Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Inferring a Personalized Next Point-of-Interest Recommendation Model with Latent Behavior Patterns

Jing He[1], Xin Li[1], Lejian Liao[1], Dandan Song[1], William K.Cheung[2]
[1]BJ ER Center of HVLIP&CC, School of Comp. Sci., Beijing Institute of Technology, Beijing, China
[2]Department of Computer Science, Hong Kong Baptist University , Hong Kong, China
{skyhejing, xinli, liaolj, sdd}@bit.edu.cn, william@comp.hkbu.edu.hk

Abstract

In this paper, we address the problem of personalized next Point-of-interest (POI) recommendation which has become an important and very challenging task in location-based social networks (LBSNs), but not well studied yet. With the conjecture that, under different contextual scenario, human exhibits distinct mobility patterns, we attempt here to jointly model the next POI recommendation under the influence of user's latent behavior pattern. We propose to adopt a third-rank tensor to model the successive check-in behaviors. By incorporating softmax function to fuse the personalized Markov chain with latent pattern, we furnish a Bayesian Personalized Ranking (BPR) approach and derive the optimization criterion accordingly. Expectation Maximization (EM) is then used to estimate the model parameters. Extensive experiments on two large-scale LBSNs datasets demonstrate the significant improvements of our model over several state-of-the-art methods.

Introduction

In recent years, there have been an increased emphasis on developing the location-based social networks(LBSNs), such as Foursquare, Gowalla, Facebook Place, and GeoLife, etc., where users can check-in at venues and share their experiences towards point-of-interest(POIs) in the physical world using mobile devices. This so called check-in behavior has become a new culture of a modern life and could be used to study life patterns of millions of LBSN users. POI recommendation is one of the most important tasks in LBSN, which is to provide recommendations of places to users, and has attracted much attention as it not only improves user viscosity to LBSN service provider but also benefits for advertising agency to provide an effective way of launching advertisement to target the potential clients.

POI recommendation has become a popular research issue and attracted much effort(Gao et al. 2015; Yuan et al. 2014). Yet achieving accurate personalized POI recommendation is challenging as the data available for each user is highly sparse. The sparsity is due to the fact that check-in interactions are conducted by the users on the voluntary basis. “Diligent” users who keep checking-in on LBSN for every venue they visited in physical world are in fact rare. Next POI recommendation is a even harder task which is to be accurate on predicting user's very next move among tens of thousands of location candidates, as the successive check-in interactions are more sparse. In the literature, there exist only few works investigating next POI recommendation. C. Cheng *et al.* propose to introduce the distance restriction to a tensor-based pairwise ranking model to solve the next POI recommendation problem(Cheng et al. 2013). The intuition behind is that the checked-in POIs are usually geographically dense(Liu et al. 2013). However, the periodicity of mobility and the user preference varying with the change of contextual scenario e.g., the type of POIs, checked-in time stamps have not yet been deeply exploited.

Human mobility has been well known for its periodic property (Eagle and Pentland 2009; Li et al. 2010; Cho, Myers, and Leskovec 2011). For next POI recommendation, we focus more on the transition periodicity of location categories. For example, people may regularly stop by coffee stalls, starbucks stores, to grab a cup of coffee on their way to work in the morning, which can be explained as a periodic transition pattern from coffee shop to workplace on weekday morning. The next POI is highly likely related to the previous POI. For example, after taking part in intense outdoor activities, e.g., hiking, running, some user may prefer to have high-protein meals in restaurants like Steakhouse rather than Juicy Bar. Fig.1(a) and Fig.1(b) plot the check-in probabilities of the top-4 most popular location categories over time of day(hours) and day of week respectively, based on the check-in data of LA, collected from Foursquare. The categorical mobility periodicity is very obvious, e.g., the outdoor places are often checked on weekends. Another interesting observation is that the check-ins of nightspots occur most often on Friday and least often on Sunday. Fig.2 plots the transition probabilities between categories along with the day of week. We observe that the transition preference shown in Fig.2(a) is significantly different from that of Fig.2(e) but somehow similar to that of Fig.2(b), which indicates that there exist several latent transition patterns and such patterns may play a key role for our next POI recommendation.

In this paper, we attempt to jointly model next POI recommendation under the influence of user's latent behavior pattern. We propose to adopt a third-rank tensor to model the successive check-in behaviors. By incorporating the soft-

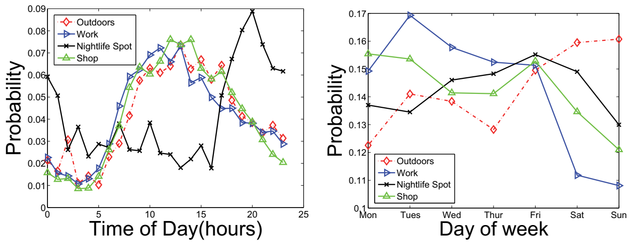


Figure 1: Check-in Periodicity Analysis

max function to fuse the personalized Markov chain with the aforementioned latent pattern's influence, we furnish a Bayesian Personalized Ranking (BPR)(Rendle et al. 2009) approach and derive the optimization criterion accordingly. In the model learning phase, the Expectation Maximization (EM)(Neal and Hinton 1998) is used to estimate the model parameters. The main contributions of this paper can be summarized as follows:

- We propose a unified tensor-based latent model to fuse the observed successive check-in behavior with latent behavior preference for each user to address a personalized next POI recommendation problem. The corresponding optimization criterion and learning steps/tricks have been carefully studied.
- We evaluate the proposed model by detailed experiments on two large-scale LBSN datasets and demonstrate that our method outperforms other state-of-the-art POI recommendation approaches by a large margin.

Related Work

Location recommendation has received intensive attention recently due to a wide range of potential applications. It was studied on GPS trajectory logs of hundreds of monitored users (Zheng et al. 2009). With the easy access of users' check-in data in LBSNs, many recent studies have been conducted for POI recommendation, which can be roughly classified into four categories: 1) ***time-aware POI recommendation*** which mainly leverages the temporal influence on POIs to enhance the recommendation performance (Yuan et al. 2013). 2) ***geographical influence enhanced POI recommendation*** which exploits the "geographical clustering phenomenon" of check-in activities to improve the POI recommendation system (Liu et al. 2013). 3) ***content-aware POI recommendation approaches*** which propose to detect users' current locations by analyzing their published tweets (Chen et al. 2013) or to rank POIs by analyzing user's comments on them (Gao et al. 2015). However semantic analysis is a very challenging research issue as most of comments in LBSN are short and contextually ambiguous. 4) ***social influence enhanced POI recommendation*** which is inspired by the intuition that friends of LBSNs tend to have more common interests. By inferring the social relations, the quality of recommendation could be enhanced. However, there are other opinions of leveraging social influence in the literature, as previous studies also report a large number of friends share nothing in terms of POI (Ye, Yin, and Lee 2010). And E. Cho *et al.* report their findings that the long-distance travel is more influenced by social relations (Cho, Myers, and Leskovec 2011).

Some very recent works have incorporated group behaviors into recommender systems for enhancing performance. T. Yuan *et al.* proposed a Group-Sparse Matrix Factorization (GSMF) approach to factorize the rating matrices for multiple behaviors into a user and item latent factor space (Yuan et al. 2014). H. Wang *et al.* proposed a group-based algorithm for POI recommendation (Henan Wang 2014) by grouping users of similar interests based on their frequently visited locations' category hierarchy.

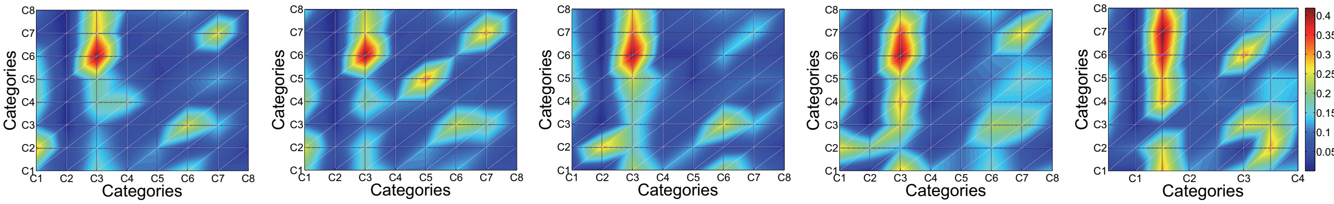
The next POI recommendation is a newly emerging task and even challenging. In the literature, there exist only few works in which the sequential influence between successive check-ins is not yet well-studied. S. Feng *et al.* proposed a personalized ranking metric embedding method (PRME) to model personalized check-in sequences for next new POI recommendation (Feng et al. 2015). C. Cheng *et al.* proposed a tensor-based FPMC-LR model by considering the order relationship between visitings (Cheng et al. 2013). However, the periodicity of check-in data and categorical influence are not well studied. Moreover, the candidate set of POIs is filtered by simply removing the venues far from the previous checked-in POI to deal with the data sparsity. The yielded smaller set leads to a lower computation cost at expense of neglecting the experience of users whose check-in behavior patterns are exclusive from the majority ones and a failure of predicting those far way POIs.

Problem Definition

Let $U = \{u_1, u_2, ..., u_M\}$ be a set of LBSN users, and $L = \{l_{1,2}, ..., l_N\}$ be a set of locations, also called POIs, where each location is geocoded by {longitude, latitude}. The set of POIs visited by user u before time t is denoted by L_u, i.e. $L_u = \{L_u^1, ..., L_u^{t-1}\}$. The contextual feature vector is defined as $\mathbf{g}(\mathbf{c}) = \{g_1(c), ..., g_F(c)\}$ which infers a specific contextual scenario $\mathbf{c}$. The contextual features include previous location, time of day, day of week, previous location's category, etc. F denotes the number of features. Assuming there are K latent behavior patterns determined by contextual scenarios, the pattern distribution can be represented as $\Pi = (\pi_1, ..., \pi_K)$, *s.t.* $\sum_{k=1}^{K} \pi_k = 1$, where π_k denotes the probability of the contextual scenario belonging to the k_{th} latent pattern. With the conjecture that the check-in behaviors are governed by the pattern-level preferences, the probability distribution over next POIs is then the mixture of each pattern-level preference towards those POIs. Our goal is to estimate the pattern distribution Π and pattern-level preference, so as to recommend top-N venues to the user u for his next move by combining the obtained pattern-level preferences.

Proposed Method

Our proposed model is to recommend next personalized POIs via the ranking of probabilities that user u will move from location i to next location l. Based on the first-order



(a) Monday (b) Tuesday (c) Friday (d) Saturday (e) Sunday

Figure 2: Statistics on Location Category Transitions Along with The Day of Week. Categories={c_1 : Arts & Entertainment, c_2: College & University, c_3 : Food, c_4 : Outdoors, c_5 : Work, c_6 : Nightlife Spot, c_7 : Shop, c_8 : Travel Spot}

Markov chain property, the probabilities is given as:

$$x_{u,i,l} = p(L_{u,l}|\mathbf{c}) \tag{1}$$

where **c** denotes the contextual scenario. Thus, each user is associated with a specific transition matrix χ^u which in total generates a transition tensor $\chi \in [0,1]^{|U|\times|L|\times|L|}$ with each $\chi_{u,i,l}$ representing the observed transition record of user u from location i to location l. To further boost the recommendation performance, here we study both personal preference and spatial preference.

Personal Preference. As the transitions of χ are partially observed, here we adopt the low-rank factorization model—a special case of Canonical Decomposition which models the pairwise interaction between all three modes of the tensor (i.e. user U, location I, next location L), to fill up the missing information, given as:

$$\hat{x}_{u,i,l} = u_{U,L} \cdot l_{L,U} + l_{L,I} \cdot i_{I,L} + u_{U,I} \cdot i_{I,U} \tag{2}$$

where $u_{U,L}$ and $l_{L,U}$ denote the latent factor vectors for users and next locations, respectively. Other notions are similarly defined. The term $u_{U,I} \cdot i_{I,U}$ can be removed since it is independent of location l and does not affect the ranking result, as shown in (Rendle, Freudenthaler, and Schmidt-Thieme 2010), which generates a more compact expression for $\hat{x}_{u,i,l}$:

$$\hat{x}_{u,i,l} = u_{U,L} \cdot l_{L,U} + l_{L,I} \cdot i_{I,L}. \tag{3}$$

Spatial Preference. Inspired by (Cho, Myers, and Leskovec 2011), human mobility is constrained geographically by the distance one can travel within a day and their preference to visit a location decreases as the geographic distance increases. Moreover, most of POIs are likely explored near to users' residence, workplace, and frequently visited POIs. Different from the existing works which simply remove locations out of the candidate list based on predefined distance threshold, we leverage on the distance constraint by defining $\rho \cdot d_{i,l}^{-1}$ as the spatial preference of user u to visit a $d_{i,l}$km far away POI, and the optimal setting of ρ will be learned during model inference phase.

Combining these two types of preference linearly, we have an updated transition probability estimation, given as:

$$\hat{x}_{u,i,l} = u_{U,L} \cdot l_{L,U} + l_{L,I} \cdot i_{I,L} + \rho \cdot d_{i,l}^{-1}. \tag{4}$$

Thus, even the locations far away from the previously checked-in location have the chance to be recommended when personal preference dominates. And some occasional long journey could be predicted.

Incorporating Pattern-Level Preference

With assumption that user mobility can be classified into some latent behavior patterns, each pattern has distinct impact to user's transition preference, which indicates that users' transition probability is pattern-sensitive. Here, we propose a novel model by introducing an intermediate latent patterns layer to capture the pattern-level preference in POI recommendation. s is the latent variable to indicate the pattern-level influence. The joint probability of $x_{u,i,l}$ and s is represented as:

$$p(L_{u,l}, s|\mathbf{c}) = p(L_{u,l}|s,\mathbf{c})p(s|\mathbf{c}) \tag{5}$$

where $p(s|\mathbf{c})$ is the mixing coefficient, i.e. π. The pattern-level preference can be defined as:

$$\begin{aligned} \hat{x}^s_{u,i,l} &= p(L_{u,l}|s,\mathbf{c}) \\ &= u^s_{U,L} \cdot l^s_{L,U} + l^s_{L,I} \cdot i^s_{I,L} + \rho^s \cdot d_{i,l}^{-1} \end{aligned} \tag{6}$$

By marginalizing out the latent variable s, the corresponding transition probability can be written as follows:

$$\hat{x}_{u,i,l} = \sum_s \hat{x}^s_{u,i,l} p(s|\mathbf{c}) \tag{7}$$

Fig.3 gives a graphical illustration of our proposed model. The upper tensor contains the historical check-in data which is in fact the transition tensor χ, where the transition probability between two locations is labeled as "1" if we observe that a transition happens between the two locations for a user, or "?" otherwise. Each user, however, may have distinct pattern-level preference under different pattern. And each entry of lower tensors denotes the pattern-level transition probability. It is noted that transition tensor χ is a mixture of the pattern-level transition tensors, and $p(s|\mathbf{c})$ is the mixing coefficient. Then, our goal is to infer the proper pattern-level transition probabilities and pattern distributions to recover the unobserved transition preference by fitting model.

We adopt a soft-max function $\frac{1}{S_c}\exp(\sum_{j=1}^{F}\alpha_j^s g_j(c))$ to infer the multi-patterns and $p(s|\mathbf{c})$. α_j^s is the weight associated with the j_{th} feature for latent pattern s and S_c is the normalization factor that scaled the exponential function to be a proper probability distribution Π, i.e. $S_c = \sum_{k=1}^{K}\exp(\sum_{j=1}^{F}\alpha_j^{s_k} g_j(c))$. In this representation, contextual scenario **c** is denoted by a bag of features $\{g_1(c), ..., g_F(c)\}$ where F is the number of features. By

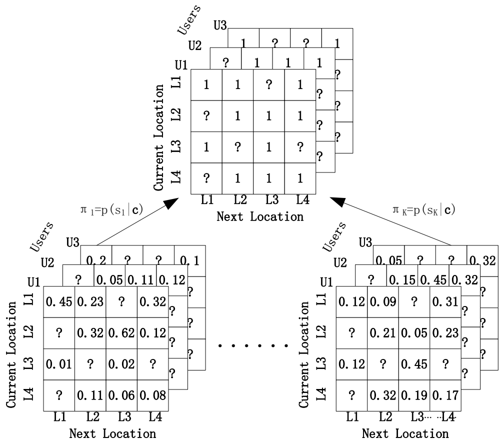


Figure 3: A graphical illustration of our proposed model

plugging the soft-max function into Eq.(7), $\hat{x}_{u,i,l}$ is rewritten as:

$$\hat{x}_{u,i,l} = \frac{1}{S_c} \sum_s \hat{x}^s_{u,i,l} \exp(\sum_{j=1}^{F} \alpha^s_j g_j(c)) \tag{8}$$

The task of next POI recommendation is to recommend top-N POIs to users, thus we can model it as a ranking $>^s_{u,i}$ over locations, where $\hat{x}^s_{u,i,l}$ denotes a personalized ranking score of transition from location i to location l for user u under pattern s.

$$m >^s_{u,i} n \Leftrightarrow \hat{x}^s_{u,i,m} > \hat{x}^s_{u,i,n} \tag{9}$$

Eq.(9) indicates user u prefers location m to location n under pattern s.

Next, we derive the sequential Bayesian Personalized Ranking (S-BPR) optimization criterion which is similar to the general BPR approach (Rendle et al. 2009). Then for user u influenced by the pattern-level preference s, the best ranking can be modeled as:

$$p(\Theta| >^s_{u,i}) \propto p(>^s_{u,i} |\Theta)p(\Theta) \tag{10}$$

where Θ is the set of model parameters, i.e. $\Theta = \{\boldsymbol{\alpha}^S, \boldsymbol{\rho}^S, U^S_{U,L}, V^S_{L,U}, V^S_{L,I}, V^S_{I,L}\}$.

Then we estimate the model by maximizing the posterior with assumption that users and their check-in history are independent:

$$\underset{\Theta}{argmax} \prod_{u\in U} \prod_{i \in L_u} \prod_{m\in L^t_u} \prod_{n \notin L^t_u} \sum_s p(m >^s_{u,i} n|\Theta) \cdot p(s|\mathbf{c})p(\Theta) \tag{11}$$

The ranking probability can be futher expressed by:

$$\begin{aligned} p(m >^s_{u,i} n|\Theta) &= p(x^s_{u,i,m} > x^s_{u,i,n}|\Theta) \\ &= p(x^s_{u,i,m} - x^s_{u,i,n} > 0|\Theta) \end{aligned} \tag{12}$$

Similar to (Rendle, Freudenthaler, and Schmidt-Thieme 2010), we use the logistic function $\sigma(z) = \frac{1}{1+e^{-z}}$ to approximate the likelihood of user's preference over location m and n:

$$p(m >^s_{u,i} n|\Theta) = \sigma(x^s_{u,i,m} - x^s_{u,i,n}) \tag{13}$$

By assuming the model parameters' prior follows a Guassian distribution $p(\Theta) \sim N(0, \frac{2}{\lambda_\Theta} I)$, the MAP estimation is now given as:

$$\underset{\Theta}{argmax} \prod_{u\in U} \prod_{i \in L_u} \prod_{m\in L^t_u} \prod_{n \notin L^t_u} \{\frac{1}{S_c} \cdot \sum_s \sigma(x^s_{u,i,m} - x^s_{u,i,n}) \exp(\sum_{j=1}^{F} \alpha^s_j g_j(c)) e^{-\frac{\lambda_\Theta}{2}||\Theta||^2}\} \tag{14}$$

Model Inference

Furthermore, Θ can also be estimated by maximizing the following log-scale objective function:

$$\underset{\Theta}{argmax} \sum_{u\in U} \sum_{i \in L_u} \sum_{m\in L^t_u} \sum_{n \notin L^t_u} \ln\{\frac{1}{S_c} \cdot \sum_s \sigma(x^s_{u,i,m} - x^s_{u,i,n}) \exp(\sum_{j=1}^{F} \alpha^s_j g_j(c)) e^{-\frac{\lambda_\Theta}{2}||\Theta||^2}\} \tag{15}$$

Here, we adopt Expectation Maximization(EM) algorithm (Dempster, Laird, and Rubin 1977) to estimate the model parameters.

In E-Step, $\gamma(s)$ the posterior distribution of s is given as:

$$\begin{aligned} \gamma(s) &= P(s| >^s_{u,i}, \Theta, \mathbf{c}) \\ &= \frac{\sigma(x^s_{u,i,m} - x^s_{u,i,n}) \exp(\sum_{j=1}^{F} \alpha^s_j g_j(c))}{\sum_s \sigma(x^s_{u,i,m} - x^s_{u,i,n}) \exp(\sum_{j=1}^{F} \alpha^s_j g_j(c))} \end{aligned} \tag{16}$$

And in M-Step, $\boldsymbol{\alpha}^S$ and $\{\Theta \setminus \boldsymbol{\alpha}^S\}$ can be derived by optimizing the Q-function of Eq.(17) and Eq.(18), respectively. The detailed algorithm and the parameter updating rules are shown in **Algorithm** 1.

$$\boldsymbol{\alpha}^S = \underset{\boldsymbol{\alpha}^S}{argmax} \sum_{u\in U} \sum_{i \in L_u} \sum_{m\in L^t_u} \sum_{n \notin L^t_u} \sum_s \gamma(s) \cdot \{\ln(\frac{1}{S_c} \exp(\sum_{j=1}^{F} \alpha^s_j g_j(c))) - \frac{\lambda_\Theta}{2}||\Theta||^2)\} \tag{17}$$

$$\{\Theta \setminus \boldsymbol{\alpha}^S\} = \underset{\{\Theta \setminus \boldsymbol{\alpha}^S\}}{argmax} \sum_{u\in U} \sum_{i \in L_u} \sum_{m\in L^t_u} \sum_{n \notin L^t_u} \sum_s \gamma(s) \cdot \{\ln \sigma(x^s_{u,i,m} - x^s_{u,i,n}) - \frac{\lambda_\Theta}{2}||\Theta||^2)\} \tag{18}$$

Experiments

In this section, we evaluate the following: (1) how is the proposed approach in comparison with other state-of-the-art recommendation techniques? (2) how does the number of latent classes affect the model accuracy? (3) how does the features perform in the POI recommendation task?

Datasets

We choose two large-scale datasets from real-world LBSNs, Foursquare and Gowalla, to conduct the experiments. *Foursquare* check-in data is within Los Angeles, provided

Algorithm 1 Our Proposed Methodology

Input: the number of patterns K, check-in data D
draw Θ from $N(0, \frac{2}{\lambda_\Theta} I)$
repeat
  E-Step:
  $S_c \leftarrow \sum_{k=1}^{K} \exp(\sum_{j=1}^{F} \alpha_j^{s_k} g_j(c))$
  $p(s|\boldsymbol{c}) \leftarrow \frac{1}{S_c} \exp(\sum_{j=1}^{F} \alpha_j^{s} g_j(c))$
  $\gamma(s) \leftarrow \frac{\sigma(x_{u,i,m}^s - x_{u,i,n}^s)\exp(\sum_{j=1}^{L} \alpha_j^s g_j(c))}{\sum_s \sigma(x_{u,i,m}^s - x_{u,i,n}^s)\exp(\sum_{j=1}^{L} \alpha_j^s g_j(c))}$
  M-Step:
  $\delta \leftarrow (1 - \sigma(x_{u,i,m}^s - x_{u,i,n}^s))$
  $u_{U,L}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (m_{L,U}^s - n_{L,U}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $i_{I,L}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (m_{L,I}^s - n_{L,I}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $m_{L,U}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot u_{U,L}^s}{\lambda_\Theta \sum_d \gamma(s)}$
  $n_{L,U}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (-u_{U,L}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $m_{L,I}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot i_{I,L}^s}{\lambda_\Theta \sum_d \gamma(s)}$
  $n_{L,I}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (-i_{I,L}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $\boldsymbol{\alpha}^s \leftarrow \frac{\sum_d \gamma(s) \cdot g(\boldsymbol{e}) \cdot (1 - p(s|\boldsymbol{c}))}{\lambda_\Theta \sum_d \gamma(s)}$
  $\boldsymbol{\rho}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (d_{i,m}^{-1} - d_{i,n}^{-1})}{\lambda_\Theta \sum_d \gamma(s)}$
until convergence
Return: Θ

by (Bao, Zheng, and Mokbel 2012), while *Gowalla* dataset is from (Cheng et al. 2012) with a complete snapshot. For both datasets, we removed the users who checked in LSBN less than 10 times, and the POIs which have been checked in less than 5 times (*Note that the categorical information of POIs are not included in Gowalla dataset*). The statistics of the two datasets are listed in Table 1.

Table 1: Dataset Statistics

	#User	#POI	#Check-in	#Avg.chenk-in
Foursquare	2823	84937	130583	46.25
Gowalla	1388	11235	301678	217.35

Evaluation Metrics

Given a top-N recommendation list $S_{N,u,rec}$ sorted in descending order of the prediction values to user u, we adapt a precision metric to evaluate the performance of our proposed next POI recommendation, given as:

$$Precision@N = \frac{1}{|U|} \sum_{u \in U} \frac{|S_{N,u,rec} \cap S_{visited}|}{|S_{visited}|} \tag{19}$$

where $S_{visited}$ are the visited locations of user u and $|U|$ denotes the number of the users, N is the size of the next POI candidate list.

Comparison

We compare the proposed model with the following methods:

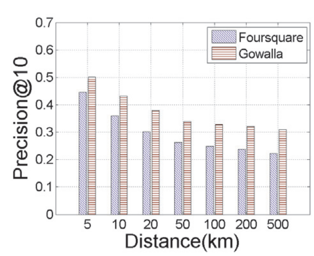
- **MF:** matrix factorization (Koren, Bell, and Volinsky 2009) is widely used in conventional recommender systems, which factorizes the user-item preference matrix.
- **PMF:** probabilistic matrix factorization is a well-known method for modeling time evolving relation data (Mnih and Salakhutdinov 2007).
- **FPMC-LR:** this method is proposed in (Cheng et al. 2013), which is the state-of-the-art personalized successive POI recommendation method.

Table 2 reports the comparison results between our model and the baseline methods. We set λ_Θ to be 1 for both FPMC-LR and our proposed model. The empirical settings of the number of latent behavior patterns are 4 and 6 for *Gowalla* dataset and *Foursquare* dataset, respectively. For other parameters, we tune them in the training sets to find the optimal values, and subsequently use them in the test set. The results show that:

- Both FPMC-LR and the proposed model outperform MF and PMF significantly, which indicates that the conventional POI recommendation algorithms are not effective for the successive POI recommendation. One possible explanation could be that MF and PMF mainly exploit the user preference rather than making use of the sequential information. More specifically, our proposed model achieves a relative improvement of at least 91% for MF and 81% for PMF respectively, while FPMC-LR also achieves an improvement compared with MF and PMF. This demonstrates that spatial influence plays an important role in next POI recommendation.
- Our proposed method consistently outperforms FPMC-LR, improving around 35% and 45% over FPMC-LR for Foursquare dataset and Gowalla dataset, respectively. It illustrates that inferring user latent behavior patterns can better capture user mobility preference in LBSNs, and therefore, help us recommend POIs to users more accurately.

User ID	Case Study of Success Next POI Recommendation
2282	From Arches National Park Visitor Center (2011.06.10, 14:24, Fri) to Patagonia Outlet (2011.6.12, 10:26, Sun), distance=304 km., time interval=44 hr.
1598	From Silvia's Hair Design(2011.05.20, 13:09, Fri) to Don Carlos(2011.05.21, 12:19, Sat), distance=0.5577 km., time interval=23.16 hr.
192	From holiday Inn Express (2011.07.03, 06:01, Sun) to Hartsfield-Jackson Atlanta International Airport (2011.07.03, 08:50, Sun), distance= 82.82 km., time interval=2.81 hr.

(a) Case Study



(b) Quantitative Evaluation of accumulated precision along with distance

Figure 4: Prediction Ability vs. Distance

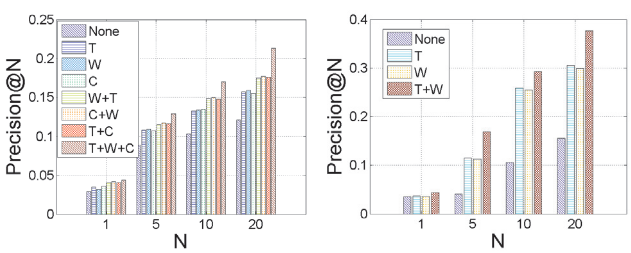
Table 2: Performance Comparison

Metrics	Foursquare				Gowalla			
	MF	PMF	FPMC-LR	Our	MF	PMF	FPMC-LR	Our
P@1	0.023	0.024	0.032	**0.044**	0.022	0.024	0.029	**0.044**
Improve	91.30%	83.33%	37.50%		100.00%	83.33%	51.72%	
P@5	0.067	0.071	0.097	**0.129**	0.086	0.093	0.116	**0.169**
Improve	92.54%	81.69%	32.99%		96.51%	81.72%	45.69%	
P@10	0.089	0.093	0.128	**0.170**	0.147	0.158	0.198	**0.293**
Improve	91.01%	82.80%	32.81%		99.32%	85.44%	47.98%	
P@20	0.108	0.116	0.155	**0.213**	0.188	0.202	0.247	**0.377**
Improve	97.22%	83.62%	37.42%		100.53%	86.63%	52.63%	

Fig. 4 shows the prediction ability vs. distance. Both case study of success next POI prediction(see in Fig. 4(a)) and quantitative results(see in Fig. 4(b)) manifest that our model is capable of predicting transitions within a localized region as well as an occasional journey with long distance.

Impact of the Contextual Features

Here, we discuss the recommendation efforts of different types of contextual information, i.e. Previous Location's Category, Time of Day, Day of Week. Figure 5 depicts the experimental results with variants of combinations of contextual information incorporated. In general, the model accuracy increases with more contextual information added in. It indicates that finer latent patterns are obtained to better capture user preference.

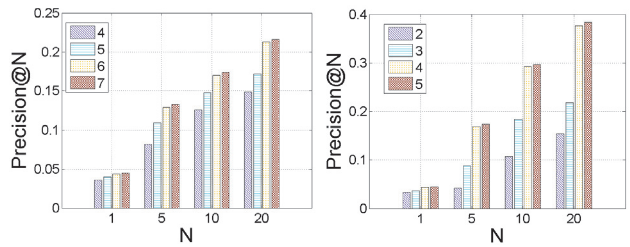


(a) Foursquare (b) Gowalla

Figure 5: Impact of Feature Combos(*T stands for time, W stands for week and C stands for*)

Impact of the Number of Latent Patterns

Figure 6 shows the experimental results with different settings of the number of latent patterns. We can see that for both datasets, the model accuracy increases with the increasing of the number of patterns. When the number of latent patterns reaches 6 for Foursquare and 4 for Gowalla, the returns diminish largely. Even the performance gained by adding one more latent pattern is minor compared to the difference between the number of patterns less than 6 for Foursquare and 4 for Gowalla. For example, P@10 on Gowalla dataset is 0.184 using 3 latent patterns, whereas the four-latent-pattern model has a P@10 of 0.293, which is a 59.2% relative improvement. Using a five-latent-pattern



(a) Foursquare (b) Gowalla

Figure 6: Number of Latent Pattern

model only increase performance by another 1.4%. Besides considering the additional computation cost of inferring preference for each pattern, we conclude that the 6 latent patterns for Foursquare and 4 latent patterns for Gowalla is rich enough to complete the task of next personalized POI recommendation.

Conclusion and Future Work

To address the personalized next POI recommendation problem, in this paper we propose a unified tensor-based latent model to capture the successive check-in behavior by exploring the latent pattern-level preference for each user. We derive a BPR-like optimization criterion accordingly and then use Expectation Maximization (EM) to estimate the model parameters. Performance evaluation conducted on two large-scale real-world LBSNs datasets shows that our proposed approach improves the recommendation accuracy significantly compared against other state-of-the-art methods. More specifically, our proposed method is capable of predicting journey of long distance and the consecutive check-ins which span a long period of time. For future work, we will soon evaluate our proposed model's ability for next new POI recommendation by redefining the transition tensor in a categorical dimension.

Acknowledgement

We would like to thank the anonymous reviewers for their insightful comments. This work has been partially supported by NSFC under Grant No. 61300178 and National Program on Key Basic Research Project under Grant No. 2013CB329605.

References

Bao, J.; Zheng, Y.; and Mokbel, M. F. 2012. Location-based and preference-aware recommendation using sparse geo-social networking data. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, 199–208. ACM.

Chen, Y.; Zhao, J.; Hu, X.; Zhang, X.; Li, Z.; and Chua, T. 2013. From interest to function: Location estimation in social media. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, Bellevue, Washington, USA.*, 180–186.

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

Cheng, C.; Yang, H.; Lyu, M. R.; and King, I. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, 2605–2611. AAAI Press.

Cho, E.; Myers, S. A.; and Leskovec, J. 2011. Friendship and mobility: User movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, 1082–1090. New York, NY, USA: ACM.

Dempster, A. P.; Laird, N. M.; and Rubin, D. B. 1977. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)* 1–38.

Eagle, N., and Pentland, A. 2009. Eigenbehaviors: Identifying structure in routine. In *Behavioral Ecology and Soc.*

Feng, S.; Li, X.; Zeng, Y.; Cong, G.; Chee, Y. M.; and Yuan, Q. 2015. Personalized ranking metric embedding for next new poi recommendation. In *Proceedings of the 24th International Conference on Artificial Intelligence*, 2069–2075. AAAI Press.

Gao, H.; Tang, J.; Hu, X.; and Liu, H. 2015. Content-aware point of interest recommendation on location-based social networks. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA.*, 1721–1727.

Henan Wang, Guoliang Li, J. F. 2014. Group-based personalized location recommendation on social networks. In *Proceeding of the 16th Asia-Pacific Web Conference, APWeb'14, Changsha, China*, 68–80.

Koren, Y.; Bell, R. M.; and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. volume 42, 30–37.

Li, Z.; Ding, B.; Han, J.; Kays, R.; and Nye, P. 2010. Mining periodic behaviors for moving objects. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '10, 1099–1108. New York, NY, USA: ACM.

Liu, B.; Fu, Y.; Yao, Z.; and Xiong, H. 2013. Learning geographical preferences for point-of-interest recommendation. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, 1043–1051. ACM.

Mnih, A., and Salakhutdinov, R. 2007. Probabilistic matrix factorization. In *Advances in neural information processing systems*, 1257–1264.

Neal, R. M., and Hinton, G. E. 1998. A view of the em algorithm that justifies incremental, sparse, and other variants. In *Learning in graphical models*. Springer. 355–368.

Rendle, S.; Freudenthaler, C.; Gantner, Z.; and Schmidt-Thieme, L. 2009. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, 452–461.

Rendle, S.; Freudenthaler, C.; and Schmidt-Thieme, L. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web*, 811–820. ACM.

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

Yuan, Q.; Cong, G.; Ma, Z.; Sun, A.; and Thalmann, N. M. 2013. Time-aware point-of-interest recommendation. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, 363–372. ACM.

Yuan, T.; Cheng, J.; Zhang, X.; Qiu, S.; and Lu, H. 2014. Recommendation by mining multiple user behaviors with group sparsity. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec City, Québec, Canada.*, 222–228.

Zheng, Y.; Zhang, L.; Xie, X.; and Ma, W.-Y. 2009. Mining interesting locations and travel sequences from gps trajectories. In *Proceedings of the 18th international conference on World Wide Web, WWW'09*, 791–800. ACM.