Where You Like to Go Next: Successive Point-of-Interest Recommendation

Chen Cheng[1,2], **Haiqin Yang**[1,2], **Michael R. Lyu**[1,2], **Irwin King**[1,2]
[1] Shenzhen Research Institute, The Chinese University of Hong Kong, Shenzhen, China
[2] Department of Computer Science and Engineering
The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
{ccheng,hqyang,king,lyu}@cse.cuhk.edu.hk

Abstract

Personalized point-of-interest (POI) recommendation is a significant task in location-based social networks (LBSNs) as it can help provide better user experience as well as enable third-party services, e.g., launching advertisements. To provide a good recommendation, various research has been conducted in the literature. However, pervious efforts mainly consider the "check-ins" in a whole and omit their temporal relation. They can only recommend POI globally and cannot know where a user would like to go tomorrow or in the next few days. In this paper, we consider the task of successive personalized POI recommendation in LBSNs, which is a much harder task than standard personalized POI recommendation or prediction. To solve this task, we observe two prominent properties in the check-in sequence: personalized Markov chain and region localization. Hence, we propose a novel matrix factorization method, namely FPMC-LR, to embed the personalized Markov chains and the localized regions. Our proposed FPMC-LR not only exploits the personalized Markov chain in the check-in sequence, but also takes into account users' movement constraint, i.e., moving around a localized region. More importantly, utilizing the information of localized regions, we not only reduce the computation cost largely, but also discard the noisy information to boost recommendation. Results on two real-world LBSNs datasets demonstrate the merits of our proposed FPMC-LR.

1 Introduction

Check-in behavior becomes a new life style of millions of users who share their locations, tips, and experience about point-of-interests (POIs) with their friends in location-based social networks (LBSNs). For example, in Foursquare alone, it was reported that there are over 20 million register users corresponding to 2 billion check-ins by April, 2012[1]. The online "check-ins" embeds abundant information of users' physical movements in daily life, users' connections to others as well as their preference on the POIs. Among various tasks in LBSNs, personalized POI recommendation is especially important since it is beneficial for users to know new POIs and explore their city while for advertisers to launch advertisements to targeted users.

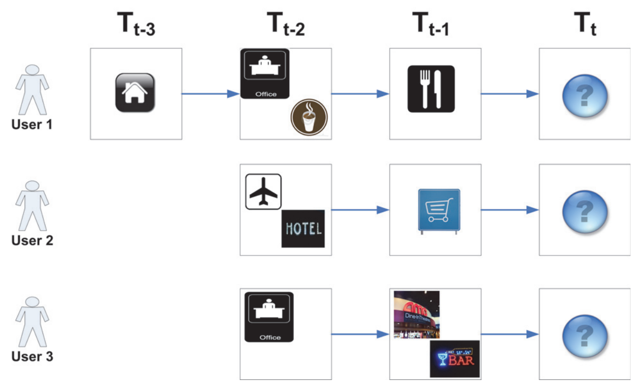
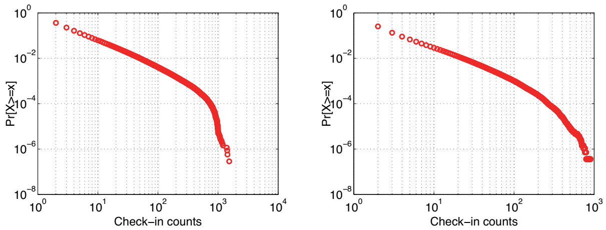


Figure 1: An example of three users' check-in sequences

Recently, POI recommendation in LBSNs has attracted much attention in both research and industry [Ye *et al.*, 2011; Sang *et al.*, 2012]. Collaborative filtering (CF) is a mainstream of algorithms to solve this task. Both memory-based and model-based CF methods have been proposed and investigated to learn users' preferences on the POIs from the user-location check-in data [Cheng *et al.*, 2012; Ye *et al.*, 2011]. However, previously proposed methods consider all check-ins in a whole and their temporal relation is usually overlooked. As the statistics shown in Fig. 2 that apart from a few routinely visiting POIs such as office and home, most POIs are visited less than 10 times, which account for 90% of total visited POIs. It indicates that most POIs are visited occasionally and they are related to users' current location. Hence, POI recommendation is very time-critical. A good POI recommender should provide good recommendation promptly based on users' current status.

Hence, in this paper, different from previous work, we consider the task of successive personalized POI recommendation in LBSNs. This task is much harder than standard personalized POI recommendation or prediction because it only

[1] http://statspotting.com/2012/04/foursquare-statistics-20-million-users-2-billion-check-ins/



(a) Foursquare (b) Gowalla

Figure 2: Check-ins probability vs. counts

recommends those locations that a user does not visit frequently or before, but he/she may like to visit it at the successive time stamp. However, this task is more significant since it can provide various personalized favorite services in LBSNs. For example, it may tell a user where to have fun after dinner, or suggest the discount information of some products in near shops when he/she is shopping, see Fig. 1 for an example. Although this task is very difficult, we believe that the collaborative information shared in users' check-in history can be further utilized to boost the recommendation. Figure. 1 also gives an intuitive example. User 3 visited a cinema and then a bar after work. It may also good to suggest user 1 to go there after the dinner. The significance of successive personalized POI recommendation in LBSNs and the promising of utilizing the collaborative information trigger our deeper study in the check-in data.

There are two main properties, personalized Markov chain and localized region constraint in the LBSN datasets, see Sec. 3 for more details. Based on these two observations, we proposed a novel matrix factorization method, namely, FPMC-LR, to include the information of the personalized Markov chain and the localized region constraint. Although our FPMC-LR borrows the idea of factoring personalized Markov chain (FPMC) for solving the task of next-basket recommendation [Rendle *et al.*, 2010], we emphasize on users' movement constraint, i.e., moving around a local region, and focus on a different problem. More specifically, we only consider the locations around users' previous check-in history which yields a much smaller set, accounting for about 0.7% and 0.3% of the set on all locations for Foursquare and Gowalla, respectively. More importantly, we not only reduce the computation cost largely, but also discard possible noisy information. We summarize our contributions in the following:

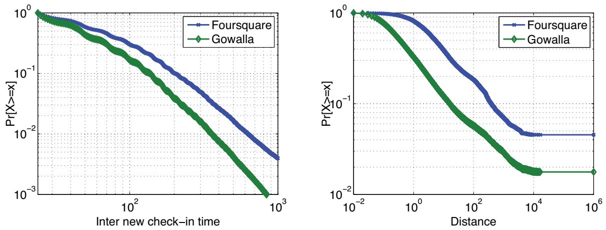
- We formally define the problem of successive personalized POI recommendation in LBSNs and analyze the spatial-temporal properties in two large-scale real-world LBSN datasets, Foursquare and Gowalla. After analyzing the dynamics of new POIs and inter check-ins, we observe two important properties: personalized Markov chain and localized region constraint.
- We propose a novel matrix factorization method, namely FPMC-LR, to incorporate these two properties. More importantly, we not only reduce the computation cost largely, but also discard noisy information.
- We conduct detailed experimental evaluation on the analyzed large-scale LBSN datasets and show that our model consistently outperforms other state-of-the-art methods.

2 Related Work

Location-based social networks have received much attention in recent years due to the new characteristics of spatial-temporal-social information embedded in the check-in data and the prevalence of various interesting real-world applications [Yuan *et al.*, 2012; Zheng *et al.*, 2011; Cheng *et al.*, 2011a]. Research topics covered in this area include user behavior study, movement pattern analysis, community detection, POI recommendation, and etc. [Cho *et al.*, 2011; Zheng and Zhou, 2012]. Among all of these topics, POI recommendation is one of the most important topics due to the high value in both research and academy.

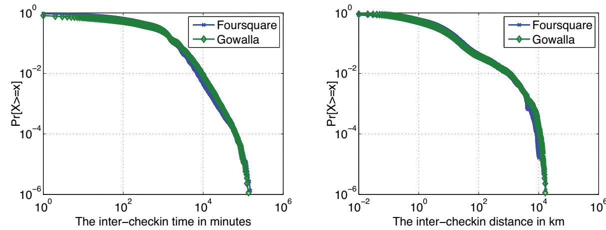
Currently, there are two line of work to solve the task of POI recommendation. One line of research is conducted based on the GPS trajectory logs [Zheng *et al.*, 2009; 2010a; 2010b; Leung *et al.*, 2011]. The GPS trajectory data usually consist of small number of users, but dense records [Zheng and Xie, 2011; Cao *et al.*, 2010]. Many collaborative filtering algorithms, e.g., collective matrix factorization [Zheng *et al.*, 2010a], tensor factorization [Zheng *et al.*, 2010b], memory-based collaborative location model (CLM) [Leung *et al.*, 2011], etc., have been proposed to solve it and deemed the locations as items in traditional recommender systems. The other line of work focuses on LBSN data, which is very sparse and large-scale [Ye *et al.*, 2010; 2011; Cheng *et al.*, 2012]. Currently, geographical influence, e.g., modeling the check-in probability to the distance of the whole check-in history by power-law distribution [Ye *et al.*, 2011], modeling users' multi-center check-in behaviors via multi-center Gaussians [Cheng *et al.*, 2012], and etc., have been addressed and fused with traditional CF algorithms.

However, the temporal relation between successive check-ins is not well-studied yet. Although some work consider the temporal information, they mainly focus on POI prediction or next place prediction on the existing locations or users' historical trajectories [Gao *et al.*, 2012a; Sadilek *et al.*, 2012; Sadilek and Krumm, 2012; Sang *et al.*, 2012; Gao *et al.*, 2012b]. On the other hand, in traditional recommender systems, temporal patterns can be utilized to boost recommendation performance [Koren, 2009]. Some other factorization methods, e.g., Bayesian Probabilistic Tensor Factorization (BPTF) [Xiong *et al.*, 2010], factorized personalized Markov chains (FPMC) [Rendle *et al.*, 2010], and etc., have been proposed and demonstrated themselves as promising methods in capturing time-evolving relational data or next-basket recommendation. The significance of the successive check-in recommendation in LBSNs and the promising solution motivate us to further investigation in this paper.



(a) The inter new check-in time in hours (b) The minimum distance of new POI to check-in history

Figure 3: The New POI Dynamics



(a) The inter check-in time in minutes (b) The inter check-in distance in km

Figure 4: The inter check-in time in minutes.

3 Successive POI Recommendation in LBSNs

3.1 Problem of Successive Personalized POI Recommendation

Let $\mathcal{U}$ be the set of users and $\mathcal{L}$ be the set of locations. $\mathcal{L}_u$ denotes the check-in history of user u. Due to the low density of the LBSNs data, we merge consecutive check-ins in T hours (in the experiment, T is set to 6 as a reasonable selection) as a slide window to construct a set of check-ins. Hence, we construct a slide window set $\mathcal{T}$ to denote the users' visiting time stamp. The check-in set of user u at time t is denoted by $\mathcal{L}_u^t$, where $t \in \mathcal{T}$. Given a sequence of check-ins, $\mathcal{L}_u^1, \ldots, \mathcal{L}_u^t$, the latitude and longitude of each location, the problem of successive personalized POI recommendation is to provide the most suitable recommendation for user u at time $t+1$.

3.2 New POIs Dynamics

New POIs are locations that a user does not visit before and will be recommended in the next time stamp. The inter check-in time and location distance on new POIs are defined as the temporal interval and distance between a new POI following the previous check-in POIs, respectively. Figure 3 shows the properties of new POIs dynamics on the time and location distance. Figure 3(a) reports how often a user would like to explore new POIs by calculating the Complementary Cumulative Distribution Function (CCDF) on the inter check-in time on new POIs. It shows that almost 70% of Foursquare users and 80% of Gowalla users would like to check-in a new POI after about 100 hours and the ratio raises up to 90% for Foursquare and 95% for Gowalla, respectively, after 200 hours. It is noted that although users would like to explore new POIs, as shown in Fig. 2, most of their check-ins are distributed among a few frequently visited places, e.g., home and office.

Figure 3(b) shows the spatial property of a new POI versus previously successive visited POIs. Obviously, users' exploration on new POIs is restricted by the geographical influence. More specifically, about 60% of Foursquare new POIs and about 88% of Gowalla are within 10 km of users' previous check-in locations while when the distance increases to 100 km, the number of new POIs accounts to about 80% for Foursquare and about 95% for Gowalla, respectively. This observation implies that users in Foursquare prefer to explore farther new POIs than Gowalla users.

3.3 Inter Check-in Dynamics

The property of inter check-in dynamics is also another key factor revealing the temporal relation of the LBSN data. We obtain similar results in [Noulas *et al.*, 2011] and observe two significant properties on the LBSN data: personalized Markov chain and localized region constraint.

Figure 4(a) shows that almost 40% and 48% successive check-ins occur in Foursquare and Gowalla, respectively, within two hours. The ratio raise to about 70% for both Foursquare and Gowalla when the inter check-in time is larger than twelve hours. After further studying the categories of two successive check-ins for a user in a short period, we find that there is strong connection between them. For example, cinemas or bars may be always visited after restaurants as users would like to relax after dinner. This is exactly a personalized Markov chain property and will motivate us to utilize the transition probability for solving the task of successive personalized POI recommendation.

Figure 4(b) shows the CCDF of inter check-in distance. It is observed that more than 75% of inter check-ins in Foursquare and more than 80% of inter check-ins in Gowalla occur within 10 km, respectively. Only less than 5% inter check-in distance is more than 100 km in both datasets. This observation is reasonable since most users' inter check-ins occur within a specific area they live or the long distance inter check-ins imply an occasional journey. Overall, users' movement is constrained by their geographical influence within a short time. Hence, when we provide successive personalized POI recommendation, we mainly consider the new POI near to a user's previous check-ins.

4 FPMC with Localized Region Constraint

4.1 Model

Our FPMC-LR is to recommend a successive personalized POI via the probability that user u will visit location l at time t, which is calculated by

$$x_{u,i,l} = p(l \in \mathcal{L}_u^t | i \in \mathcal{L}_u^{t-1}). \quad (1)$$

Based on only the first-order Markov chain property, the probability can be calculated by

$$p(l \in \mathcal{L}_u^t | \mathcal{L}_u^{t-1}) = \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} p(l \in \mathcal{L}_u^t | i \in \mathcal{L}_u^{t-1}), \quad (2)$$

where $p(l \in \mathcal{L}_u^t | i \in \mathcal{L}_u^{t-1})$ is the probability for user u moves from location i to location l.

In FPMC, all locations are considered for each user and yield a transition tensor $\mathcal{X} \in [0,1]^{|\mathcal{U}|\times|\mathcal{L}|\times|\mathcal{L}|}$. Differently, our FPMC-LR considers only the neighborhood locations. More specifically, we divide the whole earth into different square grids whose side length is d km. Then, for each location l, its neighbor locations will be those fall in one of the nine adjacent square grids:

$$N_d(\mathcal{L}_u^t) = \{l \in \mathcal{L} \setminus \mathcal{L}_u^{t-1} : D(l, l_0) \le d, \forall l_0 \in \mathcal{L}_u^{t-1}\},$$

where $D(l, l_0)$ is the distance between l and l_0 calculate by Haversine formula.

Let $N(\mathcal{L}_u^t)$ be the neighbor location set of the check-in history of user u at time t. Our FPMC-LR yields a transition tensor $\mathcal{X} \in [0,1]^{|\mathcal{U}|\times|\mathcal{L}|\times|N_d(\mathcal{L})|}$. It is noted that $|N_d(\mathcal{L})|$ is reduced largely, e.g., around hundred when $d = 40$, which accounts for less than 0.7% and 0.3% of the total locations in Foursquare and Gowalla, respectively. Hence, our FPMC-LR can save the time cost largely compared with FPMC. Since FPMC [Rendle *et al.*, 2010] provides a good framework for successive personalized POI recommendation, we adopt it in the paper, but focuses on the localized region constraint, which motivates the name of our model.

Low-rank approximation is a promising tool to recover the partially observed transition tensor $\mathcal{X}$ when it is sparse. Here, we adopt a special case of Canonical Decomposition which models the pairwise interaction between of the three modes of the tensor (i.e. user $\mathcal{U}$, last location $\mathcal{I}$ and next location $\mathcal{L}$):

$$\hat{x}_{u,i,l} = \boldsymbol{v}_u^{\mathcal{U},\mathcal{L}} \cdot \boldsymbol{v}_l^{\mathcal{L},\mathcal{U}} + \boldsymbol{v}_l^{\mathcal{L},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{L}} + \boldsymbol{v}_u^{\mathcal{U},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{U}}, \qquad (3)$$

where $\boldsymbol{v}_u^{\mathcal{U},\mathcal{L}}$ and $\boldsymbol{v}_l^{\mathcal{L},\mathcal{U}}$ model the latent features for users and the next locations, respectively. Other notations are similar defined. This gives the set of model parameters, i.e., $\Theta = \{\boldsymbol{V}^{U,L}, \boldsymbol{V}^{L,U}, \boldsymbol{V}^{U,I}, \boldsymbol{V}^{I,U}, \boldsymbol{V}^{L,I}, \boldsymbol{V}^{I,L}\}$.

Combining Eq. (2) and Eq. (3), we obtain

$$\begin{aligned}
&\hat{p}(l \in \mathcal{L}_u^t | \mathcal{L}_u^{t-1}) = \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} \hat{x}_{u,i,l} \\
&= \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} (\boldsymbol{v}_u^{\mathcal{U},\mathcal{L}} \cdot \boldsymbol{v}_l^{\mathcal{L},\mathcal{U}} + \boldsymbol{v}_l^{\mathcal{L},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{L}} + \boldsymbol{v}_u^{\mathcal{U},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{U}}) \\
&= \boldsymbol{v}_u^{\mathcal{U},\mathcal{L}} \cdot \boldsymbol{v}_l^{\mathcal{L},\mathcal{U}} + \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} (\boldsymbol{v}_l^{\mathcal{L},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{L}} + \boldsymbol{v}_u^{\mathcal{U},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{U}}).
\end{aligned}$$

Notice that the last step holds as the interaction $\mathcal{U}$ and $\mathcal{L}$ are independent of last location i.

Our goal of successive personalized POI recommendation is to recommend top-k new POIs to users, thus we can model it as a ranking $>_{u,t}$ over locations:

$$i >_{u,t} j :\Leftrightarrow \hat{x}_{u,t,i} > \hat{x}_{u,t,j}. \qquad (4)$$

A sequential BPR optimization criterion can be derived similarly to the general BPR approach [Rendle *et al.*, 2009]. Then for user u at time t, the best ranking can be modeled as:

$$p(\Theta| >_{u,t}) \propto p(>_{u,t} |\Theta) p(\Theta). \qquad (5)$$

Assuming that users and their check-in history are independent, then we can estimate the model using maximum a posterior (MAP):

$$\arg\max_{\Theta} \prod_{u \in \mathcal{U}} \prod_{\mathcal{L}_u^t \in \mathcal{L}_u} \prod_{i \in \mathcal{L}_u^t} \prod_{j \in N_d(\mathcal{L}_u^t)} p(>_{u,t} |\Theta) p(\Theta). \qquad (6)$$

The ranking probability can be further expressed by

$$\begin{aligned}
p(>_{u,t}) = p(i >_{u,t} j) &= p(\hat{x}_{u,t,i} > \hat{x}_{u,t,j} \mid \Theta) \\
&= p(\hat{x}_{u,t,i} - \hat{x}_{u,t,j} > 0 \mid \Theta)
\end{aligned} \qquad (7)$$

Using the logistic function σ defined by $p(z > 0) = \sigma(z) = \frac{1}{1+e^{-z}}$, we can repressed Eq. (7) as

$$p(i >_{u,t} |\Theta) = \sigma(\hat{x}_{u,t,i} - \hat{x}_{u,t,j}) \qquad (8)$$

Furthermore, by placing Gaussian priors on the model parameters $\Theta \sim \mathcal{N}(0, \frac{1}{\lambda_\Theta})$, we can seek the optimal solution of our FPMC-LR by

$$\begin{aligned}
&\arg\max_{\Theta} \ln p(>_{u,t} |\Theta) p(\Theta) \\
&= \arg\max_{\Theta} \ln \prod_{u \in \mathcal{U}} \prod_{\mathcal{L}_u^t \in \mathcal{L}_u} \prod_{i \in \mathcal{L}_u^t} \prod_{j \in N(\mathcal{L}_u^{t-1})} \sigma(\hat{x}_{u,t,i} - \hat{x}_{u,t,j}) p(\Theta) \\
&= \arg\max_{\Theta} \sum_{u \in \mathcal{U}} \sum_{\mathcal{L}_u^t \in \mathcal{L}_u} \sum_{i \in \mathcal{L}_u^t} \sum_{j \in N(\mathcal{L}_u^{t-1})} \ln \sigma(\hat{x}_{u,t,i} - \hat{x}_{u,t,j}) \\
&\quad - \lambda_\Theta \|\Theta\|_F^2
\end{aligned} \qquad (9)$$

To recommend a new location, we rank them based on the probability of $\hat{x}_{u,t,l}$:

$$\hat{x}_{u,t,l} = \boldsymbol{v}_u^{\mathcal{U},\mathcal{L}} \cdot \boldsymbol{v}_l^{\mathcal{L},\mathcal{U}} + \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} (\boldsymbol{v}_l^{\mathcal{L},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{L}} + \boldsymbol{v}_u^{\mathcal{U},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{U}}) \qquad (10)$$

As shown in [Rendle *et al.*, 2010], the term $\boldsymbol{V}^{U,I} \cdot \boldsymbol{V}^{I,U}$ will vanish since it does not affect the final ranking. This yields a more compact expression for $\hat{x}_{u,t,l}$:

$$\hat{x}_{u,t,l} = \boldsymbol{v}_u^{\mathcal{U},\mathcal{L}} \cdot \boldsymbol{v}_l^{\mathcal{L},\mathcal{U}} + \frac{1}{|\mathcal{L}_u^{t-1}|} \sum_{i \in \mathcal{L}_u^{t-1}} \boldsymbol{v}_l^{\mathcal{L},\mathcal{I}} \cdot \boldsymbol{v}_i^{\mathcal{I},\mathcal{L}} \qquad (11)$$

4.2 Learning Algorithm

Directly optimize the objective function in Eq. (9) is very time consuming. Even though we only consider neighbor location pairs, the number of quadruples is still huge, i.e. $\mathcal{O}(|S||\bar{N}|)$, where $S = \{(u,t,i) | u \in \mathcal{N}, t \in \mathcal{T}, i \in \mathcal{L}_u^t, \mathcal{L}_u^t \in \mathcal{L}_u\}$ and $|\bar{N}|$ is the average number of neighbor locations. We follow the strategy used in [Rendle *et al.*, 2009; 2010] to draw the quadruples independently and apply the stochastic gradient descent on the bootstrap samples. The detailed algorithm is shown in Algorithm 1.

For each parameter θ, the update procedure is performed as:

$$\theta = \theta + \alpha\left(\frac{\partial}{\partial\theta}(\ln \sigma(\hat{x}_{u,t,i} - \hat{x}_{u,t,j}) - \lambda_\theta \theta^2)\right), \qquad (12)$$

where α is the step size.

Algorithm 1 Learning Algorithm for FPMC-LR

draw $\boldsymbol{V}^{U,I}, \boldsymbol{V}^{I,U}, \boldsymbol{V}^{I,L}, \boldsymbol{V}_{L,I}$ from $\mathcal{N}(0, \sigma^2)$
repeat
  draw (u, t, i) uniformly from S
  draw location j uniformly from $N(\mathcal{L}_u^{t-1}) \backslash \mathcal{L}_u^t$
  for $f = 1 \to k_{U,I}$ **do**
    update $v_{u,f}^{U,I}, v_{i,f}^{I,U}, v_{j,f}^{I,U}$
  end for
  for $f = 1 \to k_{I,L}$ **do**
    update $v_{i,f}^{I,L}, v_{j,f}^{I,L}$
    for $l \in \mathcal{L}_u^{t-1}$ **do**
      update $v_{l,f}^{L,I}$
    end for
  end for
until convergence
return $\boldsymbol{V}^{U,I}, \boldsymbol{V}^{I,U}, \boldsymbol{V}^{I,L}, \boldsymbol{V}_{L,I}$

5 Experiments

In the experiments, we address the following questions: 1) How does our approach compare with the baseline model and other state-of-the-art methods? 2) How does the parameter of the side length d, which determines the neighbor locations, affect the model performance? 3) What is the convergence and efficiency property of our FPMC-LR?

5.1 Datasets

We evaluate the models on the two publicly available location-based online social networks: Foursquare and Gowalla[2]. Gowalla provides public APIs which allows us to crawl all users' information including all check-in history with the time stamp and location details. Although it is not possible to directly crawl Fourquare data using their APIs, part of Foursquare users link their accounts with Twitter and their check-in information can be crawled from Twitter. In this paper, we use the Foursquare dataset provided by [Cheng *et al.*, 2011b] and the Gowalla data from [Cheng *et al.*, 2012]. For both datasets, we use four month check-in history from May 2010 to August 2010. To remove outliers and clean up the data, we require that every user should have check-in at least 120 times and each location should be visited at least 5 times. The basic statistics are summarized in Table 1.

Table 1: Basic statistics of Foursquare and Gowalla dataset.

	$\#U$	$\#L$	# check-in	# avg. check-in
Foursquare	3571	28754	744055	208.36
Gowalla	4510	59355	873071	193.58

5.2 Evaluation Metrics

The experiment is tested as follows: check-ins in the last time slot is used as the test data while the previous check-in history is used as training data. Since recommending infrequently visited POIs is more meaningful, we only keep POIs which are visited less than 5 times by the user before the test period and remove them from the training set. Note that, this setting makes it much harder to recommend new POIs to a user than recommending POIs he has visited before. This can also explain why we can only get very low precision and recall values in the results. In our experiment, we use Precision@N and Recall@N to evaluate the performance:

$$\mathrm{P}@N := \frac{|S|}{N}, \ \mathrm{R}@N := \frac{|S|}{|\mathcal{L}_u^{t+1}|}, \tag{13}$$

where $|S|$ is the number of top-N recommended POIs that visited by user u at last time $t+1$. In the reported results, we set N to 10.

5.3 Comparison

In this section, we compare our method with the following state-of-the-art methods:

1. **PMF**: probabilistic matrix factorization is a well-known method in matrix factorization [Salakhutdinov and Mnih, 2007]. It is widely used in traditional recommender systems.
2. **PTF**: probability tensor factorization is introduced in [Xiong *et al.*, 2010] for modeling time evolving relation data. Due to efficiency consideration, we do not use its nonparametric version, Bayesian PTF.
3. **FPMC**: this method is proposed in [Rendle *et al.*, 2010], which is a strong baseline model embedding users' preference and their personalized Markov chain to provide next-basket item recommendation.

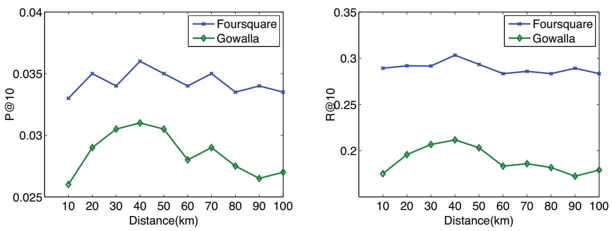
The experiment results on Foursquare and Gowalla datasets are shown in Table 2. We set the number of latent dimensions to 100 for all the compared models. For our FPMC-LR model, we set the time window size to be 6 hours and the side length d to be 40 km. We set λ_θ to be 0.03 through setting the last visitings in the training as validation set. The results show that:

- Both FPMC and FPMC-LR outperforms PMF and PTF significantly. More specifically, FPMC-LR improves PMF and PTF over at least 90% and 110%, respectively, while FPMC also beats PMF and PTF over 50%, and 60%, respectively. This implies that personalized Markov chain plays an important role when performing successive personalized POI recommendation. The location transition in short time provides valueable information on where the user would like to go in the next.
- It is a little surprising that PMF performs a little better that PTF. One possible reason may lie that PTF assumes the latent features in successive time periods are similar. However, this assumption is not always valid for LBSNs data since the features may be periodic. For example, most users have similar preference patterns on every morning or every Sunday. The poor results of PTF imply the assumption of PTF does not fit for LBSNs data.
- FPMC-LR performs much better than FPMC, improving around 30% and 40% over FPMC for precision and recall, respectively. This verifies that restricting the

[2] It has been acquired by Facebook

Table 2: Performance comparison

Metrics	Foursquare				Gowalla			
	PMF	PTF	FPMC	FPMC-LR	PMF	PTF	FPMC	FPMC-LR
P@10	0.0185	0.0170	0.0275	**0.0360**	0.0130	0.0110	0.0220	**0.0310**
Improve	94.59%	111.76%	30.91%		138.46%	181.82%	40.91%	
R@10	0.1542	0.1417	0.2325	**0.3033**	0.1040	0.0785	0.1575	**0.2116**
Improve	96.69%	114.04%	30.45%		103.46%	169.55%	34.35%	



(a) $P@10$ (b) $R@10$

Figure 5: Impact of parameter d.

comparing set to a localized region can reduce noisy information and achieve better performance compared to considering all the locations. As users' movement is constrained locally in short time, it is enough to only consider that the rank-pairs of current check-in and nearby previously visited locations.

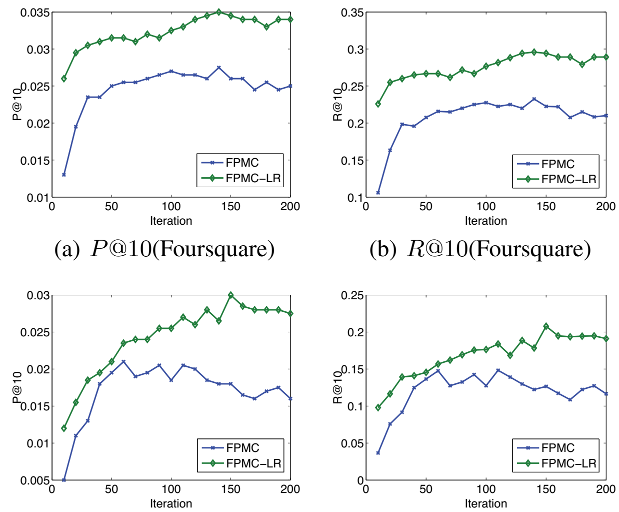
5.4 Impact of Parameter d

In FPMC-LR, the parameter d is a important factor to control the size of neighborhood check-ins history of a user at time t. This parameter affects the number of locations as well as the model performance. Figure 5 show the impact of d of both Foursquare and Gowalla datasets on $P@10$ and $R@10$. From the figure, we can see that both Foursquare and Gowalla, the model performs best when d is 40km. When d is small, we only consider a very small set of nearby locations which do not include enough information and yield suboptimal performance. While when d is large, e.g., 100 km, the model has to consider much more rank-pairs and may introduce more noisy information which yields poor performance. An extreme case is set d large enough to cover all neighbor areas in the whole earth and consider all locations, it is equivalent to the case of FPMC model. The obtained results confirm the intuition that localization constraint plays an important role in successive personalized POI recommendation.

5.5 Convergence and Efficiency Analysis

Figure 6 shows the performance change of our FPMC-LR and FPMC with respect to the number of iterations. Here, at each iteration, we draw 2×10^5 quadruples to calculate the stochastic gradient descent based on the BPR criterion. From the figures, we can see that at each iteration, FPMC-LR always performs better than FPMC and attains its best performance at around 150 iterations.

Our experiments are conducted on a PC with an Intel Pentium D CPU (3.0 GHz, Dual Core) and 2G memory. An average time for an iteration is about 30 seconds. Here, we also claim another advantage of PFMC-LR, i.e., its efficiency, in the recommendation procedure. This because that the candidate location set of PFMC-LR is only the neighbor locations of previous check-ins, whose size is much smaller than that of the whole location set used in PFMC. Hence, PFMC-LR can save much time cost in recommending a location than PFMC.



(a) $P@10$(Foursquare) (b) $R@10$(Foursquare)

(c) $P@10$(Gowalla) (d) $R@10$(Gowalla)

Figure 6: Convergence analysis.

6 Conclusions and Future Work

In this paper, we consider the task of successive personalized Point-of-Interest recommendation in LBSNs. We first investigate the spatial-temporal properties of the LBSN datasets. We then propose a novel matrix factorization model, namely FPMC-LR, to include both personalized Markov chain and localized regions for solving the recommendation task. Our experimental results on two large-scale LBSN datasets, Foursquare and Gowalla, show the effectiveness and efficiency of our model compared to several state-of-the-art methods.

There are still serval other aspects worthy of consideration in the future: 1) how can we utilize the contextual information of POIs, e.g., the location category and the activities conducted there; 2) how to incorporate the users' periodic check-in behaviors to capture users' periodic preference; 3) how to find more useful check-in sequences, e.g., higher-order Markov chain. 4) How to incorporate social information to strengthen successive personalized POI recommendation.

7 Acknowledgments

The work described in this paper was fully supported by the Basic Research Program of Shenzhen (Project No. JCYJ20120619152419087 and JC201104220300A), and the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK 413212 and CUHK 415212).

References

[Cao *et al.*, 2010] X. Cao, G. Cong, and C.S. Jensen. Mining significant semantic locations from gps data. *Proceedings of the VLDB Endowment*, 3(1-2):1009–1020, 2010.

[Cheng *et al.*, 2011a] Zhiyuan Cheng, James Caverlee, Krishna Yeswanth Kamath, and Kyumin Lee. Toward traffic-driven location-based web search. In *CIKM*, pages 805–814, 2011.

[Cheng *et al.*, 2011b] Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z. Sui. Exploring millions of footprints in location sharing services. In *ICWSM*, 2011.

[Cheng *et al.*, 2012] C. Cheng, H. Yang, I. King, and M.R. Lyu. Fused matrix factorization with geographical and social influence in location-based social networks. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[Cho *et al.*, 2011] Eunjoon Cho, Seth A. Myers, and Jure Leskovec. Friendship and mobility: user movement in location-based social networks. In *KDD*, pages 1082–1090, 2011.

[Gao *et al.*, 2012a] Huiji Gao, Jiliang Tang, and Huan Liu. Exploring social-historical ties on location-based social networks. In *ICWSM*, 2012.

[Gao *et al.*, 2012b] Huiji Gao, Jiliang Tang, and Huan Liu. gscorr: modeling geo-social correlations for new check-ins on location-based social networks. In *CIKM*, pages 1582–1586, 2012.

[Koren, 2009] Yehuda Koren. Collaborative filtering with temporal dynamics. In *KDD*, pages 447–456, 2009.

[Leung *et al.*, 2011] Kenneth Wai-Ting Leung, Dik Lun Lee, and Wang-Chien Lee. Clr: a collaborative location recommendation framework based on co-clustering. In *SIGIR*, pages 305–314, 2011.

[Noulas *et al.*, 2011] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil. An empirical study of geographic user activity patterns in foursquare. *ICWSM?11*, 2011.

[Rendle *et al.*, 2009] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *UAI*, pages 452–461, 2009.

[Rendle *et al.*, 2010] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In *WWW*, pages 811–820, 2010.

[Sadilek and Krumm, 2012] A. Sadilek and J. Krumm. Far out: Predicting long-term human mobility. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*, 2012.

[Sadilek *et al.*, 2012] A. Sadilek, H. Kautz, and J.P. Bigham. Finding your friends and following them to where you are. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 723–732. ACM, 2012.

[Salakhutdinov and Mnih, 2007] Ruslan Salakhutdinov and Andriy Mnih. Probabilistic matrix factorization. In *NIPS*, 2007.

[Sang *et al.*, 2012] Jitao Sang, Tao Mei, Jian-Tao Sun, Changsheng Xu, and Shipeng Li. Probabilistic sequential pois recommendation via check-in data. In *SIGSPATIAL/GIS*, pages 402–405, 2012.

[Xiong *et al.*, 2010] L. Xiong, X. Chen, T.K. Huang, J. Schneider, and J.G. Carbonell. Temporal collaborative filtering with bayesian probabilistic tensor factorization. In *Proceedings of SIAM Data Mining*, volume 2010, 2010.

[Ye *et al.*, 2010] M. Ye, P. Yin, and W.C. Lee. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 458–461. ACM, 2010.

[Ye *et al.*, 2011] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik Lun Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *SIGIR*, pages 325–334, 2011.

[Yuan *et al.*, 2012] Jing Yuan, Yu Zheng, and Xing Xie. Discovering regions of different functions in a city using human mobility and pois. In *KDD*, pages 186–194, 2012.

[Zheng and Xie, 2011] Yu Zheng and Xing Xie. Learning travel recommendations from user-generated gps traces. *ACM TIST*, 2(1):2, 2011.

[Zheng and Zhou, 2012] Yu Zheng and Xiaofang Zhou, editors. *Computing with Spatial Trajectories*. Springer, 2012.

[Zheng *et al.*, 2009] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW*, pages 791–800, 2009.

[Zheng *et al.*, 2010a] Vincent Wenchen Zheng, Yu Zheng, Xing Xie, and Qiang Yang. Collaborative location and activity recommendations with gps history data. In *WWW*, pages 1029–1038, 2010.

[Zheng *et al.*, 2010b] V.W. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang. Collaborative filtering meets mobile recommendation: A user-centered approach. In *Proceedings of the 24rd AAAI Conference on Artificial Intelligence*, 2010.

[Zheng *et al.*, 2011] Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei-Ying Ma. Recommending friends and locations based on individual location history. *TWEB*, 5(1):5, 2011.