

PCRM: Increasing POI Recommendation Accuracy in Location-Based Social Networks

Lianggui Liu¹, Wei Li¹, Lingmin Wang¹, Huiling Jia¹

¹School of Information Science and Technology, Zhejiang Sci-Tech University
Hangzhou, 310018, China.
[e-mail: weil0812@126.com]

*Corresponding author: Lianggui Liu

*Received November 22, 2017; revised March 21, 2018; accepted June 5, 2018;
published November 30, 2018*

Abstract

Nowadays with the help of Location-Based Social Networks (LBSNs), users of Point-of-Interest (POI) recommendation service in LBSNs are able to publish their geo-tagged information and physical locations in the form of sign-ups and share their experiences with friends on POI, which can help users to explore new areas and discover new points-of-interest, and promote advertisers to push mobile ads to target users. POI recommendation service in LBSNs is attracting more and more attention from all over the world. Due to the sparsity of users' activity history data set and the aggregation characteristics of sign-in area, conventional recommendation algorithms usually suffer from low accuracy. To address this problem, this paper proposes a new recommendation algorithm based on a novel Preference-Content-Region Model (PCRM). In this new algorithm, three kinds of information, that is, user's preferences, content of the Point-of-Interest and region of the user's activity are considered, helping users obtain ideal recommendation service everywhere. We demonstrate that our algorithm is more effective than existing algorithms through extensive experiments based on an open Eventbrite data set.

Keywords: Location-Based Social Networks, Point-of-Interest recommendation, geographical area, user's preference

1. Introduction

In recent years, with the constant development of mobile Internet technology, such as satellite communications, GPS devices, wireless sensor networks and Internet communications, the positioning function of smart terminals provided by people is increasingly accurate and convenient. In this context, Location-Based Social Networks (LBSNs) are rapidly growing and related location-based service are being loved by the majority of users [1, 2]. In LBSNs, users are able to publish their geo-tagged information and physical locations in the form of sign-ups and share their experiences with friends on points-of-interest (e.g., shopping malls, restaurants, museums, entertainment venues, hotels, etc.) , which can help users to explore new areas and discover new points-of-interest, and promote advertisers to push mobile ads to target users, making location-based social networks more attractive. User's sign-in data is critical to personalized recommendations for location-based social networks. Recent success stories include Meetup [3], Plancast [4], and Eventbrite [5]. For example, Eventbrite can help users access to the geographical location in the real life, and to publish information about the location (such as comments, photos and video) through the sign-in function, share and connect with other users. With the help of LBSNs, we can collect a spatiotemporal database containing hundreds of millions of users' sign-in records. From this database, we are able to obtain a lot of useful information, such as which activity the user is interested in, where the user likes to go, and so on. Thus, a variety of recommendation services can be deployed based on the location information obtained in LBSNs, which include friends recommendation [6-8], POI recommendation [9-13] and so on. Among them, research on POI recommendation is more extensive, and it is responsible for recommending a series of unknown points of interest to the target users to enhance the user's experience. In general, POI recommendation includes both places (such as restaurants, convention centers, shops, parks and cinemas, etc.) and activities (such as concerts, public welfare activities, etc.). With the difference between the conventional e-commerce site commodity recommendation systems, POI recommendation in LBSNs has three characteristics: weak semantic and sparsity of data [14], affection of geographical location on POI recommendation [15] and less social impact [16], which brought much challenge into the POI recommendation approach design. Existing POI recommendation algorithms based on the location of social network mainly utilize user's personal information data, location metadata and the user's historical location data to recommend. According to the type of data used for POI recommendation, they can be summarized into three categories: user's sign-in data based [17], geographical factor based [18] and social impact based [19] recommendation algorithms. However, due to the sparsity of users' activity history data set and the aggregation characteristics of sign-in area, most of the existing recommendation algorithms usually suffer from low efficiency.

To address these challenges, this paper presents a new model named Preference-Content-Region Model, hereinafter referred to as PCRM. The main contributions of our work are summarized below:

- (1). It shows that three aspects, that is, user's preferences, the content of POI and the region of the user's activity, which are complement each other, can overcome the dependence on the user's sign-in data simply caused by the sparsity of data to a certain degree.

(2). Based on the implicit content and observation mentioned above, we propose a novel model which considers not only the user's preferences and the content of POI, but also the region of the user's activity. The model combines the various factors with a unified framework.

(3). We evaluate the recommendation accuracy of the proposed method through extensive experiments and results show that it performs better than other methods.

The remainder of this paper is organized as follows. We first give a brief introduction to representative relevant work in Section 2. Then in Section 3 we give the description of Preference-Content-Region Model consisting of model introduction, formal definition of the model, parameter estimation of the model and model recommendation. In Section 4, we present our experiment settings and analyze the performance of different methods based on the obtained results of extensive experiment. We conclude this paper and derive future research direction in Section 5.

2. Related Work

Recently with the rapid growth of LBSNs, like Meetup, Plancast, and Eventbrite, etc., recommending locations (i.e., POIs) for users becomes prevalent. In general, there are three main categories for existing POI recommendation approaches:

(1). Social impact based recommendation algorithms. For this kind of methods, the social network of "friends" have the same hobbies and were combined with access history which friends sign-in to recommend. In [14], similar users were used to recommend and ignore other users in LBSNs, which greatly improves the efficiency of the algorithm. In [20], social relations were directly integrated into the Probabilistic Matrix Factorization (PMF) Algorithm, but experiments show that social relations have little effect on recommendation accuracy.

(2). User's sign-in data based recommendation algorithms. This kind of methods consider that the number of times that a user signed-in for a POI can represent his or her preference, construct the user-interest point sign matrix and use the conventional recommendation algorithm to recommend. Representative algorithms include user and POI based Hybrid Collaborative Filtering Algorithm [14] and Probabilistic Matrix Factorization (PMF) Algorithm [16]. The essence of these methods is to try to improve the recommendation model, to overcome the challenge brought by the sparsity of data. However, because the data itself are very sparse, the recommendation quality is not high.

(3). Geographical factor based recommendation algorithms. In LBSNs, geographical distance between user and POI is an important factor in recommendation, which is the important feature different from commodity recommendation. This kind of algorithms integrate geographic information into the model. In [21], the POI which the user signed in the geographical position is consistent with power-law distribution. It is proven true that modeling of geographic information can help to improve the effectiveness of recommendation.

There are also methods which consider geographical locations, user preferences, and social relationships. In [22] an user-based collaborative filtering UPS (User, Proximity and Social-Based CF) algorithm was proposed, taking into account the social impact and user's preference. After calculating users' similarity, the social impact-based collaborative filtering algorithm is combined with the user-based collaborative filtering algorithm. Experiments show that the recommendation effect of the algorithm is not very good when the data set is sparse, and the algorithm does not consider the influence of the user on the final recommendation in different geographical locations. Algorithm in literature [23] takes into

account user's preference, social influence and impact of geography, and presents a USG (User, Social and Geographical influence) recommendation algorithm using a linear integration frame work to integrate these three factors. Each factor was used to generate the location of the recommendation list of fusion to improve the accuracy of the recommendation algorithm. The algorithm considers the geographical factors. However, it only considers the characteristics of the users' residence area, the recommendation location is the area near residence, and the algorithm parameters cannot be adjusted adaptively. Thus, the method can overcome the sparse data and cold start problem, but since it completely ignored the user's preferences, its recommendation accuracy is typically low.

Table 1. Symbolic Description of Model Parameters

parameter	definition
d, u, w, l, x, ζ, r	d for the document, u for the user's preference preference, w for the word, l for the POI, x for the control switch, ζ for the content, r for the region
K, U, D, W, R, L	number of topics, number of user's preference preference, number of documents, number of words, number of regions, number of POI
N_d	number of words in document d
$C_{kd,-i}^{KD}$	regardless of the current instance, the number of times the subject k appears in document d
$C_{wk,-i}^{WK}$	regardless of the current instance, the number of times the word v is assigned to the subject k
$C_{lk,-j}^{LK}$	regardless of the current instance, the times of POI l is assigned to topic k
$C_{ku,-i}^{KU}$	regardless of the current instance, the times of the user u assigns the POI as k
$n_{l_j, preference, -j}$	regardless of the current instance, the number of occurrences of POI l_j that the user is interested in.
$n_{l_j, document, -j}$	regardless of the current instance, the number of occurrences of POI l_j from the document
$n_{l_j, region, -j}$	regardless of the current instance, the number of occurrences of POI l_j from the region content
λ_l	multinomial probability distribution based on the POI's content selection
$\zeta, \sigma, \varsigma, l, \tau$	normalizing parameters
$\alpha_u, \alpha_d, \alpha_w, \alpha_r, \beta_w, \beta_l, \eta$	hyper-parameters of Dirichlet Distribution.

3. Preference-Content-Region Model

3.1 Model introduction

Preference-Content-Region Model is a probability generation model that can truly reflect the user's sign-in process in LBSNs. The model considers that there are three factors that play a decisive role in the decision-making process of a POI:

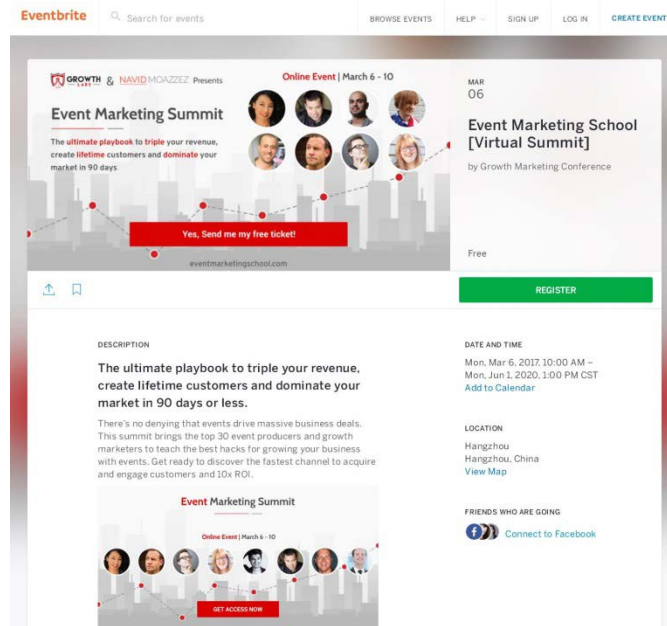


Fig. 1. An example of introduction to the content of POI

(1). User's preference: Only when the user is interested in an activity, he will be willing to go there and produce a sign-in behavior. For example, football fans may see the CBA league, and music lovers may go to the concert. In collaborative filtering recommendation algorithm, interest of users is regarded as same as other similar users, and then similar activities will be recommended to the users who may be interested.

(2). Content of POI: There are many introductory documents about POI in the current LBSNs. As shown in **Fig. 1**, introduction to the activity of Eventbrite, in addition to giving the specific time and place of the event, the page also details the subject of the event. When a user browses to the introductory document about the POI, he may be attracted by a feature (or subject) of the activity, resulting in a final sign-in behavior. At present, plenty of commodity recommendation systems are combined with collaborative filtering and mixed content recommendation technology to improve accuracy rate, and overcome the challenge brought by the sparsity of data to a certain extent. The introduction to the content of POI here is also the case.

(3). Region of the user's activity: In general, firstly, user tends to be active in the area around his resident area, and less participates in activities away from the resident area. So the distance between POI and the user is an important determinant. Secondly, when the user goes out to a new location, he knows nothing about the region, and cannot get information from his "similar users". He should consider more customs and habits of the region when making decisions, such as local's interest preferences, or the local famous cultural attractions and so on.

Preference-Content-Region Model presented in this paper considers the three factors described above.

3.2 Formal definition of the model

In the proposed PCRM, a user's sign-in action at a POI can be described as follows. Before the user signs at the POI l_u , l_u 's description of the document d already exists. Thus, a word generating process in the document has nothing to do with the POI, which is the

content distribution of the document and the content distribution of the word can be calculated independently. When the user signs in the POI, he should determine the l_u 's content z at first. There are three choices for z , which include the appeared content in the introductory document d , the user's interest content, and the content of the area to which the POI belongs. We use a selection variable X to control the source of the POI z , and X satisfies polynomial distribution, and its values include user's preferences, content of POI, and place of user's activity.

In PCRM, several sets should be introduced, which include the user's preference set U , the POI set L , the POI corresponding to the introduction of the document set D , the vocabulary set W , the area set R , and the implied content set K . The user's sign-in process of POI can be regarded as a probabilistic process [24], which is detailed as follows:

1) For each document d in the document set D , we get the document d 's subject distribution $\theta^{(d)}_k$ based on the $DIR(\alpha_d)$ distribution.

2) For each content k in the content set K , we get the probability distribution $\phi^{(w)}_k$ of the word w on the subject k , according to $DIR(\beta_w)$ distribution; then we derive the probability distribution $\phi^{(l)}_k$ of the POI l on the subject k according to $DIR(\beta_l)$ distribution.

3) For each user u in the preference set U , the user u 's subject distribution $\theta^{(u)}_k$ can be obtained according to $DIR(\alpha_u)$ distribution; then we derive the region r 's subject distribution $\theta^{(r)}_k$ based on $DIR(\alpha_r)$ distribution; finally we obtain the distribution $\theta^{(ur)}_R$ of the user u on the region R based on $DIR(\alpha_{ur})$ distribution.

4) The generation process of each word W_j in document d is as follows: (Note that the length of document is N_d .)(a). First, we derive the content Z_i according to the polynomial distribution $\theta^{(d)}_K$ of the document based on the content sampling.(b).Then, we derive the word W_i according to the polynomial distribution $\phi^{(w)}_{z_i}$ of the word based on the content Z_i sampling.

5) The user u 's process of signing the POI is as follows:

(a) we derive λ_i according to the $DIR(\eta)$ sampling, then the control value x can be obtained according to the $Multinomial(\lambda_i)$ distribution;

(b) If $x = \text{document}$, the content of the POI is generated by the topic of the introductory document: find the content set $\{Z_{W_1}, Z_{W_2}, \dots, Z_{W_{Nd}}\}$ for all words in the document. These contents are subject to uniform distribution $Uniform(Z_{W_1}, Z_{W_2}, \dots, Z_{W_{Nd}})$, based on which we get the content Z_l of the POI;

(c) If $x = \text{preference}$, then the content of interest points generated based on the user's interest. Then we get the content Z_l of POI according to the content distribution $\theta^{(u)}_K$ that the user is interested in;

(d) If $x = \text{region}$, at this time the content of the POI is generated by the region's own content. We can get region r according to the user u 's distribution $\theta^{(ur)}_R$ in the region R , and get content Z_l according to region r 's distribution $\theta^{(r)}_K$ in the content K ;

(e) Get the POI l according to POI's distribution $\phi^{(l)}_{z_i}$ in the content Z_l .

3.3 Parameter Estimation of the Model

The joint probability distribution of all known and potential variables of the model is:

$$\begin{aligned}
 p(u, r, l, z, x, W) &= p(W) p(l | z, x, r, u) p(z, r, u | x) p(x) \\
 &= \iota p(W) (p(l | z, x) (p(z | x, \theta^u) p(x = \text{preference}) \\
 &\quad + p(l | z, x) p(z | x, \theta^d) p(x = \text{document}) \\
 &\quad + p(l | z, x) p(r | u) p(z | x, \theta^r) p(x = \text{region}))
 \end{aligned} \tag{1}$$

In the parameter estimation of the LDA model, the Gibbs algorithm does not directly calculate $\theta^{(d)}$ and $\phi^{(w)}$ as parameters, but iteratively compute the posterior probability $p(z | w)$ of the word on the content, and then perform sampling to obtain the value of $\theta^{(d)}$ and $\phi^{(w)}$. It can be regarded as a state changing process of the Markov chain. Till the posterior probability of the word on the content is convergent, the state will be stable. This model can be regarded as a concurrent execution of two LDA models [25]. It simulates the generation of POI while simulating the generation of words. The related sampling equations are listed as follows (Corresponding specific parameters are described in Table 1):

1) Firstly, we calculate the posterior probability of the word on the content, as known in Formula (2), and then, sample the content of the word according to the posterior probability,

$$p(z_i = k | w_i, z_{-i}, w_{-i}, \alpha_d, \beta_w) \propto \frac{C_{kd, -i}^{KD} + \alpha_d}{\sum_k C_{kd, -i}^{KD} + K\alpha_d} \cdot \frac{C_{wk, -i}^{WK} + \beta_w}{\sum_w C_{wk, -i}^{WK} + V\beta_w} \tag{2}$$

2) Secondly, we calculate the posterior probability of the POI. The process includes three cases,

a. When variable $x = \text{preference}$ is selected, the sampling equation is:

$$\begin{aligned}
 p(z_j = k, x = \text{preference} | l_j, z_{-j}, \alpha_u, \beta_l, \eta) &\propto \\
 \frac{C_{ku, -j}^{KU} + \alpha_u}{\sum_k C_{ku, -j}^{KU} + K\alpha_u} \cdot \frac{C_{lk, -j}^{LK} + \beta_l}{\sum_l C_{lk, -j}^{LK} + L\beta_l} \cdot \frac{\eta_{\text{preference}} + n_{l_j, \text{preference}, -j}}{L + \sum_{x \in \{\text{preference}, \text{region}, \text{document}\}} \eta_x}
 \end{aligned} \tag{3}$$

b. When variable $x = \text{document}$ is selected, the sampling equation is:

$$\begin{aligned}
 P(z_j = k, x = \text{document} | l_j, z_{-j}, \beta_l, \eta) &\propto \\
 \frac{C_{kd}^{KD}}{N_d} \cdot \frac{C_{lk}^{LK} + \beta_l}{\sum_l C_{lk}^{LK} + L\beta_l} \cdot \frac{\eta_{\text{document}} + n_{l_j, \text{document}, -j}}{L + \sum_{x \in \{\text{preference}, \text{region}, \text{document}\}} \eta_x}
 \end{aligned} \tag{4}$$

c. When variable $x = \text{region}$ is selected, the sampling equation is:

$$\begin{aligned}
 P(z_j = k, r_j = r, x = \text{region} | l_j, z_{-j}, \alpha_u, \alpha_{ur}, \beta_l, \eta) &\propto \\
 \frac{C_{ru, -j}^{RU} + \alpha_{ur}}{\sum_r C_{ru, -j}^{RU} + R\alpha_{ur}} \cdot \frac{C_{kr, -j}^{KR} + \alpha_r}{\sum_k C_{kr, -j}^{KR} + K\alpha_r} \cdot \frac{C_{lk, -j}^{LK} + \beta_l}{\sum_l C_{lk, -j}^{LK} + L\beta_l} \cdot \frac{\eta_{\text{user}} + n_{l_j, \text{region}, -j}}{L + \sum_{x \in \{\text{preference}, \text{region}, \text{document}\}} \eta_x}
 \end{aligned} \tag{5}$$

The state is stable when Formula (2)(3)(4)(5) are iterated for a certain number of times, and related parameters of the model can be approximately calculated by:

$$\theta^{(d)} = \frac{C_{kd}^{KD} + \alpha_d}{\sum_k C_{kd}^{KD} + K\alpha_d} \tag{6}$$

$$\phi^{(w)} = \frac{C_{wk}^{WK} + \beta_w}{\sum_w \zeta C_{wk}^{WK} + W\beta_w} \quad (7)$$

$$\theta^{(u)} = \frac{C_{ku}^{KU} + \alpha_u}{\sum_k \sigma C_{ku}^{KU} + K\alpha_u} \quad (8)$$

$$\phi^{(l)} = \frac{C_{lk}^{LK} + \beta_l}{\sum_l \zeta C_{lk}^{LK} + L\beta_l} \quad (9)$$

$$\theta^{(ur)} = \frac{C_{ru}^{RU} + \alpha_{ur}}{\sum_r \zeta C_{ru}^{RU} + R\alpha_{ur}} \quad (10)$$

$$\phi^{(r)} = \frac{C_{kr}^{KR} + \alpha_r}{\sum_k \sigma C_{kr}^{KR} + K\alpha_r} \quad (11)$$

$$\lambda_{l, \text{preference}} = \frac{\eta_{\text{preference}} + n_{l, \text{preference}}}{L + \sum_{x \in \{\text{preference}, \text{region}, \text{document}\}} \zeta \eta_x} \quad (12)$$

$$\lambda_{l, \text{document}} = \frac{\eta_{\text{document}} + n_{l, \text{document}}}{L + \sum_{x \in \{\text{preference}, \text{region}, \text{document}\}} \sigma \eta_x} \quad (13)$$

$$\lambda_{l, \text{region}} = \frac{\eta_{\text{region}} + n_{l, \text{region}}}{L + \sum_{x \in \{\text{preference}, \text{region}, \text{document}\}} \zeta \eta_x} \quad (14)$$

3.4 Model recommendation

The parameter estimation process of the above model is completed in the offline phase. In the online recommendation phase, for a given user u , the POI l and its corresponding introductory document d , use Formula (15) to calculate the probability that the user sign in the POI l :

$$p(l | d, u) = p(x = d) \sum_{r=1}^K \sum_{k=1}^K p(l | z_k) p_{\text{test}}(z_k | d) + p(x = u) \sum_{r=1}^K \sum_{k=1}^K p(l | z_k) p(z_k | u) + p(x = r) \sum_{r=1}^R \sum_{k=1}^K p(r | u) p(z_x | r) \quad (15)$$

where $p(l | z_k)$, $p(x)$, $p(z_k | u)$, $p(r | u)$, $p(z_k | r)$, $p_{\text{test}}(z_k | d)$ can be obtained through model training. If it is a new POI l , we can calculate the content distribution $p_{\text{test}}(z_k | d)$ of its corresponding document d online, that is, document d will be added to the test set D , and it can be obtained by reusing Gibbs sampling method. If it is a new user, p can directly be calculated according to Formula (16). Thus, the problem of cold start of users or resources can be solved.

$$p(l | d, u) = \sum_{r=1}^K \sum_{k=1}^K p(l | z_k) p_{\text{test}}(z_k | d) \quad (16)$$

4. Experimental Results and Analysis

4.1 Data Set

Eventbrite provides an online platform for release of the event and participation for city organizers and personal user. Users can use the mobile client to search for various activities,

and express the willingness to participate and sign-in. The data set includes 100,000 users, 300,000 events and 3500,000 sign-in records. Specifically, the data set users include the user's number, user's name and user's city. Activity information includes activity number, activity name, activity location, the introduction of the activity, and the type of activities. Sign-in record includes user number and activity number. In this paper, we select the 15000 users, 150000 sign-in records of 14500 activities as the experimental data set after processing.

4.2 Experimental Evaluative Methods and Comparative Methods

In order to measure the accuracy of the algorithm, we choose the evaluation method proposed in [26]. This method is based on the recall rate commonly used in the field of information retrieval. Based on this evaluation method, the corresponding changes are as follows:

- (1) 90% of the data set randomly was selected as the training set S to build the recommendation model, the remaining 10% was used as the test data set T .
- (2) For any user u , the test randomly selects 200 activities which the user has not signed in yet to constitute a set of E . (note that we assume that the users are not interested in these activities).
- (3) Add any activity e in the user's test set to E to form 201 sets of activities, and use the designed recommendation algorithm to select the highest score from the top 200 activities as top-200 recommendation list. If the activity e is in the recommendation list, the increment of constant hits will be 1, otherwise the value of hits will be unchanged.

We chose six algorithms for comparison:

- (1) IKNN [27] uses the historical activity sign-in matrix to calculate the user's similarity, and then recommends the activities in which the neighboring users is interested, then filter the activities according to the distance between the location of the activity and the user, and prefer the activity that is closer to the user.

(2) CIKNN [28] maps the user's hobbies to different specific categories and then calculate the user's similarity according to the weight of each category and the user's sign-in record. The essence of the algorithm is collaborative filtering.

(3) In [23], user interest, social and geographical influence are taken into consideration, and a USG (User interest, Social and Geographical influence based recommendation) recommendation algorithm is proposed, using a linear fusion framework to integrate these three factors, and combining the location recommendation list generated by each factor. The core of the algorithm is collaborative filtering. In addition, two other factors are added to reduce the impact brought by the sparsity of data.

(4) Preference-Content Model (PCM) and Preference-Region Model (PRM) can be regarded as two sub-models of the PCRM, respectively. When $\lambda_{i,document} = 0$, the model ignores the content information about the POI's introductory document, and PCRM will degrade into PRM. When $\lambda_{i,region} = 0$, the model ignores the content information of the area where the POI is located, and PCRM will degrade into PCM.

4.3 Experimental Results

This model has nine hyper-parameters that need to be set in advance. For the recommendation model, the value of the hyper-parameter has little effect on the final output, but it affects the convergence rate of the model. Here we set each one of $(\alpha_u, \alpha_d, \alpha_{ur}, \alpha_r)$ 0.1, each one of (β_w, β_l) 0.05, and η 0.01.

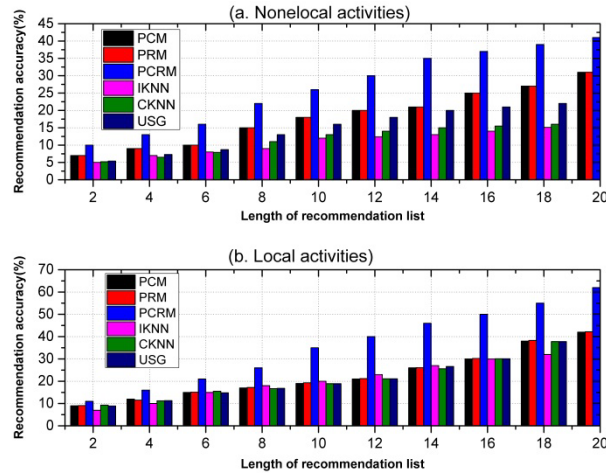


Fig. 2. Comparison of recommendation accuracy in nonlocal/local activities

We evaluated the recommendation accuracy of the six algorithms, considering that the sign-in behavior of users with geographical clustering. Focused activities considered in the test are divided into two categories: (a) the user's local activities, (b) the user's nonlocal activities. **Fig. 2** shows the top- N recommendation accuracy of the above six algorithms in different cases. In the experiment, the length N of the recommendation list is changed from 2 to 20.

It can be observed from **Fig. 2a**, with the increasing of N , the accuracy of referenced algorithms is constantly improving. For the nonlocal activity recommendation, the method based on hidden subject model (including PCRM, PCM, PRM) is superior to the memory-based method (USG algorithm, CKNN algorithm, IKNN algorithm) which are based on cooperative filtering. Because the possibility of user's sign-in in this region is little, the sparse data make the user's similarity or POI's similarity is not accurate in these algorithms. Because the USG algorithm also considers the influence of the user's social network, the recommendation accuracy is slightly higher in the CKNN algorithm and the IKNN algorithm. The implicit subject model is less affected by the sparsity of the data. The implicit content of the POI in the model is also affected by the distribution of the user's interest, the distribution of the content of POI and the distribution of the content of the area where the POI belongs. This information is a useful supplement to the user's sign-in data. Both PCM and PRM consider only two of these effects, so the recommendation accuracy is not better than PCRM.

As can be observed from **Fig. 2b**, for the recommendation of the local activities, this method is also superior to other methods in the paper. But the advantage is not obvious, even the accuracy of the USG algorithm and CKNN algorithm is comparable to that of PCM and PRM. Because users have more data on local sign-in activities, the use of collaborative filtering algorithm can accurately calculate the users' preferences. At the same time, it shows that the USG algorithm has little difference with CKNN algorithm and IKNN algorithm, indicating that the user's social impact is not significant for local activities recommendation.

5. Conclusion and Future Directions

In this paper, a new POI recommendation model, Preference-Content-Regional Model is proposed, taking user's preference, content of POI and region of the user's activity into consideration. In PCRM sign-in behavior of the user is modeled using an integrated

probability graph. Experiments results show that the model can overcome the influence of the sparsity of data and improve the recommendation accuracy compared with other methods. In the future work, we will integrate time, mood and other context factors into the model, making the recommendation more accurate and efficient.

Acknowledgement

This work is supported by the National Natural Science Foundation of China and Civil Aviation Administration of China under grant No. U1533133, National Natural Science Foundation of China under grant No. 61002016 and No. 61711530653, the Humanities and Social Sciences Research Project of Ministry of Education of China under grant No. 15YJCZH095, the China Scholarship Council under grant No. 201708330439 and the 521 Talents Project of Zhejiang Sci-Tech University.

References

- [1] M. Elhamshary, A. Basalmah, and M. Youssef, "A Fine-Grained Indoor Location-Based Social Network," *IEEE Transactions on Mobile Computing*, vol. 16, no. 5, pp. 1203-1217, 2017. [Article \(CrossRef Link\)](#)
- [2] T. Stepan, J. M. Morawski, S. Dick, and J. Miller, "Incorporating Spatial, Temporal, and Social Context in Recommendations for Location-Based Social Networks," *IEEE Transactions on Computational Social Systems*, vol. 3, no. 4, pp. 164-175, 2016. [Article \(CrossRef Link\)](#)
- [3] [Article \(CrossRef Link\)](#).
- [4] [Article \(CrossRef Link\)](#).
- [5] [Article \(CrossRef Link\)](#).
- [6] L. Liu and H. Jia, "Trust Evaluation via Large-Scale Complex Service-Oriented Online Social Networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 11, pp. 1402-1412, 2015. [Article \(CrossRef Link\)](#)
- [7] C. Liu, J. Liu, and Z. Jiang, "A multiobjective evolutionary algorithm based on similarity for community detection from signed social networks," *IEEE transactions on cybernetics*, vol. 44, no. 12, pp. 2274-2287, 2014. [Article \(CrossRef Link\)](#)
- [8] W. Wang, J. Jiang, B. An, Y. Jiang, and B. Chen, "Toward efficient team formation for crowdsourcing in noncooperative social networks," *IEEE transactions on cybernetics*, 2016. [Article \(CrossRef Link\)](#)
- [9] G. Xu, B. Fu, and Y. Gu, "Point-of-Interest Recommendations via a Supervised Random Walk Algorithm," *IEEE Intelligent Systems*, vol. 31, no. 1, pp. 15-23, 2016. [Article \(CrossRef Link\)](#)
- [10] B. Hu, M. Jamali, and M. Ester, "Spatio-temporal topic modeling in mobile social media for location recommendation," in *Proc. of Data Mining (ICDM), 2013 IEEE 13th International Conference on*, pp. 1073-1078: IEEE, 2013. [Article \(CrossRef Link\)](#)
- [11] J.-D. Zhang and C.-Y. Chow, "TICRec: A probabilistic framework to utilize temporal influence correlations for time-aware location recommendations," *IEEE Transactions on Services Computing*, vol. 9, no. 4, pp. 633-646, 2016. [Article \(CrossRef Link\)](#)
- [12] H. Gao, J. Tang, X. Hu, and H. Liu, "Exploring temporal effects for location recommendation on location-based social networks," in *Proc. of Proceedings of the 7th ACM conference on Recommender systems*, pp. 93-100: ACM, 2013. [Article \(CrossRef Link\)](#)
- [13] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proc. of Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pp. 325-334: ACM, 2011. [Article \(CrossRef Link\)](#)

- [14] J. D. Zhang and C. Y. Chow, "CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations," *Information Sciences*, vol. 293, pp. 163-181, 2015. [Article \(CrossRef Link\)](#)
- [15] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A General Geographical Probabilistic Factor Model for Point of Interest Recommendation," *IEEE Transactions on Knowledge & Data Engineering*, vol. 27, no. 5, pp. 1167-1179, 2015. [Article \(CrossRef Link\)](#)
- [16] J. Zhang, C. Chowmember, and Y. Li, "iGeoRec: A Personalized and Efficient Geographical Location Recommendation Framework," *IEEE Transactions on Services Computing*, vol. 8, no. 5, pp. 701-714, 2015. [Article \(CrossRef Link\)](#)
- [17] O. Kwon and J. Kim, "Concept lattices for visualizing and generating user profiles for context-aware service recommendations," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1893-1902, 2009. [Article \(CrossRef Link\)](#)
- [18] B. Liu, Y. Fu, Z. Yao, and H. Xiong, "Learning geographical preferences for point-of-interest recommendation," in *Proc. of Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1043-1051: ACM, 2013. [Article \(CrossRef Link\)](#)
- [19] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model-based collaborative filtering for personalized POI recommendations," *IEEE Transactions on Multimedia*, vol. 17, no. 6, pp. 907-918, 2015. [Article \(CrossRef Link\)](#)
- [20] C. Cheng, H. Yang, I. King, and M. R. Lyu, "Fused Matrix Factorization with Geographical and Social Influence in Location-Based Social Networks," in *Proc. of the Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012. [Article \(CrossRef Link\)](#)
- [21] C. Chen, D. Zhang, B. Guo, and X. Ma, "TripPlanner: Personalized Trip Planning Leveraging Heterogeneous Crowdsourced Digital Footprints," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1259-1273, 2015. [Article \(CrossRef Link\)](#)
- [22] Q. Yuan, G. Cong, and A. Sun, "Graph-based point-of-interest recommendation with geographical and temporal influences," in *Proc. of Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pp. 659-668: ACM, 2014. [Article \(CrossRef Link\)](#)
- [23] M. G. Ozsoy, F. Polat, and R. Alhajj, "Making recommendations by integrating information from multiple social networks," *Applied Intelligence*, pp. 1-19, 2016. [Article \(CrossRef Link\)](#)
- [24] R. Linn Jr and M. Uyar, "Testing probabilistic and nondeterministic processes," in *Proc. of Protocol Specification, Testing and Verification, XII: Proceedings of the IFIP TC6/WG6. 1. Twelfth International Symposium on Protocol Specification, Testing and Verification, Lake Buena Vista, Florida, USA, 22-25 June, 1992*, vol. 8, p. 47: Elsevier, 2016. [Article \(CrossRef Link\)](#)
- [25] R. Fu, B. Qin, and T. Liu, "Open-categorical text classification based on multi-LDA models," *soft computing*, vol. 19, no. 1, pp. 29-38, 2015. [Article \(CrossRef Link\)](#)
- [26] P. Cremonesi, Y. Koren, and R. Turrin, "Performance of recommender algorithms on top-n recommendation tasks," in *Proc. of the Proceedings of the fourth ACM conference on Recommender systems*, Barcelona, Spain, 2010. [Article \(CrossRef Link\)](#)
- [27] P. Cremonesi, Y. Koren, and R. Turrin, "Performance of recommender algorithms on top-n recommendation tasks," in *Proc. of ACM Conference on Recommender Systems, Recsys 2010, Barcelona, Spain, September*, pp. 39-46, 2010. [Article \(CrossRef Link\)](#)
- [28] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proc. of the Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, Redondo Beach, California, 2012. [Article \(CrossRef Link\)](#)



Lianggui Liu received the PhD degree in Communications and information system from Nanjing University of Posts & Telecommunications, Nanjing, China. He is now an associate professor with the School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou, China. He has been a Visiting Associate Professor with Cornell University, Ithaca, NY, USA since 2011. His current research interests focus on social networks, network communications, network security and natural computation.



Wei Li received the Bachelor degree from Zhongyuan University of Technology. He is now a postgraduate student with Zhejiang Sci-Tech University. His research focuses on social networks.



Lingmin Wang is now a postgraduate student with Zhejiang Sci-Tech University. Her current research interests include social networks and complex networks.



Huiling Jia received the PhD degree in Communications and Information System from Zhejiang University. She is now a lecturer with the School of Information, Zhejiang Sci-Tech University, Hangzhou, China. Her current research interests include online social networks, wireless sensor networks, heterogeneous wireless networks, and radio resource management.