2018 19th IEEE International Conference on Mobile Data Management

Next Check-in Location Prediction via Footprints and Friendship on Location-Based Social Networks

Yijun Su*†‡ , Xiang Li†‡, Wei Tang†‡, Ji Xiang†, Yuanye He†
†*Institute of Information Engineering,Chinese Academy of Sciences*
‡*School of Cyber Security,University of Chinese Academy of Sciences*
Beijing, China
{suyijun, lixiang9015, tangwei, xiangji, heyuanye } @iie.ac.cn

***Abstract*—With the thriving of location-based social networks, a large number of user check-in data have been accumulated. Tasks such as the prediction of the next check-in location can be addressed through the usage of LBSN data. Previous work mainly uses the historical trajectories of users to analyze users' check-in behavior, while the social information of users was rarely used. In this paper, we propose a unified location prediction framework to integrate the effect of history check-in and the influence of social circles. We first employ the most frequent check-in model (MFC) and the user-based collaborative filtering model (UCF) to capture users' historical trajectories and users' implicit preference, respectively. Then we use the multi-social circle model (MSC) to model the influence of three social circles. Finally, we evaluate our location prediction framework in the real-world data sets, and the experimental results show that our model performs better than the state-of-the-art approaches in predicting the next check-in location.**

***Index Terms*—Location-based social networks, Location Prediction, Historical Trajectories, Social Circle**

I. Introduction

Recent years have witnessed the popularity of smart mobile devices and the development of location-acquisition techniques, which make users' location information much easier to obtain than ever before. This development triggers the emergence of *location-based social networks* (LBSNs) [1] platforms such as Facebook Places, Gowalla, and Foursquare, and so on. On these platforms, users can establish social links and share their text content, photos, experience on the Points-of-Interests (POIs). These user-generated activities, which are closely related to some places, are called "check-in". In a LBSN, the application provides location service for user check-in at real physical places. As a consequence, a large amount of user activity data can be obtained, especially the posts of check-in in a LBSN. One of the essential tasks in LBSNs is to utilize the user's check-ins to predict the user's next check-in location, which not only benefits a series of location-based services such as urban computing, POIs recommendation [2], but also makes a better understand human mobility patterns [3].

As check-in data contains rich information, more and more researchers use them to study the problem of location prediction in the present stage. Traditional researches mainly focus on utilizing the individual historical footprints to predict the next check-in location [4], [5], [6], while social friendship information is rarely used. The success of these methods is mainly based on the observation of the user's historical information. However, the prediction performance only using historical records of the user is limited. Therefore, it is a meaningful and challenging task for improving performance in location prediction. Social connection is an indispensable part of the location-based social network. In papers [7], Ye and Cho revealed that users' movement is usually affected to a certain extent by their social relations, such as having dinner with families in a famous restaurant, travelling by following friends' recommendations, and so on. Undoubtedly, social relationships are important to users. Thus, the availability of social networks provides an opportunity to solve check-in prediction. Motivated by these, the goal of this paper is to develop effective algorithms to predict the next check-in location of the user from the perspective of the footprint and the friendship.

In this paper, we propose a unified location prediction framework to integrate the user's history check-ins and the influence of social circles. Specifically, this framework is divided into two modules: (1) personal historical behavior pattern (2) the influence of social circles. Our work first presents the comprehensive study of social relation and then solves next check-in location prediction from the perspective of a new social relationship. Finally, We further evaluate our model in the real-world dataset. The experimental results show that our model performs well in predicting the next check-in location.

In summary, the main contributions of this paper are divided into four aspects, as follows:

- We study the problem of predicting the check-in location with footprint and social relation of the user in location-based social network.
- We put forward the concept of the social circle and model the influence of social circles on user check-in.
- We propose the most frequent check-in model (*MFC*) and the user-based collaborative filtering model (*UCF*) to capture user's historical trajectories and users' implicit preference, respectively. And we use the multi-social circle model (*MSC*) to model the impact of social relationships on user check-in. We integrate the above modules together and present a unified framework (*MUC*)

* Corresponding author.

2375-0324/18/$31.00 ©2018 IEEE
DOI 10.1109/MDM.2018.00044

to predict user's next check-in locations.

- We evaluate our *MUC* prediction framework on large-scale real-world datasets for predicting next check-in location. The experimental results outperform state-of-the-art methods in accuracy aspect.

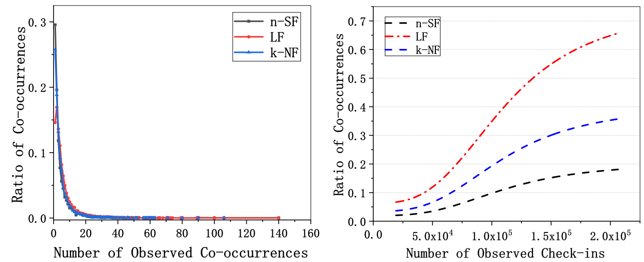
The rest of this paper is organized as follows. Section II reviews the related work. Section III introduces the concept of social circles. Section IV formalizes our research problem and presents the proposed model in details. Section V reports the experimental results. Finally, we draw some conclusions of this study in Section VI.

II. RELATED WORK

One of the main research tasks in location-based social networks is to predict the location of the user. Early work relies mainly on the GPS data collected by cell phones and bluetooth devices [8]. However, some of the obvious drawbacks of the data are poor social relationships and low accuracy. Currently, owing to the development of location-based social services, the way to obtain LBSN data is simpler and more efficient. The LBSN data contains four information layers: a content layer, a social layer, a geographical layer, a temporal layer. The content layer is user-generated photos, audios, videos and words, which record the behavior of the user in interesting locations. The social layer contains the social relationship information about the user, the geographical layer contains the location information checked in by the user, and the temporal layer represents the time stamps information of the user's check-in action. In the following, we review some main researches in these aspects mentioned above.

The content-based approach focuses on the location information mentioned in the post. In papers [9], [10], authors resorted to the content of posts to estimate a user's location. The availability of the temporal layer information presents different views to study the user's location. The temporal information of users' check-in action has two properties in [11], [12]: strong temporal cyclical, short-term effect. However, the above two methods have limited performance without the user history check-in data.

Efforts have also been made to utilize users' historical records for improving the performance of location prediction. In paper [4], Chang et al. found that users always like to check in where they often visit and proposed a logistic regression model to predict users next check-in locations. The difference from our paper is that although the method takes into account users' historical trajectory, it doesn't consider users' implicit preference. In paper [7], studies on human mobility pattern showed that the user's behavior can be affected to a certain extent by his social relations. Social relation of the user plays a key role for predicting location. Many scholars began to investigate the relationship between space trajectory and social networks, and hoped to use mixed information to predict the check-in location in [12], [13]. Compared with the papers we mentioned above, our division of users' friendship is more comprehensive in this paper.



(a) Ratio of co-occurrences for each social circle

(b) Distribution of co-occurrences for each social circle

Fig. 1. Characteristics on social circle on Foursquare

III. SOCIAL CIRCLES ON LBSNS

In this section, we first define the concept of the social circle. Then, to better explain the influence of the user's social circle, we need to investigate the characteristics of the three social circles.

A. Social Circle Definition

When we analyse check-in behavior of the user, there are two situations: checked in at locations alone, or with friends and family. Users always like to do something with their friends. By analyzing datasets (more details about the datasets in section V), we find that people who have checked in the same location with the user are social friends, geo-neighbors, and unrelated strangers. Therefore, these friends are most likely to affect the user's check-in behavior. We consider these friends as three different social circles for the user. Defined as follows:

Definition 1 (n-hop social friends): The social friends are the set of users who have socially connected with the user in LBSNs. According to the structure of the LBSN network, we further expand to n-hop social friends, which is expressed in n-SF, and n denotes n-th hop of the user nodes on social network, for example, 1-SF are friends directly connected with the user, 2-SF denote friends of 1-SF of the user.

Definition 2 (cocheck-in location friends): The common check-in location friends are the set of users who check-in the locations same with the user, which is denoted as LF.

Definition 3 (k-nearest neighbor friends): The k-nearest neighbor friends are the set of users who are geographically close to the user's home, which is denoted as k-NF, and k is the number of friends.

B. Characteristics On Social Circle

To model the influence of the social circle, we investigate check-in habits of users' friends. Due to space limit, we only present analysis on Foursquare, and the Gowalla is similar to it. For any user, we first count the number of co-occurrences between friends in each social circle. Then, the ratio of co-occurrences for each social circle can be shown in Fig. 1(a). The x-axis represents the number of observed co-occurrences in an ascending order, and the y-axis denotes the ratio of co-occurrences. Note that the n and k are set to 1 and 20 here, respectively. The graph shows that the percentage of pairs of friends who have common check-in locations within 50 times is about 90%. The above percentage is approximately the same for each social circle. We guess that the cause of a large number of correlated check-in behaviors is that users might share POIs to friends or check in with friends.

How much of the impact each social circle has the user? We further investigate the check-in patterns of correlations between the user and his social circles. We plot the distribution of co-occurrences as the number of observed check-ins increases in Fig. 1(b). From the figure we can see that as the number of observed check-ins increases, the ratio of co-occurrences also increases, and eventually becomes stable. The reason for this trend may come from two parts: (1) the user hasn't social circles and historical records when he starts to use this application; (2) as time goes on, the user's social circle friends and check-ins also increase. Note that the three social circles may overlap.

IV. Proposed Model

A. Problem Definition

For convenience, we difine $U = \{u_1, u_2, \ldots, u_n\}$ as a set of users, $L = \{l_1, l_2, \ldots, l_m\}$ as a set of locations, where n, m represents the number of users and locations. User u_i check-in at location l_j at time t_k can be expressed as user's check-in trajectory C, where $C = \{\langle u_1, l_1, t_1\rangle \ldots \langle u_i, l_j, t_k\rangle\}$. Let $n\text{–}SF_u, LF_u, k\text{–}NF_u$ denote n-hop social friends, cocheck-in location friends, k-nearest neighbor friends of the user u, and $HC_{u,t} = \{\langle u_i, l_1, t_1\rangle \ldots \langle u_i, l_j, t_k\rangle \mid u_i = u, t_k < t\}$ be the set of historical check-ins of the user u before the time t. In the same way, $SC_{u,t} = \{\langle u_i, l_1, t_1\rangle \ldots \langle u_i, l_j, t_k\rangle \mid u_i \in F_u, t_k < t\}$ be the set of historical check-ins of the user's friends before the time t, where F_u is a set of friends of the user's social circle.

The user's next check-in location $l_?$ at time t is mainly affected by two aspects: personal historical behavior pattern and social circle friends' influence. Hence, we formalize the problem of predicting the user's next check-in location as follows. Given the $HC_{u,t}$ and $SC_{u,t}$, our aim is to calculate the probability of the next location visited by the user at time t. Based on the above explanation, we define the probability as:

$$P_u^t(l) = P_u^t(l|HC_{u,t}, SC_{u,t}) \tag{1}$$

Based on the common assumption of modeling human movement behavior in LBSNs [14], [15], we consider personal footprint and social influence as two independent modules and propose a combination of methods similar to [15]:

$$\begin{aligned} P_u^t(l|HC_{u,t}, SC_{u,t}) =& \alpha P_u^t(l|HC_{u,t}) \\ &+ (1-\alpha) P_u^t(l|SC_{u,t}) \end{aligned} \tag{2}$$

where α is a constant controls parameter.

B. Personal Historical Behavior Pattern

In this module, we use the most frequent check-in model and the user-based collaborative filtering model to capture the user's historical trajectories and implicit preferences. By means of weighted methods, the probability $P_u^t(l|HC_{u,t})$ can be further written as:

$$\begin{aligned} P_u^t(l|HC_{u,t}) =& \beta P_u^{MFC}(l|HC_{u,t}) \\ &+ (1-\beta) P_u^{UCF}(l|HC_{u,t}) \end{aligned} \tag{3}$$

where β is the constant parameter that controls the weight of users' historical trajectories and implicit preferences.

1) MFC Model: The check-in frequency of a location in the user's history check-ins is an important indicator of location prediction [4]. In this study, we prefer to use the most frequent check-in model (MFC) to capture the user's historical preferences. It can be expressed in the following formula:

$$P_u^{MFC}(l|HC_{u,t}) = \frac{\left|\left\{c_k \mid c_k \in C_u, c_k = l, t' < t\right\}\right|}{|C_u|} \tag{4}$$

where c_k is current location of the user, and C_u is a set of the user's check-ins.

2) UCF Model: Based on common sense that similar users have similar preferences, the user-based collaborative filtering method (UCF) can capture users' implicit preference. Let R represents a check-in matrix of user-location. The $r_{i,j}$ is an entry of the check-in matrix R, and $r_{i,j}$ = 1 or 0, where 1 denotes that the user $u_i \in U$ checked in at this location $l_j \in L$ and 0 is the oppsite. We represent the probability by $P_u^{UCF}(l|HC_{u,t})$, and denote as follows.

$$P_u^{UCF}(l|HC_{u,t}) = \frac{\sum_{u_k} w_{i,k} \cdot r_{k,j}}{\sum_{u_k} w_{i,k}} \tag{5}$$

where $w_{i,k}$ is the similarity weight between u_i and u_k. In this study, we adopt cosine similarity measure and denote as follows.

$$w_{i,k} = \frac{\sum_{j \in L} r_{i,j} \cdot r_{k,j}}{\sqrt{\sum_{j \in L} r_{i,j}^2}\sqrt{\sum_{j \in L} r_{k,j}^2}} \tag{6}$$

C. The Multi-Social Circle

In this section we calculate the social circle coefficient and measure social influence strength. Based on weighted methods, the multi-social circle model can be further written as:

$$\begin{aligned} P_u^t(l|SC_{u,t}) =& \psi_1 P_u^t(l|LF_{u,t}) + \psi_2 P_u^t(l|k\text{–}NF_{u,t}) \\ &+ \psi_3 P_u^t(l|n\text{–}SF_{u,t}) \end{aligned} \tag{7}$$

where ψ_1, ψ_2, and ψ_3 are three correlation coefficients, $P_u^t(l|LF_{u,t})$, $P_u^t(l|k\text{–}NF_{u,t})$ and $P_u^t(l|n\text{–}SF_{u,t})$ are social influence strength.

TABLE I
CHECK-IN FEATURES IN THE SOCIAL CIRCLES

Feature	Description
N_X^f	Number of friends in X social circle
N_X^c	Number of check-in in X social circle
N_X^{cc}	Number of cocheck-in in X social circle
N_X^{nc}	Number of new check-in in X social circle

1) Social Circle Coefficient: It can be seen from Fig. 1(b) that the distribution of co-occurrences increases with the increase of the number of observed check-ins at the beginning, and eventually tends to stabilize. Based on the characteristics of the figure, we find that the trend of the curve is similar to the sigmoid function. Inspired by the paper [16], we set correlation coefficients ψ_1, ψ_2, ψ_3 as sigmoid activation functions, which consider a set of features capturing social friends' influence. We take ψ_1 for example, and the other two correlation coefficients are the same.

$$\psi_1 = \frac{1}{1 + e^{-\left(\mathbf{w_1}^T \mathbf{f}_{u,t}^1 + b_1\right)}}, 0 \leq \psi_1 \leq 1 \tag{8}$$

where $\mathbf{f}_{u,t}^1$ is a check-in feature vector of user's friends in LF social circle, $\mathbf{w_1}$ is a weight vector of $\mathbf{f}_{u,t}^1$ and b_1 is the bias. In this study, here are four significant features defined for each user's social circle in the Table I. Note that $\mathbf{f}_{u,t}^1$ is the feature vector before the time t.

2) Social Influence Strength: The potential check-in locations of the user may be related to check-in locations of friends. In this study, social influence strength based on social friends can be realized by the friend-based collaborative filtering (FCF) method. We take LF for example and define the $P_u^t\left(l|LF_{u,t}\right)$ as follows.

$$P_u^t\left(l|LF_{u,t}\right) = \frac{\sum_{u_k \in F_u} SI_{i,k} \cdot r_{k,j}}{\sum_{u_k \in F_u} SI_{i,k}} \tag{9}$$

where $SI_{i,k}$ is the friend similarity between u_i and u_k, and F_u is a set of friends of LF social circle. In this study, we adopt the friend similarity that mixed social links and check-in locations [17].

$$SI_{i,k} = \eta \frac{|F_i \cap F_k|}{|F_i \cup F_k|} + (1 - \eta) \frac{|L_i \cap L_k|}{|L_i \cup L_k|} \tag{10}$$

where η is a tuning parameter ranging within $[0, 1]$, F_k and F_i are the set of friends of social circle, L_k and L_i are the set of check-ins of u_k and u_i, respectively.

D. Unified Framework

We propose the unified framework (MUC) based on personal history module and social circle module, to predict a user's next check-in location.

Therefore, according to (3) and (7), the final framework can be expressed as follows:

$$\begin{aligned} P_u^t(l) = &\alpha\Big(\beta P_u^{MFC}\left(l|HC_{u,t}\right) + (1-\beta) P_u^{UCF}\left(l|HC_{u,t}\right)\Big) \\ &+ (1-\alpha)\Big(\psi_1 P_u^t\left(l|LF_{u,t}\right) + \psi_2 P_u^t\left(l|k\text{–}NF_{u,t}\right) \\ &+ \psi_3 P_u^t\left(l|n\text{–}SF_{u,t}\right)\Big) \end{aligned} \tag{11}$$

E. Parameter Inference

According to (11), we can see that only social module need parameter learning, which greatly simplifies our work. Thus, the product of the probability on the whole set of data can be defined as follows:

$$P\left(C|\Theta\right) = \prod_{(u,l,t) \in C} P_u^t(l) \tag{12}$$

where $\Theta = \{\mathbf{w_1'}, \mathbf{w_2'}, \mathbf{w_3'}\}$ denotes all parameters to be estimated, and the $\mathbf{w_1'}$ contains $\mathbf{w_1}$, b_1, the $\mathbf{w_2'}$ is composed of $\mathbf{w_2}$ and b_2, the $\mathbf{w_3'}$ consists of $\mathbf{w_3}$, b_3. The problem can be further converted to the following minimization problem. And, all parameters are learned by maximum likelihood.

$$\begin{aligned} &min \sum_{u,l,t \in C} - \ln P\left(C|\Theta\right) \\ &+ \lambda\left(\|\mathbf{w_1'}\|_2^2 + \|\mathbf{w_2'}\|_2^2 + \|\mathbf{w_3'}\|_2^2\right) \end{aligned} \tag{13}$$

where λ is the regularization term that avoids overfitting. The λ is set to 0.05 in our study.

V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we first introduce datasets. Then we discuss the location prediction performance of MUC.

TABLE II
STATISTICAL INFORMATION OF THE TWO DATASETS

statistical item	Foursquare	Gowalla
Number of users	11,326	107,092
Number of locations	182,968	1,280,969
Number of chekc-ins	1,385,223	6,442,890
Number of social links	47,164	950,327

A. Datasets Description

In this study, we perform our experiment on the public dataset. We choose Foursquare [16] and Gowalla [7] to evaluate the performance of the MUC framework. The statistics of the datasets are shown in Table II. Note that 1-hop social friends of users are also provided. Since the user's k-nearest neighbor friends were used in the model, we use recursive grid method [12] to estimate the home location of the user. We empirically select users who have at least 80 check-ins and remove POIs that have fewer than 20 check-ins. In our experiments, we divide the dataset into training set and testing set in terms of the user's check-in time instead of choosing a random partition method. Hence, 70 % check-ins of each user are selected for training, and 30% for testing, in a chronological order.

B. Experiment Setup

According to the characteristics of the three social circles, we respectively set the parameter η to 0.2, 0.5, 0.7, corresponding to the LF, the k–NF, the n –SF. The parameters $\psi 1$, $\psi 2$, $\psi 3$ can be learned by training the model. For the parameters α, β, n, k, we will discuss the influence of them in the section V-F.

C. Evaluation Metrics

We use prediction accuracy metric to evaluate the performance of the model. We calculate a probability for each candidate location and return the top-N highest ranked locations as predictions for the user. As long as the actual check-in locations are in top-N predicted in the training, we consider the prediction is correct. We employ the Accuracy@N represent the prediction accuracy of different N. In our experiment, N = 1, 2, 3.

D. Baseline Methods

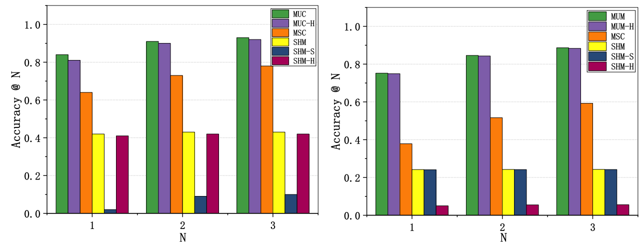
To illustrate the performance of our proposed location prediction framework, We thus introduce the following baseline methods to compare.

- **SHM** [13] is proposed by Gao et al. It integrates historical and social effects to predict the location of the user.
- **SHM-H** is the historical model of the SHM. The Hierarchical PitmanYor process is used in this model to capture power-law distribution and short term effect of check-in behavior.
- **SHM-S** is the social model of the SHM. Unlike our model, the model only uses binary friends information.
- **MUC-H** is the historical model of our framework that only uses the information from personal historical trajectories.
- **MSC** is the multi-social circle model that only uses the information from social circle friends.

E. Performance Comparison

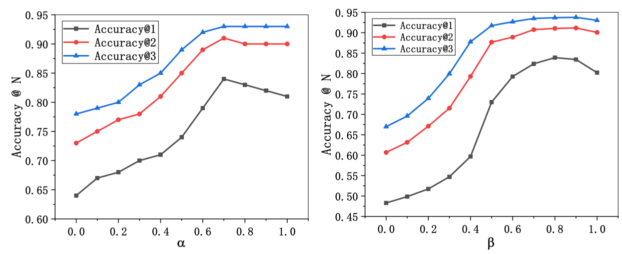
The Fig. 2 shows the prediction results on both datasets. Based on the result of Foursquare, we can see that MUC has the best performance across all models. From the Fig. 2(a) we observe that our model MUC has an average accuracy of 89% accuracy, but the SHM algorithm is less than 43%. Our historical model (MUC-H) performs better than the baseline method (SHM-H) and improves performance by 50%. We guess the reason for the performance gap is that we use the user-based collaborative filtering model to capture users' implicit preference in the history module. MSC also presents high accuracy in all social modules. In addition, our social model (MSC) is worse than the MUC-H model, but better than the SHM-S.

From the Fig. 2(b) we can see that the MUC also performs best on Gowalla. The SHM-S is the lowest accuracy among all the models. The other situations are similar to the performance on the Foursquare. From the above social module comparison we know that the social relationship has a certain impact on the user's check-in behavior, but not the dominant. Users' implicit preferences also play an important role in predicting the next check-in location.



(a) Performance of the MUC on Foursquare (b) Performance of the MUC on Gowalla

Fig. 2. Performance comparison in terms of prediction accuracy



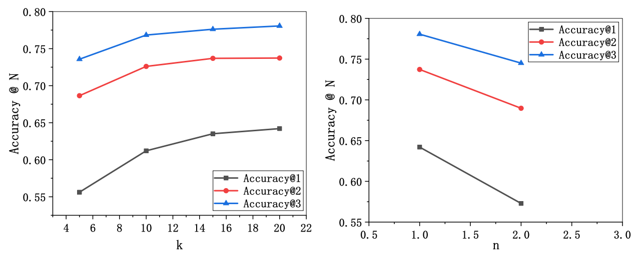
(a) Performance of different weight α (b) Performance of different weight β

Fig. 3. Performance of different weight on Foursquare

F. Parameter Selection

Due to limited space, we only present the parameter selection process on Foursquare data. For α and β, each of which varies from 0 to 1 with an increment step of 0.1, we carry out 100 tests. By observing the results, we find that when α= 0.7 and β = 0.8 hits the best performance. Increasing the parameter β from 0 to 1 with an increment step of 0.1 and fixing α = 0.7, we can see that when β = 0.8 hits the highest accuracy from the Fig. 3(b). It shows that users' implicit preference is very important in historical module. The parameter α is used to control the weight of historical and social module. Setting α from 0 to 1 with an increasing step of 0.1 and fixing β = 0.8, we can observe the changes in performance from the Fig. 3(a). Some interesting insights can be observed:

- When $\alpha = 0$, only the influence of the social circle is considered. Its prediction accuracy is the worst. It shows that only using social information is not good enough to predict user behavior.
- When $\alpha = 0.7$, the performance hits the highest prediction accuracy. It is the best weight of the social module and the historical module. We observe that the historical module has a higher weight, indicating that the



(a) Performance of the MSC on different k

(b) Performance of the MSC on different n

Fig. 4. k of k-NF and n of n-SF

user's historical module is more important than the social module.

- When $\alpha = 1$, this situation only considers the historical trajectories of users, without the influence of the social circle. Its accuracy is not the best, suggesting that social influence is also important.

For the k-nearest neighbor friends of the user, we set k as a variable parameter. So we set k to 5, 10, 15, 20, respectively. In our experiment, we choose the MSC to observe the performance of the prediction. The Fig. 4(a) shows the change in the prediction accuracy at different number k. From the figure we can observe that the performance of the MSC is increasing at different top-N with the increase of the k. We can draw the conclusion that the impact of neighbors' friends on users is obvious. For the n-hop social friends of the user, we set n to 1, 2, respectively. According to Dunbar's number, we know that the number of social friends of any person is limited. Users keep a close contact with a part of social friends, so these close friends have a greater impact on users. As shown in the Fig. 4(b), we observe that the performance of the MSC is descreasing as n increases. We guess the reason for this phenomenon is that too many social friends weaken the impact of close friends.

VI. Conclusion

In this paper, we proposed a unified location prediction framework (MUC) to integrate the user's history check-ins and the influence of the social circle. In the historical behavior pattern, we employed the most frequent check-in model (MFC) and the user-based collaborative filtering (UCF) model to capture user's historical trajectories and implicit preference, respectively. In the social circle module, we used the multi-social circle model (MSC) to model the impact of social circles. We evaluated our location prediction framework in the real-world datasets. The experimental results showed that our model performs well in predicting the next check-in location.

Acknowledment

This work was partially supported by National Natural Science Foundation of China No. U163620068, and National Key R&D Plan No. 2016QY02D0401.

References

[1] L. Humphreys, "Mobile social networks and social practice: A case study of dodgeball," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 341–360, 2007.

[2] H. Li, Y. Ge, and H. Zhu, "Point-of-interest recommendations: Learning potential check-ins from friends," in *Proceedings of the 22th ACM SIGKDD conference on Knowledge discovery and data mining. ACM*, 2016.

[3] Y. Liu, C. Liu, X. Lu, M. Teng, H. Zhu, and H. Xiong, "Point-of-interest demand modeling with human mobility patterns," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13 - 17, 2017*, 2017, pp. 947–955.

[4] J. Chang and E. Sun, "Location 3 : How users share and respond to location-based data on social networking sites," *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pp. 74–80, 01 2011.

[5] W. Mathew, R. Raposo, and B. Martins, "Predicting future locations with hidden markov models," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, ser. UbiComp '12. New York, NY, USA: ACM, 2012, pp. 911–918.

[6] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *Data mining (ICDM), IEEE 12th international conference on*. IEEE, 2012, pp. 1038–1043.

[7] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD conference on knowledge discovery and data mining*. ACM, 2011, pp. 1082–1090.

[8] D. Zhang, D. Zhang, H. Xiong, L. T. Yang, and V. Gauthier, "Nextcell: Predicting location using social interplay from cell phone traces," *IEEE Transactions on Computers*, vol. 64, no. 2, pp. 452–463, Feb 2015.

[9] B. Hecht, L. Hong, B. Suh, and E. H. Chi, "Tweets from justin bieber's heart: The dynamics of the location field in user profiles," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '11. ACM, 2011, pp. 237–246.

[10] D. Flatow, M. Naaman, K. E. Xie, Y. Volkovich, and Y. Kanza, "On the accuracy of hyper-local geotagging of social media content," in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, ser. WSDM '15. New York, NY, USA: ACM, 2015, pp. 127–136.

[11] M. Ye, K. Janowicz, C. Mülligann, and W.-C. Lee, "What you are is when you are: The temporal dimension of feature types in location-based social networks," in *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. GIS '11. New York, NY, USA: ACM, 2011, pp. 102–111.

[12] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, "Exploring millions of footprints in location sharing services." in *ICWSM*. The AAAI Press, 2011.

[13] H. Gao, J. Tang, and H. Liu, "Exploring social-historical ties on location-based social networks." in *ICWSM*, J. G. Breslin, N. B. Ellison, J. G. Shanahan, and Z. Tufekci, Eds. The AAAI Press, 2012.

[14] J. Tang, H. Gao, H. Liu, and A. Das Sarma, "etrust: Understanding trust evolution in an online world," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ser. KDD '12. New York, NY, USA: ACM, 2012, pp. 253–261.

[15] H. Gao, J. Tang, X. Hu, and H. Liu, "Modeling temporal effects of human mobile behavior on location-based social networks," in *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, ser. CIKM '13. New York, NY, USA: ACM, 2013, pp. 1673–1678.

[16] J. Gao, Huiji andTang and H. Liu, "gscorr: modeling geo-social correlations for new check-ins on location-based social networks," in *21st ACM International Conference on Information and Knowledge Management, CIKM'12, Maui, HI, USA, October 29 - November 02, 2012*, 2012, pp. 1582–1586.

[17] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ser. SIGIR '11. New York, NY, USA: ACM, 2011, pp. 325–334.