

Research Article

Joint Modeling of User Behaviors Based on Variable-Order Additive Markov Chain for POI Recommendation

RuiChang Li 

Beijing Key Lab of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, and the School of Computer Science (National Pilot Software Engineering School), Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to RuiChang Li; lrc@bupt.edu.cn

Received 20 August 2021; Revised 8 October 2021; Accepted 25 October 2021; Published 23 November 2021

Academic Editor: Yuyu Yin

Copyright © 2021 RuiChang Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The POI recommendation system has become an important means to help people discover attractive and interesting places. Based on our data analysis, we observe that users pay equal attention to conservatism and curiosity. In particular, adopting analysis corresponding to different time intervals, we find that users lean towards old POIs in the short term and look for new POIs with the increase of the time interval. However, existing approaches usually neglect users' conservatism and curiosity preferences. Therefore, they are confronted with a bottleneck of depicting accurate user needs, making it difficult to improve the recommendation performance further. Besides, we further find that the number of user daily check-ins has uneven distribution, which is not conducive to capture the accurate transition patterns of user behaviors. In light of the above, we design a single POI sequential method. On this basis, we propose a recommendation method of the variable additive Markov chain. We consider the user sequential preferences, especially liking old and pursuing new features. In addition, our model exploits the geographical tendency of user behaviors. Finally, we conduct abundant experiments on four cities in the two real datasets, i.e., Foursquare and Jiebang. The experimental results show its superiority over other competitors.

1. Introduction

Recommender systems are valuable tools that play a crucial role in mitigating information overload problems. Today, such systems are used in many application domains [1–3]. With the increasing popularity of WSN [4] and location-based social networks (LBSNs), such as Foursquare, Gowalla, and Yelp, unlimited possibilities are provided for users to share their highlights. Users not only explore location-aware information but also write reviews and share their experiences [5]. In LBSNs, user-generated trajectory data contains rich information, such as POI location, category, content, visited time, and trajectory sequence, which can be used to exploit user's preferences for providing personalized POI recommendation for the target user. There are huge business opportunities. So the academic and industry has invested a great deal of enthusiasm and energy in studies of recommendation, such as location-based activity recommendation [6], friend recommendation [1, 7], and location rec-

ommendation [8, 9]. In these studies, providing location recommendations becomes an important application with the rapid emergence of LBSNs, such as POIs recommendation [10–16] and routes recommendation [17, 18].

Existing researches mainly study how to use time influence, geographical influence, social relationship, and other characteristics for POI recommendation. For example, a user is more likely to choose a restaurant for dinner at noon. A POI closer to the current location is more popular with the user, and friends may have similar preferences. However, these studies mainly provide POI recommendations based on common sense analysis and fail to notice that users' past experiences that will affect their attitudes towards POI. For example, the pursuit of familiarity and novelty is also an important factor. Through data analysis and reasoning, we found some interesting phenomena to consider POI recommendations from a new perspective.

In a limited geographic space, users tend to like the old and pursue the new. We define POIs that users have visited

as old POIs and that have not been visited in the previous period as new POIs. Due to time and distance constraints, the user choice space and the number of new POIs are limited. For a user, he chooses a POI considering his available time and the acceptable transfer distance. For example, if the user only has half an hour, POIs that take longer than an hour are not considered. If the user likes to walk, then POIs within 1 km are more popular. Long-distance movement requires more time and financial support, which limits users' desire to visit it. The POIs that users frequently visit are not evenly distributed geographically but clustered around limited centers [19, 20]. For example, users have more check-ins near their homes, workplaces, and popular areas. Taylor's first law of geography also confirms this phenomenon [21]. Obviously, the category and number of POIs are limited in a limited geographic area. After repeated exploration by users, new POIs will gradually become old. Therefore, users can only choose POIs on the premise that the new is less and the old is more.

The POI alternates between the old and the new, driving users to like the old and pursue the new POIs. As we all know, users often visit the same POI multiple times, and each visit has a different experience. For example, it is impossible for the user to taste all dishes of a restaurant at one time. For scenic spots, users will find discoveries during repeated visits. For example, climbing the Great Wall from different roads has different sceneries, and the view in the morning and evening is also different. In addition, the POI is also changing. For example, POIs may be updated irregularly, and the surrounding environment may also vary. Just as it is impossible for a person to step into the same river, all the POIs constantly change. Changes of the POI and the surrounding environment will make the old POI glow with new appeal. Intuitively, users will prefer the new POI to the old one. For example, if a restaurant leaves a good memory to the user, the user's experience of going to this restaurant again for dinner will not be worse than the last time and may be better. The user will not wander in the corridor as he revisited the museum the first time, he will be more confident, and the second visit will be smoother than the initial visit [22].

Sparse trajectories affect the acquisition of POI relationships. We define a sequence containing multiple POIs as a dense trajectory, and correspondingly, the other holding a single POI is called a sparse trajectory. Intuitively, a user sometimes generates multiple check-ins a day. Occasionally, they check in only one POI, and even worse, they do not do that for several days. Some studies believe that there is an association relationship between POIs that a user visits continuously, and this relationship reflects the law of co-occurrence between them. Intuitively, due to the continuity of user behavior, the smaller the time interval between check-ins, the more the influential impact on the next behavior. In practice, there are studies that use hours, days, weeks, months, etc., as time intervals [23]. No matter what time interval is used, there will be dense and sparse trajectories coexisting. Currently, two methods are usually used to deal with sparse trajectories: (1) filter out the check-in records of a single POI and only retain the dense trajectories

that visit multiple POIs in a day and (2) increase the time threshold between adjacent POIs to expand the coverage of a single trajectory. For example, we can adjust the time interval so that POIs of several consecutive days are formed a trajectory. However, both of them have drawbacks. The first method will make user data more sparse, which is not conducive to accurately obtaining user preferences. Correspondingly, although the second method reduces data loss, it introduces some additional noise. Obviously, additional noise may interfere with the accuracy of user preferences.

In this paper, we propose a variable-order additive Markov chain based on sequential patterns, the influence of liking the old and pursuing the new, and geographical restriction (SONG). In SONG, the model consists two components: POI transition and variable influence. The POI transition probability model is used to obtain the one-step transition probability between any POIs. The variable influence model is used to extract the joint influence of three factors. We first construct a POI-POI transition graph based on the user-POI interaction sequence contained in check-in trajectories of all users. The transition probabilities can be deduced by dividing transition degrees by outgoing degrees of the graph vertex. Then, we jointly model the variable effects of liking old, pursuing new, and geographical restriction based on the mere exposure effect [24], depreciation theory [25], and Taylor's first law [21], respectively. We finally provide TOP-N POI recommendations. The main contributions of this paper can be summarized:

- (1) User check-in behavior has an apparent feature of liking the old and pursuing the new. Currently, there is no research on this feature. We found that it is mainly due to the unbalanced distribution of interest POIs and user likes to pursue familiarity and novelty, which provides new ideas for POI recommendations
- (2) The check-in number of a user is very uneven every day. For example, there may be only one or no record in a day. This imbalanced situation influences acquiring the sequence pattern implied by the user's check-in behavior. We propose a method to serialize sparse trajectories, which alleviates the problems caused by the neglect of a single POI or the introduction of additional noise by extending the sequence and better obtain user preferences
- (3) Based on the above discovery and the results of single POI serialization, we propose a variable-order additive Markov chain model to capture the influence of historical sequences on subsequent POIs. Then, we use the mere exposure effect and depreciation theory to jointly model the users' preference of preferring old, pursuing new, and combine the geographical restriction to provide personalized recommendations
- (4) We conducted extensive experiments on four city subsets of real trajectory datasets, Foursquare and Jiebang. Experimental results show that our model outperforms other state-of-the-art methods.

The remainder of the paper is organized as follows. In Section 2, we review related work on POIs recommendation. In Section 3, we first describe and define the task we are tackling. Then, we make data analysis on four subdatasets of two real-world datasets. In Section 4, we present our proposed model SONG for POI recommendation in detail. In Section 5, we describe our experiment settings for evaluating the performance of SONG against the state-of-the-art POI recommendation techniques. Finally, Section 6 concludes this paper.

2. Related Work

POI recommendation based on trajectories is a hot topic in LBSNs, and it has attracted increasing attention in both academia and industry. Therefore, we briefly introduce three lines of researches related to our task: (a) POI recommendation based on temporal; (b) recommendation based on geographical influence; and (c) sequential POI recommendation. All of these factors are closely related to acquiring user preferences [26].

2.1. POI Recommendation Based on Temporal. Temporal influence plays a vital role in POI recommendation. Yuan et al. [10] found out that most users tend to visit different POIs at a different time in a day, and the check-in behaviors between neighbor time slots are similar. In addition to that, Zhang et al. [27] pointed out that POI may be not available in all time, for example, POIs are only accessible during their opening hours. He et al. [28] investigated the temporal popularity of a POI and the temporal check-in trends to provide personalized POI recommendation. He et al. [28] propose a spatial-temporal topic model (STM), which embedded the temporal and spatial patterns in users check-in activities. Oppokhonov et al. [29] develop a recommendation system based on a directed graph. The algorithm of the system considers both the temporal factor and the distance for recommending a new POI for next hours. Gao et al. [30] put forward four temporal aggregation strategies, such as sum, mean, maximum, and voting, to integrate a user's check-in preferences of different temporal states. The method using the strategy greatly improves the POI recommendation performance. Ji et al. [31] proposed a social-period-aware topic model (SPATM) to learn the influence weights of both user interests and her social preferences on making-decision for each check-in time automatically. Actually, the time law of users' behaviors include two different patterns, i.e., periodic and aperiodic. For example, a user may like to find a restaurant for dinner at noon, which is periodic. Since the historical experience will affect the user's choice of the target POI, this influence is generally believed to be related to the interval between visiting two POIs. For convenience, we use the interval hops between POIs instead of specific time to model aperiodic effects.

2.2. Recommendation Based on Geographical Influence. In LBSNs, the geographical influence is an important factor that distinguishes the POI recommendation from other rec-

ommendations because physical interactions are required for users to visit POIs [32]. For example, users prefer to select POIs near to their homes or offices and also may be fond of exploring the nearby POIs of their current locations. Several studies have attempted to leverage the geographical influence to improve POI recommendation systems [33, 34]. Ye et al. [35] proposed a power-law distribution model to capture the geographical influence, and proposed a collaborative POI recommendation algorithm based on geographical influence via naive Bayesian. Because it is difficult to find an anchor point to derive a reasonable distance for the new POI, Zhang et al. [32] develop a kernel function to model the geographical influence. The kernel function on two-dimensional is more reasonable than the one-dimensional distance power-law distribution. For users' check-ins is unevenly distribution. Some exiting studied assume that user checked locations conform to the Gaussian distribution of multiple centers. Chenget al. [19] proposed a multicenter Gaussian model to capture the geographical influence for POI recommendation. In addition to directly using geographical influences, there are some joint model methods. Liu et al. [13] develop a general geographical probabilistic factor model (Geo-PFM) to capture the geographical influence on user mobility behaviors, and then combine the influence with Bayesian nonnegative matrix factorization (BNMF) to model user preferences. Finally, POIs are recommended for users by combining the effects of multiple factors. Yin et al. [12] propose a joint probabilistic generative model to integrate geographical-social influence, temporal and word-of-mouth effect for solving check-in data sparsity in the out-of-town recommendation scenario. Griesner et al. [15] propose a augmenting matrix factorization model (GeoMF-TD) for POI recommendation by combining geographical and temporal influences. Li et al. [16] put forward a ranking based geographical factorization method (Rank-GeoFM) model to provide POIs prediction. Specifically, the proposed model can easily combine geographical influence and temporal influence. In this paper, we use power-law distribution to capture the influence of distance on the user's choice of POI.

2.3. Sequential POI Recommendation. In the real world, users' behaviors usually happen in succession, and the next action is often related to the previous one. In recent years, studies have focused on various sequential recommendation tasks, such as next POI recommendation [29, 36]. Early studies were typically based on the Markov chain models for sequential recommendation [32, 36]. For example, He and McAuley [36] proposed fossil that fuses similarity-based models with the Markov chains to predict personalized sequential behavior. Cheng et al. [8] developed a matrix factorization method, which embeds the personalized Markov chains and the localized regions for solving the sequential recommendation task. The order of the Markov chain model decides the scale-free parameters. A higher order will increase the computation cost. Following the development of machine learning, a lot of complex models have been proposed: recurrent neural networks (RNNs) [37], convolutional neural networks (CNNs) [38, 39], translation-based

methods [40], and attention mechanism [41]. Approaches based RNN are popular models for sequential recommendation due to their performance on sequential works. Liu et al. [42] extend RNN and propose a novel method to model temporal and spatial contexts for different time intervals and distance-specific transition matrices for different geographical distances. However, existing RNN methods neglect some details of users preferences, thus making the recommendation results unreliable. To address the above limitations, Sun et al. [43] proposed a method named LSTPM for next POI recommendation. Li et al. [44] introduce a novel neural network model named TMCA that employed the LSTM-based encoder-decoder framework for the next POI recommendation. In this paper, we use a variable high-order Markov chain to model the relationship between preceding POI to the posterior in trajectory.

3. Task Definition and Data Analysis

In this section, we first describe and define the task we are tackling. Then, we make data analysis on four subdatasets of two real-world datasets, which serve as the foundation of our model.

3.1. Task Definition. A notation U denotes the collection of users $U = \{u_1, u_2, \dots, u_{|U|}\}$, and the notation V represents the collection of POIs $V = \{v_1, v_2, \dots, v_{|V|}\}$. We use the v to represent a POI identifier and l_v to denote its corresponding geographical attribute in terms of longitude and latitude coordinates. For each user $u \in U$, we sort her historical records by time and subdivided these into trajectories with different lengths relying on the suitable time interval, i.e., $S_u = \{s_u^1, s_u^2, \dots, s_u^n\}$, where n denotes the latest trajectory. The subtrajectory $s_u^j = \{v_1, v_2, \dots, v_t\}$ is the j th trajectory of user u . Given a trajectory set of target user u , our goal is to recommend Top- N POIs, namely $\Pr(v | s_u^j) = \Pr(v | v_n, v_{n-1}, v_2, v_1)$.

3.2. Dataset Description. We conduct analysis on four real-world datasets from the Foursquare [45] and the JiePang [11]. The Foursquare check-in dataset is from 12 April 2012 to 16 February 2013 in New York (NYC) and Tokyo (TKY), while JiePang contains Beijing and Shanghai from April 2011 to April 2013. For all of them, we filter unpopular POIs with less than five user visits.

3.3. Data Analysis. In theory, users can check any locations. However, user behavior is affected by various factors such as available time, geographic distance, and economic support. Users often look for interesting POIs near their living and working centers. Intuitively, users' repeated exploration might make POIs less appealing, but users can always find pleasure in these POIs. It is worth exploring that makes these POIs so appealing.

We divide POIs into two categories. POIs that appeared in the last time period are called old POIs, and the corresponding ones that did not appear are called new POIs. Figure 1 shows the new and old changes of POIs by different time intervals. We use different time intervals to analyze the new and old changes of visited POIs. Traditionally, we

divide a year into 52 weeks. In the figure, "2012-02" indicates the second week of 2012. Figure 1 contains four cities information, and every city has two figures. The left takes one week as an interval, and the right is five weeks. As we can see from the left figure of every city, the old POI is predominant. Especially, the Foursquare dataset shows this more clearly. Comparing the left and right, it is clear that the ratio of new POIs is predominant in JiePang. Although the proportion of old POIs exceeds the new POIs, the new increases and old decreases.

According to the above analysis, we can come to the following conclusion: Users are conservative in the short term and show more curiosity characteristics with the increase of the time interval. Making full use of users' preference features of liking old and pursuing new simultaneously may improve the recommendation performance.

4. The Proposed Model

In this section, we present our proposed model SONG in detail. SONG mainly consists of three parts, which are the sequential recommendation modeling and the old and new influence weight modeling. Our main contribution lies in joint modeling liking the old and pursuing the new influence, the geographical influence for sequential recommendation in a unified way.

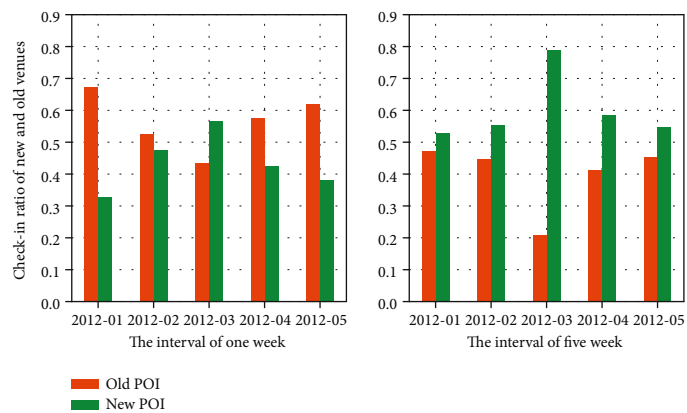
4.1. Modeling Sequential Recommendation Based on Trajectory. For sequential recommendation, the order-based models are used to derive order patterns from items sequences of users', such as using the Markov chains [36, 46]. Inspired by the successful cases of the Markov method in capturing sequence features and the high compatibility [46] with the problem in our paper, we use the high-order Markov chain model to model the sequential recommendation task. The mathematical form of this model is shown by the following equation.

$$\Pr(v_r | s_u^j) = \Pr(v_r = v_{t+1} | v_t, v_{t-1}, \dots, v_2, v_1). \quad (1)$$

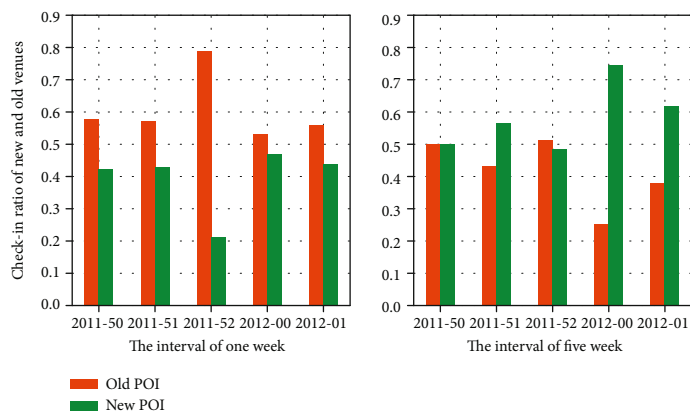
However, as [47] points out, if the k -order Markov chain has n possible states, it usually leads to exponential expansion on the number of states, e.g., $(n-1)n^k$. Moreover, high-order Markov chains also suffer from the sparsity of transitions in the given dataset, which discourages people from using a higher-order Markov chain directly and leads to ineffectiveness in modeling the behaviors of users. Inspired by [32, 47], a higher-order Markov chain model involves only one additional parameter for each one-step transition probability. The model can be written as follows:

$$\Pr(v_r | s_u^j) = \Pr(v_r = v_{t+1} | s_u^j) \approx \sum_{i=1}^t w_u \Pr(v_{t+1}, v_{t+1-i}), \quad (2)$$

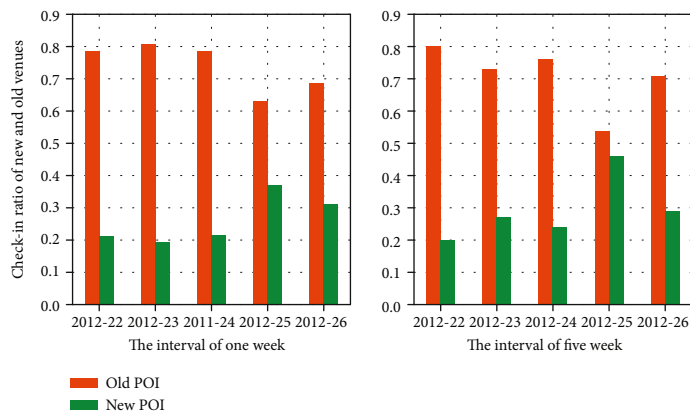
where $\Pr(v_{t+1}, v_{t+1-i})$ denotes the one-step transition probability from POI v_{t+1-i} to POI v_{t+1} and w_u is the additional parameter.



(a)

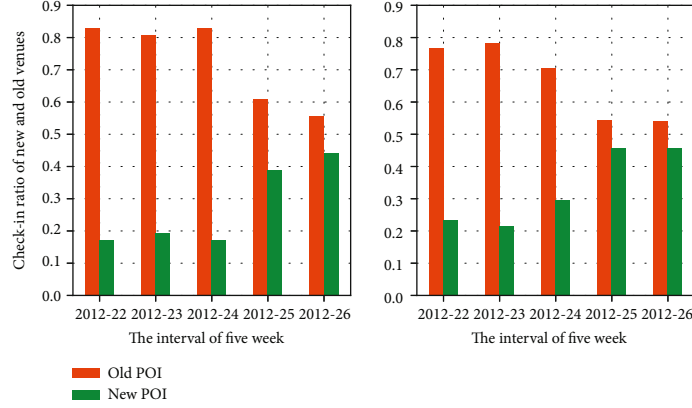


(b)



(c)

FIGURE 1: Continued.



(d)

FIGURE 1: Changes of the ratio of new and old POIs on four datasets at different time intervals. (a) Beijing (Jiepan). (b) Shanghai (Jiepan). (c) NYC (Foursquare). (d) TKY (Foursquare).

Input: Training trajectories of user u : $S_u, s_u^j \in S_u, u \in U$.
Output: Dense trajectories set of user u : S'_u

- 1: for ($u = 0; u < |U|; u++$) **do**
- 2: Get the trajectories count of u : $|S_u|$;
- 3: Set trajectory index: $p = 0$;
- 4: **while** $p < |S_u|$ **do**
- 5: Get the length of trajectory $|s_u^p|$;
- 6: **if** $|s_u^p| > 2$ **then**
- 7: $p \leftarrow p + 1$
- 8: **else if** $|s_u^p| \leq 1$ and $p = 0$ **then**
- 9: Add the current trajectory to the next one
- 10: **else if** $|s_u^p| \leq 1$ and $p < (|S_u| - 1)$ **then**
- 11: Get the time interval of $T_{s_u^p - s_u^{p-1}}$ and $T_{s_u^{p+1} - s_u^p}$.
- 12: **if** $T_{s_u^p - s_u^{p-1}} \leq T_{s_u^{p+1} - s_u^p}$ **then**
- 13: Insert s_u^p to s_u^{p-1} , $p \leftarrow p + 1$
- 14: **else**
- 15: Insert s_u^p to s_u^{p+1} , $p \leftarrow p + |s_u^{p+1}| + 1$
- 16: **end if**
- 17: **else**
- 18: Insert s_u^p into s_u^{p-1} , $p \leftarrow p + 1$
- 19: **end if**
- 20: **end while**
- 21: **end for**
- 22: Return S'_u .

ALGORITHM 1: Serialization algorithm.

4.1.1. One-Step Transition Probability. The number of POIs that a user visits every day is unevenly distributed, and sparse trajectories containing a single POI are sometimes ignored in mining POIs transition patterns. In order to make reasonable use of these sparse trajectories, we propose a method to serialize sparse trajectories. The core idea is as follows. We first split the user's training data by one day and then determine whether it needs to be serialized according to the length of the trajectory. When the user's trajectory

length of a day is greater than 2, this trajectory does not need to be processed. The processing flow is shown in the pseudocode on lines 5-7. When the trajectory length is equal to 1, it needs to be dealt with in three cases: (1) If it is the first trajectory of the user, insert this one into the next one; (2) if it is the last trajectory of the user, add this trajectory to the previous one; and (3) in other cases, it needs to judge the time interval to the previous and the subsequent trajectories and insert the current trajectory into the closer trajectory. The insert operation is relatively simple, it only needs to modify the inserted trajectory's date to the date of the target trajectory, and the time remains unchanged. For example, the date of the target trajectory is October 6, 2016, and the date and time of the trajectory to be inserted is October 7, 2016, 13:23:11. It only needs to change the date to October 6, 2016. Comparing S'_u with S_u , only the date of some trajectories in S'_u has changed, and the time remains the same to retain the user's time-related preferences. The time remains unchanged at 13:23:11. The specific operation process is shown in the following algorithm.

Inspired by [32], the one-step transition probabilities between two POIs are derived by employing the first-order Markov chain. They are supposed to be fixed, which can be represented by the following equation:

$$\Pr(v_{t+1} | s_u^j) = \Pr(v_{t+1} | v_t) = \frac{\text{ToDegree}(v_t, v_{t+1})}{\text{AoDegree}(v_t)}, \quad (3)$$

when $\text{AoDegree}(v_t) > 0$, where $\text{AoDegree}(v_t)$ is out degree of POI v_t and $\text{ToDegree}(v_t, v_{t+1})$ is the degree from v_t to v_{t+1} . When $\text{AoDegree}(v_t) = 0$, $\Pr(v_{t+1} | v_t)$ denotes the following:

$$\Pr(v_{t+1} | v_t) = \begin{cases} 1, & v_{t+1} = v_t \\ 0, & v_{t+1} \neq v_t. \end{cases} \quad (4)$$

4.1.2. Modeling the Influence Weight of User Pursuing Old. It is common that the more exposure something is in front of

us, the more it will strengthen our attitude towards it. Users' attitudes towards things can be divided into three categories: like, neutral, and dislikes. Since the neutral attitude does not change the user's opinion, it can be classified into the other two categories, or ignored. According to this assumption, users' attitudes can be divided into two categories: like and dislike. A user's check-in record is positive feedback. For example, a user who checks in a restaurant indicated that he likes the restaurant.

If a user repeatedly checks in a POI, it indicates that the user likes the POI more. Similarly, if the user does not like it, the record of it will not exist. We assume that if the user checks in to the current POI multiple times, the probability of the POI being checked in again will increase. Based on this hypothesis, we use the following equation to express the exposure effect in the subtrajectory of user u . Multiple exposures of POI v may make the user's loving old preference to be enhanced.

$$\Gamma(u, v_\kappa)_{s_u^j} = 1 + \left(\frac{\sum_{r=1}^{|s_u^j|} \mathbb{1}_{v_r=v_\kappa}}{R} \right)_{v_\kappa \in s_u^j}^{\phi_u}, \quad (5)$$

where $\mathbb{1}_{\text{cond}}$ is the indicator function, and it returns 1 if cond is satisfied, or otherwise returns 0. R is the maximum number that the user checked in repeatedly in all trajectories. ϕ_u is u 's personalized parameter for exposure. The exposure influence $\Gamma(u, s_u^j, v)$ monotonically increases as the POI is visited repeatedly.

4.1.3. Modeling the Influence Weight of User Pursuing New. In real life, visiting a POI is a consumable activity. For example, users need to spend time or move a certain distance to enjoy the happiness brought by POI-related activities [18]. The memory of the user's access to the POI will affect the novelty of the subsequent POI, and the impact is related to the time interval. Like depreciation in economics, the residual value is smaller when the time is longer, and the possibility of purchasing new equipment is greater [25]. We assume that the POI currently checked in by the user will affect the utility of subsequent POIs, and this effect is related to the time interval. The larger the interval, the weaker the impact, making users more likely to choose a new POI. For convenience, we use interval hops instead of time intervals. For example, for the sequence "a, b, c, d," the number of hops between ab, ac, and ad are 1, 2, and 3. We assume that the influence of the current POI on the rear target POI is inversely proportional to the total number of hops between them, and we use the following equation to express the pursuit of newness.

$$\Psi(u, \kappa)_{s_u^j} = \frac{c_u}{\kappa^{\alpha_u} e^{\beta_u \kappa}}, \quad (6)$$

where c_u , α_u , and β_u are personalized parameters and κ refers to the intervals between the current POI to the target POI visited by the same user.

4.1.4. Modeling Geographical Influence. Some existing studies show that the geographical proximities of POIs have a significant influence on users' check-in behavior. To better understand this geographical influence on users, Ye et al. [35] performed a spatial analysis on user check-in activities records collected from the Foursquare. A significant percentage of two consecutive POIs visited by the same user appear to be within a short distance. There are three tendencies: (1) people tend to select POIs near their homes or workplaces; (2) people may prefer to explore POIs from current location; and (3) the POI visited by the user conforms to the Gaussian distribution of multiple centers, which means that the POI visited by the user is not evenly distributed but concentrated in several areas [48]. Based on the above conclusions, it can be seen that distance is the core of geographic influence. There are many mature models for modeling geographic influence, for example, power-law, exponential, and hyperbolic [34]. Thus, we also use power-law distribution to model the check-in probability to the distance between two POIs visited by the same user.

$$Y(d) = a_u \times d^{-b_u}, \quad (7)$$

where a_u and b_u are parameters of a power-law distribution and d refers to the distance between the current POI to the target POI visited by the same user.

4.2. Integrating Model and Inferring Model Parameters. According to the above submodel representation, we change Equation (2) to Equation (8). The weight w_u^i becomes a combination of distance, liking the old, and pursuing the new.

$$\begin{aligned} \Pr(v_r | s_u^j) &\approx \sum_{i=1}^t w_u \Pr(v_{t+1}, v_{t+1-i}) \\ &= \sum_{i=1}^t Y(d) \cdot \left[\mu \cdot \Gamma(u, v_i)_{s_u^j} + (1 - \mu) \Psi(u, i)_{s_u^j} \right] \\ &\quad \cdot \Pr(v_{t+1}, v_{t+1-i}), \end{aligned} \quad (8)$$

where μ is the parameter of liking old and $1 - \mu$ is the pursuing new, respectively.

Let θ denote the set of parameters in the personalized SONG framework. Then, we define our optimization problem to obtain the optimal θ^* as follows:

$$\Theta^* = \operatorname{argmax}_{\Theta} \prod_u \prod_{s_u, s_u^j \in S_u} P_r(v_r | s_u^j; \Theta). \quad (9)$$

The goal of this problem is aimed at maximizing the probability of predicting the last POI given the rest of an observing sequence s_u . To solve this optimization problem, we employ the maximum a posteriori (MAP) estimation on the following log-likelihood function:

$$\begin{aligned}
\Theta^* &= \underset{\Theta}{\operatorname{argmin}} \mathcal{L} = -\sum_u \sum_{s_u, s_u^j \in \mathcal{S}_u} \ln P(v | s_u^j; \Theta) \\
&= -\sum_u \sum_{s_u} \ln \sum_{i=1}^t Y(d) \cdot \left[\mu \cdot \Gamma(u, v_i)_{s_u^j} + (1 - \mu) \Psi(u, i)_{s_u^j} \right] \\
&\quad \cdot \Pr(v_{t+1}, v_{t+1-i}).
\end{aligned} \tag{10}$$

Furthermore, the gradient of the log-likelihood L with respect to the model parameters is given by

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial \theta} &= -\sum_u \sum_{s_u} \frac{\sum_{i=1}^t \left[Y(\cdot) \cdot \mu \left(\frac{\partial \Gamma(\cdot)_{s_u^j}}{\partial \theta} \right) \right] \Pr(\cdot)}{\sum_{i=1}^t Y(\cdot) \cdot \left[\mu \cdot \Gamma(\cdot)_{s_u^j} + (1 - \mu) \Psi(\cdot)_{s_u^j} \right] \Pr(\cdot)}, \theta \in \{\phi_u\}, \\
\frac{\partial \mathcal{L}}{\partial \theta} &= -\sum_u \sum_{s_u} \frac{\sum_{i=1}^t \left[Y(\cdot) (1 - \mu) \left(\frac{\partial \Psi(\cdot)_{s_u^j}}{\partial \theta} \right) \right] \Pr(\cdot)}{\sum_{i=1}^t Y(\cdot) \cdot \left[\mu \cdot \Gamma(\cdot)_{s_u^j} + (1 - \mu) \Psi(\cdot)_{s_u^j} \right] \Pr(\cdot)}, \theta \in \{\alpha_u, c_u, \beta_u\}, \\
\frac{\partial \mathcal{L}}{\partial \theta} &= -\sum_u \sum_{s_u} \frac{\sum_{i=1}^t Y(\cdot) \cdot \left(\Gamma(\cdot)_{s_u^j} - \Psi(\cdot)_{s_u^j} \right) \Pr(\cdot)}{\sum_{i=1}^t Y(\cdot) \cdot \left[\mu \cdot \Gamma(\cdot)_{s_u^j} + (1 - \mu) \Psi(\cdot)_{s_u^j} \right] \Pr(\cdot)}, \theta \in \{\mu\}, \\
\frac{\partial \mathcal{L}}{\partial \theta} &= -\sum_u \sum_{s_u} \frac{\sum_{i=1}^t (\partial Y(d) / \partial \theta) \cdot \left[\mu \cdot \Gamma(\cdot)_{s_u^j} + (1 - \mu) \Psi(\cdot)_{s_u^j} \right] \Pr(\cdot)}{\sum_{i=1}^t Y(d) \cdot \left[\mu \cdot \Gamma(\cdot)_{s_u^j} + (1 - \mu) \Psi(\cdot)_{s_u^j} \right] \Pr(\cdot)}, \theta \in \{a_u, b_u\}.
\end{aligned} \tag{11}$$

In order to make the formula easier to understand and save space, $Y(d)$, $\Gamma(u, v_i)_{s_u^j}$, $\Psi(u, i)_{s_u^j}$, and $\Pr(v_{t+1}, v_{t+1-i})$ are abbreviated as $Y(\cdot)$, $\Gamma(\cdot)_{s_u^j}$, $\Psi(\cdot)_{s_u^j}$, and $\Pr(\cdot)$, respectively.

In addition, the set of parameters θ in our method consists of all the personalized parameters in the definitions of liking old and pursuing new POIs, such as ϕ_u , α_u , β_u , c_u , a_u , b_u , and μ . α_u is randomly drawn from absolute $N(0, 0.1)$. μ is initialized to 0.5. The other parameters are randomly initialized in the range $[0.0, 1.0]$.

Given a training set containing observing sequences of agents, we can iteratively update the parameters θ using the gradient ascent method. Once the (near) optimal Θ is obtained, our SONG framework allows for personalized recommendations based on Equation (8).

5. Experiments Evaluation

In this section, we describe our experiment settings for evaluating the performance of SONG against the state-of-the-art POI recommendation techniques.

5.1. Experiment Settings. Each dataset is divided into the training and testing set in terms of visit time rather than random partition methods in the experiments. In practice, we can only utilize the past data to predict future check-in events. For each user, we select 80% of her check-ins to constitute a training dataset for learning the parameters of the

proposed model. The remaining portions are used as ground truth for testing.

5.2. Evaluation Metrics. To study the effectiveness of the methods, we use two popular metrics, i.e., Recall (Rec@ k) and normalized discounted cumulative gain (NDCG@ k). While the Recall is defined as the ratio of recovered POIs to the number of POIs generated by the recommendation model, NDCG@ N is a measure to indicate the ranking quality of the recommendation models.

To produce a TOP- k recommendation list for a query user, we compute a preference score for each POI and sort them by score. The Recall@ k for each user is defined as

$$\operatorname{Rec}@k = \frac{tp}{tp + tn}, \tag{12}$$

where tp is the number of POIs visited by a user u and also in the TOP- k recommendations and tn is the number of POIs visited by u but not in the TOP- k recommendations. The NDCG@ k for each user is defined as

$$\operatorname{NDCG}@k = \frac{\operatorname{DCG}@k}{\operatorname{IDCG}@k}, \tag{13}$$

where

$$\operatorname{DCG}@k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i+1)}, \tag{14}$$

and rel_i refers to the graded relevance of the result ranked at the position. We use the binary relevance in our work, i.e., $rel_i = 1$ if the result is in the test set, and 0, otherwise. $\operatorname{IDCG}@k$ is the $\operatorname{DCG}@k$ value when the recommended POIs are ideally ranked. The average of recall and NDCG values overall users are reported as the final $\operatorname{Rec}@k$ and $\operatorname{NDCG}@k$ ($k = 1, 5, 10, 15$). These two metrics are both in the range $[0, 1]$, and a higher value means better results.

5.3. Comparison Methods

5.3.1. FPMC-LR. Cheng et al. [8] observed two prominent properties in the check-in sequence: personalized Markov chain and region localization. Then, they proposed a matrix factorization method, which embeds the personalized Markov chains and the localized regions for solving the recommendation task.

5.3.2. TMCA. Li et al. [44] proposed an encoder-decoder-based neural network model named TMCA to capture the complex spatial and temporal dependencies among historical check-in activities automatically. The model leverages the embedding method to incorporate heterogeneous contextual factors to boost recommendation performance. Furthermore, they introduce the temporal and multilevel context attention mechanisms to dynamically select the relevant check-ins and discriminative contextual factors for predicting the preferences over POIs to visit next.

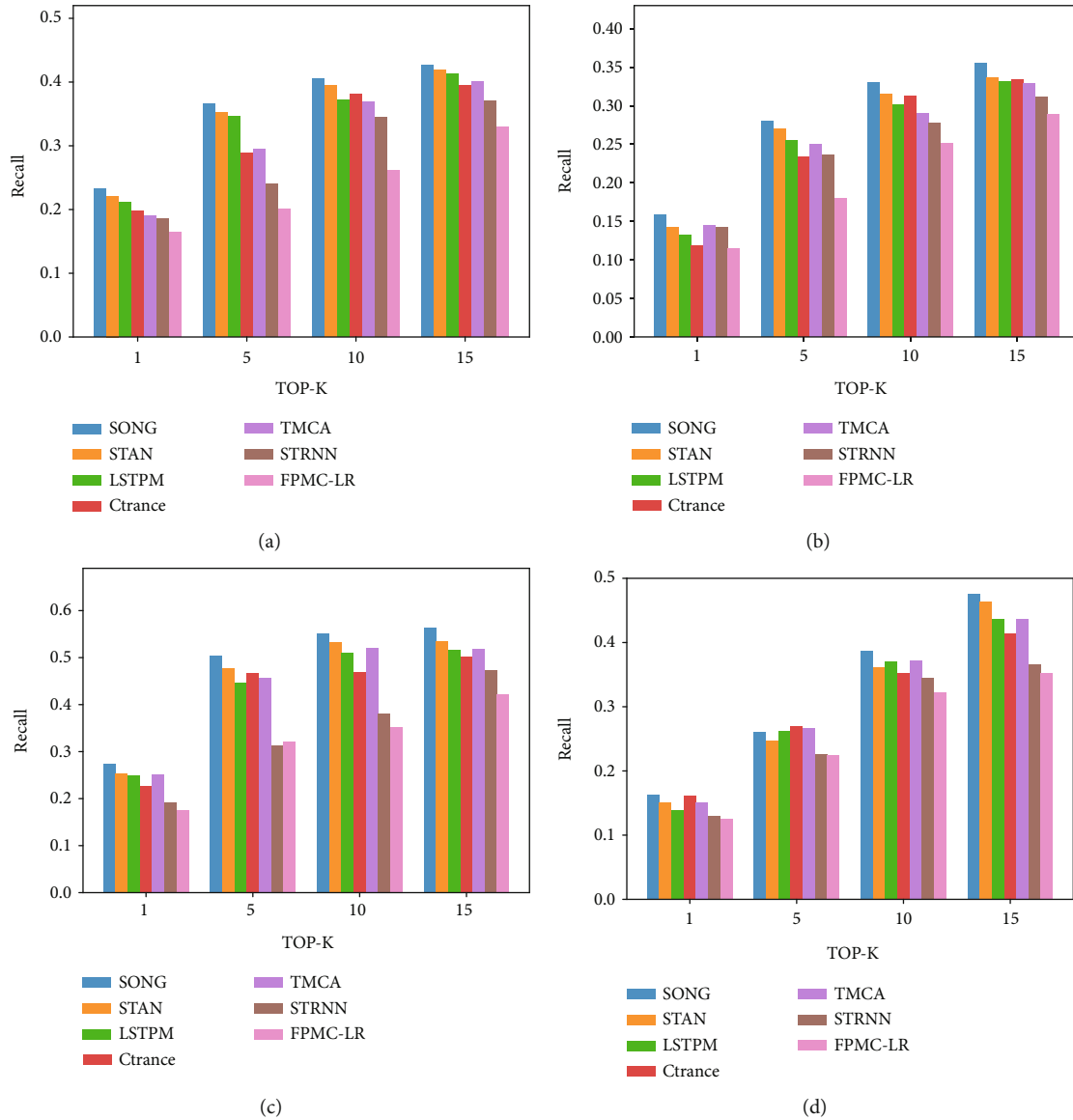


FIGURE 2: Recall on four datasets. (a) Beijing (Jiepan). (b) Shanghai (Jiepan). (c) NYC (Foursquare). (d) TKY (Foursquare).

5.3.3. *STRNN*. Liu et al. [42] extend RNN and propose the spatial-temporal recurrent neural networks (STRNN). To capture time interval and geographical distance information, they replace the single-transition matrix in RNN with time-specific transition matrices and distance-specific transition matrices.

5.3.4. *CTransRec*. CTransRec [40] is a translation-based recommender for complex users, which utilizes auxiliary information (item category and timestamp) and category-specific projection and temporal dynamic relaxation in recommender systems to improve the performance of sequential recommendation.

5.3.5. *LSTPM*. Sun et al. [43] proposed a novel method named long-term and short-term preference modeling for

the next POI recommendation. The proposed model consists of a context-aware nonlocal network for long-term preference modeling and a geodilated RNN for short-term preference learning.

5.3.6. *STAN*. Luo et al. [41] adopt a spatiotemporal attention network (STAN) for location recommendation. This allows a point-to-point interaction between nonadjacent locations and nonconsecutive check-ins with explicit spatiotemporal effect. STAN uses a bilayer attention architecture to learn the explicit spatiotemporal correlations.

5.4. *Comparison with Baselines*. We present the comparison results on the four datasets in terms of Recall and NDCG in Figures 2 and 3, respectively. We can make the following important observations.

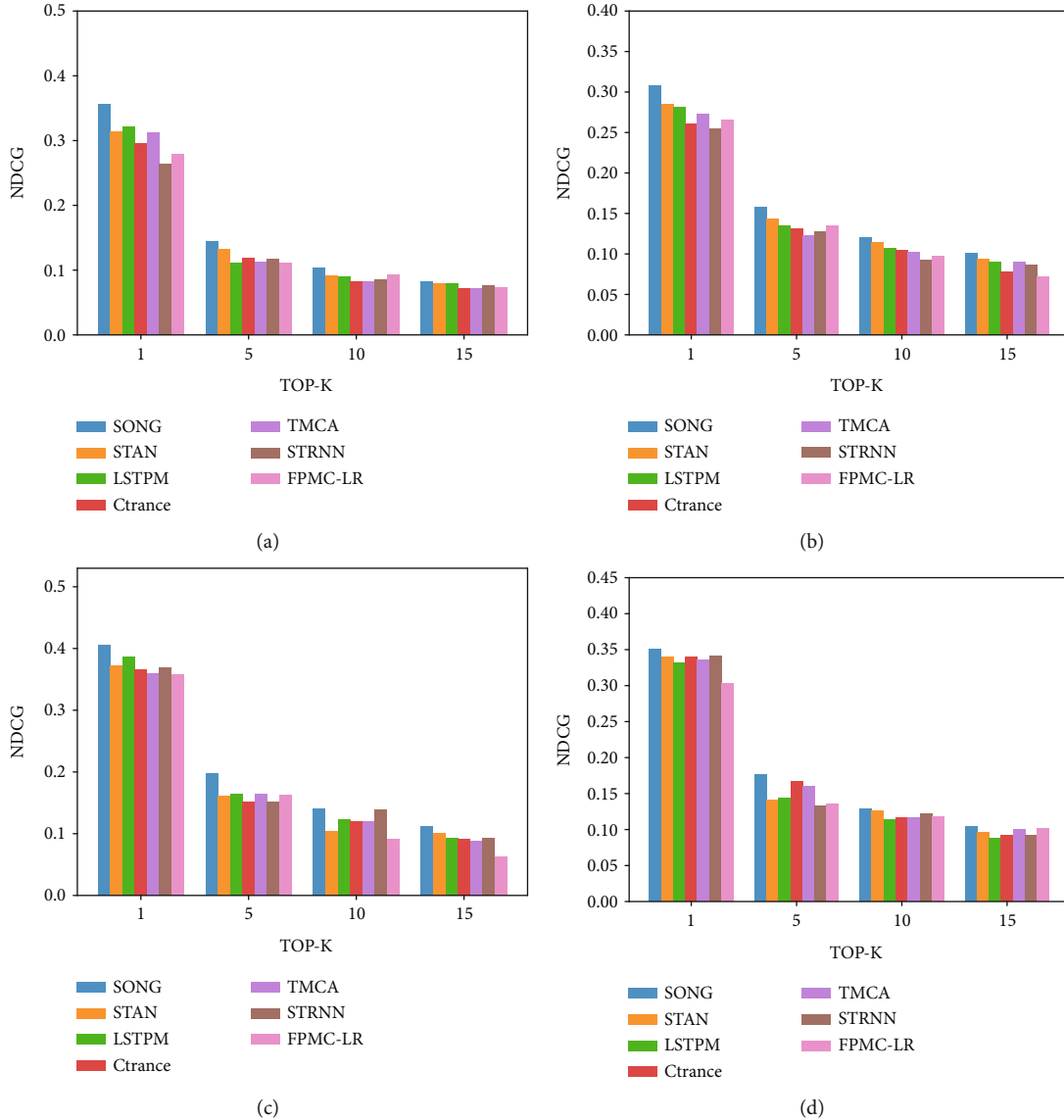


FIGURE 3: NDCG on four datasets. (a) Beijing (JiePang). (b) Shanghai (JiePang). (c) NYC (Foursquare). (d) TKY (Foursquare).

We model users' preference for liking old, assume that repeated check-in POI was the basis of users' liking old, and converted the check-in times into liking old influence. New POIs will bring users unprecedented experience. We calculate the residual value of POIs according to the time interval between check-in POIs, which will serve as the basis for users to pursue new POIs. The lower the residual value, the stronger the incentive to pursue something new. Finally, the combination of liking the old and chasing the new can better reflect the influence of precursor POIs on successor POIs.

Our proposed SONG consistently and significantly outperforms all baselines in terms of every metric on four datasets. For example, on Beijing, Shanghai, NYC, and TKY, compared with the second best method STAN, SONG improves the Recall@1 by 5.22%, 11.17%, 8.22%, and 7.62%, and the NDCG@1 13.55%, 8.17%, 8.88%, and

3.24% by, respectively. The quantitative evaluation clearly demonstrates the superior effectiveness of our method.

Among the baseline methods, STAN performs the best on all datasets. It may be because STAN established the association between nonadjacent locations and nonconsecutive check-ins with explicit spatiotemporal effect, which has alleviated data sparsity. Secondly, the performances of LSTPM and TMCA are better. Specifically, LSTPM performed well on BJ and SH of the JiePang, and TMCA performed well on NYC and TKY of the Foursquare. LSTPM requires dense trajectory support, which is not very good for users with sparse trajectories. In comparison, TMCA adopted an embedding method to incorporate heterogeneous information for mitigating sparsity.

Ctrance used item category and timestamp, and STRNN used time interval and geographic distance. Both methods use time, which shows that time played a role in the

TABLE 1: Performance of SONG_N and SONG on four datasets.

Data	Metric	SONG_N	SONG	Inc
BJ	Rec@1	0.2049	0.2327	14%
	Rec@5	0.3208	0.3662	14%
	Rec@10	0.3577	0.4055	13%
	Rec@15	0.3710	0.4270	15%
SH	Rec@1	0.1488	0.1581	6%
	Rec@5	0.2622	0.2800	7%
	Rec@10	0.3094	0.3297	7%
	Rec@15	0.3291	0.3555	8%
NYC	Rec@1	0.2337	0.2740	17%
	Rec@5	0.4361	0.5042	16%
	Rec@10	0.4800	0.5523	15%
	Rec@15	0.4909	0.5639	15%
TKY	Rec@1	0.1769	0.1994	13%
	Rec@5	0.3364	0.3724	11%
	Rec@10	0.3909	0.4318	10%
	Rec@15	0.4192	0.4623	10%
BJ	NDCG@1	0.3256	0.3564	9%
	NDCG@5	0.1341	0.1440	7%
	NDCG@10	0.0959	0.1030	7%
	NDCG@15	0.0771	0.0832	8%
SH	NDCG@1	0.2978	0.3084	4%
	NDCG@5	0.1522	0.1580	4%
	NDCG@10	0.1170	0.1209	3%
	NDCG@15	0.0972	0.1008	4%
NYC	NDCG@1	0.3580	0.4056	13%
	NDCG@5	0.1785	0.1973	11%
	NDCG@10	0.1280	0.1407	10%
	NDCG@15	0.1024	0.1124	10%
TKY	NDCG@1	0.3248	0.3506	8%
	NDCG@5	0.1649	0.1765	7%
	NDCG@10	0.1199	0.1283	7%
	NDCG@15	0.0977	0.1044	7%

recommendation. Comparing the performance of them, the former is slightly better than the latter. This result may indicate that the item category better reflects the user's preferences.

5.5. Comparison of SONG_N and SONG. Due to the uneven distribution of the number of daily check-ins, dense and sparse trajectories coexist. In order to make full use of user data, this paper proposes a method to serialize sparse trajectories. We test the proposed SONG on the original and the serialized data, and the test results are shown in Table 1. It is found that the effects of Recall and NDCG after serialization (SONG) are significantly higher than those of no serialization (SONG_N). It shows in Table 1. For example, on BJ, SH, NYC, and TKY, Rec@ k ($k = 1, 5, 10, 15$) has increased by more than 13%, 6%, 15%, and 10%, respectively. Comparing

the growth rates of Recall and NDCG of each dataset, NYC and BJ have the best growth rates, followed by TKY and SH. The growth trends of the two indicators are consistent. From this point of view, the sparse trajectory data serialization method we proposed is effective.

6. Conclusions

In this paper, we find out that users' check-in behaviors have an apparent characteristic of liking the old and pursuing the new. After studying the problem of POI recommendation, we proposed a novel sequential prediction method based on the Markov chain model, i.e., SONG. We jointly model users' behaviors and geographical influence with a variable-order additive Markov chain. The experimental results demonstrate that our proposed approach substantially improves the recommendation performance compared with the state-of-the-art methods.

Data Availability

The data used to support the findings of this study are included within the supplementary information file.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Supplementary Materials

We used two datasets in this paper. They are Foursquare [45] and JiePang [11]. Anyone can freely obtain the dataset from the following URLs, Foursquare (<https://sites.google.com/site/yangdingqi/home/foursquare-dataset>) and JiePang (https://pan.baidu.com/s/1ntXTnuT?_at_=1637459298206). (*Supplementary Materials*)

References

- [1] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W. Y. Ma, "Recommending friends and locations based on individual location history," *ACM Transactions on the Web (TWEB)*, vol. 5, no. 1, pp. 1–44, 2011.
- [2] H. Gao, X. Qin, R. J. D. Barroso, W. Hussain, Y. Xu, and Y. Yin, "Collaborative learning-based industrial IoT API recommendation for software-defined devices: the implicit knowledge discovery perspective," *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–11, 2020.
- [3] Y. Xu, Y. Wu, H. Gao, S. Song, Y. Yin, and X. Xiao, "Collaborative APIs recommendation for Artificial Intelligence of Things with information fusion," *Future Generation Computer Systems*, vol. 125, pp. 471–479, 2021.
- [4] H. Gao, C. Liu, Y. Li, and X. Yang, "V2VR: reliable hybrid-network-oriented V2V data transmission and routing considering RSUs and connectivity probability," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 6, pp. 3533–3546, 2021.
- [5] P. Kefalas, P. Symeonidis, and Y. Manolopoulos, "A graph-based taxonomy of recommendation algorithms and systems in LBSNs," *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 604–622, 2016.

- [6] J. Ye, Z. Zhu, and H. Cheng, "What's your next move: user activity prediction in location-based social networks," in *Proceedings of the 2013 SIAM International Conference on Data Mining*, pp. 171–179, Texas, USA, 2013.
- [7] S. Kim, "Friend recommendation with a target user in social networking services," in *2015 31st IEEE International Conference on Data Engineering Workshops*, pp. 235–239, Seoul, Korea (South), 2015.
- [8] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: successive point-of-interest recommendation," in *Twenty-Third international joint conference on Artificial Intelligence*, vol. 13, pp. 2605–2611, Beijing, China, 2013.
- [9] W. He, D. Li, T. Zhang, L. An, M. Guo, and G. Chen, "Mining regular routes from gps data for ridesharing recommendations," in *Proceedings of the ACM SIGKDD International Workshop on Urban Computing ACM*, pp. 79–86, Beijing, China, 2012.
- [10] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann, "Time-aware point-of-interest recommendation," in *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval ACM*, pp. 363–372, Dublin, Ireland, 2013.
- [11] D. Lian, C. Zhao, X. Xie et al., "Joint geographical modeling and matrix factorization for point-of-interest recommendation," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining ACM*, pp. 831–840, New York, USA, 2014.
- [12] H. Yin, X. Zhou, Y. Shao, H. Wang, and S. Sadiq, "Joint modeling of user check-in behaviors for point-of-interest recommendation," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pp. 1631–1640, Melbourne, Australia, 2015.
- [13] B. Liu, H. Xiong, S. Papadimitriou, Y. Fu, and Z. Yao, "A general geographical probabilistic factor model for point of interest recommendation," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 5, pp. 1167–1179, 2015.
- [14] R. Baral and T. Li, "Maps: a multi aspect personalized poi recommender system," in *Proceedings of the 10th ACM Conference on Recommender Systems*, pp. 281–284, Boston Massachusetts, USA, 2016.
- [15] J. B. Griesner, T. Abdesslem, and H. Naacke, "POI recommendation: towards fused matrix factorization with geographical and temporal influences," in *Proceedings of the 9th ACM Conference on Recommender Systems ACM*, pp. 301–304, New York, USA, 2015.
- [16] X. Li, G. Cong, X. L. Li, T. A. N. Pham, and S. Krishnaswamy, "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*, pp. 433–442, New York, USA, 2015.
- [17] D. Chen, C. S. Ong, and L. Xie, "Learning points and routes to recommend trajectories," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management ACM*, pp. 2227–2232, Indianapolis Indiana, USA, 2016.
- [18] Y. Ge, H. Xiong, A. Tuzhilin, and Q. Liu, "Cost-aware collaborative filtering for travel tour recommendations," *ACM Transactions on Information Systems*, vol. 32, no. 1, pp. 1–31, 2014.
- [19] C. Cheng, H. Yang, I. King, and M. Lyu, "Fused matrix factorization with geographical and social influence in location-based social networks," *Proceedings of the AAAI conference on artificial intelligence*, vol. 26, 2012.
- [20] E. Cho, S. A. Myers, and J. 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*, pp. 1082–1090, San Diego California, USA, 2011.
- [21] W. R. Tobler, "A computer movie simulating urban growth in the Detroit region," *Economic geography*, vol. 46, pp. 234–240, 1970.
- [22] E. O'Brien, "A mind stretched: the psychology of repeat consumption," *Consumer Psychology Review*, vol. 4, no. 1, pp. 42–58, 2021.
- [23] M. Xie, H. Yin, H. Wang, F. Xu, W. Chen, and S. Wang, "Learning graph-based poi embedding for location-based recommendation," in *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management ACM*, pp. 15–24, Indianapolis Indiana, USA, 2016.
- [24] R. B. Zajonc, "Mere exposure: a gateway to the subliminal," *Current Directions in Psychological Science*, vol. 10, no. 6, pp. 224–228, 2001.
- [25] H. Hotelling, "A general mathematical theory of depreciation," *Journal of the American Statistical Association*, vol. 20, no. 151, pp. 340–353, 1925.
- [26] Y. Huang, H. Xu, H. Gao, X. Ma, and W. Hussain, "SSUR: an approach to optimizing virtual machine allocation strategy based on user requirements for cloud data center," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 2, pp. 670–681, 2021.
- [27] C. Zhang, H. Liang, and K. Wang, "Trip recommendation meets real-world constraints," *ACM Transactions on Information Systems*, vol. 35, no. 1, pp. 1–28, 2016.
- [28] T. He, H. Yin, Z. Chen, X. Zhou, S. Sadiq, and B. Luo, "A spatial-temporal topic model for the semantic annotation of POIs in LBSNs," *ACM Transactions on Intelligent Systems and Technology*, vol. 8, no. 1, pp. 1–24, 2016.
- [29] S. Oppokhonov, S. Park, and I. K. Ampomah, "Current location-based next POI recommendation," in *Proceedings of the International Conference on Web Intelligence ACM*, pp. 831–836, Leipzig, Germany, 2017.
- [30] H. Gao, J. Tang, X. Hu, and H. Liu, "Exploring temporal effects for location recommendation on location-based social networks," in *Proceedings of the 7th ACM conference on Recommender systems ACM*, pp. 93–100, Hong Kong, China, 2013.
- [31] W. Ji, X. Meng, and Y. Zhang, "SPATM: a social period-aware topic model for personalized venue recommendation," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1–14, 2020.
- [32] J. D. Zhang, C. Y. Chow, and Y. Li, "Lore: exploiting sequential influence for location recommendations," in *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems ACM*, pp. 103–112, Dallas Texas, USA, 2014.
- [33] S. Zhao, I. King, and M. R. Lyu, "A survey of point-of-interest recommendation in location-based social networks," <http://arxiv.org/abs/1607.00647>.
- [34] H. Wang, H. Shen, W. Ouyang, and X. Cheng, "Exploiting POI-specific geographical influence for point-of-interest recommendation," in *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pp. 3877–3883, Stockholm, Sweden, 2018.

- [35] M. Ye, P. Yin, W. C. Lee, and D. L. Lee, "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*, pp. 325–334, 2011.
- [36] R. He and J. McAuley, "Fusing similarity models with Markov chains for sparse sequential recommendation," in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pp. 191–200, Barcelona, Spain, 2016.
- [37] K. Zhao, Y. Zhang, H. Yin et al., "Discovering subsequence patterns for next POI recommendation," in *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pp. 3216–3222, Yokohama, Japan, 2020.
- [38] J. Xiao, H. Xu, H. Gao, M. Bian, and Y. Li, "A weakly supervised semantic segmentation network by aggregating seed cues: the multi-object proposal generation perspective," *ACM Transactions on Multimedia Computing Communications and Applications*, vol. 17, no. 1s, pp. 1–19, 2021.
- [39] Y. Yin, Z. Cao, Y. Xu, H. Gao, R. Li, and Z. Mai, "QoS prediction for service recommendation with features learning in mobile edge computing environment," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 4, pp. 1136–1145, 2020.
- [40] H. Li, Y. Liu, N. Mamoulis, and D. S. Rosenblum, "Translation-based sequential recommendation for complex users on sparse data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 8, pp. 1639–1651, 2020.
- [41] Y. Luo, Q. Liu, and Z. Liu, "STAN: spatio-temporal attention network for next location recommendation," in *Proceedings of the Web Conference*, pp. 2177–2185, Ljubljana, Slovenia, 2021.
- [42] Q. Liu, S. Wu, L. Wang, and T. Tan, "Predicting the next location: a recurrent model with spatial and temporal contexts," in *Thirtieth AAAI conference on artificial intelligence*, pp. 194–200, Phoenix, Arizona, USA, 2016.
- [43] K. Sun, T. Qian, T. Chen, Y. Liang, Q. V. H. Nguyen, and H. Yin, "Where to go next: modeling long-and short-term user preferences for point-of-interest recommendation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 214–221, 2020.
- [44] R. Li, Y. Shen, and Y. Zhu, "Next point-of-interest recommendation with temporal and multi-level context attention," in *2018 IEEE International Conference on Data Mining (ICDM)*, pp. 1110–1115, Singapore, 2018.
- [45] Dingqi Yang, Daqing Zhang, V. W. Zheng, and Zhiyong Yu, "Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 45, no. 1, pp. 129–142, 2015.
- [46] W. K. Ching, E. S. Fung, and M. K. Ng, "Higher-order Markov chain models for categorical data sequences," *Naval Research Logistics*, vol. 51, no. 4, pp. 557–574, 2004.
- [47] A. E. Raftery, "A model for high-order Markov chains," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 47, no. 3, pp. 528–539, 1985.
- [48] C. Su, J. Wang, and X. Xie, "Point-of-interest recommendation based on geographical influence and extended pairwise ranking," in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 966–971, Toronto, ON, Canada, 2020.