2016 2nd International Conference on Open and Big Data

Mining Semantic Location History for Collaborative POI recommendation in Online Social Networks

Luepol Pipanmekaporn
Department of Computer and Information Science
King Mongkut's University of Technology
North Bangkok, Bangkok, Thailand 10800
e-mail: luepol.p@sci.kmutnb.ac.th

Suwatchai Kamolsantiroj
Department of Computer and Information Science
King Mongkut's University of Technology North
Bangkok, Bangkok Thailand 10800
e-mail: suwatchai.k@sci.kmutnb.ac.th

Abstract— **Location-based social networks (LBSNs) have recently attracted millions of mobile users to explore attractive locations and share their visited experiences. As the increasing use of the online social networks, people demand personalized service to recommend places of interests (POIs) based on their personal preferences. Among POIs recommendation approaches, collaborative filtering that predicts POIs of the user based on the geospatial location and users' opinions is suite for LBSNs. Despite this, it is still a challenge to infer the similarity between users because of the unique characteristics of spatial items in LBSN. In this paper, we propose an effective POI recommendation method for LBSNs based on collaborative filtering. Our method focuses on mining interest similarity of users based on their check-in activities in LBSN. Since the geospatial locations cannot capture user's interests, we perform to mine semantic features of user's check-in history based on semantic location descriptions to discover the user's interests. We finally perform recommending nearby places to a particular user by fusing opinions from similar users according to the user's current location. Experimental results with two real-world datasets collected from Foursquare show that our proposed method can achieve satisfying precision and recall of recommended places.**

Keywords— Location based Social Networks, POI recommendations, User Location History, Collaborative Filtering

I. INTRODUCTION

With the rapid deployment of GPS-enabled devices, location-based services (LBS) have grown extremely popular. LBS uses a real-time geospatial position (i.e., latitude and longitude) of a mobile user to provide nearby places around the user. This enables the user to explore attractive locations such as restaurants, shopping malls, cafes and movie theaters within a distance range [1]. Some of the online services (e.g. Foursquare and Facebook places) include social networking features, known as "*location-based social network (LBSN)*" [2], that have attracted millions of mobile users to share their location experiences at anytime and anywhere.

Despite the growing popularity of the online social networks, people often demand personalized recommendation service to help them explore new places or locations that the users may be interested in. This service is often referred as a point-of-interest (POI) recommendation system [3, 7]. As shown in Figure 1, the POI recommender service has played an important role in the online social networks to recommend POIs to a particular user based on the user's current location and their personal preferences or interests.

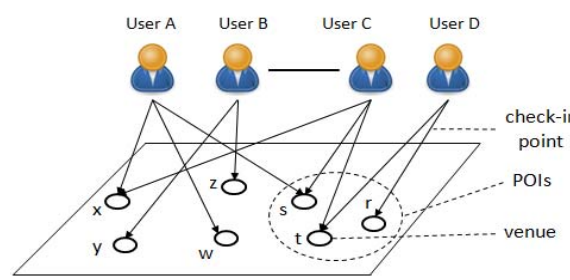


Figure 1 POIs recommendation service in LBSNs

Over the few years, POIs recommendation algorithms were proposed [2] [4] [7-11]. Among these recommendation algorithms, user-based collaborative filtering is a common and effective approach to POIs recommendation in LBSNs [3] [10]. Despite this, traditional collaborative filtering techniques are not suitable for POIs recommendations for some reasons. The first reason is that traditional recommender systems do not take the effect of geographical locations, which is very important for discovering POIs of the user. The second reason involves the unique characteristics of user-location check-in data in LBSNs [3]. This data typically includes only the frequency of users' location check-ins. The large range of frequency values can have negative impact on calculation of user similarity compared to explicit user-rating in the traditional recommendation systems. Moreover, the sparsity of user-location check-in frequency matrix is extremely higher than that of user-item rating matrix. The above reasons require a specific solution for POIs recommendations in the online social networks.

In this paper, we devise a novel effective POIs recommendation algorithm for LBNSs. Our proposed method is based on a user-based collaborative filtering that fuses opinions from similar users to recommend POIs for a particular user. As GPS-based locations cannot describe user's check-in behavior, we extract semantic features of the user's visited locations based on semantic location descriptions (e.g. "food", "nightlife spot", "movie theater" and "museum") to model the

978-1-5090-4054-4/16 $31.00 © 2016 IEEE
DOI 10.1109/OBD.2016.12

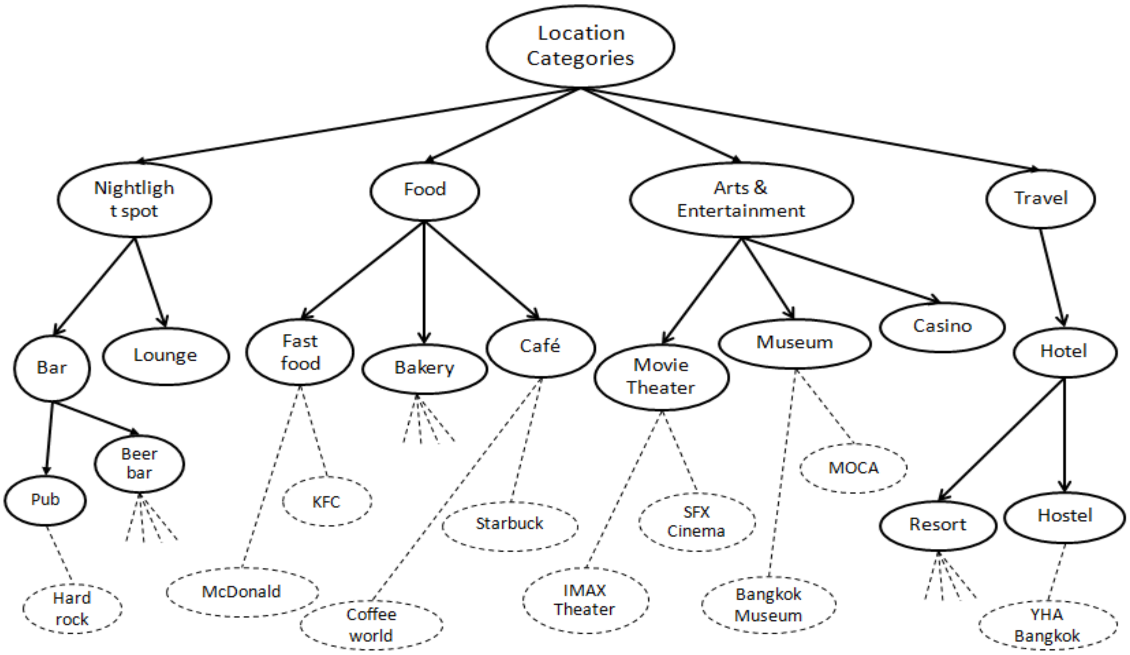


Figure 2 A part of the three-level structure of categories of places in Foursquare

TABLE 1 USER LOCATION HISTORY FROM FOURSQUARE

User	Local ID	Category Name	Latitude	Longitude	Offset	UTC time
470	49bbd6c0f964a520f4531fe3	Arts & Crafts Store	40.719810375488535	74.00258103213994	240-	Tue Apr 03 18:00:09 2012 0000+
704	4cb4b31cb4b0a35df33969ce	American Restaurant	40.73067679262482	74.06567180055883-	240-	Tue Apr 03 18:11:35 2012 0000+
161	4ea2c05d0cd6fc5af3c0928b	Seafood Restaurant	40.70964742200671	74.01051729615469-	240-	Sat May 26 02:56:54 2012 0000+
161	43bba61df964a520eb2c1fe3	Seafood Restaurant	40.70964742200671	7-4.01108529090852	240-	Sun May 27 23:28:30 2012 0000+

user's interests. By using the semantics, the similarity between users who may never overlap their visited locations because they may live in different cities can be inferred. To recommend POIs to a particular user, we fuse opinions of users who share similar interests according to the user's current location. We also conduct experiments with a real dataset of Foursquare database. The experimental results demonstrate that our proposed method can achieve satisfying precisions and recalls of the top-N recommended POIs. The contributions of this work are as follows:

- Our proposed method alleviates the quality concerns of the social network data that can affect inferring user similarity in POIs recommendation. Instead of the direct use of the frequency data, we compute semantic features of the user's visited locations based on high-level descriptions of the physical locations.
- Our recommendation method can infer the interest similarity between users who may never overlap their visited locations. This is known as the travel locality which is still a challenge issue in POIs recommendation.
- Finally, we evaluate the proposed recommendation method over two real-world datasets from Foursquare. Experimental results show that our user-based collaborative recommendation technique achieves superior POI recommendation performance against other CF approaches.

The rest of the paper is organized as follows. In Section II, we summarize the related works on location recommendation. In Section III, we explain basic concepts of our proposed approach. Section IV and V describe the details of our recommendation method. The experimental settings and results are shown in Section VI and VII. Finally, we make conclusions in Section VIII.

II. Related work

Recently, location recommendation has attracted a lot of attentions from researchers. Several efforts have been invested in the past few years. We survey and summarize the related works into two categories. The first category is based on predicting specific POI that a user will visit next and the second category is to infer user's preference for recommending unvisited POI.

To predict the user's next locations, various studies have been proposed. For example, the authors in [3] proposed to extract intermediate locations between start and destination locations of each user based on GPS trajectories and then train hidden Markov model (HMM) to predict user's next POI. The study in [4] introduced the notion of spatiotemporal-periodic (STP) patterns to capture user's mobility and then propose an efficient algorithm to mine these patterns for predicting next places. The work [5] proposed to extract user specific features and global mobility features. The authors then built a prediction model for next visits. While the previous works focused on the problem of predicting the user's next places, our work focuses on inferring the user's preference from location history to suggest POIs.

Unlike the above efforts, most research works target on predicting unvisited POIs to users by identifying their user's preferences. One of the earliest works in this direction is to ask a user to manually provide his personal interests by categories [6]. However, manually specifying the personal preferences is not effective. A brunch of the recent works learns a user's interests from users' check-in data for POI recommendation [7-11]. For example, in [7] the authors developed an item-based collaborative recommendation algorithm for spatial items (i.e., venues) in LBSNs. In [8], the authors proposed user-based and item-based collaborative POI recommendation algorithms based on geographical and social (i.e., friends) influences and then demonstrated that user-based POI approach performs better than item-based one. In [9] the authors tackled new venue recommendation problem by estimating user preference on unvisited POIs. Some of these works focus on handling data sparseness problem by using model-based recommendation techniques, such as the works in [10][11] applied SVD to a user-location matrix to obtain a semantic space of user's location history. Some of these works also combine additional information available in LBSN, such as text-based tips [12], POI categories [13, 14, 15] and spatial-temporal data [16], with user's check-in data to enhance the performance of POI recommendation. Differing from these works, our work focuses on discovering user's preferences from semantic location history that captures more meaningful user's interests than that of using a pure user location matrix.

III. Prelimenary

In this section, we explain basic concepts that are necessary for our proposed approach.

A. User's Location History

In location-based social networks (LBSNs), such as Foursquare and Facebook Places, a user can maintain personal information, such as name, age, gender and hometown. The user can also mark a place of interest and leave some comments to share experience with his/her friends and other users. This activity is called as "check-in" activity. All of the user's check-ins recorded by LBSN is known as user's "location history". As shown in Table 1, each row of the user location history contains user id, location id, category name, latitude, longitude, time zone offset and UTC time. Some users, such as user 161, have multiple visited at the same location.

B. User-based Collaborative Filtering

Collaborative Filtering (CF) is one of the most popular techniques used in recommendation systems. The CF methods produce user specific recommendations of items based on patterns of user ratings. In CF, users' implicit preference can be discovered by aggregating the opinions of peer users, i.e., other users those have similar tastes in candidate items. Let U and L denote the user set and the POI set in a LBSN, which keeps record of user check-ins. The check-in activity of a user $u_i \in U$ has at a POI $l_j \in L$ is denoted as $c_{i,j}$ where $c_{i,j} = 1$ indicates user u_i has a check-in at l_j and $c_{i,j} = 0$ means there is no record of u_i. These recorded user check-in activities are used to infer a user's implicit preference of a POI, which can be represented as a probability to predict how likely the user would like to have a check-in at an unvisited place.

We denote the prediction by $\hat{s}_{i,j}$ and obtain this predicted check-in score of u_i to l_j as follows:

$$\hat{s}_{i,j} = \frac{\sum_{u_k} w_{i,k} * c_{k,j}}{\sum_{u_k} w_{i,k}} \quad (1)$$

where $w_{i,k}$ indicates the similarity weight between users u_i and u_k.

Table 2 ten top-level categories in Foursquare database

Category Name	#Sub-Categories
Art& Entertainment	17
College& University	23
Event	9
Food	78
Nightlife Spot	20
Outdoors & Recreation	28
Professional & Other Places	94
Residence	15
Shop & Service	45
Travel & Transport	14

C. Semantic Location Description

As mentioned in previous sections, a physical location does not have enough information to describe underlying interests of a user. This is because these locations may be related in some ways. For example, "Starbucks" and "Coffee World" are related in relation to "Café". Instead of using the physical locations, we utilize their semantic descriptions to discover user's interests. In this work, we utilize the three-level hierarchical structure of predefined categories for venues in Foursquare [20] to obtain the semantics of user's check-ins. Figure 2 illustrates a part of the hierarchical structure of venue categories. The bottom level categories assign specific descriptions for venues whilst the upper level categories describe more general descriptions. Table 2 gives all of the top-level categories and their number of sub-categories.

IV. Mining Location History

To discover personal interests of a particular user, we mine location history in LBSN by incorporating the categorical information of his/her visited locations.

A. Top Specific Locations

Although there are many locations that a user visits, but the user may visit only a few locations frequently. The locations visited more by the user can infer the user's preferences better than the locations visited less. In this work, only the top-m significant locations of each user are used to infer the user similarity. To achieve this goal, we define the significant score based on the tf-idf weighting scheme [19]. For each location l_j that user a has visited, we calculate the significant score of the user's location as follows.

$$\hat{r}_a(l_j) = \frac{f_a(l_j)}{f_a(.)} \times log\frac{|U|}{|U_j|} \quad (2)$$

where $f_a(l_j)$ is the number of the user's visited location l_j
$f_a(.)$ is the total number of the user's visits locations
$|U|$ is the total number of users in LBSN
$|U_j|$ is the number of users that visit the location l_j

According to the above equation, the location significance of user a consists of combining two terms. The first term refers the probability of user's visited location l_j whilst the second term reflects how popular this location is. This is because popular locations, such as "train station" and "shopping complex", tend to be meaningless to capture the user's preferences compared to less popular ones. By using Eq.(2), the location history of user a can be represented by the weighted vector of his/her top-m significant locations.

$$R_a = < \hat{r}_a(1), \hat{r}_a(2), \dots, \hat{r}_a(m) >$$

B. Computing Semantic Space

As mentioned previous sections, the physical locations cannot capture the user's interests. Instead of using the low-level locations, we map the significant scores of the user's top-m locations into categorical space obtained by LBSN. For each category $c \in C$, the propagated weight is assigned as the following equation.

$$w_a(c) = \frac{\sum_{j\in c} \hat{r}_a(l_j)}{\sum_{k\in C} \hat{r}_a(l_k)} \times log\frac{|U|}{|U_c|} \quad (2)$$

where $\hat{r}_a(l_j)$ is the significant score of location l by user a
$|U_c|$ is the number of users that visits all of the locations belonging to category c

Similar to the significant scores of physical locations, we also compute the appropriate weights for each category in the hierarchy. The motivation is to aggregate the weights of user's physical locations into related categories for inferring the user similarity.

The user's weighted category graph provides the following advantages: 1) handle the data sparseness problem of user location history and 2) capture more semantics for computing the similarity between users those do not have share any physical locations. Figure 3 illustrates an example of computing the categorical weights in the hierarchy.

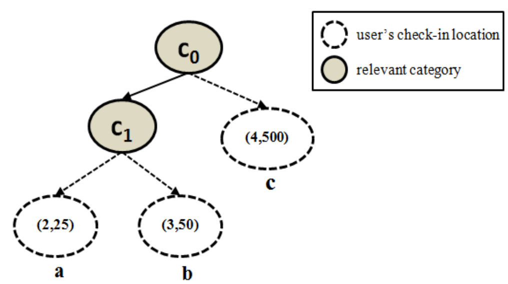


Figure 3 Calculation of categorical weights

Figure 3 shows a part of the hierarchical structure of categories. The dash nodes indicate the user's check-in locations that we assume that the places a, b and c respectively. Each of the visited locations contains a pair of values (x, y) where x is the frequency of visiting the location by the user and y is the number of users that visit the location. The nodes C_1 and C_0 mean the categories of these places. Let us give an example of the weight calculation. Assume that the total number of users in the system is 10,000, the significant score of the user's place a can be calculated by Equation (2) as follows: $\frac{2}{2+3+4} \times log\frac{10000}{25} = 1.918$ while the significant score of the place b is $\frac{3}{2+3+4} \times log\frac{10000}{50} = 2.548$.

To obtain the appropriate weights for the category C_1, we aggregate the significant scores of the places a and b belonging to this category by using Equation (3) as the follow $w(C_1) = \frac{\hat{r}(a)+\hat{r}(b)}{\hat{r}(a)+\hat{r}(b)+\hat{r}(c)} = \frac{1.918+2.548}{1.918+2.548+0.658} \times \log\frac{10000}{75} = 6.152$. Finally, the weight of category C_0 is $w(C_0) = \frac{\hat{r}(a)+\hat{r}(b)+\hat{r}(c)}{\hat{r}(a)+\hat{r}(b)+\hat{r}(c)} = \frac{1.918+2.548+0.658}{1.918+2.548+0.658} \times \log\frac{10000}{575} = 4.120$. For each category, we perform to calculate the weights corresponding to the user visited locations.

C. User Similarity Calculation

To infer the similarity between users, we perform matching their common interests based on the categorical space. The idea is that two users are more similar if they share more common categories. To achieve this, we introduce the similarity between users U_a and U_b based on their common categories as follows:

$$Sim(U_a,\ U_b) = \sum_{j \in C} \min\{w_a(c_j), w_b(c_j)\} \qquad (3)$$

where $w_i(c_j)$ means the weight of category c_j assigned for user U_i by using Equation (2). By using the above equation, the similarity between the two users can be calculated as the total sum of the minimum weights of their common categories. We keep the minimum weights of the shared categories as their common interests. The higher the score is, the more likely the users are.

V. POI Recommendation Algorithm

Algorithm 1 describes the overall procedure of the POIs recommendation algorithm. The algorithm starts by retrieving a collection of candidate venues V' from an online social network service based on a limited distance of the user's current location l_c. In Line 2, we retrieve a set of local users those have visited these venues V' and their location check-in activities. In Line 3-4, we perform to construct a categorical graph for the user based on his/her top-m significant locations computed by using the equations (1) and (2). Each categorical node in the graph is assigned by the weights of top-m significant locations to represent the user's interests. Line 5-7 are performed to extract the categorical graph for each local user u. Line 8 performs to match the similarity between the user and local ones on the categorical space. Finally, the algorithm returns top- k selected users U_k and a set of candidate venues $V^{'}$ for POI recommendations.

Algorithm 1: POI recommendation algorithm

Input: 1) The user location l_c of user U_c and 2) a location region r
Output: 1) A set of selected k local users U_k and 2) A set of candidate locations $V^{'}$

1) Retrieve places V' within the radius r from the center location l_c
2) Retrieve local users U who have visited V'
3) Compute significant scores (Eq. (1)) for each visited location of U_c
4) Construct a categorical graph for U_c based on his top-m significant locations.
5) **for** each user $U_j \in U$ **do**
6) Compute significant scores for each visited location of U_j
7) Construct a categorical graph for U_j based on his top- m significant locations
8) Compute the pairwise similarity between U_c and U_j (Eq. (3))
9) **end for**
10) **return** top-k selected users U_k and candidate places $V^{'}$.

To recommend a location l_j, we fuse the opinions of k selected local users by aggregating the similarity between the user and these local users who have visited the location as follows:

$$S_a(l_j) = \frac{\sum_{x \in U_k} Sim(U_a, U_x) \times f_x(l_j)}{\sum_{x \in U_k} Sim(U_a, U_x)} \qquad (4)$$

where $f_x(l_j)$ indicates the frequency of user x in a set of K selected users for the place l_j. The function $Sim(a, x)$ means the similarity between user a and user x obtained by using the equation (3). The higher score assigned to the place l_j indicates that this user would like enjoying the place. Finally, the candidate places are ranked corresponding to the score function and only the top ranked places are recommended to the user.

VI. Experiment Settings

In this section, we describe the settings of experiments including the dataset, baseline approaches and the evaluation method. After that, we report on major results on the effectiveness.

A. Data Collection

To evaluate our proposed model, we use two datasets collected from Foursquare [16]. These datasets mainly consist of collections of Foursquare check-ins for ten months (from 12 April 2012 to 16 February 2013) in two big cities, i.e., Tokyo and New York. Each venue of the user's check-ins is also assigned by the most appropriate category of Foursquare. Table 3 shows statistics of the selected datasets for the experiments.

TABLE 3 DATASET STATISTIC

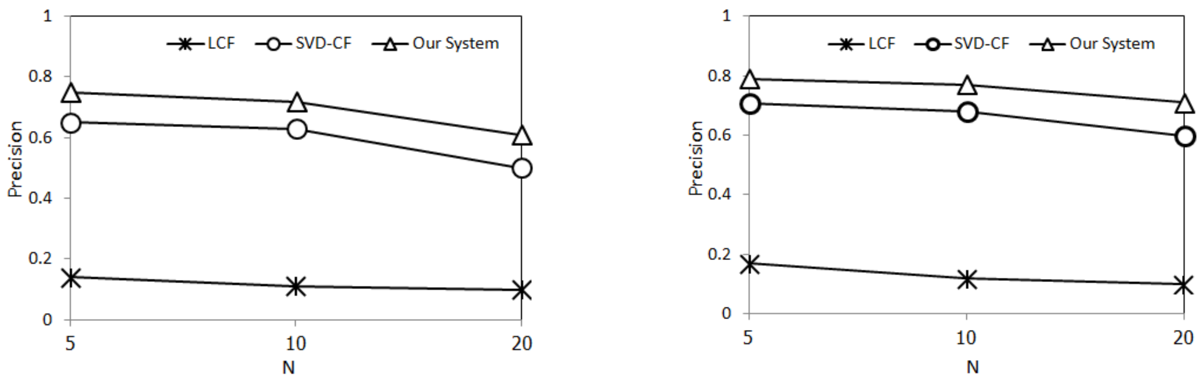
Dataset	Users	Venues	Check-ins	Avg. number of visited venues/user
New York (Foursquare)	824	38,336	227,428	31.39
Tokyo (Foursquare)	1,939	61,858	573703	38.37

B. Baseline approaches

We compare our proposed approach with the following baseline approaches.

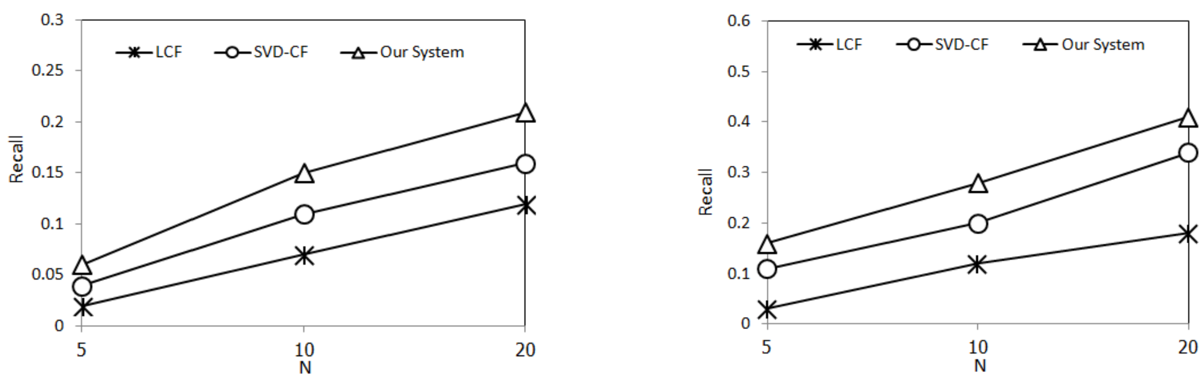
1) Location-based Collaborative filtering (LCF)

Collaborative filtering (CF) is one of the most widely used approaches in location recommender systems because traditional CF methods can be directly applied over the venues. Location-based collaborative filtering (LCF) basically utilizes the users' location history with a user-venue matrix and then applies CF algorithms to make recommendations of interesting places. The cosine similarity between two users' location vectors is used as the similarity between two users.



(a) Tokyo city (b) New York city

Figure 4 Average precision@N of different methods



(a) Tokyo city (b) New York city

Figure 5 Average recall@N of different methods

Finally, the locations in the user-specified range and having a relatively inference score will be recommended.

2) SVD-based CF recommendation (SVD-CF)

According to the study [12], matrix factorization based collaborative filtering has been successfully in recommender systems. This is because of its ability in dealing with very large sparse user-item rating matrix. Recently, the collaborative filtering has gained a more popularity for POI recommendation in LBSNs. To examine the effectiveness of our proposed method, Singular Value Decomposition (SVD), a well-known matrix factorization method, is applied to make POI recommendations in collaborative filtering.

In SVD-CF, we perform z-scores to normalize the frequency values in user-location check-in matrix. SVD is then applied to discover a meaningful latent space to transform high-dimensional user's location vectors into lower dimensional space for inferring the user similarity.

C. Evaluation methods

To evaluate the prediction accuracy of the proposed method, we examine the precision and recall of recommended venues and compare the accuracy with LCF and SVD-CF methods.

In the evaluation, we select 20 testing users in the users' location history and then hide some of their already visited places for examining the recommendation method. After that, we use the recommendation method to recommend top-5, top-10 and top-20 venues to each user. The more recommended places that a user truly visited, the more effective the recommendation method is. By using this method, we can calculate the precision and recall of the recommendation algorithm for each user using the following equations (5) and (6)

$$precision = \frac{tp}{tp + fp} \tag{5}$$

$$recall = \frac{tp}{tp + fn} \tag{6}$$

where tp is the number of recommended places that were truly visited (hidden) by the user. fp is the number of recommended places that were not visited by the user. fn is the number of truly visited places that were not recommended by the system. Finally, tn is the number of unvisited places that are not recommended by the system. Based on the precision and recall, the F-score is also calculated as

$$F - score = \frac{2 \times precision \times recall}{precision + recall} \quad (7)$$

We finally compute the average precision, recall, and F-scores of all the querying users.

VII. EXPERIMENTAL RESULTS

Figure 4 and 5 demonstrate the average precision and recall of different methods varying in the number of top-N recommended places. The most important findings revealed in these figures are that our proposed method outperforms the baseline approaches over all the standard measures. We can also see that the lowest performance results are obtained by LCF. This is due to the limited information provided by sparse datasets. Compared to LCF, the significant improvements of SVD-CF and our CF method are caused by the advantages of exploiting semantic information to model a user's location history. Moreover, we compared our CF method with SVD-CF and found that the proposed method outperforms SVD-CF. As shown in Figure 4 and 5, the best performance of our CF method is achieved when $N = 20$ in Tokyo (F-score = 0.312) and New York (F-score = 0.551) datasets. These results support the effective use of using location categories for modeling user's activity preferences.

VIII. DISCUSSIONS

The performance of our recommendation algorithm is affected by the following two major factors: 1) the number of visited locations in user's location history and 2) the number of local users in a user's query range. In this section, we discuss these issues.

A) User's Specific Locations

Our proposed method uses only the top specific locations of a user to model user's preferences. We believe that these locations are meaningful to capture the user's interests. We experimentally show that the considering only top-m visited locations of a particular user is better than that of all the visited locations of the user. As shown in Figure 6, the F-scores of using all visited locations to recommend POIs perform dropping behind using the selected locations. These results demonstrate that the large part of user's location history in user's check-in matrix is meaningless to model the user's preferences.

As seen in the above figure, the average F-score performs dropping after selecting top-20 and top-30 specific locations in Tokyo and New York datasets respectively. In addition, the benefit of the selection of visited locations improves the efficiency of POI recommendations.

B) Users Opinion Selection

User-based collaborative filtering performs to recommend POIs based on fusing the opinions of similar users. Our proposed method selects the top-k users those have visited the candidate places based on a geospatial range. Intuitively, a larger range will include more venues and user candidates. We examine the effects of the candidate selection within 10 miles. Table 4 shows the average precision@10 and recall@10 of selecting the candidate users to recommend POIs.

As seen in this table, our CF method performs close performance on the different number of the selected users compared to using all the candidates. These results can be described by the fact that the users have sparse check-in activities in the candidate venues. This would become the benefit for our proposed method in term of the efficiency performance due to only considering a small part of local users employed to recommend high-quality POIs to a user.

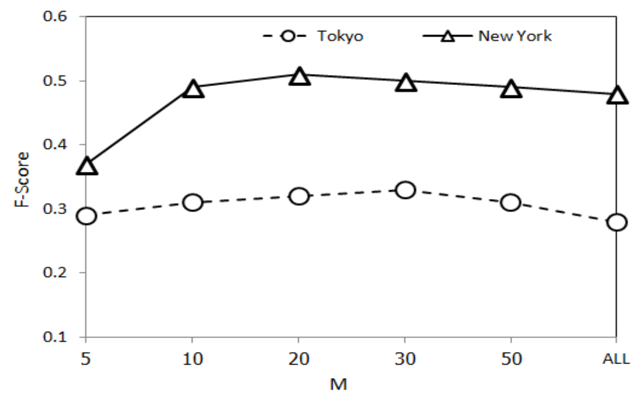


Figure 6 Average F-score w.r.t. various top-m specific locations

TABLE 4 AVERAGE PRECISION@10 & RECALL@10 ON VARIOUS K LOCAL USERS

Candidate users	Precision@10		Recall@10	
	Tokyo	New York	Tokyo	New York
5	0.78	0.81	0.16	0.24
10	0.78	0.81	0.17	0.26
15	0.79	0.80	0.18	0.25
20	0.79	0.80	0.18	0.25
ALL	0.79	0.80	0.17	0.26

IX. CONCLUSIONS

In this paper, we proposed a novel user-based collaborative POI recommendation algorithm for location-based social networks. Our method focuses on inferring user's check-in behaviors from user's location history to make a high-quality recommendation of POIs to a user based on the current location and opinions of similar users. Unlike existing recommendation methods that purely uses original user-location matrix, we learn the user's preferences from semantic location history obtained by using the location category hierarchy. By using the category information, the similarity between users who may have never overlapping their visited locations can be semantically estimated. We evaluated our recommendation technique by conducting experiments on two real-world datasets collected from Foursquare. According to the experimental results, our proposed approach outperforms some state-of-the-art collaborative location recommendation method, including LCF and SVD-CF methods. For these two datasets, the best performance of our proposed method is achieved by recommending top-10 POIs.

As a future work, we would explore other information such as user generated tags and temporal features into the recommendation system to get more accurate results.

X. REFERENCES

[1] Rao, Bharat, and Louis Minakakis. "Evolution of mobile location-based services." *Communications of the ACM* 46.12 (2003): 61-65.

[2] Wang, Hao, Manolis Terrovitis, and Nikos Mamoulis. "Location recommendation in location-based social networks using user check-in data." *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2013.

[3] Yu, Yonghong, and Xingguo Chen. "A Survey of Point-of-Interest Recommendation in Location-Based Social Networks." *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*. 2015

[4] Zhang, Chao, Lidan Shou, Ke Chen, Gang Chen, and Yijun Bei. "Evaluating geo-social influence in location-based social networks." In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pp. 1442-1451. ACM, 2012.

[5] Lee, Sungjun, et al. "Next Place Prediction Based on Spatiotemporal Pattern Mining of Mobile Device Logs." *Sensors* 16.2 (2016): 145.

[6] Cho, Sung-Bae. "Exploiting machine learning techniques for location recognition and prediction with smartphone logs." *Neurocomputing* 176 (2016): 98-106.

[7] Gambs, Sébastien, Marc-Olivier Killijian, and Miguel Núñez del Prado Cortez. "Next place prediction using mobility markov chains." *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*. ACM, 2012.

[8] Park, Moon-Hee, Jin-Hyuk Hong, and Sung-Bae Cho. "Location-based recommendation system using bayesian user's preference model in mobile devices." *Ubiquitous Intelligence and Computing*. Springer Berlin Heidelberg, 2007. 1130-1139.

[9] Levandoski, Justin J., Mohamed Sarwat, Ahmed Eldawy, and Mohamed F. Mokbel. "Lars: A location-aware recommender system." In Data Engineering (ICDE), 2012 IEEE 28th International Conference on, pp. 450-461. IEEE, 2012.

[10] Ye, M., Yin, P., Lee, W. C., & Lee, D. L. (2011, July). 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* (pp. 325-334). ACM.

[11] Bao, Jie, Yu Zheng, and Mohamed F. Mokbel. "Location-based and preference-aware recommendation using sparse geo-social networking data." *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*. ACM, 2012.

[12] Zheng, Vincent W., Yu Zheng, Xing Xie, and Qiang Yang. "Collaborative location and activity recommendations with gps history data." In *Proceedings of the 19th international conference on World wide web*, pp. 1029-1038. ACM, 2010.

[13] Lian, Defu, Cong Zhao, Xing Xie, Guangzhong Sun, Enhong Chen, and Yong Rui. "Geomf: Joint geographical modeling and matrix factorization for point-of-interest recommendation." In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 831-840. ACM, 2014.

[14] Yang, Dingqi, et al. "A sentiment-enhanced personalized location recommendation system." *Proceedings of the 24th ACM Conference on Hypertext and Social Media*. ACM, 2013.

[15] Li, Quannan, Yu Zheng, Xing Xie, Yukun Chen, Wenyu Liu, and Wei-Ying Ma. "Mining user similarity based on location history." In *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, p. 34. ACM, 2008.

[16] Wang, Z., Zhang, D., Zhou, X., Yang, D., Yu, Z., & Yu, Z. (2014). Discovering and profiling overlapping communities in location-based social networks. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, *44*(4), 499-509.

[17] Yang, D., Zhang, D., Zheng, V. W., & Yu, Z. (2015). Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *Systems, Man, and Cybernetics: Systems, IEEE Transactions on*, *45*(1), 129-142.

[18] Xiao, Xiangye, et al. "Finding similar users using category-based location history." Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM, 2010.

[19] Hiemstra, Djoerd. "A probabilistic justification for using tf× idf term weighting in information retrieval." *International Journal on Digital Libraries* 3, no. 2 (2000): 131-139.

[20] Li, Yanhua, Matthias Steiner, Limin Wang, Zhi-Li Zhang, and Jie Bao. "Exploring venue popularity in foursquare." In *INFOCOM, 2013 Proceedings IEEE*, pp. 3357-3362. IEEE, 2013.