navigation satellite system. It is the generic term for satellite navigation systems that provide autonomous geo-spatial positioning with global coverage. For Example, GPS( Global Positioning Service ), GLONASS( Globanaya Navigazionnaya Sputnikovaya Sistema ) ( Russian version of the American GPS ), and China’s BEIDOU. India is also developing the IRNSS( Indian Regional Navigation Satellite System ) with an operational name of NAVIC( Navigation with Indian Constellation ).

GIS

GIS stands for Geographic Information System. It is a system that is designed to capture, store, manipulate, analyse, manage and visualize the spatial or geographic data. For Example ArcGIS, QGIS, GRASS GIS are most popular.