Next location prediction

Introduction

As a result of using smartphones GPS sensors, the location based services has raised (google now, foursquare), these services support the end-user with better recommendations depending on the context (location alarm, recommend nearby places, weather updates related to the location), with the use of new features like predicting the next location of the user, which he/she expected to visit, or next place recommendations, a new level of applications can be offered in advertising new places or even taking a decision based on the expected activity the user is going to take (traffic jams avoidance, offering new restaurants and make hotel reservations).

Datasets: For our research two main types of datasets are available online: first type, is the raw GPS datasets (i.e. longitude, latitude, timestamp) with a small sampling rates of 5 to 10 seconds which can be considered as continuous data, second type, is chick-ins data, which are the places a user my checked, with difference hours between two check-ins, it is sparse data.

Applications: Location data can be used in a large segment of applications like context aware applications (google now) which can recommend services based on the user location (location alarms, recommend places, location temperature), traffic recognition applications that allow users avoid congestions and recommend better routes, also recommender systems for advertising (ex. A restaurant recommender can promote a list of restaurants based on spatial and temporal features of the user, it can then make a reservation while a restaurant still has spaces, also location advertising application (ex. Yellow Pages) which can promote local businesses based on user current and next location.

Main research points: location related applications using raw dataset (Geolife) are mainly depend on: 1- Identifying the points of interest, 2- Identifying place and route popularity (popular places, travel sequence), 3- Identifying home and work location, 4-Identifying semantic places, 5-identifying user demographics (gender, age, education) (Zhong 2015), 6- Predicting next location, 7- Recommender system to propose new locations, 8- Decision taking applications based location.

Papers

1. Huang et al. (2015) [1], using Geolife dataset they solved the problem of identifying stay points using DBSCAN with distance threshold of 200 meters, they considered the time threshold of 20 minutes, overcoming the problem of overlapping path by exclude the points within the same spatial region if they exceed the temporal threshold.

They enhanced the human movement prediction accuracy by solving the activity change problem, which is the problem of changing home, work or a routine activity, they solved this problem by calculating the similarity between the changed activities and if a similarity measure passes a threshold the activities considered as one.

They used Markov chain to predict the user’s next location, they solved the problem of changing activity to improve the prediction coefficient from 0.295 to 0.762 for high changing activity users, and from 0.965 to 0.971 for low changing activity users.

1. Yang et al. (2015) [2], using Geolife they solved the problem of identifying stay points problem by using a DBSCAN variant algorithm which they proposed using the distance and speed as score for relation between points, if this score exceeded a threshold it would means the points are related to the same cluster.

They showed that the use of their clustering technique is more powerful than fixed time and space algorithms and extract more stay points, also the use of Variable Order Markov Model enhanced the prediction precision and solved two issues: zero frequency problem and high space complexity.

References

1. Wei Huang, Songnian Li, Xintao Liu & Yifang Ban “Predicting human mobility with activity changes”, International Journal of Geographical Information Science, 29:9, 1569-1587, 2015.
2. Yang, Jie, et al. “Predicting next location using a variable order Markov model.” Proceedings of the 5th ACM SIGSPATIAL International Workshop on GeoStreaming. ACM, 2014.