International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 02 | Feb 2020 www.irjet.net p-ISSN: 2395-0072

Next Location Prediction

Ashvini M. Gupta[1], Pratiksha M. Khedkar[2], Nikita R. Choudhari[3], Amar G. Waghade[4]

[1,2,3]Student, Dept. of Computer Science & Engineering, DES'S COET, Dhamangaon Rly
[4]Professor, Dept. of Computer Science & Engineering, DES'S COET, Dhamangaon Rly

---***---

Abstract - *Location information of Web pages plays an important role in location-sensitive tasks such as Web search ranking for location-sensitive queries. However, such information is usually ambiguous, incomplete or even missing, which raises the problem of location prediction for Web pages. Meanwhile, Web pages are massive and often noisy, which pose challenges to the majority of existing algorithms for location prediction. While this problem has been explored earlier in the batch data set-up, we propose in this paper new solutions in the streaming data set-up. We examine four incremental learning methods using a Damped window model namely, Multivariate multiple regression, Spherical-spherical regression, Randomized spherical K-NN regression and an Ensemble of these methods for their effectiveness in solving the destination prediction problem. This application allows the user to plan their travel ahead by providing a route for the destination and facilities near the location and also routes to these locations with the help of Google Maps.*

***Key Words*: Location Prediction, Data Analysis, k-means, K-NN.**

1. INTRODUCTION

Location-based Service (LBS) is a kind of information service that provides users geographical positions located by mobile devices and wireless network. There exists wealth information in the location data, such as user's interests, user's hobbies and user's behavior pattern. LBS may be employed in a number of applications, including: location-based advertising, personalized weather services, entertainment, personal life and so on. An effective location prediction or recommendation can make the users have good experience. Location-based social network services (LBSNs) have been booming during the past five years. Nowadays, it is common for a user to attach his location when he publishes a photo or a status using his online social network account. Moreover, users may just share their locations, called check in, to tell their friends where they are or to engage in social games as in Foursquare.

Since these large amount of location and social relation data become available, studying human mobility and its connection with social relationships becomes quantitatively achievable. Understanding human mobility can lead to compelling applications including location recommendation urban planning, immigration patterns, etc. Previous works, including, show that human mobility is influenced by social factors. However, there is one common shortcoming: they all treat friends of users equally. Similar to other social behaviors, in most cases mobility is influenced by specific communities but not all friends. For example, the aforementioned colleagues can influence the place a user goes for lunch but probably have nothing to do with his weekend plans. Meanwhile, where a user visits on weekends largely depends on his friends or family, but not his colleagues.

Therefore, the impact on a user's mobility should be considered from the perspectives of communities instead of all friends. In a broader view, community is arguably the most useful resolution to study social networks. Real-time location prediction can have other interesting applications for the transport dispatch system such as in vehicle allocation for the future rides (Powell, 1986, 1987), diversion in real time (Ichoua et al., 2000), ride-sharing (Tran, 2015) and reduction of the total idle time (Miao et al., 2015). Powell (1986) is a key reference on the dynamic vehicle allocation problem which has seen many extensions and applications over the years. In recent years, we have seen that several on-demand transportation services such as Uber (i.e. UberPool), Ola (i.e. OLA Share), etc. provide ride-sharing as well. Tran (2015) worked on the real-time ride-sharing schedule and dispatch problem by using the location information of the passengers who requested a pickup and the location of drivers in the near-by region. Ichoua et al. (2000) worked on diversion issues in real-time vehicle dispatching system. Miao et al. (2015) worked on the taxi dispatch problem using real-time sensor information to reduce the taxi idle driving time.

Overall, real-time location prediction helps the transport dispatch system in increasing the reliability and efficiency of their services which finally leads to an increase in profits. Data generated by GPS can be put to use for tasks such as prediction of human movement which can lead to support for more advanced applications like adaptive mobile services. Existing location based applications like Google Maps provide various services like satellite images, real – time traffic data, street view maps, satellite imagery and planning of routes for travelling by car, two-wheelers or public transport or by walking. Most of these location based systems store the user's location history and their tracking data. Some of these applications provides other useful information such as turn by turn data for reaching a destination, travel times and route information that is provided by users. Location dependent information is downloaded over a mobile network.

2. SOFTWARE TOOLS SPECIFICATION

At the same time, mobile user location prediction is also a very difficult and challenging task. First of all, due to the occlusion of high buildings in the city and the user's random switching device positioning function, it is easy to cause the loss of some important data and there is a lack of training data. Second, the activities of mobile users are uncertain, and the mobile modes exhibited by the same user may be quite different in different time periods. Furthermore, the user's trajectory data is both discrete and massive, and researchers cannot directly use the raw data for location prediction. The existing research methods usually divide location prediction into three steps:

(1) Preprocessing of the original trajectory data and screening out of meaningless and abnormal location points;

(2) In the remaining location points, important locations with special significance to users are constructed according to point density, visit frequency and other information, such as the user's home address, work unit and other areas inseparable from their daily life;

(3) According to the training model of user's historical access information to location points, the next location of users can be predicted given the current and historical location points of users. In terms of trajectory preprocessing and important locations construction, only articles using GPS data and check-in data as experimental data contains these two steps, articles using Wi-Fi location data directly take each access point as an important location. Similarly, articles using data recorded by base stations also take each base station as an important location, while articles using data recorded by traffic bayonets take each traffic bayonet as an important location.

The main idea of Infomap can be summarized as follows: information flow in a network can characterize the behavior of the whole network, which consequently reflects the structure of the network. A group of nodes among which information flows relatively fast can be considered as one community. Therefore, Infomap intends to use information flow to detect communities in a network. In the beginning, Infomap simulates information flow in a network with random walks. Then the algorithm partitions the network into communities and exploits Huffman coding to encode the network at two levels. At the community level, the algorithm assigns a unique code for each community based on the information flow among different communities; at the node level, the algorithm assigns a code for each node based on the information flow within the community.

Although the potential location information in non location terms are ignored. Here this paper considers both location terms and non location terms and search queries. We will find how many times that the location is displayed in that particular meta data content. In twitter data we have location-tagged tweets but there will be no information about location of web pages. The stopwords are removed by using the steps stemming rooting etc.

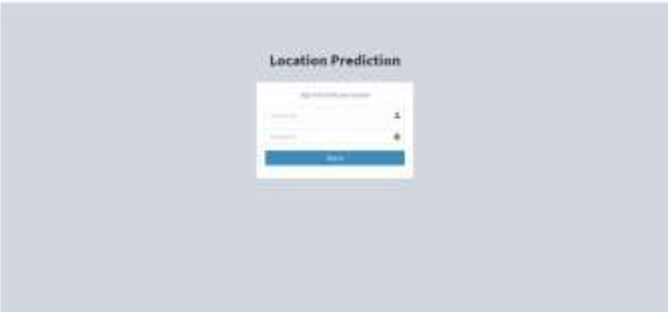
Researchers often cluster trajectory sampling points based on their density to extract important locations. However, commonly used density clustering algorithms usually have the shortcomings that the clustering result is too sensitive to the input parameters. How to determine the appropriate input parameters has become an urgent problem to be solved. In this paper, we propose an improved density peak clustering algorithm called ADPC to cluster the feature points and extract the important locations. The density peak clustering algorithm was a new density clustering algorithm proposed by Rodriguez et al..

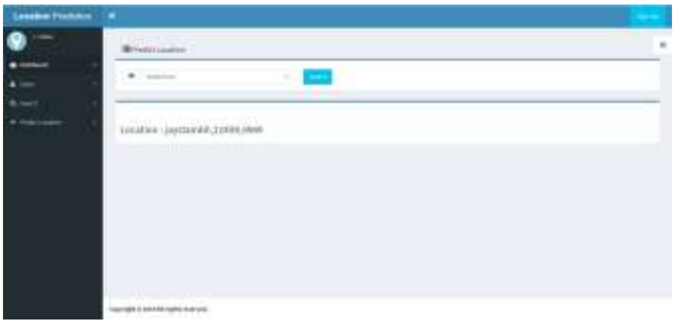
Compared with the classical density clustering algorithms, the algorithm is simple in principle, requires only one input parameter, and does not require iteration, and it can cluster data of various shapes. However, it is necessary to select the clustering center manually through the decision graph, which not only increases the redundancy of the algorithm, but also causes subjective troubles. The improved density peak clustering algorithm ADPC incorporates the self-adaptive improvement of the cut-off distance and the clustering center, so that it can automatically determine the best input parameters and clustering center according to the distribution of sampling points.

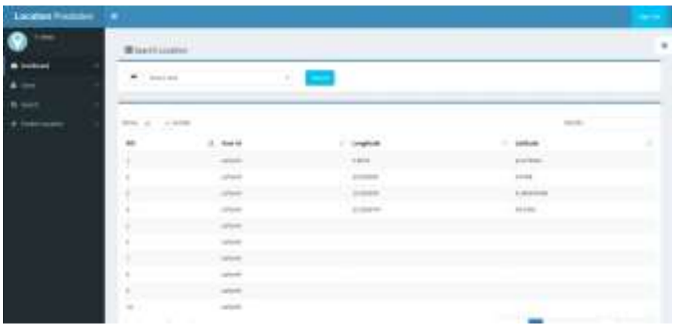
3. IMPLEMENTATION

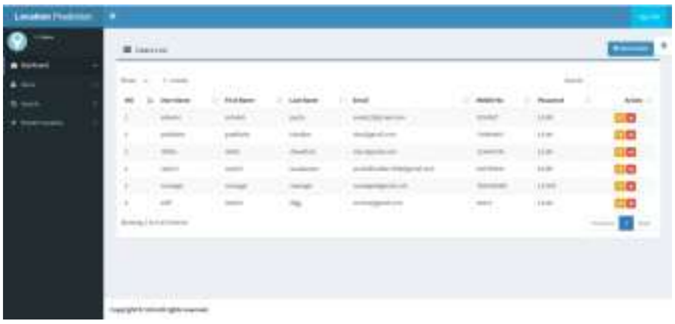
The Proposed system is a connected app to Moves that predicts the destination of the user with the location history available from Moves and provides facilities like stores, restaurants, local government offices etc. close to the destination (in the case that a correct prediction has been made) or the current location (in the case that either the prediction is incorrect or cannot be made) within a given radius. The authorization code obtained from the callback url is used to get an access token through an authenticated server-side call. Access token is required to make Moves API calls. When a user authorizes.

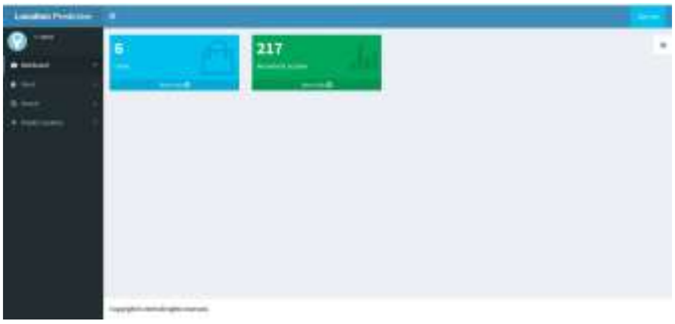
To access their data, an access token is generated. However an access token has an expiry date. For this reason a refresh token is sent with the access token in the json response. The refresh token is used to refresh the access token when it's validity expires. Using Moves API daily data is obtained and saved to the database and with each row of location data, the refresh token of the user's Moves account is also saved. In view of the above problems, the first-order Markov model only utilizes the current interest region of the user, and the utilization rate of the historical trajectory data is not high; while the high- order Markov model has problems of difficulty in matching and high sparse rate, the historical trajectory data cannot be fully utilized as the same. Therefore, combining low-order models with high-order models is a feasible method. Then, reasonably determining the influence coefficient of each model with different order is an important prerequisite for ensuring good prediction performance. From a macro perspective, high-order Markov model usually has higher prediction accuracy, and should have greater "discourse power" in the fusion model. In order to determine the specific influence coefficient of each model with different order, a mobile user location prediction method based on Adaboost-Markov model is proposed in this paper. As a representative lifting method, Adaboost algorithm can combine a plurality of weak classifiers by changing the probability distribution of training data and the weight coefficient of weak classifier to generate a strong classifier [44]. In this paper, the fusion model order k is adaptively determined, and the first-order to k th-order Markov model are used as k weak predictors. The probability distribution of user trajectory data and the weight coefficients of each Markov model are changed by Adaboost algorithm, and finally a multi-order fusion Markov is generated, which is used to predict user importan











4. CONCLUSION

The application of location based service becomes more and more popular and it is important to provide superior service for the user. The most common method for location prediction is Markov Model. But it only considers the user's location movement sequence and it is impossible to give a prediction related to the time information. So to overcome this problem our web-based app 'Next location prediction' predicts the intended destination of user using location history and machine learning algorithms and providing route for the same. The main goal of the web application is to allow the user to plan their travel even before they start. This application is very user friendly that helps to solve user travelling problem.

REFERENCES

[1] Li Y, Guo A, Liu S, Gao Y, Zheng YT. A location based reminder system for advertisement. International Conference on Multimedea. 2010;1501–1502.

[2] Guo B, Fujimura R, Zhang DQ, Imai M. Design-in-play: improving the variability of indoor pervasive games. Multimedia Tools & Applications, 2012; 59(1):259–277.

[3] Zheng Y, Xie X, Ma WY. GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. Bulletin of the Technical Committee on Data Engineering. 2011; 33(2):32–39.

[4] Zheng Y. Trajectory Data Mining: An Overview. ACM Transaction on Intellignet Systems and Technology. 2015;6(3):41.

[5] Zheng K, Zheng Y, Xie X, Zhou XF. Reducing Uncertainty of Low-Sampling-Rate Trajectories.

IEEE, International Conference on Data Engineering. 2012;1144–1155.

[6] Xue AY, Zhang R, Zheng Y, Xie X, Huang J, Xu ZH. Destination prediction by sub-trajectory synthesis and privacy protection against such prediction. IEEE, International Conference on Data Engineering. 2013;254–265

[7] M.I. L�ez, J.M Luna, C. Romero and S. Ventura, "Classification via clustering for predicting final marks based on student participation in forums".

[8] M. Young, The Technical Writer's Handbook. Mill Valley, CA:University Science, 1989

[9] Economou, G., Pothos, V., and Ifantis, A. (2004). Geodesic distance and mst based image segmentation. In In XII European Signal Processing Conference (EUSIPCO 2004), pages 941–944. Ellis, B. (2014). Real - Time Analytics: Techniques to analyze and visualize streaming data. Wiley Publishing, 1st edition. isbn: 1118837916, 9781118837917.

[10] Fisher, N. I., Lewis, T., and Embleton, B. J. J. (1993). Statistical Analysis of Spherical Data. Cambridge university press.