International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 25 Issue 4 – APRIL 2018.

LOCATION FORECASTING FROM CHECK-IN DATA OF LOCATION BASED SOCIAL NETWORKS

R. Karthiga[#1] and R. Angelin Preethi[*2]

[#]*M.Sc (Computer Science), Department of Computer Science, Kamban College of arts and science for women, India.*

[*] *Assistant Professor, Department of Computer Science, Kamban College of arts and science for women, India.*

Abstract- With the ubiquity of GPS-enabled devices and location-based social network services, research on human mobility becomes quantitatively achievable. Understanding it could lead to appealing applications such as city planning and epidemiology. In this paper, we focus on predicting whether two individuals are friends based on their mobility information. Intuitively, friends tend to visit similar places, thus the number of their co-occurrences should be a strong indicator of their friendship. Besides, the visiting time interval between two users also has an effect on friendship prediction. By exploiting machine learning techniques, we construct two friendship prediction models based on mobility information. The first model focuses on predicting friendship of two individuals with only one of their co-occurred places' information. The second model proposes a solution for predicting friendship of two individuals based on all their co-occurred places. Experimental results show that both of our models outperform the state-of-the-art solutions

Keywords: GPS, Social networks, Friendship, Machine Learning systems and Prediction models.

I. INTRODUCTION

Mobility is one of the most common human behaviors, understanding it can result in many appealing applications, such as urban planning, public transportation system design, epidemiology, etc. It is evident that social relationships can affect human mobility, for example, friends tend to visit similar places or one visits some places recommended by his friends. On the other hand, human mobility also has influence on social connections, e.g., two people are more likely to become friends if their mobility profile is similar. In the past, obtaining people's mobility information is considered as an obstacle for related study. Researchers have recruited a group of people to monitor their GPS-enabled devices [1] or conducted questionnaires [2]. These methods always end up with a biased dataset because of the limited number of people or an imprecise dataset considering people's memory pattern. With the development of GPS enabled devices, such as smart phones and tablets, people begin to share more of their mobility information on their social networks. Moreover, a new type of social network services has emerged, namely Location-based social networks (LBSNs). In LBSNs, a user can share his location information (called check-in) to get some reductions and engage in social games. Popular LBSNs include Yelp, Instagram and Foursquare.

Intuitively, friends tend to visit same places due to similar interests. This is known as social homophily [2]. Friends may visit same places together or separately. The former can refer to friends hanging out together while the latter may be an evidence of place recommendation. If two people visit many same places, it may indicate that they are probably friends. Similarly, if the number of visits for two people together to places is large, it is also a good indication that they are friends. On the other hand, the visiting time interval of two people can also have influence on their relationship. If two check-ins happen at roughly the same time, the corresponding users probably visit the place together with intention. If the check-in time interval is about a short time period (e.g., one or two months), these two visits can be considered to be linked because of place recommendations between friends. Based on these intuitions, we develop two models for friendship prediction.

Due to the continuity of space and time, trajectory is not suitable to be directly imported to a prediction model. Before using prediction models, each of the points in a trajectory is first preprocessed in order to convert the real continuous values associated to the geospatial coordinates of latitude and longitude, into discrete codes associated to specific regions. Traditional prediction methods usually start with clustering trajectories into frequent regions or stay points, or simply partition trajectories into cells. Trajectories are transformed into clusters or grids with discrete codes, then pattern mining or model building techniques are utilized to find frequent patterns along the clusters [3]. For example, the historical trajectories of a person show that he always go to the restaurant after the gym. If the person is now in the gym, it is a distinct possibility that the next place he will visit is the restaurant.

The rest of the paper is organized as follows: Section II presents the related work; Section III presents the proposed work; Section IV presents the experimental analysis and finally concludes in Section V.

II. RELATED WORK

This section presents the prior works of the location based social networks. Many works aiming at predicting friendship from spatial-temporal information have been published during the last several years. Li et al. [4] extracted users' visiting trajectories and stay points from location information and represented the set of stay points as a hierarchical graph where each layer clusters the stay points into several spatial clusters divisively. The pair of users who share similar spatial clusters on a lower layer has stronger similarity, and similarity is used to indicate friendships between them. Along this direction, Chen et al. have used trajectory pattern to represent user mobility profiles and proposed several metrics to measure the similarity among user mobility profiles with a tool support [5].

In [6], conducted a study to observe 94 students and faculty on their mobile phones for nine months. Through the analysis on the dataset, they found out that two people visiting the same place at roughly the same time is a strong indicator that they are friends. Particularly, the indicator becomes even stronger when the visits happen at non-working time and locations. The higher the probability that they are friends. In addition, they proposed a probabilistic model to predict friendship. However, the model does not fit the real life scenario since they made the assumption that each user only has one friend.

Then, the study is formalized the problem into a binary classification and extracted a large number of features including the spatial and temporal range of the set of co-locations, location diversity and specificity, and structural properties to train the friendship predictor. In addition, they propose a notion namely location entropy to characterize a location's popularity which we will use in our work. They also utilized machine learning classifier for friendship prediction. In their problem set, they only have one common location's information that two users have been to. In their solution, they only considered very simple features. We tackle the same problem in our first model. By considering more meaningful features, our model outperforms theirs significantly. The authors proposed an entropy-based model (EBM) to estimate social strength which also leads to friendship prediction. They extracted two factors for each pair of users to train their model. The result in [7] shows that EBM outperforms all the above mentioned models. We tackle the same problem in our second model. By considering time in a more general way, we are able to achieve better result than EBM. Some recent works on friendship prediction based on location information include.

Movement pattern mining techniques find the regularity of movements of objects and combine current movements with historical data for prediction. By transforming trajectories into cells, Jiang [8] studied trajectories of taxis and found they move in flight behaviors. The author in [9] utilized an improved a priori algorithm to find association rules with support and confidence. These frequent patterns reveal the co-occurrences of locations. The author in [10] developed a modified PrefixSpan algorithm to discover both the relevance of locations and the order of location sequences. Moreover, sequential pattern methods can be improved by adding temporal information.

III. PROPOSED WORK

This section presents the proposed work of issues on location based social networks. The main objectives of the study are:

- To propose a new feature fusion approach, to cope with the variety problem in location prediction.
- To improve the applicability of location prediction approach,
- To utilize several kinds of features and discuss their different characteristics in the variety of check-in scenarios.
- To introduce intuitive ways to model these check-in features and then formalize a combination framework to deliver the predicted target places to end users.

The proposed location based social networks composes of three phases, namely,

A) *Context Feature*

It refers to the spatial dimension. Users' check-in activities are distributed in a spatial scope. Nearby places can contribute to the representation of users' check-in records, especially when the users visit a focused set of places. Density closeness of users' check-in logs from the spatial perspective has received a lot of attention. It has shown advantage over traditional spatial modeling, both in flexibility and accuracy. It also follows a similar motivation and provides our own design to extract users' preference from the spatial aspects.

B) *Collaboration Feature*

The commonly used collaborative feature extraction method usually delivers a reasonable performance in many scenarios, it does not perform well in the check-in usage, which is revealed in our initial study.

The causes for the unsatisfying performance are two-folds. First, the check-in matrix *R* is very sparse, *i.e.*, it has many zero items. Zero check-in record means that the user never visits the corresponding place. This is probably because he/she is not a fan of that place or his/her location scope is rather focused. Second, for some rated places, the frequency information is not enough. In the counterpart representation of movie or product recommendation, users not only reveal their favorite items with high ratings, but also their least favorite items with low rating. However, this situation does not hold in the case of location prediction, frequency of check-ins merely imply the confident level of users' preference for the corresponding location, which makes them barely serve as explicit ratings for locations given from users.

C) Content Feature

It refers to the place dimension. Places are not merely visited by users. These places have inherent attributes, *i.e.*, categories, text descriptions and other kinds of annotations. We discuss how the transition between places reveals users' interest/preference over time. In further, the transition patterns benefit the closeness extraction of places for better prediction. In this paper, we use the categories of the location as its attribute description. We discuss a general way to obtain the content description features. POI, *i.e.*, places in check-in records are usually annotated with categories or attributes. Users' transition pattern from one place to another one shows the interest flow between these places. The extracted closeness from these transition patterns can be used to predict the potential locations that user will take a visit next.

The following are the advantages of the proposed study are:

- The predicted candidates based on the combination of local district, local city and state scales. The weights of each scale are learned from training data.
- An extensive study over several real datasets reveals the improvement and advantage of our approach.

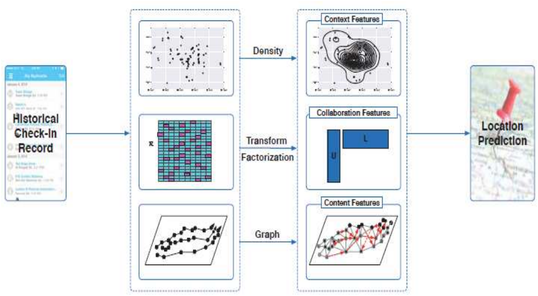

Fig.3.1 System Architecture

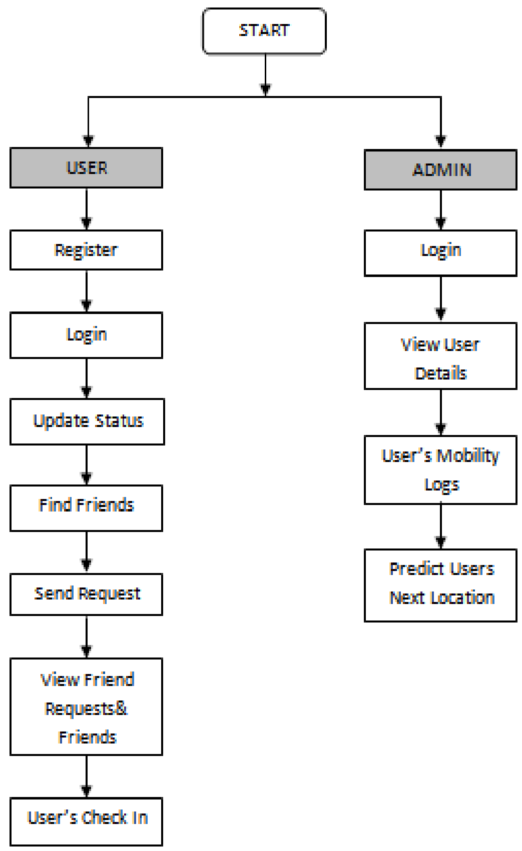

Fig.3.2 Proposed workflow

IV. EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the experimental analysis of the proposed study.

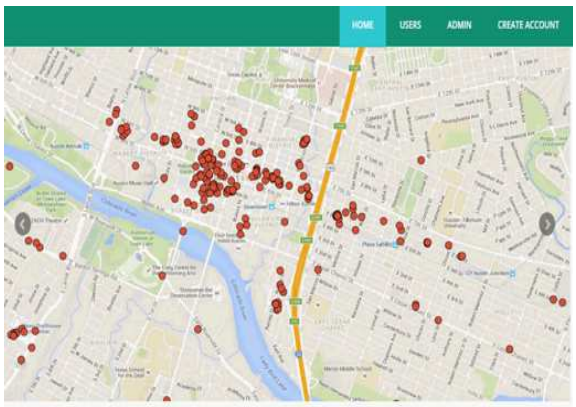

Fig.4.1 Start page of the systems

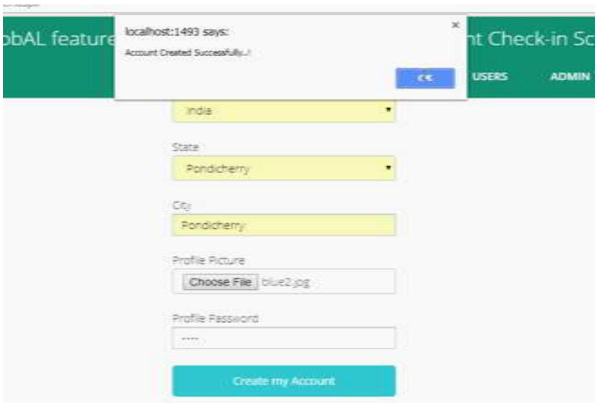

Fig.4.2 Creating the account

Fig.4.3 Author's login

Fig.4.4 Status updating process

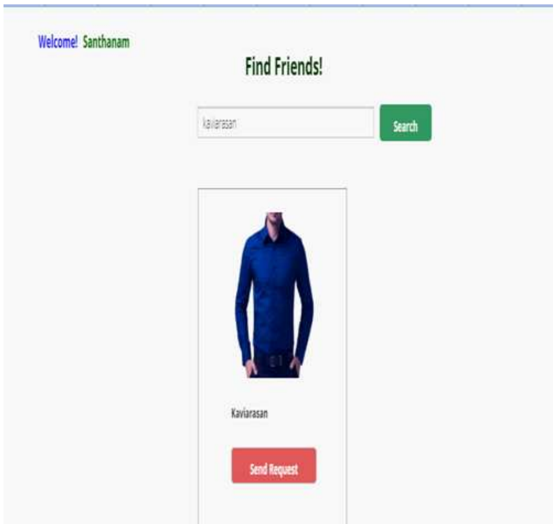

Fig.4.5 Finding the friends

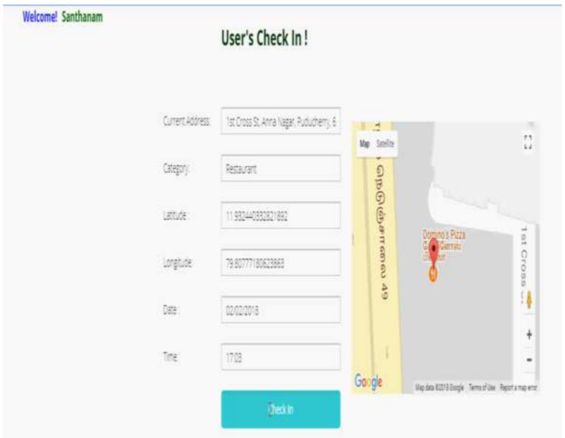

Fig.4.6 Check-in details of User's

Fig.4.7 User's Details viewed by the Admin

Fig. 4.8 Predicting the user's next location

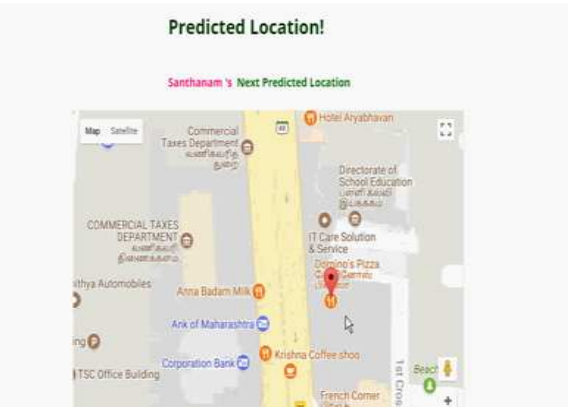

Fig.4.9 Successfully predicting the next location

V. CONCLUSION

In this paper, we have focused on friendship prediction from LBSN dataset. We proposed two friendship prediction models. In model I, we studied the problem of predicting whether two people are friends under the situation that only the check-ins happened at one certain location can be obtained. Compared with the state-of-the-art CS model [6], we take check-in time interval and location entropy into consideration, which leads to a more effective friendship prediction. In model II, differently, we focus on utilizing all the check-in information that belongs to any co-location of two users to predict their relationship. We consider five elements that would make a difference on friendship prediction - the weighted number of co-occurrences, the weighted number of co-locations, the average time interval, the minimum time interval and maximum time intervals. Experimental analysis has shown the efficiency of the proposed system.

REFERENCES

[1] Yuxing Han et al, "GALLOP: GlobAL feature fused Location Prediction for Different Check-in Scenarios", IEEE Transactions on Knowledge and Data Engineering, 2017.

[2] Xiong, L.; Chen, X.; Huang, T.K.; Schneider, J.G.; Carbonell, J.G. Temporal collaborative filtering with bayesian probabilistic tensor factorization. In Proceedings of the Siam International Conference on Data Mining (SDM 2010), Columbus, OH, USA, 29 April–1 May 2010; pp. 211–222.

[3] Zheng, V.W.; Zheng, Y.; Xie, X.; Yang, Q. Collaborative location and activity recommendations with gps history data. In Proceedings of the International Conference on World Wide Web (WWW 2010), Raleigh, NC, USA, 26–30 April 2010; pp. 1029–1038.

[4] Bahadori, M.T.; Yu, R.; Liu, Y. Fast multivariate spatio-temporal analysis via low rank tensor learning. Adv. Neural Inf. Process. Syst. 2014, 4, 3491–3499.

[5] Zhuang, J.; Mei, T.; Hoi, S.C.H.; Xu, Y.Q.; Li, S. When recommendation meets mobile:Contextual and personalized recommendation on the go. In Proceedings of the Ubiquitous Computing, International Conference (UBICOMP 2011), Beijing, China, 17–21 September 2011; pp. 153–162.

[6.] Monreale, A.; Pinelli, F.; Trasarti, R.; Giannotti, F. Wherenext: A location predictor on trajectory pattern mining. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 28 June–1 July 2009; pp. 637–646.

[7] Jiang, B.; Yin, J.; Zhao, S. Characterizing the human mobility pattern in a large street network. Phys. Rev. E Stat. Nonlinear Soft Matter Phys. 2009, 80, 1711–1715

[8] Jeung, H.; Liu, Q.; Shen, H.T.; Zhou, X. A hybrid prediction model for moving objects. In Proceedings of the IEEE 24th International Conference on Data Engineering (ICDE 2008), Cancun, Mexico, 7–12 April 2008; pp. 70–79.

[9] Yava¸s, G.; Katsaros, D.; Ulusoy, Ö.; Manolopoulos, Y. A data mining approach for location prediction in mobile environments. Data Knowl. Eng. 2005, 54, 121–146.

[10] Morzy, M. Mining frequent trajectories of moving objects for location prediction. In Proceedings of the International Conference on Machine Learning and Data Mining in Pattern Recognition (MLDM 2007), Leipzig, Germany, 18–20 July 2007; pp. 667–680