

Analysing mobility patterns of people to determine the best transportation method

J. M. D. Senanayake* and W. M. J. I. Wijayanayake

*Department of Industrial Management, Faculty of Science,
University of Kelaniya, Sri Lanka
janaka.senanayake@gmail.com*

With the technological enhancements related to Internet, Wireless Communication, Big Data Analytics, Sensor-based Data, and Machine Learning; new paradigms are enabled for processing large amount of data which are collected from various sources. In the past decades, both coarse and fine-grained sensor data had been used to perform location-driven activity inference. In recent years, GPS phone and GPS enabled PDA become prevalent in people's daily lives. With such devices people become more capable than ever of tracing their outdoor mobility and using location-based applications. Based on the collected data from these GPS enabled devices with the help of IoT related to user mobility lots of research areas are opened. In this research the data related to user locations when users do any outdoor movements is collected using the mobile devices that are connected to the Internet and is mined using data mining techniques and come up with an algorithm to model & analyse those big data to identify mobility pattern, traffic prediction, transportation method satisfaction etc.

The data for this research will be collected using a mobile application which has to be installed in smart devices like smart phones, tablet PCs etc. In this application the user has to enter the activity that he or she currently doing and the method of transportation & the users' opinion on the transportation method if he is doing some sort of travelling. The GPS coordinates (longitude & latitude) as GPS trajectories along with the time stamp and the date will be automatically acquired from the users' IoT device. A cloud based storage will be used to store collected data.

Since the dataset is going to be a huge one, there can be data which contains outlier values due to the uncertainty of the mobile network coverage and the GPS coverage of the devices. Therefore, these data should be properly cleaned when doing data mining activities otherwise these data will lead to incorrect results such as wrong traffic prediction in certain places if several users are stuck in the same GPS coordinates for a while. Not only that but also when it comes to the user satisfaction, it might lead to generate incorrect outcome if the users in the sample will not enter their satisfaction accurately. This can be avoided by comparing cluster wise users with the consideration of the location and the transportation method. We can get the average opinion of the users and take it as the satisfaction of the transportation method in that cluster.

Using the final results of this research the government can also be benefited if we selected the sample users well with mixing all the types of people and by providing necessary information for planning smart cities.

Keywords: Big Data, IoT, Data mining, GPS