Next and Next New POI Recommendation via Latent Behavior Pattern Inference

XIN LI, DONGCHENG HAN, JING HE, and LEJIAN LIAO, Beijing Institute of Technology, China
MINGZHONG WANG, The University of the Sunshine Coast, Australia

Next and next new point-of-interest (POI) recommendation are essential instruments in promoting customer experiences and business operations related to locations. However, due to the sparsity of the check-in records, they still remain insufficiently studied. In this article, we propose to utilize personalized latent behavior patterns learned from contextual features, e.g., time of day, day of week, and location category, to improve the effectiveness of the recommendations. Two variations of models are developed, including GPDM, which learns a fixed pattern distribution for all users; and PPDM, which learns personalized pattern distribution for each user. In both models, a soft-max function is applied to integrate the personalized Markov chain with the latent patterns, and a sequential Bayesian Personalized Ranking (S-BPR) is applied as the optimization criterion. Then, Expectation Maximization (EM) is in charge of finding optimized model parameters. Extensive experiments on three large-scale commonly adopted real-world LBSN data sets prove that the inclusion of location category and latent patterns helps to boost the performance of POI recommendations. Specifically, our models in general significantly outperform other state-of-the-art methods for both next and next new POI recommendation tasks. Moreover, our models are capable of making accurate recommendations regardless of the short/long duration or distance.

CCS Concepts: • **Information systems** → **Collaborative filtering**; • **Applied computing** → *Sociology*;

Additional Key Words and Phrases: Next POI recommendation, next new POI recommendation, latent behavior patterns

ACM Reference format:
Xin Li, Dongcheng Han, Jing He, Lejian Liao, and Mingzhong Wang. 2019. Next and Next New POI Recommendation via Latent Behavior Pattern Inference. *ACM Trans. Inf. Syst.* 37, 4, Article 46 (September 2019), 28 pages.
https://doi.org/10.1145/3354187

1 INTRODUCTION

Location-based social networks (LBSNs) enable Internet users worldwide to track and share their locations in the physical world. The fusion of physical and virtual world in LBSNs has attracted great interest from both the industry and users [16, 39]. Various LBSNs have been deployed, such

This work has been partially supported by National Key R&D Program of China under Grant No. 2017YFB0803300, and NSFC under Grant No. 61772074.
Authors' addresses: X. Li, D. Han, J. He, and L. Liao, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing, 100081, China; emails: {xinli, dchan, skyhejing, liaolj}@bit.edu.cn; M. Wang, The University of the Sunshine Coast, 90 Sippy Downs Drive, Sippy Downs, Queensland, Australia; email: mwang@usc.edu.au.

1046-8188/2019/09-ART46 $15.00
https://doi.org/10.1145/3354187

as Foursquare,[1] Gowalla, Facebook Place,[2] and GeoLife. They allow users to check in at a physical location, or point-of-interest (POI), with mobile devices and share their journeys and experiences online. For example, as of May 2018, Foursquare recorded 12B check-ins at more than 105M venues by more than 50M Foursquare users worldwide.[3] A vast amount of check-in data from millions of LBSN users can serve as a precious treasure to study user behaviors and patterns, and in turn to boost the efficiency of businesses and societies.

POI recommendation, which recommends places to users, becomes one of the most important tasks in LBSNs, since it can help to improve user viscosity to LBSN service providers and to support personalized advertisement to potential clients. In the broader context, accurate POI recommendations are an essential component of computing systems dealing with users and locations, such as urban computing [56, 63], behavior informatics [6, 52, 62], and disease control [21].

POI recommendation has attracted extensive interests in both academia and industry [24, 55]. However, personalized POI recommendation still remains a challenging task, because the data collected for each user is highly sparse. That is, the number of locations in LBSNs is too large for an individual user, and each user can only visit a limited number of locations; therefore, the corresponding check-in records account for an extremely small proportion in the connections between a user and all the sites. What is more, users usually do not do check-ins on LBSNs for every venue they have visited, which makes the data is even sparser.

Collaborative filtering (CF), which is commonly used in recommendation systems, has been extended to predict users' preferences to POIs with the check-in data. Related solutions can be categorized as memory-based [49, 50, 54] and model-based [12, 13, 33]. Memory-based CF solutions are apparently disadvantaged by the data sparsity problem due to their reliance on the user-user or item-item similarities from the check-in records. In comparison, model-based CF solutions try to learn the patterns or models with user ratings for the task of POI recommendations. For example, Cheng et al. [12] introduced a multi-center Gaussian model to recommend POIs, following the observation that the majority of check-in records can be attributed to a few centers. However, most models overlook the consecutive relationships between check-ins, which are in fact important in POI recommendation, as human movements are naturally sequential. That is, an effective POI recommendation system should incorporate the temporal dependencies of locations.

Consequently, next POI recommendation, or successive POI recommendation, was proposed to predict the next move of a user with respect to his/her "current" location. The features and sparsity—which indicate there is insufficient check-in records for a specific user—of check-in data prevent the direct application of conventional recommendation solutions onto next POI recommendation. First, unlike a conventional five-star rating system that captures both "like" and "dislike" feedback from users, check-in records only represent positive or "like" samples. A POI without check-in might be a piece of negative feedback (the location is unattractive to the user) or a missing value (the user might visit the location in the future). As a result, the requirement of inferring user preferences from the implicit feedback makes POI recommendation a harder task, which also brings the same problem to next POI recommendation. Second, considering the users' visiting history as a user-POI check-in matrix, its sparsity is dramatically higher than the user-item rating matrix in Netflix data [53]. Moreover, next POI recommendation needs to model the successive check-in behaviors that add a new dimension of temporal constraint, making the data sparsity even more troublesome.

[1] https://foursquare.com.
[2] https://www.facebook.com/places.
[3] https://foursquare.com/about.

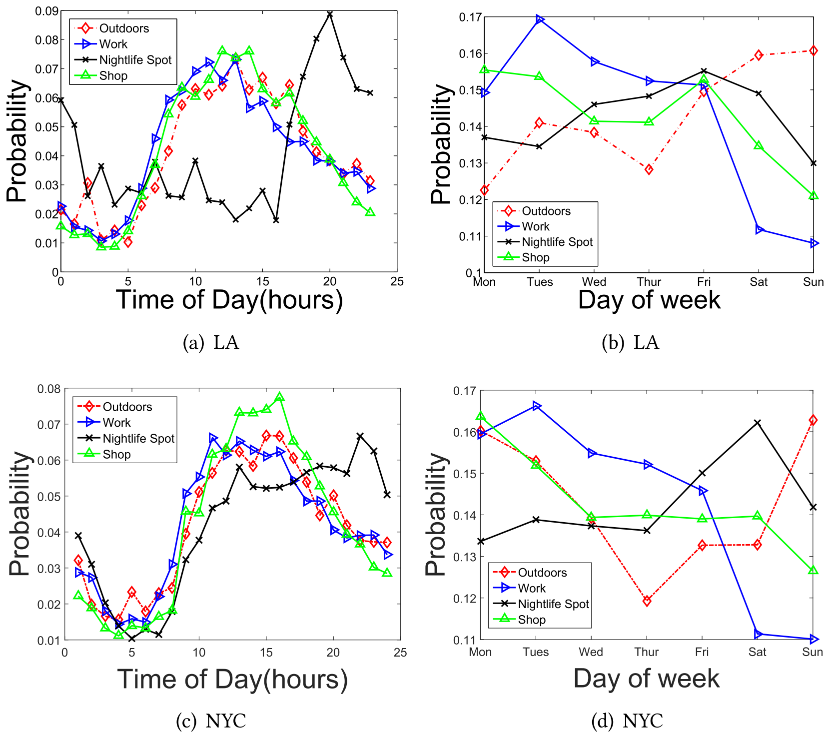
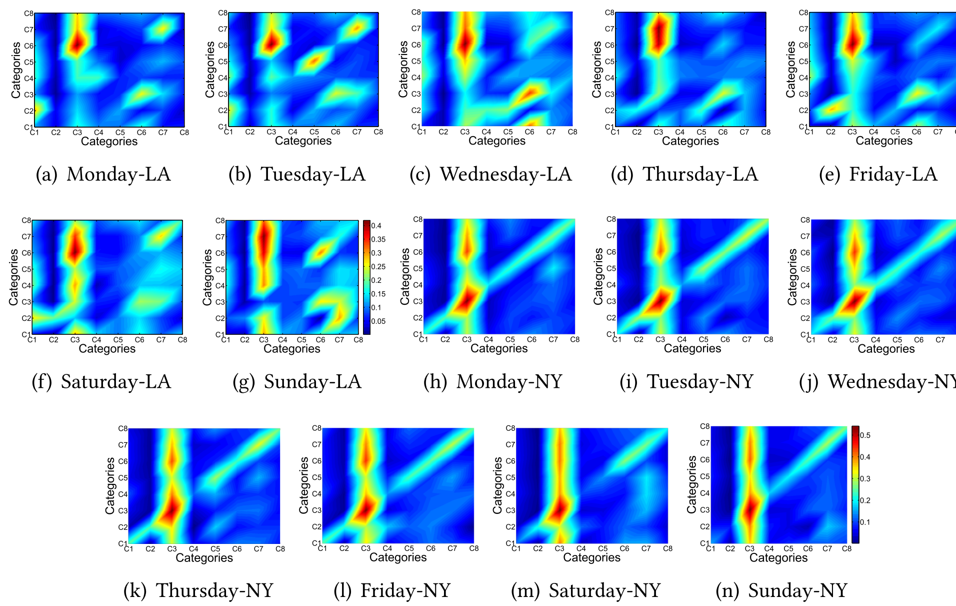


Fig. 1. Temporal probabilities for categorical check-ins.

Various studies have shown that human mobility has periodic or co-related characteristics [14, 20, 34]. For example, people regularly stop by coffee shops on their way to work in the morning, which can be explained as a transition pattern from coffee shop to workplace on weekday mornings. Or, after intense outdoor activities, such as hiking and running, users are more likely to have high-protein meals in restaurants, such as a steakhouse rather than a juice bar. Therefore, we propose to utilize the transition patterns based on contextual features, such as time of day, day of week, and location category, to address the aforementioned challenges in next POI recommendation. Observations on the Foursquare check-in data in LA and NY (details are provided in Section 3) reveal that there are some latent behavior patterns on top of contextual features in check-in data, and they might be helpful to boost the effectiveness of POI recommendation.

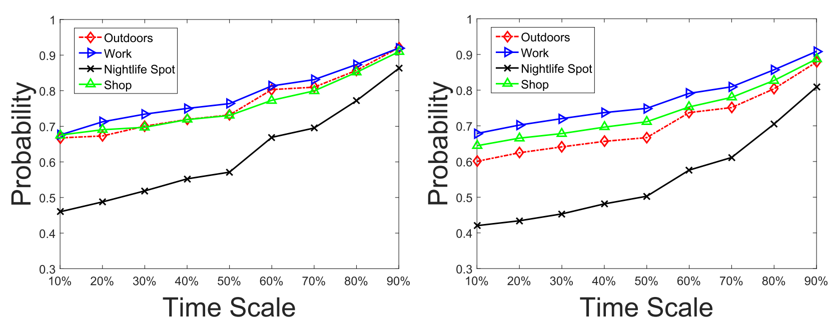
The first observation focuses on the probability of visiting an individual location category at a different time of the day or week. Figures 1(a), 1(c), and Figures 1(b), 1(d) illustrate the probabilities of visiting the top-four most popular location categories during the day and the week, respectively. Some evident patterns can be found in the figures. For example, the probabilities for check-ins in LA and NY reveal that people often choose to work, shop, or do outdoor activities during the daytime and visit nightlife places at night. For another example, people often visit workplaces on weekdays and go to nightspots and other outdoor places at weekends. The findings suggest it may be beneficial to learn latent patterns from the dependency between the time and the categories of locations.

The second observation focuses on the transition probabilities between different location categories. Figures 2(a)–2(g) and Figures 2(h)–2(n) demonstrate the transition probabilities between location categories in LA and NY, respectively, on each day in a week. Similarities in transition preferences are evident across days in each city, indicating that there are some latent transition



(a) Monday-LA (b) Tuesday-LA (c) Wednesday-LA (d) Thursday-LA (e) Friday-LA (f) Saturday-LA (g) Sunday-LA (h) Monday-NY (i) Tuesday-NY (j) Wednesday-NY (k) Thursday-NY (l) Friday-NY (m) Saturday-NY (n) Sunday-NY

Fig. 2. Transition Probabilities between Location Categories. 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}.



(a) LA (b) NYC

Fig. 3. Accumulated probability of new POIs for location categories.

patterns that might play a key role for next POI recommendation. However, the transition preferences of LA users are apparently different from those of NY users, suggesting that different user groups, such as being identified by the city, have their own features. Therefore, local patterns, which are learned from users sharing the same contextual scenarios, should be incorporated into global patterns, which are learned from all users' visit history. That is, POI recommendation needs to address not only global patterns but also personalized patterns.

The third observation shows that the next POI has a considerable probability to be new for the user, resulting in a more challenging task of next new POI recommendation. In fact, this task becomes increasingly popular and useful, since it not only helps users to explore interesting new places in the city but also creates the opportunities for businesses to increase their revenues by attracting and discovering potential customers. Figure 3 shows the accumulated probability regarding the time that the next POI is new for the top-four most popular location categories. For example, the category of Nightlife Spot has the highest ratio (the increment between two time points) for new POIs, indicating that users are more likely to visit new nightspots. Along with the time, each category has a distinguishable ratio of new POIs. Traditional recommendation systems have limited support for next new POI recommendation, because they rely on past routine visits. In contrast, patterns learned on categorical preferences and transitions can help to predict new POIs.

Following these observations, we propose a method to utilize latent behavior patterns learned from contextual features, e.g., time of day, day of week, and location category, to improve the effectiveness of next and next new POI recommendations. In comparison with current solutions, we introduce a third-rank tensor to model the dependencies in check-in behaviors. A soft-max function is applied to integrate the personalized Markov chain with the latent patterns, and a sequential Bayesian Personalized Ranking (S-BPR) [42] approach is used to derive the optimization criterion accordingly. Expectation Maximization (EM) [41] is deployed to estimate the model parameters during the model learning phase.

This article extends our previous Global Pattern Distribution Model (GPDM) in Reference [25] and proposes a new Personalized Pattern Distribution Model (PPDM) to learn the unique pattern distribution for each user, thus providing personalized next POI recommendations. GPDM captures the users' implicit preferences in Next POI assuming all users globally share the same pattern distribution. However, the global pattern distribution may not be appropriate for an individual user, because LBSN users vary largely with respect to diverse features, such as age, gender, home city, and occupation. Therefore, PPDM is proposed to address the fact that different users may have personalized pattern distribution under the same contextual scenario. Moreover, in comparison with Reference [25], this article also studies Next New POI recommendation on top of GPDM and PPDM, in addition to next POI recommendation. Next New POI recommendation studies the challenge of recommending new POIs to users, instead of selecting from users' past visits.

The main contributions of this article include:

- Two variations of models are developed to utilize latent behavior patterns for POI recommendation, including GPDM [25], which learns a fixed pattern distribution for all users; and PPDM, which learns personalized pattern distribution for each user. The corresponding optimization criteria and learning steps/tricks have been carefully studied.
- In addition to next POI recommendation task, the proposed models can achieve next new POI recommendation task simultaneously, instead of just selecting from users' past visits. This feature makes our models more practical in real applications, because discovering new destinations is a common practice in human behavior and provides valuable guidance to business operations.
- The models have been extensively experimented on three large-scale LBSN data sets, showing that they in general significantly outperform other state-of-the-art methods for both next and next new POI recommendation tasks. The experiments also show that our models are capable of making accurate recommendations regardless of the short/long duration or distance.

2 RELATED WORK

Location recommendation is used to analyze GPS trajectory logs collected from monitored users [45–47, 64]. However, the recent studies have shifted the focus to POI recommendation due to the easy access of check-in data in LBSNs, which can be roughly classified into four major categories in terms of the main influencing factors.

(1) ***Temporal POI recommendation*** studies the temporal influence on POIs to improve the recommendation performance. Reference [1] classified the contextual factors depending on what

the recommender system knows about them and whether they change over time or not. Reference [23] studied the features of temporal non-uniformness and consecutiveness to utilize the cyclic patterns in users' behavior. Reference [8] managed to balance between the intrinsic interest and the temporal context in user behaviors. Reference [54] incorporated time influence in the user-based collaborative filtering model. Reference [51] explicitly modeled the dynamics of user interests and temporal context to capture users' changing interests. Reference [19] proposed the Recurrent Marked Temporal Point Process to simultaneously model time and marker information. Reference [48] utilized human mobility data to profile the temporal popularity of POIs and developed a user-POI temporal matching model to conclude users' temporal regularity. However, much existing work neglects the temporal orders or periodic features between past visits and is incapable of recommending new POIs. In comparison, we utilize temporal features, including the time of day and the day of week, to learn personalized latent behavior patterns, and our model addresses both next and next new POI recommendations.

(2) ***Geographical influence enhanced POI recommendation*** studies the "geographical clustering phenomenon" in human activities to improve the POI recommendation system. Reference [29] is the CF-based recommender achieved by matrix factorization (MF). Reference [50] also proposed a user-based CF model considering geographical influence; however, it is not generic in incorporating other factors such as sequential patterns. Reference [60] proposed to incorporate the geographical influence into the pairwise preference ranking method. Reference [2] modeled the location's context information; however, their capacity limitation also lies in dealing with other kinds of contexts. Reference [35] proposed a geographical probabilistic factor analysis framework including Gaussian distribution and Bayesian non-negative MF. Reference [32] proposed a ranking-based MF model incorporating users' geographical preferences through using a latent matrix. Reference [57] used two-dimensional Kernel Density Estimation with an adaptive bandwidth to model personalized geographical influence. However, factorization methods neglect the implicit feedback and the underlying properties/implicit patterns of users' sequential behaviors. In comparison, we consider both personal and spatial preferences along with the integration of geographical influence to estimate the transition probability for each user, thus improving the recommendation performance.

(3) ***Content-aware POI recommendation*** matches up the attributes of a user profile, in which preferences and interests are stored, with the attributes of each POI to recommend interesting locations to a user [37]. Reference [11] proposed to mine user interests from short text and map between user interests and locations. Reference [24] modeled sentiment indications, user interests and POI properties under a unified framework to improve the performance of POI recommendations. However, semantic analysis has limited application in LBSNs, because most texts/comments in LBSNs are short and contextually ambiguous.

(4) ***Social influence enhanced POI recommendation*** studies the impact of social relationships on POI recommendation with the assumption that friends tend to share more common interests. However, some previous study reported that a large number of friends share nothing in terms of POI [49]. It was also reported that the social relationships explain only 10% to 30% of all human movement [15]. Therefore, an extension is to consider the more general concept of groups, instead of merely friend relationships. Reference [55] introduced Group-Sparse MF to convert the rating matrices of multiple behaviors into a user and item latent factor space. Reference [27] grouped users with similar interests according to their frequently visited locations' category hierarchy.

Methodologically, both traditional collaborative filtering technologies and neural network–based technologies have contributed to the recent advance of POI recommendation.

Neural network–based POI recommendation exploits the application of deep networks to boost the performance of POI recommendations. Reference [26] integrated visual signals into personalized recommendations by utilizing visual features extracted from images using (pre-trained) Deep Convolutional Neural Network (Deep CNN). Reference [38] proposed a Deep Recurrent Collaborative Filtering framework (DRCF) model including RNN layers to encapsulate dynamic user preferences for venue recommendation.

Collaborative Ranking (CR)–based POI recommendation predicts the ranking of potential POIs for recommendations. Reference [4] proposed point-wise models and pair-wise models for the task. Reference [17] considered three formulations based on collaborative p-norm push, infinite push, and reverse-height push. Reference [28] proposed an ordinal classification framework that decomposes the task into binary classification problems. Reference [3] proposed CR algorithms based on users' preferences for POI recommendation. However, these methods do not address the cold-start and data sparsity problems well. Furthermore, they neglect the sequential influence of check-in activities and cannot model users' preferences accurately. In comparison, we introduce a third-rank tensor to model the dependencies in check-in behaviors, with the consideration that an effective POI recommendation system should incorporate the temporal dependencies of locations.

In comparison with POI recommendation, next POI recommendation, which targets at predicting the exact next visit based on a user's prior location, is still an emerging task that is also more challenging. Some methods designed ranking models for next POI recommendation by considering sequential transition and user preference [22], spatial-temporal influence [61], or location category [9]. References [31, 36, 59] used neural networks to model contextual factors for next POI recommendation, for example, Reference [36] applied Recurrent Neural Networks (RNN) [58] by modeling spatial temporal contexts in each layer. The embedding models are also used for recommending next POI, in which Reference [10] cast objects, locations, and time slots to a low-dimensional latent space, Reference [7] modeled two layers to capture the geographical influence and the characteristics information of a POI. However, the periodicity of check-in data and categorical influence are still not well studied. Moreover, to deal with the data sparsity challenge, most solutions simply remove the venues far from the previous checked-in POI to get a valid candidate set of POIs. Hence, they cannot provide personalized POI recommendations when the check-in records of a user have many venues missing.

The tensor-based FPMC-LR model [13] including the personalized Markov chains and the localized region constraint was proposed to recommend a successive personalized POI. However, FPMC-LR neglects the experience of users whose check-in behavior patterns are different from the majority and cannot predict POIs that are not close. Differently from these studies, our models consider both spatial and temporal influences and learn personalized latent pattern distribution for each user. Moreover, as we use the transition probability estimation in the models, we can recommend locations far from the current location when the personal preference outweighs the spatial preference.

3 DATA DESCRIPTION AND FEATURES

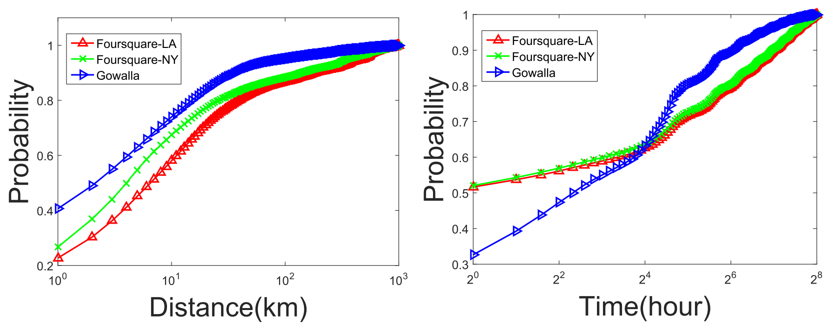
This section performed empirical analysis on three real-world LBSN data sets to explore key features in users' successive check-in behaviors, including spatial influence, temporal influence, check-in counts, and exploration for new locations. They provide key concepts and definitions to understand the proposed model.

3.1 Data Sets

We selected three large-scale check-in data sets from real-world LBSNs, including Foursquare Los Angeles and Foursquare New York City from Reference [5], and Gowalla from Reference [12] with a complete snapshot.

Table 1. Data Set Summary

	#User	#POI	#Check-in	#Avg.check-in	Sparsity
Fours.-LA	2,823	84,937	130,583	46.25	5.81×10^{-10}
Fours.-NY	2,579	97,013	157,404	61.03	5.85×10^{-9}
Gowalla	1,388	112,351	301,678	217.35	1.01×10^{-6}



(a) Geographical distance (b) Time interval

Fig. 4. CDF of geographical distance and time interval of two consecutive check-ins.

For each data set, users who checked in less than 10 times were removed. Thereafter, each data set was split into two non-overlapping sets by using the first 80% check-ins of each user as the training set and the remaining 20% check-ins as the test set to evaluate the performance of different algorithms. Table 1 lists the statistic summary of each data set.

Gowalla data set does not have categorical information associated with POIs. Foursquare data sets provide categorical descriptions for POIs at different levels of granularity. For example, "Arts & Entertainment" category consists of subcategories, such as "Art Museum" and "Jazz Club." For the purpose of this article, we focus on eight top-level categories following Reference [7].

3.2 Spatial and Temporal Influence

Figure 4(a) shows the cumulative distribution function (CDF) regarding the geographical distance of two consecutive check-ins. The CDF curve increases faster when the distance is shorter, which suggests that most users' movements occur within a limited area, and the probability of being the next POI is inversely proportional to its geographic distance.

Figure 4(b) shows the CDF regarding the time interval of two consecutive check-ins. In addition to our experience that the selection of next POI is highly correlated to the time interval, many consecutive check-ins occur in a longer time. For more than 10% of consecutive check-ins, their time differences are larger than 48 hours. Moreover, the study of temporal dependencies between the location categories of successive POIs also demonstrates a strong correlation between them. For example, Figure 2 shows that Food is commonly visited after Shop, as users would like to have lunch/dinner after shopping. This observation suggests that transition probability between categories, and Markov chain can be used to improve next personalized POI recommendation.

3.3 Check-in Counts

Figure 5 shows the CDF regarding the number of check-ins at each POI, which demonstrates that up to 90% of POIs are visited less than four times in each data set. There are over 23% of Foursquare

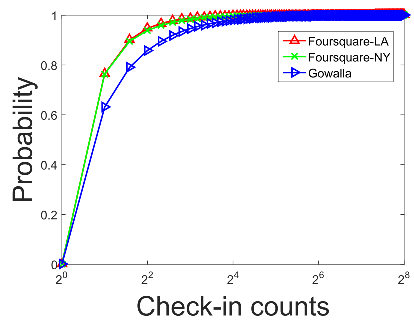


Fig. 5. Check-ins probability vs. counts.

POIs and 35% of Gowalla POIs are checked more than once, which suggests that users' check-in activities have some periodic patterns. Based on the observations in Figure 4(a), Figure 4(b), and Figure 5, it can be concluded that most POIs are visited occasionally within a short distance and within a short time interval. Therefore, both spatial and temporal influences should be combined for next POI recommendation.

4 PROBLEM DEFINITION

Let $U = \{u_1, u_2, \ldots, u_M\}$ be the user set of LBSN, and $L = \{l_{1,2}, \ldots, l_N\}$ be the POI or location set. Each POI is denoted as $l_i = \{longitude_i, latitude_i\}$. Let L_u^i be the set of POIs visited by user u at time i, and L_u be the set of POIs visited by user u before time t, then $L_u = L_u^1 \cup \cdots \cup L_u^{t-1}$. The contextual feature vector is defined as $\mathbf{g}(\mathbf{c}) = \{g_1(c), \ldots, g_F(c)\}$, which represents a specific contextual scenario $\mathbf{c}$. The contextual features include previous location, time of day, day of week, previous location's category, and so on. F denotes the number of features.

Assuming there are K latent behavior patterns determined by contextual scenarios, the pattern distribution can be represented as $\Pi = (\pi_1, \ldots, \pi_K)$, *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 preferences to recommend top-N next POIs to user u.

5 PROPOSED METHOD

To recommend next personalized POI, the proposed model computes and ranks the conditional probability of each potential destination regarding the current context. Based on the first-order Markov chain property, the probability that user u will move from location i to l can be defined as:

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

where $\mathbf{c}$ denotes the contextual scenario.

A transition tensor $\chi \subseteq [0, 1]^{|U| \times |L| \times |L|}$ with each element $\chi_{u,i,l}$ representing the observed transition record of user u from location i to location l can be generated from check-in records. χ^u is used to denote the transition matrix for user u. To improve the recommendation performance, both the personal and spatial preference are considered.

Personal Preference. Tucker Decomposition (TD) is a linear factorization model that is generally used to estimate the transition tensor χ:

$$\hat{\chi} = C \times U \times I \times L, \tag{2}$$

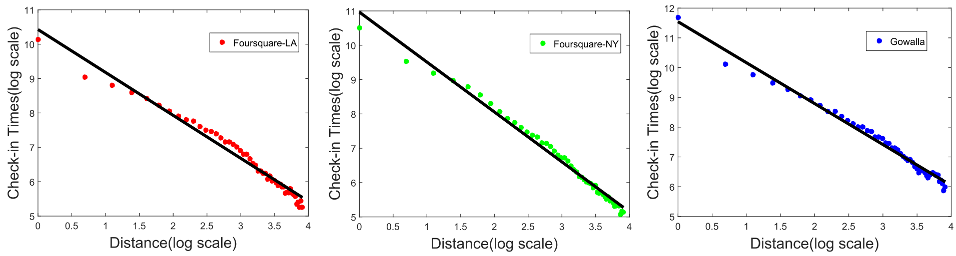


Fig. 6. Spatial preference on check-ins.

in which C is a core tensor, U represents the feature matrix of users, I is the feature matrix of source locations, and L is the feature matrix of target locations. As χ is partially observed, a special case of Canonical Decomposition, which is a low-rank factorization model capable of modeling the pairwise interaction between all three tensor modes (U, I, and L), is applied to estimate the missing information, computed as:

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

where $u_{U,L}$ and $l_{L,U}$ represent the latent factor vectors for users and target locations, respectively. $l_{L,I}$, $i_{I,L}$, $u_{U,I}$, and $i_{I,U}$ have similar meanings. The component $u_{U,I} \cdot i_{I,U}$ can be ignored, because it has no dependency on l and no impact on the ranking result [43]. Therefore, $\hat{x}_{u,i,l}$ can be simplified as:

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

In comparison with TD, Equation (4) has noticeably lower complexity for prediction and learning. Moreover, even though TD is also capable of modeling pairwise interactions, it cannot be properly identified by standard regularization estimation procedures [44].

Spatial Preference. Human mobility is evidently constrained by the distance one can travel within a day, and the probability of visiting a place decreases if the distance to that place increases [14]. Moreover, most POIs that a user visits are close to the POIs that the user frequently visits, such as the home and office.

Figure 6 plots the number of successive check-ins regarding the distance between them in the log scale for each data set. A clear power-law distribution between the successive check-ins and their distance can be observed, which proves that the locations that a user visited are geographically dense.

The spatial preference, $sp(d_{i,l})$, of visiting a POI with $d_{i,l}$ km distance in our model is defined as

$$sp(d_{i,l}) = a \times d_{i,l}^{k}, \tag{5}$$

in which a and k represent the parameters in the power-law distribution.

In comparison, GeoSoCa [57] applied power-law distributions to model all three aspects, including Geographical, Social, and Categorical correlations for POIs. Specifically, for geographical correlations, GeoSoCa modeled the check-in distribution with the adaptive kernel estimation method, in which the key parameters in the power-law distribution are check-in frequency and density. Our model is much simpler by using the distance between the current position and next POI as the parameter in the power law.

Equation (5) can be equally converted into the linear format by applying logarithmic on both sides, which in turn can be used to learn a and k with the least-square regression. That is,

$$\log sp(d_{i,l}) = \log a + k \cdot \log d_{i,l}. \tag{6}$$

Our experiments learned that $a = 10.5$ and $k = -1.25$ for Foursquare-LA, $a = 11.0$ and $k = -1.45$ for Foursquare-NY, and $a = 11.5$ and $k = -1.37$ for Gowalla data set. Therefore, $a = 11$ and $k = -1$ were chosen as the empirical parameters for Equation (5), resulting in:

$$sp(d_{i,l}) = 11 \times d_{i,l}^{-1}. \tag{7}$$

Since both personal and spatial preferences should be considered, they are combined linearly in our model to define the transition probability estimation:

$$\hat{x}_{u,i,l} = u_{U,L} \cdot l_{L,U} + l_{L,I} \cdot i_{I,L} + \rho \cdot d_{i,l}^{-1}. \tag{8}$$

ρ is used as a weighing factor to join the two preferences and a in Equation (5) becomes part of ρ. The optimal setting of ρ is learned during model inference phase. Equation (8) is capable of recommending locations far from the current location when the personal preference outweighs the spatial preference.

5.1 Incorporating Pattern-level Preferences

The observations discussed in the introduction show that user mobility is also impacted by the pattern-level preferences, which represent the impact of latent behavior patterns in addition to personal and spatial preferences. Therefore, the proposed model introduces a latent pattern layer to capture the users' implicit preferences in POI recommendation. Let s be the latent variable to represent the pattern-level influence. Then, the joint probability of $x_{u,i,l}$ and s is defined as:

$$p(L_{u,l}, s|\mathbf{c}) = p(L_{u,l}|s, \mathbf{c})p(s|\mathbf{c}), \tag{9}$$

where $p(s|\mathbf{c})$ is the mixing coefficient, i.e., π.

Based on Equation (8), the pattern-level preference for a latent variable s is defined as:

$$\begin{aligned} \hat{x}^s_{u,i,l} &= p(L_{u,l}|s, \mathbf{c}) \\ &= u^s_{U,L} \cdot l^s_{L,U} + l^s_{L,I} \cdot i^s_{I,L} + \rho^s \cdot d_{i,l}^{-1}. \end{aligned} \tag{10}$$

Considering all possible latent variables, the transition probability is defined as:

$$\hat{x}_{u,i,l} = \sum_s \hat{x}^s_{u,i,l} p(s|\mathbf{c}). \tag{11}$$

Figure 7 provides a visual explanation to the proposed model. The upper tensor is, in fact, the transition tensor χ that represents the transition probability based on the check-in records. For each user in χ, the cell (i, j) is "1" if the user visits location j after i. Otherwise, it is "?" In comparison, each block in the lower tensors denotes the transition probabilities regarding one latent variable. Each user may have distinct pattern-level transition tensors. It should be noted that transition tensor χ is a mixture of the pattern-level transition tensors, and $p(s|\mathbf{c})$ is the mixing coefficient. Therefore, the goal is to infer the proper pattern-level transition probabilities and pattern distribution to estimate the unobserved transition preferences.

A soft-max function $\frac{1}{S_c}\exp(\sum_{j=1}^F \alpha_j^s g_j(c))$ is applied to infer multi-patterns and $p(s|\mathbf{c})$. α_j^s is the weight associated with the jth feature for latent pattern s and S_c is the normalization factor that scales the exponential function to be a proper probability distribution Π, i.e., $S_c = \sum_{k=1}^K \exp(\sum_{j=1}^F \alpha_j^{s_k} g_j(c))$. Contextual scenario $\mathbf{c}$ is expressed as a bag of features $\{g_1(c), \ldots, g_F(c)\}$

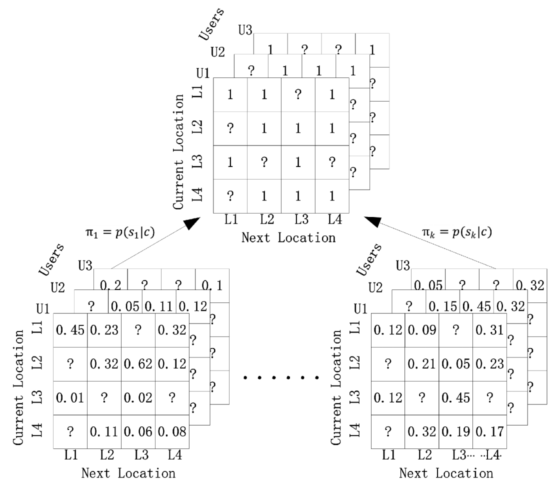


Fig. 7. Visual explanation to global pattern distribution model.

where F is the number of features. Applying the soft-max function into Equation (11), $\hat{x}_{u,i,l}$ is rewritten as:

$$\hat{x}_{u,i,l} = \frac{1}{S_c} \sum_s \hat{x}^s_{u,i,l} \exp\left(\sum_{j=1}^{F} \alpha^s_j g_j(c)\right). \tag{12}$$

Because the learned pattern distribution is identical for all users, this model is also named as global pattern distribution model (GPDM) in this article.

5.2 Optimization Criterion

Next POI recommendation aims at finding top-N POIs that are ranked by the probability that a user is likely to visit after the current location. In this way, the ranking order between candidate POIs is more relevant than the accurate probability values of choosing them. Therefore, we define a ranking operator $>^s_{u,i}$ between locations:

$$m >^s_{u,i} n \Leftrightarrow \hat{x}^s_{u,i,m} > \hat{x}^s_{u,i,n}. \tag{13}$$

In Equation (13), $m >^s_{u,i} n$ denotes that user u at location i prefers location m to location n as the next visit regarding pattern s. It is equivalent to $\hat{x}^s_{u,i,m} > \hat{x}^s_{u,i,n}$, which tells the transition probability for user u from location i to m regarding s is higher than to n.

Thereafter, the optimization criterion based on sequential Bayesian Personalized Ranking (S-BPR), which is similar to the general BPR approach [42], is applied. Mathematically, the best ranking for user u influenced by the pattern-level preference s is modeled as:

$$p\left(\Theta| >^s_{u,i}\right) \propto p\left(>^s_{u,i} |\Theta\right) p(\Theta), \tag{14}$$

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

The model parameters can be learned by maximizing the posterior with the assumption that the check-in history of each user is independent:

$$\underset{\Theta}{argmax} \prod_{u \in U} \prod_{i \in L_u} \prod_{m \in L^t_u} \prod_{n \notin L^t_u} \sum_s p\left(m >^s_{u,i} n|\Theta\right) p(s|\mathbf{c}) p\left(\Theta\right). \tag{15}$$

With Θ, the ranking probability between two locations is defined as:

$$p\left(m >_{u,i}^{s} n|\Theta\right) = p\left(x_{u,i,m}^{s} > x_{u,i,n}^{s}|\Theta\right) = p\left(x_{u,i,m}^{s} - x_{u,i,n}^{s} > 0|\Theta\right). \tag{16}$$

Following Reference [43], the logistic function $\sigma(z) = \frac{1}{1+e^{-z}}$ is used to model the ranking probability between location m and n for user u:

$$p\left(m >_{u,i}^{s} n|\Theta\right) = \sigma\left(x_{u,i,m}^{s} - x_{u,i,n}^{s}\right). \tag{17}$$

With the assumption that the prior distribution of the model parameters is Gaussian $\Theta \sim N(0, \frac{2}{\lambda_\Theta}I)$, the maximum a posterior (MAP) estimator for Θ can finally be defined as:

$$\underset{\Theta}{argmax} \prod_{u\in U} \prod_{i\in L_u} \prod_{m\in L_u^t} \prod_{n\notin L_u^t} \left\{\frac{1}{S_c} \cdot \sum_{s} \sigma\left(x_{u,i,m}^{s} - x_{u,i,n}^{s}\right) \exp\left(\sum_{j=1}^{F} \alpha_j^s g_j(c)\right) e^{-\frac{\lambda_\Theta}{2}||\Theta||^2}\right\}. \tag{18}$$

5.3 Model Inference

The logarithm format of Equation (18) can be used as the objective function to compute Θ:

$$\underset{\Theta}{argmax} \sum_{u\in U} \sum_{i\in L_u} \sum_{m\in L_u^t} \sum_{n\notin L_u^t} \ln\left\{\frac{1}{S_c} \cdot \sum_{s} \sigma\left(x_{u,i,m}^{s} - x_{u,i,n}^{s}\right) \exp\left(\sum_{j=1}^{F} \alpha_j^s g_j(c)\right) e^{-\frac{\lambda_\Theta}{2}||\Theta||^2}\right\}. \tag{19}$$

Expectation Maximization(EM) algorithm [18] is applied to learn the model parameters.

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

$$\gamma(s) = P\left(s| >_{u,i}^{s}, \Theta, \mathbf{c}\right) = \frac{\sigma(x_{u,i,m}^{s} - x_{u,i,n}^{s}) \exp(\sum_{j=1}^{F} \alpha_j^s g_j(c))}{\sum_s \sigma(x_{u,i,m}^{s} - x_{u,i,n}^{s}) \exp(\sum_{j=1}^{F} \alpha_j^s g_j(c))}. \tag{20}$$

And in the M-Step, $\boldsymbol{\alpha}^S$ and $\{\Theta \setminus \boldsymbol{\alpha}^S\}$ are computed by optimizing the Q-function of Equation (21) and Equation (22), respectively.

$$\boldsymbol{\alpha}^S = \underset{\boldsymbol{\alpha}^S}{argmax} \sum_{u\in U} \sum_{i\in L_u} \sum_{m\in L_u^t} \sum_{n\notin L_u^t} \sum_{s} \gamma(s) \cdot \left\{\ln\left(\frac{1}{S_c} \exp\left(\sum_{j=1}^{F} \alpha_j^s g_j(c)\right)\right) - \frac{\lambda_\Theta}{2}||\Theta||^2\right\}, \tag{21}$$

$$\{\Theta \setminus \boldsymbol{\alpha}^S\} = \underset{\{\Theta\setminus\boldsymbol{\alpha}^S\}}{argmax} \sum_{u\in U} \sum_{i\in L_u} \sum_{m\in L_u^t} \sum_{n\notin L_u^t} \sum_{s} \gamma(s) \cdot \left\{\ln \sigma\left(x_{u,i,m}^{s} - x_{u,i,n}^{s}\right) - \frac{\lambda_\Theta}{2}||\Theta||^2\right\}. \tag{22}$$

Algorithm 1 lists the key steps and parameter updating rules for computing Θ.

5.4 Personalized Pattern Distribution Model

Previous sections explain the learning process for the global pattern distribution model (GPDM), in which a fixed pattern distribution is learned with the global perspective for all users. However, one size does not fit all, and the global pattern distribution may not be appropriate for an individual user, because LBSN users vary largely with respect to diverse features, such as age, gender, home city, and occupation. Even with the same contextual scenario, different users may choose different next POIs, exhibiting personalized pattern distribution.

Therefore, instead of inferring a fixed pattern distribution Π for all users, personalized pattern distribution model (PPDM), a variation of GPDM, is proposed to learn the unique pattern distribution for each user. Figure 8 provides a visual explanation of PPDM. In comparison with GPDM,

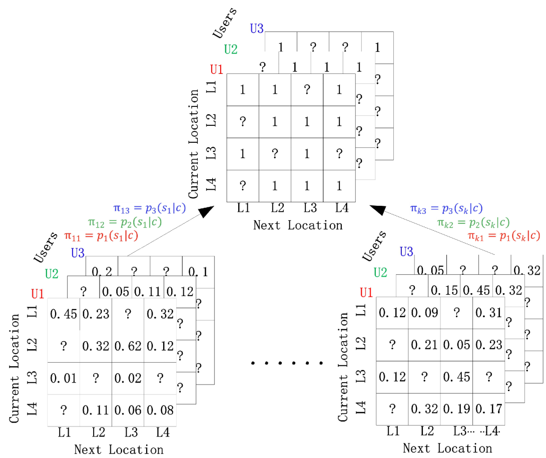


Fig. 8. Visual explanation of personalized pattern distribution model.

which has only $\pi_1, \ldots, \pi_k$ without considering individual users, PPDM decomposes each π_i into $\pi_{i1}, \ldots, \pi_{i|U|}$, representing a distinct pattern distribution for each user. The mixing coefficients, i.e., $p_u(s|\mathbf{c})$, are also personalized in the figure.

By inferring $\alpha^s_{u,j}$ as the personalized weight associated with the j_{th} feature for latent pattern s, the corresponding transition probability $\hat{x}_{u,i,l}$ for PPDM can be rewritten as

$$\hat{x}_{u,i,l} = \frac{1}{S_{u,c}} \sum_s \hat{x}^s_{u,i,l} \exp\left(\sum_{j=1}^{F} \alpha^s_{u,j} g_{u,j}(c)\right), \tag{23}$$

and the optimization function of $\boldsymbol{\alpha}^S$ in Equation (21) is rewritten as

$$\boldsymbol{\alpha}^S_u = \underset{\boldsymbol{\alpha}^S_u}{argmax} \sum_{i \in L_u} \sum_{m \in L^t_u} \sum_{n \notin L^t_u} \sum_s \gamma(s) \cdot \left\{ \ln\left(\frac{1}{S_c} \exp\left(\sum_{j=1}^{F} \alpha^s_j g_j(c)\right)\right) - \frac{\lambda_\Theta}{2} ||\Theta||^2 \right\}. \tag{24}$$

The equation for updating the parameters $\{\Theta \setminus \boldsymbol{\alpha}^S\}$ remains the same as in Equation (22).

Accordingly, the EM algorithm for PPDM is identical to Algorithm 1 with only Line 16 replaced with

$$\boldsymbol{\alpha}^s_u \leftarrow \frac{\sum_{u,d} \gamma(s) \cdot g(\mathbf{c}) \cdot (1 - p(s|\mathbf{c}))}{\lambda_\Theta \sum_{u,d} \gamma(s)}. \tag{25}$$

6 EXPERIMENTS

Extensive experiments were executed to answer the following questions:

(1) How do the proposed approaches, GPDM and PPDM, perform in comparison with other state-of-the-art next POI and next new POI recommendation solutions?
(2) When should GPMD or PPDM be selected?
(3) How do the proposed approaches perform regarding the time or distance constraint?
(4) How do the contextual features affect the model performance?
(5) How do the latent patterns affect model accuracy?

ALGORITHM 1: EM Algorithm for Model Parameter Learning

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:
  $\delta \leftarrow (1 - \sigma(x_{u,i,m}^s - x_{u,i,n}^s))$
  $u_{U,L}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (m_{L,U}^s - n_{L,U}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $i_{I,L}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (m_{L,I}^s - n_{L,I}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $m_{L,U}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot u_{U,L}^s}{\lambda_\Theta \sum_d \gamma(s)}$
  $n_{L,U}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (-u_{U,L}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $m_{L,I}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot i_{I,L}^s}{\lambda_\Theta \sum_d \gamma(s)}$
  $n_{L,I}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (-i_{I,L}^s)}{\lambda_\Theta \sum_d \gamma(s)}$
  $\boldsymbol{\alpha}^s \leftarrow \frac{\sum_d \gamma(s) \cdot g(\boldsymbol{c}) \cdot (1 - p(s|\boldsymbol{c}))}{\lambda_\Theta \sum_d \gamma(s)}$
  $\boldsymbol{\rho}^s \leftarrow \frac{\sum_d \delta \cdot \gamma(s) \cdot (d_{i,m}^{-1} - d_{i,n}^{-1})}{\lambda_\Theta \sum_d \gamma(s)}$
until convergence
return : Θ

6.1 Evaluation Metrics

The output of the recommendation system is $S_{N,u,rec}$, which is a list of top-N recommended POIs for user u in descending order of the probability. To evaluate the performance of next POI recommendation, the recall metric is applied:

$$Recall@N = \frac{1}{|U|} \sum_{u \in U} \frac{|S_{N,u,rec} \cap S_{visited}|}{|S_{visited}|}, \tag{26}$$

in which N denotes the size of the recommendation list, $|U|$ is the total number of users, and $S_{visited}$ denotes the set of locations visited by user u next to his/her prior location. For next POI recommendations, $S_{visited}$ has only one element for a given user and a given location.

Similarly, the recall metric for next new POI recommendation evaluation is defined as:

$$Recall@N_{new} = \frac{1}{|U|} \sum_{u \in U} \frac{|S_{N,u,rec} \cap S_{visited}^{new}|}{|S_{visited}^{new}|}, \tag{27}$$

in which $S_{visited}^{new}$ denotes the set of locations visited by user u for the first time next to his/her prior location. For next new POI recommendations, $S_{visited}^{new}$ has only one element for a given user and a given location.

Precision, which is a classical evaluation metric for recommendations, is not applied in the experiments. As there is only one correct answer in next and next new POI recommendation tasks, the numerator in the precision equation will always be 1 even if the recommendation is successful. That is, the precision keeps on decreasing with the increase of N, and the precision cannot be higher than 1/N. Therefore, we exclude Precision in the experiments, as it cannot reflect the effectiveness of the tasks and models.

Instead, we apply Normalized Discounted Cumulative Gain (NDCG@N) as the second evaluation metric. To evaluate the performance of next POI recommendation, it is defined as:

$$NDCG@N = \frac{1}{|U|}\sum_{u\in U}\frac{DCG@N(u)}{IDCG@N(u)}, \tag{28}$$

$$DCG@N(u) = \sum_{n=1}^{N}\frac{2^{Rel_u}-1}{\log_2(Ind_u+2)}, \tag{29}$$

in which Rel_u equals 1 if $|S_{N,u,rec} \cap S_{visited}| = 1$ and equals 0 otherwise. Ind_u is the index of $S_{visited}$ in $S_{N,u,rec}$, which ranges from 0 to $N-1$. $IDCG@N(u)$ is the ideal $DCG@N(u)$, which means the index value of $S_{visited}$ is 0.

Similarly, for next new POI recommendation evaluation, the NDCG metric is defined as:

$$NDCG@N_{new} = \frac{1}{|U|}\sum_{u\in U}\frac{\sum_{n=1}^{N}\frac{2^{Rel_u^{new}}-1}{\log_2(Ind_u^{new}+2)}}{IDCG@N(u)_{new}}, \tag{30}$$

in which Rel_u^{new} equals 1 if $|S_{N,u,rec} \cap S_{visited}^{new}| = 1$ and equals 0 otherwise, Ind_u^{new} is the index of $S_{visited}^{new}$ in $S_{N,u,rec}$, and $IDCG@N(u)_{new}$ means Ind_u^{new} is 0.

Due to the experiment setup, recall@1 and NDCG@1 have the same results. Therefore, we only report NDCG@N with N ∈ {5, 10, 20}.

6.2 Experiment Settings

To evaluate the proposed models, GPDM and PPDM, the following solutions are included for comparison:

- **MF** (matrix factorization) [30] factorizes the user-item preference matrix. It is included to represent conventional recommendation solutions.
- **PMF** (probabilistic matrix factorization) [40] extends MF with better support for data evolving with time.
- **FPMC-LR** [13] embeds both the personalized Markov chains and the localized regions for the task of successive personalized POI recommendation.
- **PRME-G** [22] combines two Euclidean distances in the latent space: the distance between the current and next location and the distance between the user and next location.
- **GeoSoCa** [57] exploits geographical, social, and categorical correlations between users and POIs.
- **Rank-GeoFM** [32] learns the factorization by fitting the ranking for POIs and incorporates different types of context information.
- **ST-RNN** [36] is a state-of-the-art model for next POI recommendation, which utilizes time-specific transition matrices and distance-specific transition matrices to model different time intervals and geographical distances.

In the experiments, three data sets were used for the evaluation. The detailed explanations to the data sets are provided in Section 3.

In the experiments, $\lambda_\Theta = 1$ for both FPMC-LR and our models, and the time window size was set to be 6h for FPMC-LR and PRME-G. The regularization term $\lambda = 0.03$ and component weight

$\alpha = 0.2$ for PRME-G following Reference [22]. The empirical setting of the number of latent behavior patterns was 4 for *Gowalla* data set and 6 for *Foursquare*.

The number of latent dimensions was set as 60 for all models in the experiments. The number was selected because the experiments showed that:

- our models achieved the best performance with 15 or more dimensions,
- FPMC-LR and PRME-G gained little improvement after 60 dimensions, and
- Rank-GeoFM, ST-RNN, and other models achieved their best results at 60 dimensions.

For other parameters, they were tuned on the training sets to find the optimal values and subsequently used in the test sets.

Tables 2–4 report the experiment results for next POI and next new POI recommendation on each data set.

6.3 Performance of Next POI Recommendation

The left side of Tables 2–4 lists *recall*@1, *recall*@5, *recall*@10, *recall*@20, *NDCG*@5, *NDCG*@10, and *NDCG*@20 for next POI recommendations achieved by each model.

The results show that:

- GeoSoCa, Rank-GeoFM, ST-RNN, and the proposed models outperform MF and PMF significantly and have better results than FPMC-LR and PRME-G. In terms of the metric of recall@N, GPDM has 91%–300% improvement on MF and 81%–200% on PMF. The results demonstrate that the spatial influence is a key factor for next POI recommendation. Conventional POI recommendation algorithms are not effective for next POI recommendation, because they focus on the user preference and have inadequate use of sequential dependencies between locations.
- In terms of the metrics of recall@N and NDCG@N, on average, both GPDM and PPDM consistently outperform Rank-GeoFM with 24% improvement and outperform ST-RNN with 17% improvement. The results demonstrate that the use of users' latent behavior patterns helps to better capture user mobility preferences in LBSNs, thus making next POI recommendation more accurate.

6.4 Performance of Next New POI Recommendation

The right side of Tables 2–4 list *recall*@1, *recall*@5, *recall*@10, *recall*@20, *NDCG*@5, *NDCG*@10, and *NDCG*@20 for next new POI recommendations achieved by each model.

The results show that:

- GeoSoCa, Rank-GeoFM, ST-RNN, and the proposed models outperform MF and PMF significantly and have better results than FPMC-LR and PRME-G. In terms of the metric of recall@N, GPDM has 123%–800% improvement on MF and 116%–646% on PMF. Conventional POI recommendation algorithms focus on tuning the latent factor vectors of users and locations to explain observed check-ins and recover unobserved check-ins. They assign low credit or weight to new POIs that have little preference from users in comparison with past visited POIs. Therefore, they cannot functionally recommend new POIs to users without extra information.
- Both FPMC-LR and PRME-G achieve better performance than MF and PMF due to their consideration of geographical influence. Moreover, PRME-G is better than FPMC-LR, since PRME-G has been customized to predict next new POIs by representing each POI as one point in latent space rather than two independent vectors. In addition, GeoSoCa and

Table 2. Performance Comparison on Foursquare-NY

Metrics	Next POI Recommendation									Next New POI Recommendation								
	MF	PMF	FP.-LR	PR.-G	G.SoCa	R.G.FM	ST-RNN	GPDM	PPDM	MF	PMF	FP.-LR	PR.-G	G.SoCa	R.G.FM	ST-RNN	GPDM	PPDM
R@1	0.011	0.015	0.022	0.034	0.030	0.025	0.036			0.004	0.006	0.030	0.032	0.027	0.024	0.030		
IMP.-G	300.0%	193.3%	100.0%	29.41%	46.67%	76.00%	22.22%	**0.044**	**0.045**	750.0%	466.7%	13.33%	6.250%	25.93%	41.67%	13.33%	**0.034**	**0.036**
IMP.-P	309.1%	200.0%	104.5%	32.40%	50.00%	80.00%	25.00%			800.0%	500.0%	20.00%	12.50%	33.33%	50.00%	20.00%		
R@5	0.034	0.041	0.068	0.092	0.073	0.086	0.091			0.021	0.030	0.095	0.109	0.096	0.105	0.114		
IMP.-G	261.8%	200.0%	80.88%	33.70%	68.49%	43.02%	35.16%	**0.123**	**0.129**	590.5%	383.3%	52.63%	33.03%	51.04%	38.10%	27.19%	**0.145**	**0.154**
IMP.-P	279.4%	214.6%	89.71%	40.22%	76.71%	50.00%	41.76%			633.3%	413.3%	62.11%	41.28%	60.42%	46.67%	35.09%		
R@10	0.051	0.062	0.096	0.115	0.105	0.128	0.138			0.040	0.057	0.131	0.139	0.140	0.152	0.177		
IMP.-G	219.6%	162.9%	69.79%	41.74%	55.24%	27.34%	18.12%	**0.163**	**0.169**	385.0%	240.4%	48.09%	39.57%	38.57%	27.63%	9.605%	**0.194**	**0.210**
IMP.-P	231.4%	172.6%	76.04%	46.96%	60.95%	32.03%	22.46%			425.0%	268.4%	60.31%	51.08%	50.00%	38.16%	18.64%		
R@20	0.072	0.083	0.124	0.137	0.141	0.187	0.156			0.076	0.095	0.156	0.164	0.169	0.181	0.195		
IMP.-G	180.6%	143.4%	62.90%	47.45%	43.26%	8.021%	29.49%	**0.202**	**0.212**	228.9%	163.2%	60.26%	52.44%	47.93%	38.12%	28.21%	**0.250**	**0.265**
IMP.-P	194.4%	155.4%	70.97%	54.74%	50.35%	13.37%	35.90%			248.7%	178.9%	69.87%	61.59%	56.80%	46.41%	35.90%		
N.@5	0.019	0.022	0.054	0.060	0.047	0.062	0.075			0.015	0.020	0.057	0.066	0.052	0.055	0.071		
IMP.-G	268.4%	204.5%	64.81%	48.33%	89.36%	43.55%	18.67%	**0.089**	**0.091**	373.3%	230.0%	50.88%	30.30%	65.38%	56.36%	21.13%	**0.086**	**0.094**
IMP.-P	278.9%	213.6%	68.52%	51.67%	93.62%	46.77%	21.33%			426.7%	270.0%	64.91%	42.42%	80.77%	70.91%	32.39%		
N.@10	0.023	0.028	0.070	0.073	0.071	0.089	0.093			0.019	0.023	0.067	0.074	0.076	0.090	0.092		
IMP.-G	243.5%	164.3%	45.71%	39.73%	43.66%	14.61%	9.677%	**0.102**	**0.107**	363.2%	265.2%	59.70%	44.59%	40.79%	18.89%	16.30%	**0.107**	**0.112**
IMP.-P	265.2%	182.1%	52.86%	46.58%	50.70%	20.22%	15.05%			389.5%	287.0%	67.16%	51.35%	47.37%	24.44%	21.74%		
N.@20	0.026	0.031	0.077	0.078	0.083	0.108	0.099			0.021	0.025	0.073	0.079	0.075	0.093	0.097		
IMP.-G	234.6%	164.5%	46.75%	44.87%	36.14%	4.630%	14.14%	**0.113**	**0.116**	381.0%	288.0%	67.12%	54.43%	62.67%	31.18%	25.77%	**0.122**	**0.125**
IMP.-P	246.2%	174.2%	50.65%	48.72%	39.76%	7.407%	17.17%			395.2%	300.0%	71.23%	58.23%	66.67%	34.41%	28.87%		

*FP.-LR: FPMC-LR, PR.-G: PRME-G, G.SoCa: GeoSoCa, R.G.FM: Rank-GeoFM, IMP.-G: improved by GPDM, IMP.-P: improved by PPDM.

Table 3. Performance Comparison on Foursquare-LA

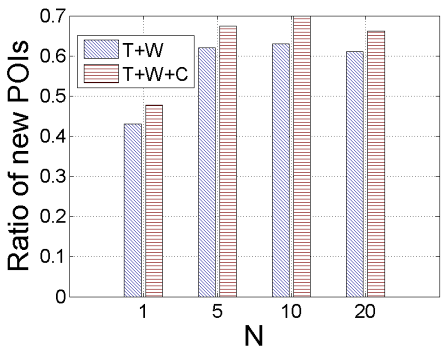
Metrics	Next POI Recommendation									Next New POI Recommendation								
	MF	PMF	FP.-LR	PR.-G	G.SoCa	R.G.FM	ST-RNN	GPDM	PPDM	MF	PMF	FP.-LR	PR.-G	G.SoCa	R.G.FM	ST-RNN	GPDM	PPDM
R@1	0.023	0.024	0.032	0.034	0.020	0.036	0.039	**0.044**	**0.044**	0.007	0.007	0.027	0.034	0.016	0.032	0.029	**0.036**	**0.036**
IMP.-G	91.30%	83.33%	37.50%	29.41%	120.0%	22.22%	12.82%			414.3%	414.3%	33.33%	5.882%	125.0%	12.50%	24.14%		
R@5	0.067	0.071	0.097	0.097	0.088	0.109	0.108	**0.129**	**0.125**	0.039	0.039	0.093	0.114	0.101	0.117	0.125	**0.147**	**0.141**
IMP.-G	92.54%	81.69%	32.99%	32.99%	46.59%	18.35%	19.44%			276.9%	276.9%	58.06%	28.95%	45.54%	25.64%	17.60%		
R@10	0.0.089	0.093	0.128	0.124	0.119	0.140	0.155	**0.170**	**0.163**	0.066	0.069	0.124	0.144	0.139	0.155	0.174	**0.202**	**0.185**
IMP.-G	91.01%	82.80%	32.81%	37.10%	42.86%	21.43%	9.677%			206.1%	192.8%	62.90%	40.28%	45.32%	30.32%	16.09%		
R@20	0.108	0.116	0.155	0.150	0.152	0.161	0.181	**0.213**	**0.199**	0.107	0.110	0.153	0.173	0.177	0.178	0.205	**0.239**	**0.229**
IMP.-G	97.22%	83.62%	37.42%	42.00%	40.13%	32.30%	17.68%			123.4%	117.3%	56.21%	38.15%	35.03%	34.27%	16.59%		
N.@5	0.036	0.043	0.055	0.053	0.044	0.061	0.060	**0.085**	**0.082**	0.012	0.014	0.051	0.060	0.043	0.069	0.077	**0.093**	**0.089**
IMP.-G	136.1%	97.67%	54.55%	60.38%	93.18%	39.34%	41.67%			575.0%	464.3%	82.35%	55.00%	116.3%	34.78%	20.78%		
N.@10	0.042	0.051	0.063	0.062	0.055	0.072	0.078	**0.097**	**0.091**	0.015	0.018	0.060	0.069	0.057	0.078	0.084	**0.111**	**0.106**
IMP.-G	131.0%	90.20%	53.97%	56.45%	76.36%	34.72%	24.36%			540.0%	416.7%	85.00%	60.87%	94.74%	42.31%	32.14%		
N.@20	0.047	0.055	0.067	0.068	0.065	0.077	0.090	**0.106**	**0.098**	0.016	0.020	0.067	0.075	0.070	0.084	0.092	**0.116**	**0.113**
IMP.-G	125.5%	92.73%	58.21%	55.88%	63.08%	37.66%	17.78%			525.0%	380.0%	73.13%	54.67%	65.71%	38.10%	26.09%		

*FP.-LR: FPMC-LR, PR.-G: PRME-G, G.SoCa: GeoSoCa, R.G.FM: Rank-GeoFM, IMP.-G: improved by GPDM.

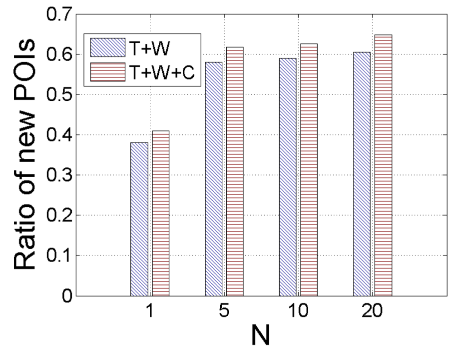
Table 4. Performance Comparison on Gowalla

Metrics	Next POI Recommendation									Next New POI Recommendation								
	MF	PMF	FP.-LR	PR.-G	G.SoCa	R.G.FM	ST-RNN	GPDM	PPDM	MF	PMF	FP.-LR	PR.-G	G.SoCa	R.G.FM	ST-RNN	GPDM	PPDM
R@1	0.022	0.024	0.029	0.040	0.038	0.036	0.033	**0.044**	**0.041**	0.006	0.006	0.028	0.045	0.022	0.019	0.027	**0.022**	**0.020**
IMP.-P	86.36%	70.83%	41.38%	2.500%	7.895%	13.89%	24.24%			233.3%	233.3%	−28.58%	−55.56%	−9.091%	5.263%	−25.93%		
R@5	0.086	0.093	0.116	0.142	0.092	0.134	0.152	**0.169**	**0.160**	0.031	0.036	0.118	0.182	0.083	0.167	0.205	**0.271**	**0.238**
IMP.-P	86.05%	72.04%	37.93%	12.68%	73.91%	19.40%	5.263%			667.7%	561.1%	101.7%	30.77%	186.7%	42.51%	16.10%		
R@10	0.147	0.158	0.198	0.195	0.184	0.217	0.180	**0.293**	**0.259**	0.059	0.067	0.182	0.249	0.210	0.263	0.284	**0.372**	**0.322**
IMP.-P	76.19%	63.92%	30.81%	32.82%	40.76%	19.35%	43.89%			445.8%	380.6%	76.92%	29.32%	53.33%	22.43%	13.38%		
R@20	0.188	0.202	0.247	0.246	0.232	0.281	0.236	**0.377**	**0.333**	0.109	0.118	0.252	0.311	0.295	0.327	0.350	**0.462**	**0.395**
IMP.-P	77.13%	64.85%	34.82%	35.37%	43.53%	18.51%	41.10%			262.4%	234.7%	56.75%	27.01%	33.90%	20.80%	12.86%		
N.@5	0.069	0.075	0.088	0.098	0.063	0.089	0.092	**0.133**	**0.125**	0.042	0.049	0.118	0.131	0.102	0.130	0.147	**0.158**	**0.152**
IMP.-P	81.16%	66.67%	42.05%	27.55%	98.41%	40.45%	35.87%			261.9%	210.2%	28.81%	16.03%	49.02%	16.92%	3.401%		
N.@10	0.091	0.104	0.121	0.119	0.098	0.117	0.120	**0.161**	**0.150**	0.063	0.075	0.164	0.170	0.155	0.171	0.174	**0.205**	**0.198**
IMP.-P	64.84%	44.23%	23.97%	26.05%	53.06%	28.21%	25.00%			214.3%	164.0%	20.73%	16.47%	27.74%	15.79%	13.79%		
N.@20	0.106	0.122	0.137	0.134	0.133	0.160	0.166	**0.182**	**0.171**	0.074	0.080	0.171	0.186	0.172	0.195	0.206	**0.223**	**0.220**
IMP.-P	61.32%	40.16%	24.82%	27.61%	28.57%	6.875%	3.012%			197.3%	175.0%	28.65%	18.28%	27.91%	12.82%	6.796%		

*FP.-LR: FPMC-LR, PR.-G: PRME-G, G.SoCa: GeoSoCa, R.G.FM: Rank-GeoFM, IMP.-P: improved by PPDM.



(a) Foursquare-LA



(b) Foursquare-NY

Fig. 9. Ratio of new POIs with Recall@N.

Rank-GeoFM have better results than FPMC-LR and PRME-G with the consideration of social relations.

- Both GPDM and PPDM consistently outperform FPMC-LR with 12%–130% improvement and outperform PRME-G with 5%–52% improvement except *recall*@1 on Gowalla. The results demonstrate that users' latent behavior patterns play an important role in next new POI recommendation.

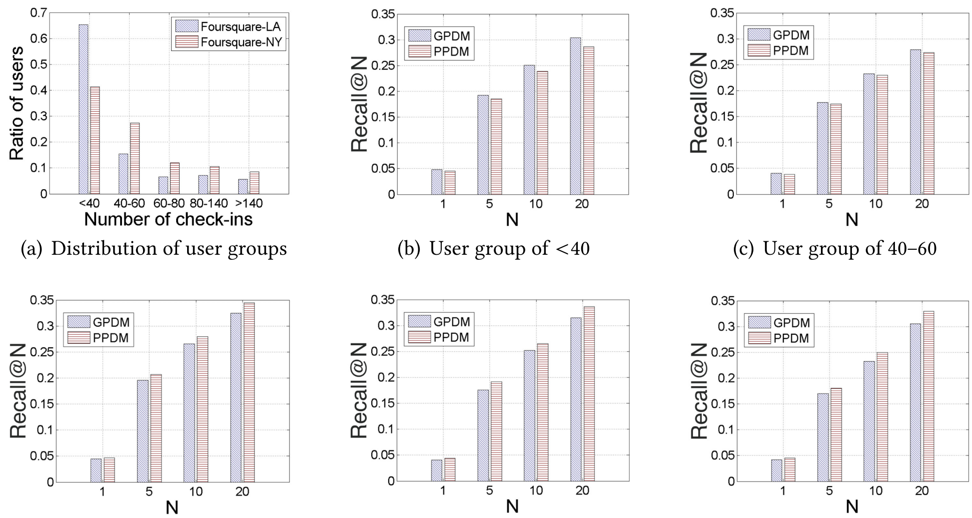
It should be noted that, for the task of next new POI recommendation, GPDM and PPDM perform worse than compared methods except MF and PMF for recall@1 on Gowalla data set. This can be attributed to the lack of categorical information for POIs in Gowalla, making the proposed models fail to integrate the categorical information to infer the latent behavior patterns. Intuitively, we conclude that the categorical information is an important factor in modeling the specific preference of a user for new POI recommendation.

To quantify the importance of categorical information, Figure 9 depicts the fraction of new POIs over all accurately predicted POIs by GPDM, which is computed as $\frac{|S_{N,u,rec} \cap S_{visited}^{new}|}{|S_{N,u,rec} \cap S_{visited}|}$. The left column of each precision unit represents the ratio of new POIs discovered by considering time of day (T) and day of week (W), and the right column considers location category (C) in addition to T and W. It is evident that the inclusion of the previous location category into the contextual features helps GPDM to achieve better performance in next new POI recommendation, suggesting that the categorical influence should be incorporated into inferring latent behavior patterns. The logic lies in that the category of a POI implicitly implies the type of activities a user can do there. In fact, human mobility demonstrates biases towards certain types or categories of POIs. For example, a foodie is more likely to visit another restaurant for a new taste, and a tourist prefers to visit other tourist attractions. Moreover, as shown in Figure 2, the transitions between location categories demonstrate some distinct patterns. Consequently, our models can improve the prediction of the relevance between the user and a new POI with the categorical information.

6.5 GPDM vs. PPDM

Both GPDM and PPDM consistently outperform other solutions in next POI and next new POI recommendation, demonstrating the effectiveness of the assumption of latent behavior patterns. In comparison, GPDM outperforms PPDM on all evaluation metrics on Foursquare-LA and Gowalla, while PPDM outperforms GPDM on Foursquare-NY.

Since PPDM differs from GPDM by considering personalized pattern distribution, the performance variation is data-related. Figure 10(a) illustrates the distribution of users regarding their



(a) Distribution of user groups (b) User group of <40 (c) User group of 40–60

(d) User group of 60–80 (e) User group of 80–140 (f) User group of >140

Fig. 10. Performance comparison on different user groups.

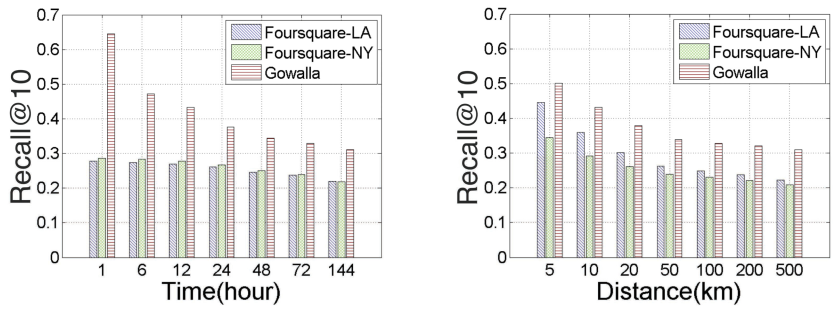
check-in frequency in Foursquare-LA and Foursquare-NY, demonstrating that users in Foursquare-NY have more personal visiting records. Figures 10(b)–10(f) illustrate the performance results on Foursquare-LA for each user group identified by check-in frequency. For users with less check-ins (e.g., <60), GPDM outperforms PPDM. However, PPDM starts to perform better than GPDM when users have more frequent check-ins. The result is intuitive that insufficient check-in records of a user make the learning of personalized latent behavior patterns inaccurate, and then GPDM makes better recommendations based on the shared distribution of latent behavior patterns in this case.

Although a user in Gowalla data set has more than 200 check-ins on average, which is much higher than 61 in Foursquare-NY, GPDM still outperforms PPDM on it. The reason lies in that the contextual features of Gowalla only include time of day and day of week, but no categorical information about POIs, which requires the model to find an accurate pattern out of thousands of potential POIs directly. In this case, the history of 200 check-ins is still not enough to learn an accurate personalized pattern. In comparison, the use of categorical information in fact groups thousands of potential POIs into a few classes, turning the recommendation process into two steps: the selection of a category, and then the selection of POI in the category. Consequently, it becomes easier to learn personalized patterns even with fewer check-in records.

Although GPDM and PPDM are designed for different purposes, they have the same computational complexity. Instead of computing $\boldsymbol{\alpha}^S$ for all users in Equation (21), PPDM learns $\boldsymbol{\alpha}_u^S$ for each user in Equation (24) and updates it for all users in the data set. Hence, GPDM and PPDM have identical computation cost, because PPDM only replaces Line 16 in Algorithm 1 with Equation (25).

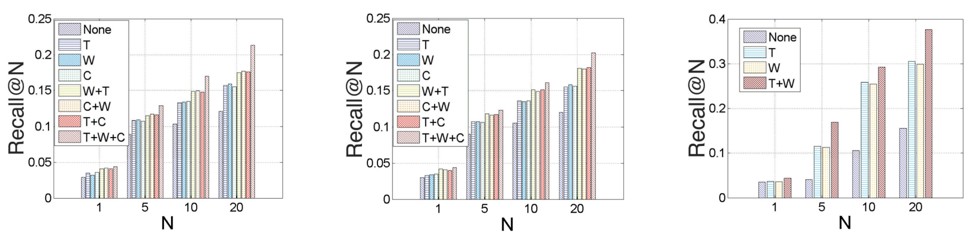
We further used independent t-tests for the compared methods and our proposed model, taking the significance level as $\alpha = 0.05$. The results indicate that the improvements achieved by GPDM and PPDM have statistical significance in all three datasets (see Tables 2–4). The difference between PPDM and GPDM are also significant, except for top-one recommendation.

In conclusion, PPDM is a better choice if users have frequent check-in records and/or POIs have categorical information provided. Otherwise, GPDM performs better, especially for users with few check-ins.



(a) Recommendation with Time Constraint (b) Recommendation with Distance Constraint

Fig. 11. Recommendation performance vs. distance and time.



(a) Foursquare-LA (b) Foursquare-NY (c) Gowalla

Fig. 12. Impact of features (*T: Time of Day, W: Day of Week, C: Category*).

6.6 Recommendation with Time or Distance Constraint

Another critical question about next POI recommendation is how the model performs regarding time or distance constraints. For example, if a model can predict POIs accurately for the next 2h or within 10km, how does it perform for the next 10h or within 50km?

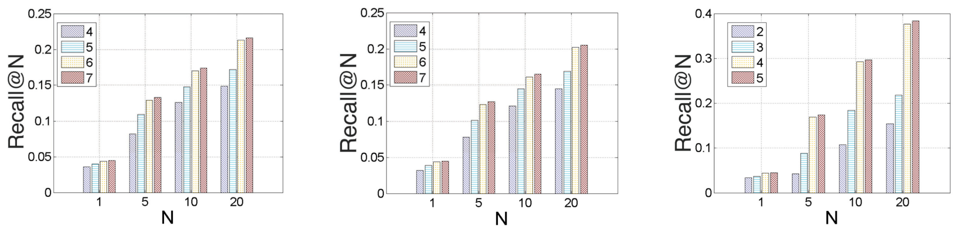
Figure 11(a) illustrates the *recall*@10 outcomes of GPDM with the increasing duration for each data set. In general, the model performs better when the duration is shorter, which is reasonable that in general there is more information available to decide the nearer future activity. GPDM has a slight decrease in performance for Foursquare-LA and Foursquare-NY, showing that it is not sensitive to the duration, and it is capable of making accurate recommendations for both large and short time intervals. The sharp decrease in Gowalla again proves the importance of categorical information.

Figure 11(b) illustrates the *recall*@10 outcomes of GPDM with the increasing distance for each data set, showing similar observations. It proves that the proposed model is capable of making accurate recommendations for both long and short distances.

6.7 Impact of the Contextual Features

The proposed models include three different types of contextual features, including time of day (T), day of week (W), and the previous location's category (C). Figure 12 depicts the experimental results with all possible combinations of three features to discover how they affect the performance.

In general, better outcomes are achieved with more contextual information included. If there is only one feature considered, then the models perform similarly no matter which feature it is. If there are two features considered, then the models perform similarly for any possible combination. And the models perform best when all three features are considered. In addition, changing the



(a) Foursquare-LA (b) Foursquare-NY (c) Gowalla

Fig. 13. Impact of number of latent pattern.

number of features F from 1 to 3 led to little difference for the execution time, which indicates that the number of features has little impact on the computational cost.

In conclusion, each feature has an even contribution to the accuracy of the recommendation, and each of them represents a distinct perspective to next POI recommendation. Thus, the combination of them helps to infer finer latent behavior patterns and to uplift the performance significantly.

6.8 Impact of the Latent Patterns

The proposed models rely on latent patterns to do the recommendation. Then the question is, what is the optimized number of latent patterns?

Figure 13 depicts the experimental results with different settings of the number of latent patterns. In general, better outcomes are achieved with more latent patterns considered until the growth becomes trivial. For example, recall@10 on Gowalla data set is 0.184 with three latent patterns, and it becomes 0.293 with four latent patterns, which is a 59.2% relative improvement. However, it only gains a further 1.4% improvement with five latent patterns. The gain after four latent patterns does not outweigh related computational overhead for Gowalla.

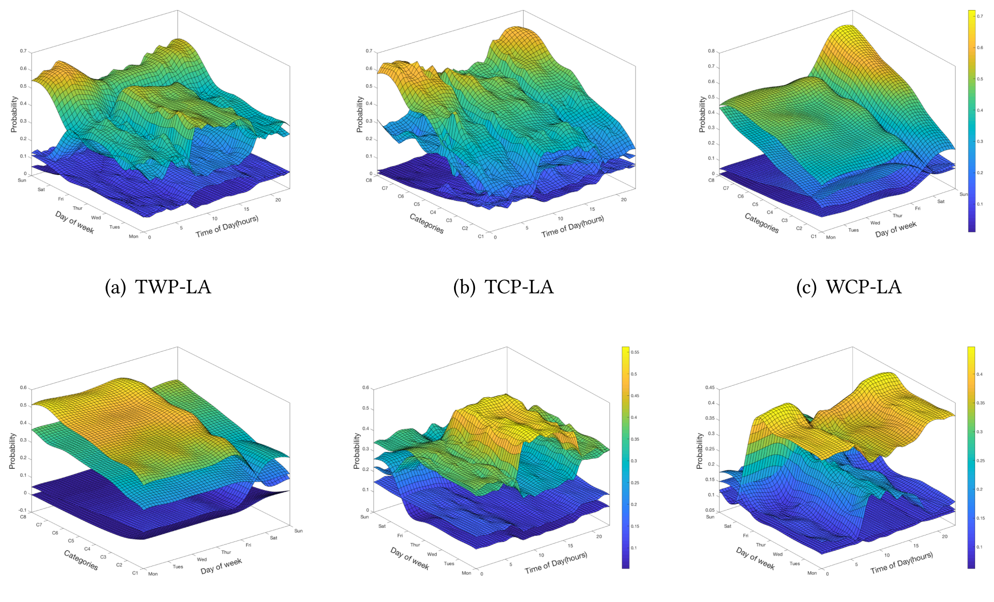
In conclusion, the optimized number of latent patterns is six for Foursquare-LA and Foursquare-NY, and four for Gowalla for the task of next personalized POI recommendation.

Furthermore, we analyzed the probabilities of latent patterns regarding different contextual features, such as time of day (hours), day of week, and location category. Our models learn a pattern-location probability matrix for each data set, which displays users' implicit preferences for different locations in distinct latent patterns. To observe the general law and facilitate the variation, we selected four latent patterns and GPDM to depict the probabilities of each latent pattern in different contextual scenarios.

Figure 14 illustrates that in all three data sets, the locations that are more frequently visited during daytime, such as universities or colleges, usually have fewer check-ins on weekends. In contrast, locations frequently visited at night, such as tourist spots and shops, may have higher check-ins on weekends. Different latent patterns portray different behavior patterns of people with their preferences for locations; the conclusions are consistent with the data set analysis before modeling, which reveals that the proposed model learns people's behavior patterns accurately.

7 CONCLUSION AND FUTURE WORK

This article achieved the tasks of next and next new POI recommendation by proposing two models: GPDM, which learns a fixed pattern distribution for all users; and PPDM, which learns personalized pattern distribution for each user. Both models utilize personalized latent behavior patterns learned from three contextual features, including time of day, day of week, and location category to improve the effectiveness of the recommendations. Technically, a soft-max function is used to infer multiple latent patterns, and sequential Bayesian Personalized Ranking (S-BPR) is applied as



(a) TWP-LA (b) TCP-LA (c) WCP-LA

(d) WCP-NY (e) TWP-NY (f) TWP-Gowalla

Fig. 14. Probabilities of latent patterns regarding contextual features.

the optimization criterion. Then, Expectation Maximization (EM) is in charge of finding optimized model parameters.

Extensive experiments on three large-scale real-world LBSN data sets show that our models in general significantly outperform other state-of-the-art methods for both next and next new POI recommendation tasks. The experiments also show that our models are capable of making accurate recommendations regardless of short/long duration or distance. Moreover, the experiments prove that the location category is an important factor for POI recommendations.

For future work, we plan to explore the theoretical connection between the number of contextual features and the corresponding number of latent patterns. In Section 6.8, we found that Foursquare, which has three features, requires six latent patterns; and Gowalla, which has two features, requires four latent patterns. We need to evaluate with more data sets to tell if the required number of patterns equals to two times the number of features. We would like to study how to integrate other contextual information into our model, e.g., social relationship and textual content of POIs, which may further enhance the recommendation performance.

ACKNOWLEDGMENTS

The authors thank the associate editor and the anonymous reviewers for their insightful and detailed comments.

REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. 2011. Context-aware recommender systems. In *Recommender Systems Handbook*. Springer, 217–253.

[2] Mohammad Aliannejadi and Fabio Crestani. 2018. Personalized context-aware point of interest recommendation. *ACM Trans. Inform. Syst.* 36, 4 (2018), 28.

[3] Mohammad Aliannejadi, Dimitrios Rafailidis, and Fabio Crestani. 2019. A joint two-phase time-sensitive regularized collaborative ranking model for point of interest recommendation. *IEEE Trans. Knowl. Data Eng.* PP (2019), 1–1. https://doi.org/10.1109/TKDE.2019.2903463

[4] Suhrid Balakrishnan and Sumit Chopra. 2012. Collaborative ranking. In *Proceedings of the 5th ACM International Conference on Web Search and Data Mining*. ACM, 143–152.

[5] Jie Bao, Yu Zheng, and Mohamed F. Mokbel. 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*. ACM, 199–208.

[6] Longbing Cao and S. P. Yu. 2012. *Behavior Computing*. Springer.

[7] Buru Chang, Yonggyu Park, Donghyeon Park, Seongsoon Kim, and Jaewoo Kang. 2018. Content-aware hierarchical point-of-interest embedding model for successive POI recommendation. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'18)*. 3301–3307.

[8] Chen Chen, Hongzhi Yin, Junjie Yao, and Bin Cui. 2013. Terec: A temporal recommender system over tweet stream. *Proc. VLDB Endow.* 6, 12 (2013), 1254–1257.

[9] Jialiang Chen, Xin Li, William K. Cheung, and Kan Li. 2016. Effective successive POI recommendation inferred with individual behavior and group preference. *Neurocomputing* 210, C (2016), 174–184.

[10] Meng Chen, Xiaohui Yu, and Yang Liu. 2018. MPE: A mobility pattern embedding model for predicting next locations. *Proceedings of the International World Wide Web Conference (WWW'18)*. 1–20.

[11] Yan Chen, Jichang Zhao, Xia Hu, Xiaoming Zhang, Zhoujun Li, and Tat-Seng Chua. 2013. From interest to function: Location estimation in social media. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence.*180–186.

[12] C. Cheng, H. Yang, I. King, and M. R. Lyu. 2012. Fused matrix factorization with geographical and social influence in location-based social networks. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*.

[13] Chen Cheng, Haiqin Yang, Michael R. Lyu, and Irwin King. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*. AAAI Press, 2605–2611.

[14] Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 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)*. ACM, New York, NY, 1082–1090.

[15] Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 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*. ACM, 1082–1090.

[16] Wen-Haw Chong and Ee-Peng Lim. 2018. Exploiting user and venue characteristics for fine-grained tweet geolocation. *ACM Trans. Inform. Syst.* 36, 3 (2018), 26.

[17] Konstantina Christakopoulou and Arindam Banerjee. 2015. Collaborative ranking with a push at the top. In *Proceedings of the 24th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 205–215.

[18] Arthur P. Dempster, Nan M. Laird, and Donald B. Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *J. Royal Stat. Soc. Series B (Methodol.)* 39, 1 (1977), 1–22.

[19] Nan Du, Hanjun Dai, Rakshit Trivedi, Utkarsh Upadhyay, Manuel Gomez-Rodriguez, and Le Song. 2016. Recurrent marked temporal point processes: Embedding event history to vector. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1555–1564.

[20] Nathan Eagle and Alex Pentland. 2009. Eigenbehaviors: Identifying structure in routine. In *Behavioral Ecology and Soc* 63, 7 (2009), 1057–1066.

[21] Stephen Eubank, Hasan Guclu, V. S. Anil Kumar, Madhav V. Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429, 6988 (2004), 180–184.

[22] Shanshan Feng, Xutao Li, Yifeng Zeng, Gao Cong, Yeow Meng Chee, and Quan Yuan. 2015. Personalized ranking metric embedding for next new POI recommendation. In *Proceedings of the 24th International Conference on Artificial Intelligence*. AAAI Press, 2069–2075.

[23] Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. 2013. Exploring temporal effects for location recommendation on location-based social networks. In *Proceedings of the 7th ACM Conference on Recommender Systems*. ACM, 93–100.

[24] Huiji Gao, Jiliang Tang, Xia Hu, and Huan Liu. 2015. Content-aware point of interest recommendation on location-based social networks. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence.*1721–1727.

[25] Jing He, Xin Li, Lejian Liao, Dandan Song, and William K. Cheung. 2016. Inferring a personalized next point-of-interest recommendation model with latent behavior patterns. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*.

[26] Ruining He and Julian McAuley. 2016. VBPR: Visual Bayesian personalized ranking from implicit feedback. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*.

[27] Jianhua Feng Henan Wang, Guoliang Li. 2014. Group-based personalized location recommendation on social networks. In *Proceeding of the 16th Asia-Pacific Web Conference (APWeb'14)*. 68–80.

[28] Jun Hu and Ping Li. 2017. Decoupled collaborative ranking. In *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 1321–1329.
[29] Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. 2010. Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In *Proceedings of the 4th ACM Conference on Recommender Systems*. ACM, 79–86.
[30] Yehuda Koren, Robert M. Bell, and Chris Volinsky. 2009. Matrix factorization techniques for recommender systems. *IEEE Comput.* 42, 8, 30–37.
[31] Ranzhen Li, Yanyan Shen, and Yanmin Zhu. 2018. Next point-of-interest recommendation with temporal and multi-level context attention. In *Proceedings of the IEEE International Conference on Data Mining (ICDM'18)*. IEEE, 1110–1115.
[32] Xutao Li, Gao Cong, Xiao-Li Li, Tuan-Anh Nguyen Pham, and Shonali Krishnaswamy. 2015. Rank-GeoFm: A ranking based geographical factorization method for point of interest recommendation. In *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 433–442.
[33] Xin Li, Mingming Jiang, Huiting Hong, and Lejian Liao. 2017. A time-aware personalized point-of-interest recommendation via high-order tensor factorization. *ACM Trans. Inform. Syst.* 35, 4 (2017), 31.
[34] Zhenhui Li, Bolin Ding, Jiawei Han, Roland Kays, and Peter Nye. 2010. Mining periodic behaviors for moving objects. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'10)*. ACM, New York, NY, 1099–1108. DOI: https://doi.org/10.1145/1835804.1835942
[35] Bin Liu, Yanjie Fu, Zijun Yao, and Hui Xiong. 2013. Learning geographical preferences for point-of-interest recommendation. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 1043–1051.
[36] Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. 2016. Predicting the next location: A recurrent model with spatial and temporal contexts. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*.
[37] Pasquale Lops, Marco De Gemmis, and Giovanni Semeraro. 2011. Content-based recommender systems: State of the art and trends. In *Recommender Systems Handbook*. Springer, 73–105.
[38] Jarana Manotumruksa, Craig Macdonald, and Iadh Ounis. 2017. A deep recurrent collaborative filtering framework for venue recommendation. In *Proceedings of the ACM Conference on Information and Knowledge Management*. ACM, 1429–1438.
[39] Stuart E. Middleton, Giorgos Kordopatis-Zilos, Symeon Papadopoulos, and Yiannis Kompatsiaris. 2018. Location extraction from social media: Geoparsing, location disambiguation, and geotagging. *ACM Trans. Inform. Syst.* 36, 4 (2018), 40.
[40] Andriy Mnih and Ruslan Salakhutdinov. 2007. Probabilistic matrix factorization. In *Proceedings of the Conference on Advances in Neural Information Processing Systems*. 1257–1264.
[41] Radford M. Neal and Geoffrey E. Hinton. 1998. A view of the EM algorithm that justifies incremental, sparse, and other variants. In *Learning in Graphical Models*. Springer, 355–368.
[42] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. 2009. BPR: Bayesian personalized ranking from implicit feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence*. 452–461.
[43] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized Markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web*. ACM, 811–820.
[44] Steffen Rendle and Lars Schmidt-Thieme. 2010. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of the 3rd ACM International Conference on Web Search and Data Mining*. ACM, 81–90.
[45] Alasdair Thomason, Nathan Griffiths, and Victor Sanchez. 2016. Context trees: Augmenting geospatial trajectories with context. *ACM Trans. Inform. Syst.* 35, 2 (2016), 14.
[46] Senzhang Wang, Xiaoming Zhang, Jianping Cao, Lifang He, Leon Stenneth, Philip S. Yu, Zhoujun Li, and Zhiqiu Huang. 2017. Computing urban traffic congestions by incorporating sparse GPS probe data and social media data. *ACM Trans. Inform. Syst.* 35, 4 (2017), 40.
[47] Cheng Yang, Maosong Sun, Wayne Xin Zhao, Zhiyuan Liu, and Edward Y. Chang. 2017. A neural network approach to jointly modeling social networks and mobile trajectories. *ACM Trans. Inform. Syst.* 35, 4 (2017), 36.
[48] Zijun Yao, Yanjie Fu, Bin Liu, Yanchi Liu, and Hui Xiong. 2016. POI recommendation: A temporal matching between POI popularity and user regularity. In *Proceedings of the IEEE 16th International Conference on Data Mining (ICDM'16)*. IEEE, 549–558.
[49] Mao Ye, Peifeng Yin, and Wang-Chien Lee. 2010. Location recommendation for location-based social networks. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 458–461.
[50] Mao Ye, Peifeng Yin, Wang-Chien Lee, and Dik-Lun Lee. 2011. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 325–334.

[51] Hongzhi Yin, Bin Cui, Ling Chen, Zhiting Hu, and Xiaofang Zhou. 2015. Dynamic user modeling in social media systems. *ACM Trans. Inform. Syst.* 33, 3 (2015), 10.

[52] Hongzhi Yin, Bin Cui, Xiaofang Zhou, Weiqing Wang, Zi Huang, and Shazia Sadiq. 2016. Joint modeling of user check-in behaviors for real-time point-of-interest recommendation. *ACM Trans. Inform. Syst.* 35, 2 (2016), 11.

[53] Yonghong Yu and Xingguo Chen. 2015. A survey of point-of-interest recommendation in location-based social networks. In *Proceedings of the Workshops at the 29th AAAI Conference on Artificial Intelligence*, Vol. 130.

[54] Quan Yuan, Gao Cong, Zongyang Ma, Aixin Sun, and Nadia Magnenat Thalmann. 2013. Time-aware point-of-interest recommendation. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM, 363–372.

[55] Ting Yuan, Jian Cheng, Xi Zhang, Shuang Qiu, and Hanqing Lu. 2014. Recommendation by mining multiple user behaviors with group sparsity. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence.* 222–228.

[56] Chenyi Zhang, Hongwei Liang, and Ke Wang. 2016. Trip recommendation meets real-world constraints: POI availability, diversity, and traveling time uncertainty. *ACM Trans. Inform. Syst.* 35, 1 (2016), 5.

[57] Jia-Dong Zhang and Chi-Yin Chow. 2015. 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.* ACM, 443–452.

[58] Yuyu Zhang, Hanjun Dai, Chang Xu, Jun Feng, Taifeng Wang, Jiang Bian, Bin Wang, and Tie-Yan Liu. 2014. Sequential click prediction for sponsored search with recurrent neural networks. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence.*

[59] Zhiqian Zhang, Chenliang Li, Zhiyong Wu, Aixin Sun, Dengpan Ye, and Xiangyang Luo. 2017. NEXT: A neural network framework for next POI recommendation. Retrieved from: *arXiv preprint arXiv:1704.04576* (2017).

[60] Shenglin Zhao, Tong Zhao, Irwin King, and Michael R. Lyu. 2017. Geo-teaser: Geo-temporal sequential embedding rank for point-of-interest recommendation. In *Proceedings of the 26th International Conference on World Wide Web Companion.* International World Wide Web Conferences Steering Committee, 153–162.

[61] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael R. Lyu, and Irwin King. 2016. STELLAR: Spatial-temporal latent ranking for successive point-of-interest recommendation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence.*

[62] Wayne Xin Zhao, Ningnan Zhou, Wenhui Zhang, Ji-Rong Wen, Shan Wang, and Edward Y. Chang. 2016. A probabilistic lifestyle-based trajectory model for social strength inference from human trajectory data. *ACM Trans. Inform. Syst.* 35, 1 (2016), 8.

[63] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. 2014. Urban computing: Concepts, methodologies, and applications. *ACM Trans. Intell. Syst. Technol.* 5, 3 (2014), 38.

[64] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th International Conference on World Wide Web (WWW'09).* ACM, 791–800.

Received December 2018; revised July 2019; accepted August 2019