

Category-aware Next Point-of-Interest Recommendation via Listwise Bayesian Personalized Ranking

Jing He, Xin Li* and Lejian Liao

BJ ER Center of HVLIP&CC, School of Comp. Sci., Beijing Institute of Technology, Beijing, China
 {skyhejing, xinli, liaolj}@bit.edu.cn

Abstract

Next Point-of-Interest (POI) recommendation has become an important task for location-based social networks (LBSNs). However, previous efforts suffer from the high computational complexity, besides the transition pattern between POIs has not been well studied. In this paper, we proposed a two-fold approach for next POI recommendation. First, the preferred next category is predicted by using a third-rank tensor optimized by a Listwise Bayesian Personalized Ranking (LBPR) approach. Specifically we introduce two functions, namely Plackett-Luce model and cross entropy, to generate the likelihood of a ranking list for posterior computation. Then POI candidates filtered by the predicated category are ranked based on the spatial influence and category ranking influence. The experiments on two real-world datasets demonstrate the significant improvements of our methods over several state-of-the-art methods.

1 Introduction

Next Point-of-Interest (POI) recommendation has become an important task for location-based social networks (LBSNs), which not only help users to explore their favorite places but also benefit for business to acquire more potential customers. Many efforts have been devoted to the task [Cheng *et al.*, 2013; Feng *et al.*, 2015; Liu *et al.*, 2016] in the literature. Yet achieving accurate personalized next POI recommendation is challenging as the check-ins data of each user is very sparse. Existing methods study users' preference on POIs by employing various context information, but usually suffer from high computational complexity as the entire check-in dataset of all users is extremely large.

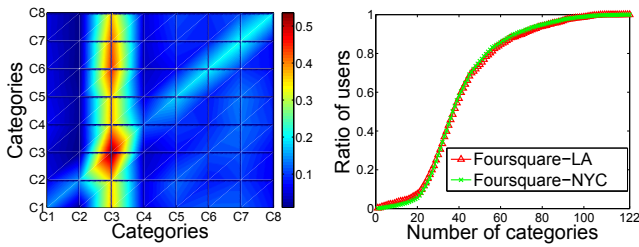
However, human daily activities usually present category-level transition patterns. Intuitively, people tend to make up their mind on the type of places they are going to first, and

then determine the specific locations according to their personal preferences. For example, a shopaholic often visits malls to purchase various goods, and a foodie usually prefers to seek good restaurants. In fact, the number of location categories is far smaller than the number of locations, the computational complexity can be largely reduced by predicting the preferred categories first. Fig.1(a) plots the transition probabilities between location categories. We observe that some categorical transitions are very obvious, for example, "Nightlife Spots" are more likely checked in after "Food places", which indicates that the next location's category is highly related to the current location's category.

In LBSNs, users are more likely to simply check-in POIs online rather than provide detailed ratings for these POIs. The check-in information implicitly reflects user's preference on POIs. To address the problems with implicit feedback, Bayesian Personalized Ranking (BPR) [Rendle *et al.*, 2009] is well known for its pioneering work. Specifically, BPR studies the personalized preference from a Bayesian analysis perspective, which can be used to predict the pairwise preference between the observed explicit feedback and the implicit feedback. Cheng *et al.* proposed a state-of-the-art approach for next POI recommendation which adopted BPR to apply on a user-location matrix but still suffered from data sparsity and high computational cost [Cheng *et al.*, 2013]. Whereas a user-category matrix can be considered as a data structure with rich information, but with smaller size. As shown in Fig.1(b), about 95% users visited more than 20 "categories" for foursquare datasets, that left only few unvisited categories can be used as negative instances in BPR. Besides, the personalized order over the visited categories cannot be learned as the conventional BPR overlooked the ordering among the observed instances. Moreover, the assumption made in BPR, that is the preference of each instance pair for a specific user is independent of the ordering, does not always hold in the real cases. And the learning process is usually managed in a way of minimizing errors in ranking instance pairs, rather than minimizing errors in ranking instance lists.

In this paper, we proposed a two-step approach for next POI recommendation: *Listwise Bayesian Personalized Ranking (LBPR) for next category prediction* and *category-based location recommendation*. Our proposed LBPR is derived from the maximum posterior estimator for optimal personalized ranking by taking category lists instead of category pairs

*Corresponding Author: Xin Li (xinli@bit.edu.cn). This work has been partially supported by NSFC under Grant No.61300178, National Program on Key Basic Research Project under Grant No.2013CB329600 and International Graduate Exchange Program of Beijing Institute of Technology.



(a) Statistics on Location Category Transitions (b) User ratio vs. Number of visited categories

Figure 1: Categorical behavior illustration. 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}

as instances during the learning process. We argue that the transition frequency list can better capture user’s preference over the location categories. Once the preferred categories are obtained from the first-step, the locations are then ranked as the candidates of the next location by considering both spatial and categorical ranking influence. For LBPR, the key is to define a listwise loss function to represent the difference between the ranking list derived from LBPR and the ranking list given as the ground truth. Specifically we introduce two probability models respectively referred to as 1) *permutation probability* by maximizing the likelihood function of each ranking list and 2) *top one probability* by using cross entropy to minimize the difference between the recommended list and the ground truth list. The main contributions of this paper can be summarized as follows:

- We propose a two-fold approach for next POI recommendation to predict user preference at the category-level first, then derive the ranking list of POI candidates based on the predicted category preference.
- We propose a novel optimization criterion LBPR with the maximum posterior estimator derived from Bayesian analysis of ranking problems and formulate the listwise loss functions on the basis of two probability models.
- We evaluate the proposed models by the experiments on two large-scale LBSN datasets and demonstrate that our proposed approach outperforms other state-of-the-art POI recommendation approaches by a large margin.

2 Related Work

POI recommendation has become an important research topic. Many recent studies had tried to improve the recommendation performance by integrating geographical influence, social influence, temporal effect and content information of POIs [Feng *et al.*, 2015; Cho *et al.*, 2011; Gao *et al.*, 2015; Zhao *et al.*, 2016b; 2016a; Li *et al.*, 2016]. Cheng *et al.* proposed a tensor-based FPMC-LR model for next POI recommendation by considering the successive relation between visitings [Cheng *et al.*, 2013]. Some works had incorporated categorical information into recommender systems for performance enhancement. Zhang *et al.* exploited geographical, social and categorical correlations among users

and POIs, and applied the bias of a user on a POI category to weigh the popularity of a POI in the corresponding category [Zhang and Chow, 2015]. He *et al.* modeled the next POI recommendation under the influence of user’s latent behavior patterns, which is determined by the contextual scenarios including temporal and categorical information [He *et al.*, 2016].

Learning to rank is a popular research topic in machine learning. The existing works of learning to rank can be roughly classified into three categories: 1) *The pointwise approach*, with absolute preference assumptions, transforms ranking into regression or classification on single objects, which takes the implicit feedback as the absolute preference score [Hu *et al.*, 2008; Pan *et al.*, 2008]. 2) *The pairwise approach*, with relative preference assumptions, transforms ranking into classification of object pairs into two categories, which takes the implicit feedback as the relative preference rather than absolute one [Rendle *et al.*, 2009]. 3) *The listwise approach* tackles the ranking problem directly by optimizing the defined loss function on a list of items, in the literature the popular listwise ranking approaches include ListNet [Cao *et al.*, 2007], ListMLE and etc. [Xia *et al.*, 2008; Lan *et al.*, 2009] which differ from each other by defining different listwise loss function.

Different from the existing listwise ranking approaches, our proposed LBPR optimization criterion for POI category prediction is derived from a Bayesian analysis perspective by using the likelihood function and the prior probability for the model parameter, which takes category lists as the instances during learning process.

3 Problem Definition

Let $U = \{u_1, u_2, \dots, u_M\}$ be a set of LBSN users, and $L = \{l_1, l_2, \dots, l_N\}$ be a set of locations, also called POIs, where each location is geocoded by {longitude, latitude}. $C = \{c_1, c_2, \dots, c_S\}$ denotes the location categories which is provided by Foursquare¹. The set of categories visited by user u before time t is denoted by C_u , i.e. $C_u = \{C_u^1, \dots, C_u^{t-1}\}$, where $C_u^{t-1} \subset C$ is the subset of categories to which the location l_u^{t-1} visited by user u at time $t - 1$ belongs. Our goal is to estimate the category-level preference first, so as to recommend top-N locations to user u for her next move based on the obtained category-level preference.

4 Proposed Model

In this section, we detail the two-fold model: *Listwise Bayesian Personalized Ranking for category prediction and category-based location recommendation*.

4.1 LBPR for Next Category Prediction

The task of category prediction is to predict next categories via the ranking of probabilities that user u will move from the current location category i to the next location category j . Based on the first-order Markov property, the probabilities are given as:

$$x_{u,i,j} = p(i \in C_u^t | j \in C_u^{t-1}) \tag{1}$$

¹<https://developer.foursquare.com/categorytree>

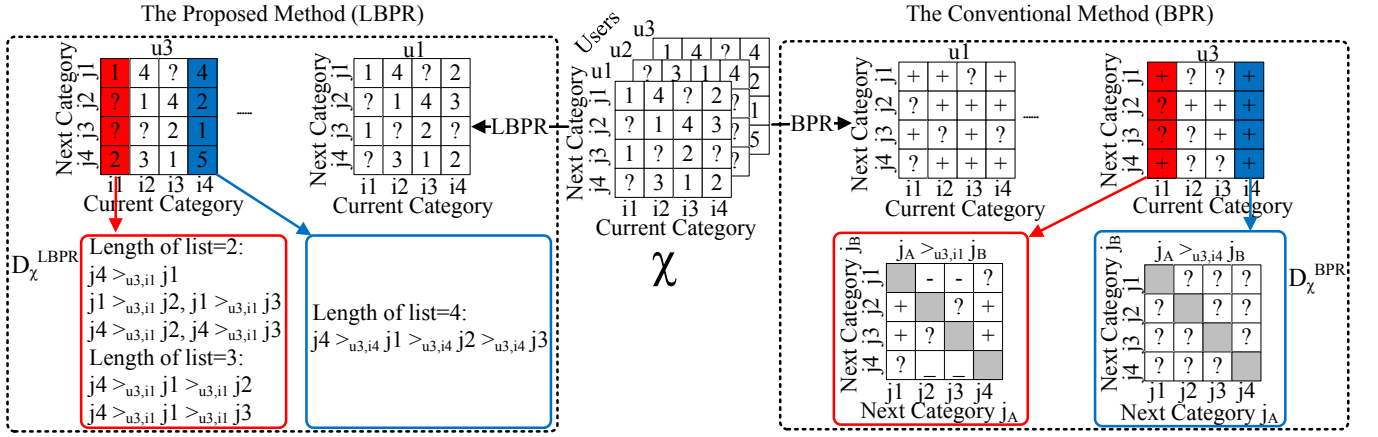


Figure 2: The observed data χ forms a ternary relation between users, current categories and next categories, from which either pairwise preferences D_χ^{BPR} (left-side) or listwise preferences D_χ^{LBPR} (right-side) among the next categories can be inferred.

Thus, each user is associated with a specific transition matrix which generates a third-rank transition tensor $\chi \in \mathbb{R}^{|U| \times |C| \times |C|}$ with each $\chi_{u,i,j}$ representing the observed transition frequency of user u from category i to category j .

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 , current category I , next category J), to fill up the missing information, given as:

$$\hat{x}_{u,i,j} = \mathbf{v}_u^{U,J} \cdot \mathbf{v}_j^{J,U} + \mathbf{v}_j^{J,I} \cdot \mathbf{v}_i^{I,J} + \mathbf{v}_u^{U,I} \cdot \mathbf{v}_i^{I,U} \quad (2)$$

where $\mathbf{v}_u^{U,J}$ and $\mathbf{v}_j^{J,U}$ denote the latent factor vectors for users and next categories, respectively. Other notions are similarly defined. The term $\mathbf{v}_u^{U,I} \cdot \mathbf{v}_i^{I,U}$ can be removed since it is independent of next category j and does not affect the ranking result, as shown in [Rendle *et al.*, 2010], which generates a more compact expression for $\hat{x}_{u,i,j}$:

$$\hat{x}_{u,i,j} = \mathbf{v}_u^{U,J} \cdot \mathbf{v}_j^{J,U} + \mathbf{v}_j^{J,I} \cdot \mathbf{v}_i^{I,J}. \quad (3)$$

Next category prediction is to obtain top- N categories and we care more about the ranking order of the candidate categories rather than the real values of transition possibilities, thus we can model it as a ranking list $>_{u,i}$ over categories:

$$j^{(1)} >_{u,i} \dots >_{u,i} j^{(k)} \Leftrightarrow \hat{x}_{u,i,j^{(1)}} > \dots > \hat{x}_{u,i,j^{(k)}} \quad (4)$$

where $j^{(k)}$ denotes the category at position k of the ranking list and $\hat{x}_{u,i,j^{(k)}}$ denotes the personalized ranking score of transition from category i to category $j^{(k)}$ for user u .

LBPR Optimization Criterion

As shown in Fig.2, the transition tensor χ contains the historical check-in data, the numerical number indicates the transition frequency $\chi_{u,i,j}$ from category i to category j for user u , “?” indicates the unobserved transitions. We take (u_3, i_1) for example (the red column), the inferred pairwise preferences of $j_1 >_{u_3,i_1} j_2, j_1 >_{u_3,i_1} j_3, j_4 >_{u_3,i_1} j_2, j_4 >_{u_3,i_1} j_3$

from BPR is actually the subset of the inferred listwise preferences from LBPR. For (u_3, i_4) with total observed transition records (the blue column), no pairwise preferences can be inferred in BPR, whereas the listwise preference of $j_4 >_{u_3,i_4} j_1 >_{u_3,i_4} j_2 >_{u_3,i_4} j_3$ can be inferred in LBPR. The two cases illustrate our motivation of proposing LBPR for recommendation tasks.

BPR typically creates a training data set D_χ^{BPR} by taking the positive instance (observed transitions) and the negative instance (unobserved transitions) as pairs, which formalizes the problem of learning to rank as a binary classification problem [Rendle *et al.*, 2009; 2010]. By contrast, we use a different approach by using category lists as training data and optimize for correctly ranking lists instead of ranking pairs to further explore the orders of observed transitions which has been overlooked in BPR. If the transition frequency $\chi_{u,i,j}$ between category i and category j for user u is higher than the transition frequency $\chi_{u,i,j'}$ between category i and category j' for user u , then we assume that user u at category i prefers category j over category j' : $j >_{u,i} j'$. Then the listwise instance of training data D_χ^{LBPR} is denoted as:

$$D_\chi^{LBPR} = \{list_{u,i,j} | j^{(1)} >_{u,i} \dots >_{u,i} j^{(k)}\} \quad (5)$$

where k denotes the length of the instance list $list_{u,i,j}$. The differences between BPR and LBPR are depicted in Fig.2.

The Bayesian formulation of finding the best personalized ranking is to maximize the following posterior probability:

$$p(\Theta | list_{u,i,j}) \propto p(list_{u,i,j} | \Theta) p(\Theta) \quad (6)$$

where Θ denotes the parameter set of an arbitrary model class (e.g. tensor factorization). It is assumed that all users act independently of each other and the ordering of each list of next categories for a user at a specific current category is independent of each other in training data, which leads to the maximum a posterior (MAP) estimator of the model parameters:

$$\underset{\Theta}{argmax} \prod_{list_{u,i,j} \in D_\chi^{LBPR}} p(list_{u,i,j} | \Theta) p(\Theta) \quad (7)$$

By assuming the model parameters' prior follows a normal distribution $p(\Theta) \sim N(0, \frac{2}{\lambda_\Theta} I)$, the maximum posterior estimator to derive the optimization criterion for personalized ranking is given as:

$$\underset{\Theta}{\operatorname{argmax}} \sum_{list_{u,i,j} \in D_x^{LBPR}} \ln p(list_{u,i,j} | \Theta) - \frac{\lambda_\Theta}{2} \|\Theta\|^2 \quad (8)$$

where $p(list_{u,i,j} | \Theta)$ denotes the likelihood of users' preference order for categories, which measures the similarity between the predicted list and the ground truth list. In this paper, we introduce two particular estimators of $p(list_{u,i,j} | \Theta)$, namely, Plackett-Luce model [Marden, 1996] and cross entropy, incorporated with which are the specialized LBPRs referred to as LBPR-MLE and LBPR-NET in the sequel.

LBPR-MLE

The adopted Plackett-Luce model [Marden, 1996] for LBPR-MLE is proposed in [Xia *et al.*, 2008], which defines a parameterized probability distribution over all the permutations given the predicted result, that is the likelihood of the ground truth list. The likelihood function is given as:

$$p(list_{u,i,j} | \Theta) = \prod_{m=1}^k \frac{\phi(\hat{x}_{u,i,j}^{(m)})}{\sum_{n=m}^k \phi(\hat{x}_{u,i,j}^{(n)})} \quad (9)$$

where k is the length of $list_{u,i,j}$. $\phi(\cdot)$ is an increasing and strictly positive function. We define $\phi(\cdot)$ as the exponential function. By incorporating Eq.(9) into Eq.(8), the objective function of LBPR-MLE is obtained as:

$$\underset{\Theta}{\operatorname{argmax}} \sum_{list_{u,i,j} \in D_x} \sum_{m=1}^k \{ \hat{x}_{u,i,j}^{(m)} - \ln \sum_{n=m}^k \exp(\hat{x}_{u,i,j}^{(n)}) \} - \frac{\lambda_\Theta}{2} \|\Theta\|^2 \quad (10)$$

LBPR-NET

Inspired by [Cao *et al.*, 2007], we define the likelihood function with the cross entropy between two parameterized probability distributions of permutations obtained from the predicted result and the ground truth, which is defined as:

$$p(list_{u,i,j} | \Theta) = \exp\left\{ \sum_{m=1}^k p_\chi(m) \ln p_{\hat{x}}(m) \right\} \quad (11)$$

where $p_\chi(m)$ and $p_{\hat{x}}(m)$ denote the *top one probability* of the transition frequency and the personalized ranking score, respectively:

$$p_\chi(m) = \frac{\phi(\chi_{u,i,j}^{(m)})}{\sum_{n=1}^k \phi(\chi_{u,i,j}^{(n)})} \quad (12)$$

$$p_{\hat{x}}(m) = \frac{\phi(\hat{x}_{u,i,j}^{(m)})}{\sum_{n=1}^k \phi(\hat{x}_{u,i,j}^{(n)})} \quad (13)$$

where the function $\phi(\cdot)$ is same as that in Eq.(9). By incorporating Eq.(11) into Eq.(8), the objective function of LBPR-NET is given as:

$$\underset{\Theta}{\operatorname{argmax}} \sum_{list_{u,i,j} \in D_x} \sum_{m=1}^k p_\chi(m) \ln p_{\hat{x}}(m) - \frac{\lambda_\Theta}{2} \|\Theta\|^2 \quad (14)$$

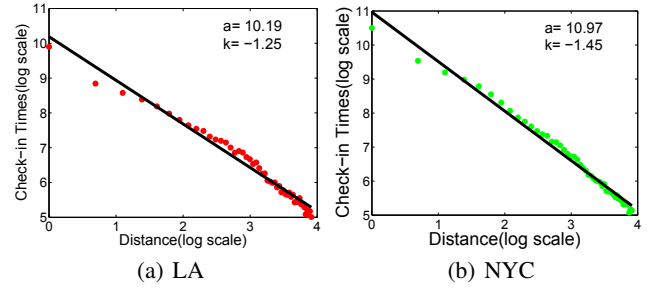


Figure 3: Spatial behavior statistics

Parameter Learning

We follow the widely used stochastic gradient decent (SGD) algorithm to optimize the objective functions (Eq.(10) and Eq.(14)) to obtain Θ for LBPR-MLE and LBPR-NET respectively. The updating procedure is performed as:

$$\Theta' = \Theta + \alpha \left(\frac{\partial}{\partial \Theta} (\ln p(list_{u,i,j} | \Theta) - \frac{\lambda_\Theta}{2} \|\Theta\|^2) \right) \quad (15)$$

where $\alpha > 0$ is the learning rate.

4.2 Category-based Location Recommendation

Given the next category prediction results, there are still tens of thousands of venues belonging the categories to be selected. In this paper, we proposed a heuristic approach to determine the location list for recommendation by incorporating the spatial influence and the category ranking influence.

Spatial Influence

Inspired by [Cho *et al.*, 2011], we investigate the relation (log scale) between the check-in counts and the distance between two successive checked-in locations for two Foursquare datasets (LA and NYC). As seen from Fig.3, it is very obvious that the user preference to visit a location decreases as the geographical distance increases, and the relation follows a power law distribution. Thus, we define $sp(d_{l',l}) = a \cdot d_{l',l}^k$ to indicate the spatial influence for a location l which is $d_{l',l}$ km away from the current location l' , where a and k can be easily obtained by using a least-square regression approach to fitting the data distribution².

Category Ranking Influence

Intuitively, the higher ranking of the obtained preferred category has more influence to the location prediction. Thus, we also incorporate category ranking influence into the personalized location recommendation model. Soft-max function $\frac{1}{Z_{\hat{x}}} \exp(\sum_{c' \in C_l} \hat{x}_{u,c',c})$ is adopted to indicate the weight of the category c according to its ranking position and $Z_{\hat{x}}$ is the normalization factor, i.e. $Z_{\hat{x}} = \sum_{s=1}^S \exp(\sum_{c' \in C_l} \hat{x}_{u,c',c_s})$, where $C_l \subset C$ is the subset of categories which the location l belongs to.

Combining the above two types of influences, we propose a heuristic location recommendation metric, given as:

$$score_{u,l} = (a \cdot d_{l',l}^k) \sum_{c \in C_l} \frac{1}{Z_{\hat{x}}} \exp(\sum_{c' \in C_l} \hat{x}_{u,c',c}) \quad (16)$$

²In our paper, the empirical settings of a and k are 10.5 and -1 respectively.

Table 1: Dataset Statistics

| | #User | #POI | #Check-in | #Avg.check-in | #Cate |
|-----|-------|-------|-----------|---------------|-------|
| LA | 1603 | 70342 | 103400 | 64.50 | 241 |
| NYC | 2579 | 97013 | 157404 | 61.03 | 241 |

5 Experiments

In this section, we evaluate the following: (1) how is the category prediction approaches optimized by LBPR in comparison with other state-of-the-art recommendation techniques? (2) how is the location recommendation model in comparison with other POI recommendation methods? (3) how does the maximum length of the list affect the prediction accuracy? We apply our approach on two real-world datasets acquired from Foursquare, which contains the check-in data within Los Angeles and New York City, respectively. The statistics of the two datasets are listed in Table 1.

5.1 Evaluation Metrics

We evaluate the performance of category prediction by defining precision as:

$$Precision@N = \frac{1}{|U|} \sum_{u \in U} \frac{|S_{N,u}^C \cap S_{visited}|}{|S_{visited}|} \quad (17)$$

where $S_{N,u}^C$ is the top-N recommended categories list sorted in descending order of the obtained $\hat{x}_{u,i,j}$, $S_{visited}$ denotes the visited categories of user u , $|U|$ denotes the number of the users and N is the size of the next category candidate list.

The precision metrics for next POI recommendation and next new POI recommendation are given as:

$$Precision@N_{POI} = \frac{1}{|U|} \sum_{u \in U} \frac{|S_{N,u}^{POI} \cap S_{visited}^{POI}|}{|S_{visited}^{POI}|} \quad (18)$$

$$Precision@N_{POI}^{new} = \frac{1}{|U|} \sum_{u \in U} \frac{|S_{N,u}^{POI} \cap S_{visited}^{newPOI}|}{|S_{visited}^{newPOI}|} \quad (19)$$

where $S_{N,u}^{POI}$ is the top-N recommended POIs list sorted in descending order of the obtained $score_{u,l}$, $S_{visited}^{POI}$ denotes the visited POIs of user u and $S_{visited}^{newPOI}$ denotes the locations that a user has not visited yet in the training set and will be visited in the testing set. Note that we split the datasets into two non-overlapping sets: for each user, the earliest 80% of check-ins as training set and the remaining 20% check-ins as testing set.

5.2 Performance Comparison on Next Category Prediction

We compare our achieved performance on category recommendation with that obtained by the following methods:

- **MF**: matrix factorization [Koren *et al.*, 2009] is widely used in the conventional recommender systems, which factorizes the user-item preference matrix to obtain the user latent feature matrix and item latent feature matrix.

- **PMF**: probabilistic matrix factorization is a well-known method for modeling time evolving relation data [Mnih and Salakhutdinov, 2007].
- **FPMC**: it is a state-of-the-art personalized recommendation method optimized by BPR, which embeds users' preference and their personalized Markov chain to provide next item recommendation [Rendle *et al.*, 2010].
- **PRME**: it obtains the predictions by learning the embeddings of the categories, in which BPR is used as the optimization criterion [Feng *et al.*, 2015].

Table 2 tabulates the detailed results. The empirical setting of the maximum length of the lists from training sets is 5 for category prediction on both datasets³. We observe that:

- FPMC, PRME, LBPR-MLE and LBPR-NET all outperform MF and PMF significantly, which demonstrates the significant improvements of learning to rank methods for recommendation problem.
- Both LBPR-MLE and LBPR-NET consistently outperforms FPMC and PRME for both datasets, which illustrates that models derived from the maximum posterior estimator of rank lists can better model user preference orders in recommender systems.

5.3 Performance Comparison on Next POI Recommendation

The performance comparisons on next POI recommendation are performed among our proposed two-fold approach and the following baselines:

- **FPMC-LR**: it employs FPMC to model the successive personalized POI recommendation and utilizes the location distance constraint [Cheng *et al.*, 2013].
- **PRME-G**: it extends PRME by incorporating the geographical influence [Feng *et al.*, 2015].
- **LBP**: it jointly models the next POI recommendation under the influence of user's latent behavior pattern and adopts a unified tensor-based latent model to capture the successive check-in behavior [He *et al.*, 2016].

Fig.4 shows the performance comparison on location recommendation. And we observe that:

- Our proposed two-fold models outperform FPMC-LR, PRME-G and LBP by a large margin (See Fig.4(a), 4(b)), which implies that the incorporation of category prediction is effective for POI recommendation. Furthermore, our experiments are ran on a machine with a dual Core i7-2.67 GHz and 12G memory and the running time is around 5.5hr while the other approaches takes more than 30 hours, which demonstrates the efficiency of our proposed approaches.
- Our proposed models and LBP have shown the improved precision over FPMC-LR and PRME-G for next new POI recommendation (See Fig.4(c), 4(d)), which

³Note that, we establish a set of list instances with the list length varying from 2 to K for next category prediction learning when setting maximum length as K.

Table 2: Performance Comparison on Category Recommendation

| Metrics | LA | | | | | | NYC | | | | | |
|---------|-------|-------|-------|-------|--------------|--------------|-------|-------|-------|-------|--------------|--------------|
| | MF | PMF | FPMC | PRME | LBPR-M | LBPR-N | MF | PMF | FPMC | PRME | LBPR-M | LBPR-N |
| P@1 | 0.013 | 0.018 | 0.027 | 0.028 | 0.045 | 0.046 | 0.012 | 0.018 | 0.028 | 0.029 | 0.046 | 0.046 |
| P@5 | 0.055 | 0.062 | 0.079 | 0.076 | 0.169 | 0.170 | 0.056 | 0.058 | 0.082 | 0.073 | 0.171 | 0.175 |
| P@10 | 0.092 | 0.107 | 0.146 | 0.153 | 0.270 | 0.272 | 0.098 | 0.099 | 0.155 | 0.151 | 0.271 | 0.276 |
| P@20 | 0.160 | 0.177 | 0.236 | 0.235 | 0.411 | 0.413 | 0.161 | 0.175 | 0.233 | 0.237 | 0.406 | 0.410 |

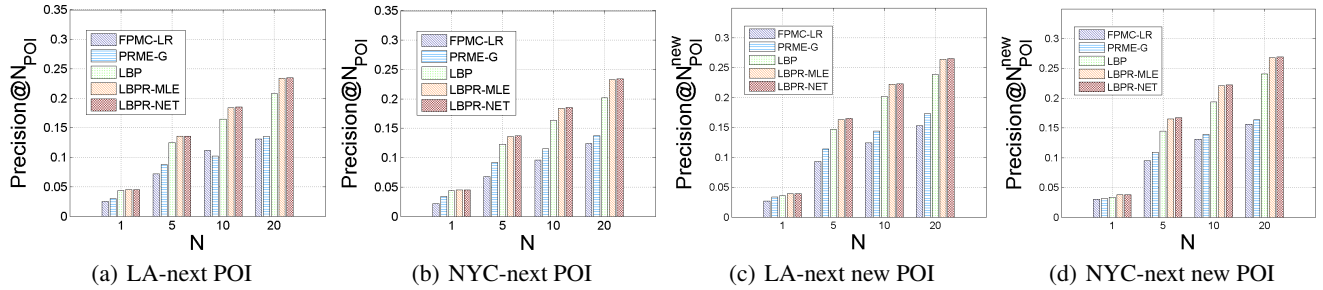


Figure 4: Performance Comparison on Location Recommendation

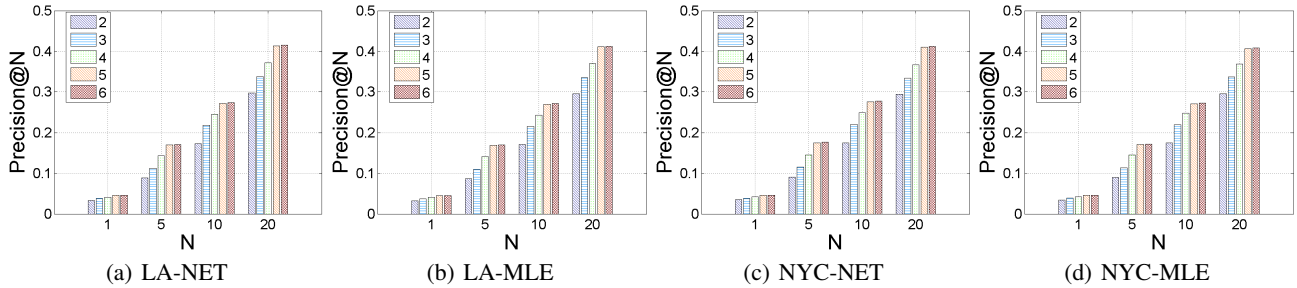


Figure 5: Maximum Length Impact of List

implies that categorical information is useful for modeling user’s specific preference in recommending new POIs. Moreover, our proposed models perform even better than LBP, which manifest the efficacy of explicitly usage of category to infer the user transition pattern in category-level. The effectiveness of recommending new POIs could benefit for advertising agency to launch a set of “categorical similar” advertisements to potential customers to help people explore more POIs.

5.4 List Length Impact

Fig.5 shows the category prediction results under different settings of the maximum length of the list. We can see that, for both datasets, the model accuracy increases with the increasing of the maximum length of list. When the maximum length reaches 6, the returns diminish. Thus 5 is the proper setting of list length for our datasets.

Note that setting the maximum length of the list as 2 does not make our model equivalent to the conventional pairwise BPR. From Table 2 and Fig.5, we can observe that our proposed models for category prediction trained by setting the maximum length of list as 2 consistently outperforms FPMC and PRME for both datasets, while FPMC and PRME adopt

conventional BPR. That is because the pairs of the observed multiple check-ins and the pairs of the observed and unobserved check-ins are all taken as the training set for LBPR, whereas BPR only consider the pairs of observed and unobserved check-ins. Furthermore, LBPR-NET performs slightly better than LBPR-MLE with much lower time complexity of $O(N \cdot K)$, whereas the time complexity of LBPR-MLE is $O(N \cdot K^2)$, N is the number of K -list instance.

6 Conclusion

In this paper, we address the problem of next POI recommendation and propose a two-fold approach to solve the recommendation task. First, we propose a Listwise Bayesian Personalized Ranking approach and apply it to optimize a categorical dimensional tensor for next category prediction. Then, successive POI recommendation is achieved by considering the spatial influence and the previously predicated category preference. Extensive experiments on real LSBN datasets demonstrate our proposed methods outperform several state-of-the-art methods by a large margin.

References

- [Cao *et al.*, 2007] Zhe Cao, Tao Qin, Tie-Yan Liu, Ming-Feng Tsai, and Hang Li. Learning to rank: from pairwise approach to listwise approach. In *Proceedings of the 24th international conference on Machine learning*, pages 129–136. ACM, 2007.
- [Cheng *et al.*, 2013] Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, pages 2605–2611. AAAI Press, 2013.
- [Cho *et al.*, 2011] Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 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, pages 1082–1090, New York, NY, USA, 2011. ACM.
- [Feng *et al.*, 2015] Shanshan Feng, Xutao Li, Yifeng Zeng, Gao Cong, Yeow Meng Chee, and Quan Yuan. Personalized ranking metric embedding for next new poi recommendation. In *Proceedings of the 24th International Conference on Artificial Intelligence*, pages 2069–2075. AAAI Press, 2015.
- [Gao *et al.*, 2015] Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. 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.*, pages 1721–1727, 2015.
- [He *et al.*, 2016] Jing He, Xin Li, Lejian Liao, Dandan Song, and William K Cheung. Inferring a personalized next point-of-interest recommendation model with latent behavior patterns. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
- [Hu *et al.*, 2008] Yifan Hu, Yehuda Koren, and Chris Volinsky. Collaborative filtering for implicit feedback datasets. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pages 263–272. Ieee, 2008.
- [Koren *et al.*, 2009] Yehuda Koren, Robert M. Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. volume 42, pages 30–37, Aug 2009.
- [Lan *et al.*, 2009] Yanyan Lan, Tie-Yan Liu, Zhiming Ma, and Hang Li. Generalization analysis of listwise learning-to-rank algorithms. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 577–584. ACM, 2009.
- [Li *et al.*, 2016] Huayu Li, Yong Ge, and Hengshu Zhu. Point-of-interest recommendations: Learning potential check-ins from friends. In *Proceedings of the 22th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2016.
- [Liu *et al.*, 2016] Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. Predicting the next location: A recurrent model with spatial and temporal contexts. In *AAAI*, pages 194–200, 2016.
- [Marden, 1996] J. I. Marden. *Analyzing and modeling rank data*. CRC Press, 1996.
- [Mnih and Salakhutdinov, 2007] Andriy Mnih and Ruslan Salakhutdinov. Probabilistic matrix factorization. In *Advances in neural information processing systems*, pages 1257–1264, 2007.
- [Pan *et al.*, 2008] Rong Pan, Yunhong Zhou, Bin Cao, Nathan N Liu, Rajan Lukose, Martin Scholz, and Qiang Yang. One-class collaborative filtering. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pages 502–511. IEEE, 2008.
- [Rendle *et al.*, 2009] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, 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 *Proceedings of the 19th international conference on World wide web*, pages 811–820. ACM, 2010.
- [Xia *et al.*, 2008] Fen Xia, Tie-Yan Liu, Jue Wang, Wensheng Zhang, and Hang Li. Listwise approach to learning to rank: theory and algorithm. In *Proceedings of the 25th international conference on Machine learning*, pages 1192–1199. ACM, 2008.
- [Zhang and Chow, 2015] Jia-Dong Zhang and Chi-Yin Chow. Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 443–452. ACM, 2015.
- [Zhao *et al.*, 2016a] Shenglin Zhao, Irwin King, and Michael R Lyu. A survey of point-of-interest recommendation in location-based social networks. *arXiv preprint arXiv:1607.00647*, 2016.
- [Zhao *et al.*, 2016b] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R Lyu, and Irwin King. Stellar: spatial-temporal latent ranking for successive point-of-interest recommendation. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, pages 315–321. AAAI Press, 2016.