

Article

Context-Aware Point-of-Interest Recommendation Based on Similar User Clustering and Tensor Factorization

Yan Zhou ^{1,2,3} , Kaixuan Zhou ^{1,3,*} and Shuaixian Chen ¹

¹ School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China

² The Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources, Shenzhen 518063, China

³ The Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou 313099, China

* Correspondence: kaixuan310@std.uestc.edu.cn

Abstract: The rapid development of big data technology and mobile intelligent devices has led to the development of location-based social networks (LBSNs). To understand users' behavioral patterns and improve the accuracy of location-based services, point-of-interest (POI) recommendation has become an important task. In contrast to the general task of product recommendation, POI recommendation faces the problems of the sparsity and weak semantics of user check-in data. To address these issues, an increasing number of studies have improved the accuracy of POI recommendations by introducing contextual information such as geographical, temporal, textual, and social relations. However, the rich context also brings great challenges to POI recommendation, such as the low utilization rate of context information, difficulty in balancing the richness of contextual information, and the complexity of the recommendation matrix. Considering that similar users have more interest preferences in common than users generally have, the check-in information of similar users has greater reference meaning. Thus, we propose a personalized POI recommendation method named CULT-TF, which incorporates similar users' contextual information into the tensor factorization model. First, we present a user activity model and a user similarity model, which integrate contextual information to calculate the user activity and similarity between users. According to user activity, the most representative active users are selected as user clustering centers, and then users are clustered based on user similarity into several similar user clusters (C). Next, we construct a third-order tensor (user-location-time matrix) for each user cluster by using user activity, POI popularity, and time slot popularity as the eigenvalues in the user (U), location (L), and time (T) dimensions, and the eigenvalue of each dimension is modeled by integrating contextual information of users' check-in behavior at the user, location, and time levels. Similar user clustering reduces the number of users in tensor modeling, reducing the U dimension. To further reduce the complexity of the recommendation matrix, the reduction of the L dimension is achieved through ROI (region of interest) clustering, and the reduction of the T dimension is achieved through time slot encoding. Then, we use tensor factorization (TF) to obtain the recommendation results. Our method decreases the complexity of the tensor matrix and integrates rich contextual information on users' check-in behavior. Finally, we conducted a comprehensive performance evaluation of CULT-TF using real-world LBSN datasets from Brightkite. The experimental results show that our proposed method performs much better than other recommendation methods in terms of precision and recall.

Keywords: point-of-interest (POI) recommendation; location-based social network; tensor factorization; context information



Citation: Zhou, Y.; Zhou, K.; Chen, S. Context-Aware Point-of-Interest Recommendation Based on Similar User Clustering and Tensor Factorization. *ISPRS Int. J. Geo-Inf.* **2023**, *12*, 145. <https://doi.org/10.3390/ijgi12040145>

Academic Editors: Hartwig H. Hochmair and Wolfgang Kainz

Received: 12 January 2023

Revised: 23 March 2023

Accepted: 27 March 2023

Published: 29 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the wide availability of mobile devices (e.g., smartphones and tablets) and advances in mobile networks have led to the development of location-based social net-

works (LBSNs). LBSNs enable users to check in to geographic locations via social networks, generating a large amount of point-of-interest (POI) data such as social, spatiotemporal, and content information. Unlike online social networks, in addition to the social relations between users, LBSNs have many geographic relations between locations and check-in relations between locations and users [1].

POI recommendation filters massive geographic information data in LBSNs by analyzing user behavior and helping users filter useless information, which plays an important role in personalized recommendation systems. POI recommendation also helps service providers promote personalized services to potential users [2,3]. However, unlike traditional product recommendation tasks that can be used to directly obtain explicit feedback through user ratings, POI check-in data cannot directly reflect users' POI preferences and are weak in semantics [4]. In addition, people's check-in activities tend to be concentrated in a few areas and are not constant throughout the day, resulting in high data sparsity [5,6]. At present, many studies have addressed the weak semantics and high sparsity by introducing contextual information, such as location [7], category [8] and text information [9] of POIs, user check-in time [10], and social relationships [11]. Although incorporating contextual information helps in the understanding of the real preferences of users' check-in behavior and improves the accuracy of POI recommendations, it is difficult to balance the richness of contextual information with the complexity of the recommendation matrix, which could affect the performance of POI recommendation [12,13].

Therefore, a personalized POI recommendation method named CULT-TF, which incorporates similar users' contextual information into the tensor factorization model, is proposed in this paper. The following contributions are provided by this work:

1. We define a user activity model and a user similarity model that can integrate contextual information of users' check-in behavior to calculate user activity and user similarity;
2. A similar user clustering method based on user activity and user similarity is presented to select the most influential active users as clustering centers based on user activity and to cluster users into several similar user clusters according to user similarity;
3. A U-L-T tensor that incorporates contextual information using user activity, POI popularity, and time slot popularity as the eigenvalues in the U, L, and T dimensions, which improves the integration of contextual information, is presented;
4. The CULT-TF recommendation method based on tensor factorization, which decreases the complexity of the matrix-integrating rich contextual information by clustering similar users, clustering POIs into regions of interest (ROIs), and encoding check-in timestamps to time slots, to realize the reduction of the U, L, and T dimension, respectively, is proposed. In this way, CULT-TF reduces the complexity of the recommendation matrix while integrating the richness of the contextual information.

The rest of this paper is organized as follows: Section 2 introduces the related work on POI recommendation methods. Section 3 summarizes the technical framework of this method. Section 4 elaborates on the principle of the CULT-TF method. Section 5 presents experiments and analysis. Section 6 is a discussion. Conclusions and future work are presented in Section 7.

2. Related Work

POI recommendation based on LBSNs has been widely studied. Much work has been performed in this area based on core ideas such as collaborative filtering (CF) and matrix factorization (MF). As one of the early trending research domains of recommendation systems, collaborative filtering has played an important role in the development of recommendation systems for years and is also the basis of many other recommendation models [14,15]. Matrix factorization can mine the implicit feature relationship between users and POIs and achieves good results in handling sparse data [16,17]. Thus, tensors are a natural choice for simulating high-level context information in POI recommendations [18]. Some researchers use tensor factorization for traditional recommendation systems with both explicit and implicit feedback [19]. Others recommend POIs by integrating time and

other contextual information into the tensor factorization framework, which is more reasonable for predicting where a user may go in a certain time period [20]. Furthermore, along with the rapid growth of data, an increasing number of researchers have paid attention to POI recommendation methods that consider one or more factors, such as geographical, temporal, social, text, and category factors. Most researchers use user check-in data and contextual information to improve recommendation effectiveness [21]. According to different types of contextual information, the existing POI recommendation methods can be classified into several main categories.

2.1. POI Recommendation Based on Geographical Influence

According to Tobler's first law of geography, everything is related to everything else, but near things are more related than distant things [22]. Ye et al. investigated the strong correlation between friend relationships and geographic locations and proposed a naive Bayes algorithm based on user social relationships and geographic locations in a recommendation model (USG) [15]. Cheng et al. incorporated social and geographical information into a generalized matrix factorization model through a polycentric Gaussian model (MGMPFM) [17]. Some studies have introduced geographic information by integrating geographic influence into matrix factorization [20,23] or by converting geographical influence into a weight of relevant similarity as input to a collaborative filtering model [14]. Due to the spatial aggregation phenomenon, Luan et al. proposed a partition-based collaborative tensor factorization (PCTF) method [24], which further improved the recommendation accuracy.

2.2. POI Recommendation Based on Temporal Influence

Users' check-in behavior tends to show periodicity, continuity, or nonuniformity in terms of temporal characteristics, which has inspired researchers to incorporate temporal influence into POI recommendations. The use of a real-time factor in the temporal characteristics can clearly portray the user's preference at the current moment, so adding this factor to the POI recommendation can yield more accurate predictions of the user's interest preferences [25]. Some studies have introduced the periodicity of user behavior, such as checking in at the office on weekdays and shopping centers on weekends [26]. Tour recommendations often include temporal continuity, where tourists tend to visit POIs in a certain geographic order. The time-series pattern of a location from the historical check-in trajectory of a tour is mined to guide the next POI recommendation in the time series [27]. Some studies have used Markov chains to capture the time-series impact of user check-in locations for POI recommendation [28]. Other studies have analyzed the influence of nonuniform temporal characteristics on user check-in behavior. For example, Li et al. proposed a fourth-order tensor factorization-based ranking methodology that considers time-varying behavioral trends while capturing nonuniform preferences [10]. In addition, Zhao et al. employed a temporal tensor factorization method subsuming these three temporal characteristics together to model check-in activity and achieve good performance [29].

2.3. POI Recommendation Based on Social Influence

Users' behavioral decisions are typically influenced by their friends or users with the same interests, and the social relationship characteristics of "followers" and "companions" are hidden in the check-in data. The social relationship between users intuitively reflects the degree of interaction influence between the users. Therefore, introducing social influence can improve recommendation accuracy. Some studies have integrated the similarity of users' social relationships based on collaborative filtering models [30] or used social relationships as regular terms or weights of the recommendation matrix to perform POI recommendations [11]. In other studies, check-in locations of users with similar characteristics are directly recommended to target users based on social relationships [31]. In addition, some studies have modeled users' social influence based on network linking methods to improve recommendation accuracy [32].

2.4. POI Recommendation Based on Text Context Influence

Text information is derived from users' comments, opinions, and views on check-in locations, which are key factors that influence other users' check-in behavior. Some studies have used LDA (latent Dirichlet allocation) models to predict user preferences and recommend local or global POIs for users [9]. Collaborative filtering techniques combined with topic models have been used by researchers such as Ye et al., who provided semantic annotations of locations with categorical tags [33], and Pennacchiotti et al., who used topic models to study user interests [34]. Although user comments are crucial for improving POI recommendation performance, not every check-in dataset includes the corresponding text information.

2.5. POI Recommendation Based on Multiple-Context Information

Contextual factors have been demonstrated to exert a substantial influence on individuals' preferences [24]. Considering more kinds of contextual information in the recommendation system helps to make more accurate recommendations. Therefore, many POI recommendation methods introduce multiple-context information to improve effectiveness. For example, Cheng et al. proposed a model (MGMPFM) that fuses matrix factorization with geographical and social influence [17]. Li et al. proposed a spatial-temporal probabilistic matrix factorization model (STPMF) based on users' general preferences as well as geographical and temporal information [35]. Some studies introduce more contextual information, including the kernel estimation method based on adaptive bandwidth, which introduces the geographical correlation between users and POIs while incorporating social and categorical correlations (GeoSoCa) [36]; the joint probabilistic generation model based on geographical, textual, social, categorical and popularity information (GTSCP) [9]; and the multigraph fusion model based on POI categories, geographical and social relations (GraphPOI) [37]. Although the abovementioned methods achieve relatively good recommendation performance by exploiting contextual information such as geographical, temporal, and social influence, the rich contextual information also increases the complexity of the recommendation matrix. We aim to solve these problems via similar user clustering, which not only helps to reduce the computational complexity but also enhances the fusion of contextual information and improves the recommendation accuracy.

3. Overview

In the CULT-TF method proposed in this paper, users are first clustered (C) based on their social relationship similarity and check-in behavior similarity to obtain several clusters of similar users; then, a third-order U-L-T tensor model containing user (U), location (L) and time (T) features is constructed for each user cluster, and user activity, POI popularity, and time slot popularity are modeled as the eigenvalues of the U, L, and T dimensions of the tensor matrix. Finally, tensor factorization (TF) based on the U-L-T tensor model is used to obtain the POI recommendation results for the user.

The recommendation framework of the CULT-TF method consists of three main components: similar user clustering, U-L-T tensor modeling, and tensor factorization, as shown in Figure 1.

3.1. Similar User Clustering

"Friends" in social networks, such as communities and interest groups, are users with similar interests. Similar users have more common interest preferences than ordinary users, and the check-in information of similar users has a greater reference value. Therefore, in our proposed method, similar user clustering is first conducted; then, the interest points of similar users are recommended to the target user, which is helpful for improving the accuracy of POI recommendations. In addition, if a large number of users, POIs, and check-in times in the LBSN are directly formed into a recommendation matrix, a large matrix size, sparse matrix data, and high computational complexity could be obtained. Therefore, similar user clustering also helps reduce the number of users in the recommendation matrix.

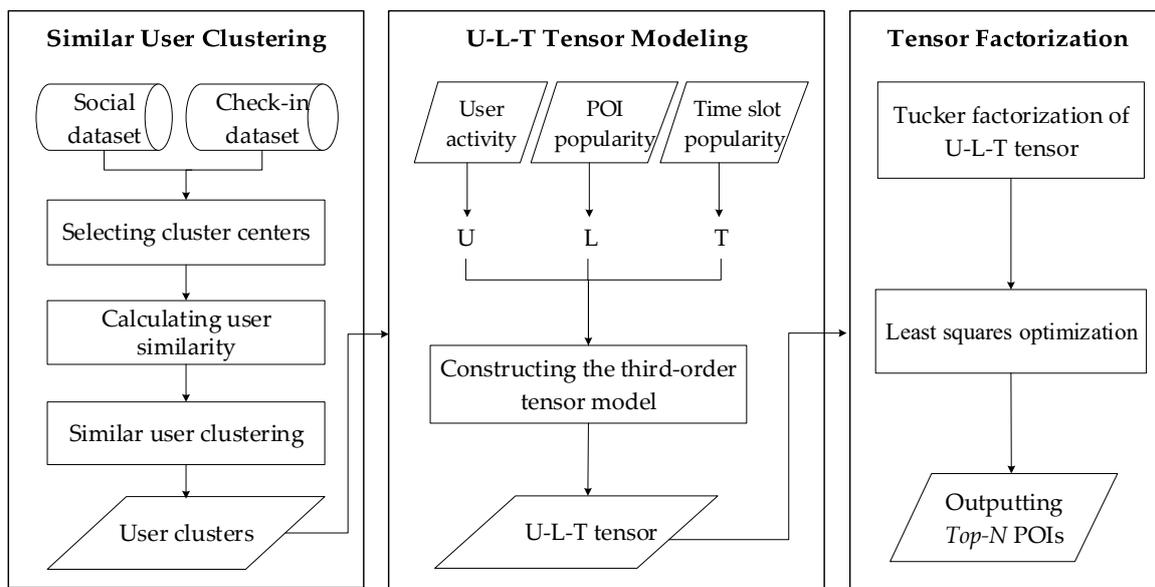


Figure 1. Framework of the CULT-TF method.

The core of similar user clustering consists of two parts: selecting clustering centers and calculating user similarity. In this paper, we select several users with the highest activity as clustering centers based on their social data and check-in data, calculate the user similarity based on their social relationship similarity and check-in behavior similarity, and then cluster the remaining users in the clustering center with the highest user similarity, as detailed in Section 4.1.

3.2. U-L-T Tensor Modeling

Based on the results of similar user clustering, a third-order U-L-T tensor model containing user (U), location (L), and time (T) dimensions is constructed for each user cluster. The advantage of the tensor model is that it can be used to explicitly model the feature information and maintain the features of the dataset in different dimensions. Unlike the POI recommendation method, which directly uses the user, location, and time information of check-in data to construct the tensor model, our proposed method models user activity, POI popularity, and time slot popularity as the eigenvalues of the U, L, and T dimensions. The eigenvalues of each dimension are modeled by integrating the contextual information of user check-in behavior at the user, location, and time levels, enhancing the fusion of contextual information. User activity reflects the activity degree of users' check-in behavior, and the check-in behavior of active users has more referential meaning than that of ordinary users. POI popularity reflects the popularity degree of an interest point, and the higher the POI popularity is, the more users check-in at this POI. Time slot popularity reflects the time distribution pattern of users' check-in behavior. See Section 4.2 for details of U-L-T tensor modeling and eigenvalue calculation.

3.3. Tensor Factorization

The sparsity of check-in data and the ambiguity of contextual information make it difficult to obtain the check-in records of the target user in a given context. Therefore, the constructed U-L-T tensor matrix is sparse, and the tensor values must be complemented to predict the check-in behavior of target users in different contexts. Tensor factorization can be used to present and maintain the structural characteristics of high-dimensional data by mapping the relationships in the original space to a low-dimensional space and extracting the potential relation between different dimensions to calculate the approximate tensor of the missing values and compensate for the matrix sparsity problem caused by missing data. In this paper, we use Tucker factorization [38] to calculate the approximate tensor

of the original tensor and the least squares method to optimize the tensor factorization results. Finally, the approximate tensor values are used to predict the check-in behavior of the target user in given contextual situations and to generate a recommended list of *Top-N* POIs for the target user.

4. The Proposed CULT-TF Method

Table 1 gives the key notation used in this paper.

Table 1. Key notations and descriptions.

Notation	Description
U	user set
L	POI set
T	time slot set
u_i	the user, with user ID $i, u_i \in U$
l_j	the location, with POI ID $j, l_j \in L$
t_k	the time slot, with time slot type ID $k, t_k \in T$
U_{u_i}	the friend set of the user u_i
L_{u_i}	the POI set checked in by the user u_i
T_{u_i}	the time slot set checked in by the user u_i
U_{l_j}	the set of users checking in at a location l_j
U_{t_k}	the set of users checking in at a time slot t_k
f_{u_i}	the number of user u_i check-ins
f_{l_j}	the number of all user check-ins at the location l_j
f_{t_k}	the number of all user check-ins at the time slot t_k
U_{u_v}	the friend set of user u_v
T_{l_j}	the set of time slot types for user check-ins at the location l_j
L_{t_k}	the set of POIs for user check-ins during the time slot t_k

4.1. Similar User Clustering Based on User Similarity

Friends with similar interests play an important role in users' decisions, so similar user clustering helps to improve the accuracy of interest point recommendations. The main process of similar user clustering is shown in Figure 2: ① Selecting user clustering centers: the user clustering centers are selected based on the user activity model described in Section 4.1.1; ② Calculating user similarity: the similarity between users and each cluster center is calculated based on the user similarity model defined in Section 4.1.2; ③ User clustering: each user is clustered to the user clustering center with the highest similarity to the user. The two cores of user clustering are user clustering center selection and user similarity calculation. These two parts are explained in detail below.

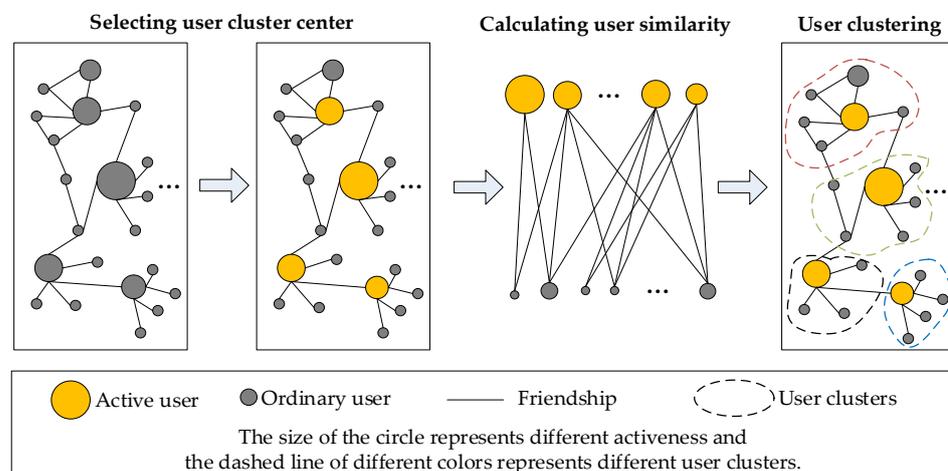


Figure 2. Similar user clustering.

4.1.1. User Activity Model

The user selected as the clustering center should be the most representative user of the cluster. Since active users have a wider social influence, user cluster centers are selected based on user activity in this paper. In LBSNs, user activity is reflected in two main aspects: the user's social activity and the user's check-in activity. The check-in activity of the user is represented by the check-in location activity of the user and the check-in time activity of the user. Thus, the user activity of user u_i is expressed as w_{u_i} , which can be calculated, as shown in Formula (1):

$$w_{u_i} = w_{u_i-u} \times w_{u_i-l} \times w_{u_i-t} \quad (1)$$

where w_{u_i-u} denotes the social activity of user u_i , which reflects the social influence of the user in the user dimension. w_{u_i-u} can be obtained from Formula (2). $|U_{u_i}|$ and $|U|$ denote the number of user u_i 's friends, and the number of users, respectively: the more friends a user has, the higher the social activity and the greater the social influence of the user.

$$w_{u_i-u} = \frac{|U_{u_i}|}{|U|} \quad (2)$$

w_{u_i-l} denotes the check-in location activity of user u_i , which reflects how active the user is in visiting POIs in the location dimension. The greater the number of user check-ins and POI types, the more active the user is in the real world. w_{u_i-l} can be obtained from Formula (3). $\sum_{l_j \in L_{u_i}} |U_{l_j}|$ calculates the number of users who have checked in at locations at which user u_i has checked in. $\sum_{l_j \in L_{u_i}} f_{l_j}$ calculates the number of check-ins of all the users at different locations.

$$w_{u_i-l} = \frac{\sum_{l_j \in L_{u_i}} |U_{l_j}|}{|U|} \times \frac{|L_{u_i}|}{|L|} \times \frac{f_{u_i}}{\sum_{l_j \in L_{u_i}} f_{l_j}} \quad (3)$$

w_{u_i-t} is the check-in time activity of user u_i , which reflects the activity of users' check-in behavior in the time dimension. The greater the number of users' check-ins and the check-in time diversity indicates that users' check-in behavior is more active at different times. w_{u_i-t} can be obtained from Formula (4). $\sum_{t_k \in T_{u_i}} |U_{t_k}|$ calculates the total number of users who have check-in time slots at the same time as user u_i . $\sum_{t_k \in T_{u_i}} f_{t_k}$ calculates the number of check-ins of all users at different time slots. In this paper, to reduce the complexity of the time dimension of the recommendation matrix, we encode a user's check-in timestamp to a particular time slot ID based on the time encoding method [39]. User check-in activities may be concentrated in some time slots or distributed in different time slots, and the time slots reflect the specific temporal preferences indicated by user check-in behavior. A user's check-in time activity is related to the number of user check-in time slots and the number of times the user checks in during each time slot.

$$w_{u_i-t} = \frac{\sum_{t_k \in T_{u_i}} |U_{t_k}|}{|U|} \times \frac{|T_{u_i}|}{|T|} \times \frac{f_{u_i}}{\sum_{t_k \in T_{u_i}} f_{t_k}} \quad (4)$$

To facilitate the analysis of the temporal distribution characteristics of user check-in behavior, the month, day, and period time granularities are used to encode check-in timestamps to 96 types of time slots in this paper. The time slot type IDs range from 0 to 95, and each time slot corresponds to a continuous period. Taking the timestamp of user check-in "9 July 2010 14:33:27", which belongs to time slot ID 62, as an example, Figure 3 illustrates the process of encoding the timestamp of the user check-in to the time slot ID.

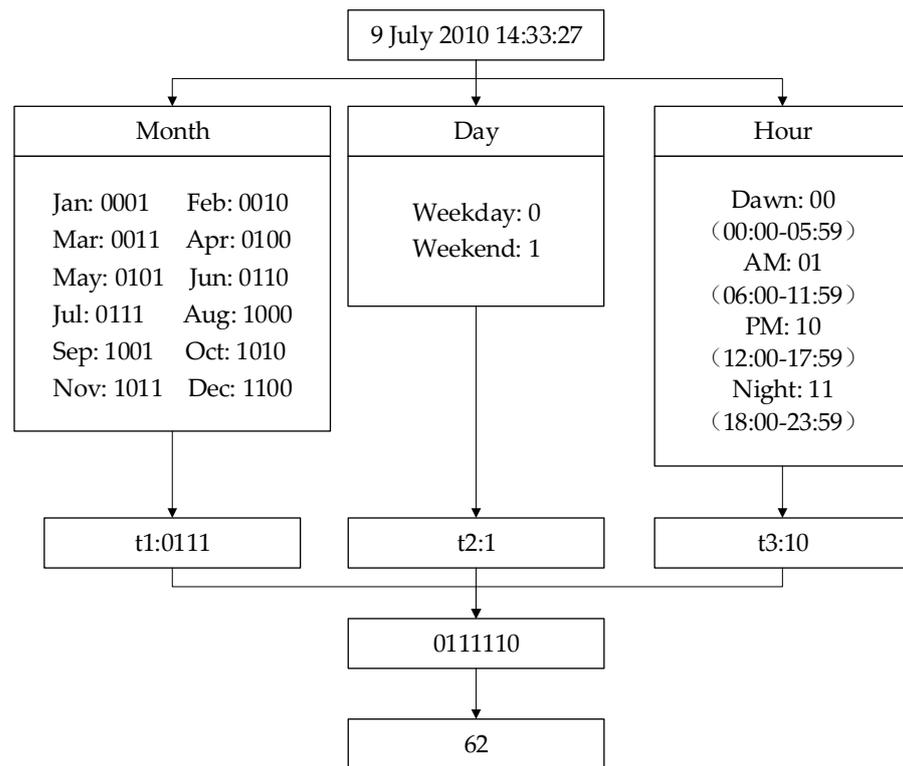


Figure 3. Time slot encoding illustration (96 time slots).

In summary, according to Formula (1), the activity of each user can be calculated; then, the C users with the largest user activity are selected as user clustering centers.

4.1.2. User Similarity Model

User similarity is indicated by the similarity of social relationships and the similarity of check-in behavior between users. The similarity of check-in behavior is represented by the check-in location similarity and the check-in time similarity between users. Therefore, the user similarity between user u_i and user u_v is expressed as $sim_{u_i-u_v}$, and its calculation is given in Formula (5).

$$sim_{u_i-u_v} = sim_{u_i-u_v} + siml_{u_i-u_v} + simt_{u_i-u_v} \quad (5)$$

$sim_{u_i-u_v}$ denotes the similarity of social relationships between user u_i and user u_v . The greater the number of common friends between the users, the higher the similarity of their social relationships. In this paper, the Jaccard coefficient [40] is used to define the social similarity of users, as shown in Formula (6) and $sim_{u_i-u_v} \in [0, 1]$.

$$sim_{u_i-u_v} = \frac{|U_{u_i} \cap U_{u_v}|}{|U_{u_i} \cup U_{u_v}|} = \frac{|U_{u_i} \cap U_{u_v}|}{|U_{u_i}| + |U_{u_v}| - |U_{u_i} \cap U_{u_v}|} \quad (6)$$

$siml_{u_i-u_v}$ is the similarity of check-in locations between user u_i and user u_v , which reflects the similarity of users' check-in behavior at the location level. The more often two users check in at the same location and the more similar the number of check-ins at the same location, the higher the similarity of their check-in behavior. In real life, a geographic location may cover many POIs, and the POIs that users check in at when they visit the same geographic location may not be the same. The similarity of check-in locations obtained only by judging whether users visit the same POIs may be close to zero. Considering the sparsity of check-in data and the spatial aggregation of users' check-in behavior, we cluster all POIs where users check in as ROIs based on K-means [41] clustering and then measure the similarity of check-in locations between users based on their check-in ROIs.

Clustering a large number of user check-in POIs into a limited number of ROIs can reduce the complexity of the recommendation matrix location dimension. The check-in location matrix of all users visiting ROIs is defined as $\mathbf{UL}_{M \times K}$, as shown in (7).

$$\mathbf{UL}_{M \times K} = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,j} & \cdots & r_{1,K} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,j} & \cdots & r_{2,K} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{i,1} & r_{i,2} & \cdots & r_{i,j} & \cdots & r_{i,K} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{M,1} & r_{M,2} & \cdots & r_{M,j} & \cdots & r_{M,K} \end{bmatrix} \quad (7)$$

where M is the number of users, K is the number of ROIs, and $r_{i,j}$ denotes the number of normalized check-ins of user u_i at the j th ROI. Representing the normalized check-in number of users at K ROIs as a vector, the check-in location vectors of user u_i and user u_v are $\mathbf{UL}_i = (r_{i,1}, r_{i,2}, \dots, r_{i,j}, \dots, r_{i,K})$ and $\mathbf{UL}_v = (r_{v,1}, r_{v,2}, \dots, r_{v,j}, \dots, r_{v,K})$, respectively. We use the cosine angle between vectors to measure the similarity of check-in locations between users. The similarity of check-in locations between user u_i and user u_v can be obtained from Formula (8).

$$\text{siml}_{u_i-u_v} = \cos(\mathbf{UL}_i, \mathbf{UL}_v) = \frac{\mathbf{UL}_i \cdot \mathbf{UL}_v}{\|\mathbf{UL}_i\| \|\mathbf{UL}_v\|} \quad (8)$$

$\text{simt}_{u_i-u_v}$ denotes the check-in time similarity between users u_i and u_v , which reflects the similarity of the users' check-in behavior at the time level. The more similar two users' check-in time slots are, the more similar the distribution of check-in time slots, indicating a higher similarity of their check-in behavior at the temporal level. As described in Section 4.1.1, we encode users' check-in timestamps to 96 time slots, and the check-in time matrix of all users based on these 96 time slots is defined as $\mathbf{UT}_{M \times 96}$, as shown in (9).

$$\mathbf{UT}_{M \times 96} = \begin{bmatrix} p_{1,0} & p_{1,1} & \cdots & p_{1,k} & \cdots & p_{1,95} \\ p_{2,0} & p_{2,1} & \cdots & p_{2,k} & \cdots & p_{2,95} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{i,0} & p_{i,1} & \cdots & p_{i,k} & \cdots & p_{i,95} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ p_{M,0} & p_{M,1} & \cdots & p_{M,k} & \cdots & p_{M,95} \end{bmatrix} \quad (9)$$

where M is the number of users and $p_{i,k}$ denotes the normalized check-in number of user u_i at the k th time slot. Representing the normalized check-in number of the user in 96 time slots as a vector, the check-in time vectors of user u_i and user u_v are $\mathbf{UT}_i = (p_{i,0}, p_{i,1}, \dots, p_{i,k}, \dots, p_{i,95})$ and $\mathbf{UT}_v = (p_{v,0}, p_{v,1}, \dots, p_{v,k}, \dots, p_{v,95})$, respectively. The check-in time similarity between user u_i and user u_v can be obtained from Formula (10).

$$\text{simt}_{u_i-u_v} = \cos(\mathbf{UT}_i, \mathbf{UT}_v) = \frac{\mathbf{UT}_i \cdot \mathbf{UT}_v}{\|\mathbf{UT}_i\| \|\mathbf{UT}_v\|} \quad (10)$$

In summary, user clustering is achieved by calculating the user similarity between each user and C active users separately according to Formula (5) and clustering each user into the cluster of the active user with the highest similarity to the user.

4.2. U-L-T Tensor Modeling with the Integration of Contextual Information

Based on the clustering results of similar users, a third-order tensor \mathbf{Z} is constructed for each user cluster, as shown in Figure 4. U, L, and T represent the user (U), location (L), and time (T) dimensions, respectively, and the eigenvalues of each dimension reflect the check-in characteristics of that dimension. In this paper, user activity, POI popularity,

and time slot popularity are modeled as the eigenvalues of the U, L, and T dimensions in tensor \mathbf{Z} . The U-dimension eigenvalues (user activity) are described in Section 4.1, and the L-dimension eigenvalues and T-dimension eigenvalues are introduced below.

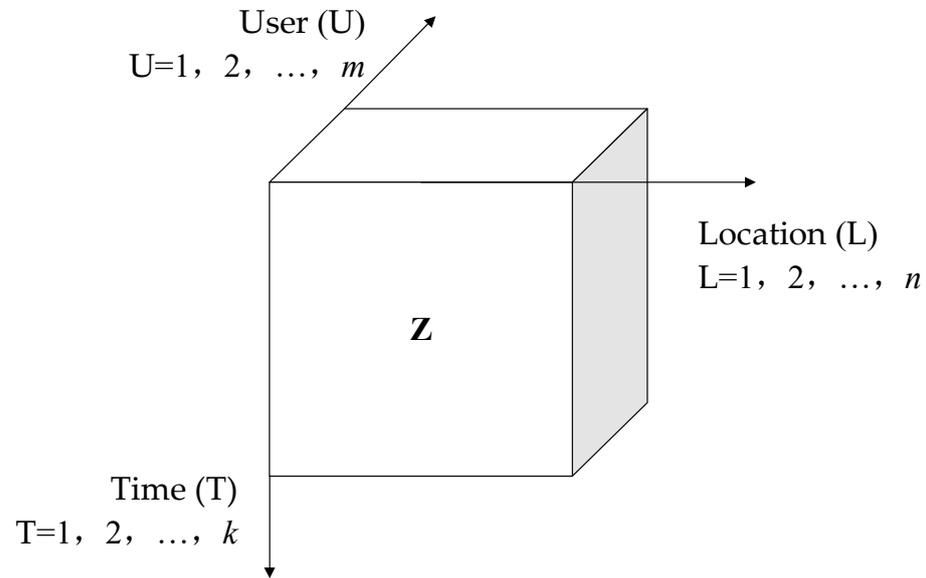


Figure 4. U-L-T third-order tensor.

The eigenvalues of the L dimension are the POI popularity values. At the user level, POI popularity is expressed as the ratio of the number of users who checked in at this POI to the total number of users; at the location level, POI popularity is represented by the ratio of the number of user check-ins at this POI to the total number of user check-ins at all POIs; at the time level, POI popularity is the ratio of the number of time slot types at which users check in at this POI to the total number of time slot types. By integrating the influences of user, location, and time contextual information on the check-ins of interest points, we express the POI popularity at location l_j as p_{l_j} , and its calculation Formula is (11).

$$p_{l_j} = \frac{|U_{l_j}|}{|U|} \times \frac{f_{l_j}}{\sum_{l_j \in L} f_{l_j}} \times \frac{|T_{l_j}|}{|T|} \quad (11)$$

The eigenvalues of the T dimension are the time slot popularity values. At the user level, the time slot popularity is expressed as the ratio of the number of users checking in at this time slot to the total number of users; at the location level, the time slot popularity is the ratio of the number of POIs at which users check in during this time slot to the total number of POIs; at the time level, the time slot popularity is represented by the ratio of the number of users' check-ins during this time slot to the total number of users' check-ins at all time slots. By integrating the influences of user, location, and time contextual information on the check-in time slots of users, we express the popularity of time slot t_k as v_{t_k} , and its calculation formula is (12).

$$v_{t_k} = \frac{|U_{t_k}|}{|U|} \times \frac{|L_{t_k}|}{|L|} \times \frac{f_{t_k}}{\sum_{t_k \in T} f_{t_k}} \quad (12)$$

To unify the magnitudes of the eigenvalues of each dimension, maximum-minimum normalization, shown in Formula (13), is used to normalize the eigenvalues. z^* denotes the

normalized eigenvalue, where $z^* \in [0, 1]$; z is the original value of the eigenvalue, z_{\min} is the minimum value of the eigenvalue, and z_{\max} is the maximum value of the eigenvalue.

$$z^* = \frac{z - z_{\min}}{z_{\max} - z_{\min}} \quad (13)$$

When the given context is user u_i , location l_j , and time slot t_k , the corresponding tensor element is expressed as z_{u_i, l_j, t_k} , and its calculation formula is (14).

$$z_{u_i, l_j, t_k} = w_{u_i}^* + p_{l_j}^* + v_{t_k}^* \quad (14)$$

where $w_{u_i}^*$ denotes the normalized user activity of user u_i ; $p_{l_j}^*$ is the normalized POI popularity of location l_j ; and $v_{t_k}^*$ represents the normalized time slot popularity of time slot t_k . The tensor element z_{u_i, l_j, t_k} calculated from the normalized eigenvalues is in the range $[0, 3]$. Ultimately, the eigenvalues of each dimension integrate the contextual information of user check-in behavior at the user, location, and time levels, enhancing the integration of contextual information.

4.3. TOP-N POIs Based on Tensor Factorization

Due to the sparsity of check-in data, tensor elements may be missing. The missing tensor element values can be filled in by tensor factorization. The approximate tensor value obtained by tensor factorization, which can be used as the predicted value of user check-in behavior, contains the check-in behavior features of the original tensor in the U, L, and T dimensions. This paper uses the Tucker method [38] for tensor factorization, and the least squares method is used to optimize the results.

In the Tucker method, the tensor is split into the product of a kernel tensor and each dimensional matrix, expressed as $\mathbf{Z} \approx \mathbf{G} \times_1 \mathbf{U} \times_2 \mathbf{L} \times_3 \mathbf{T}$, as shown in Figure 5. $\mathbf{U} \in \mathbb{R}^{m \times r_1}$, $\mathbf{L} \in \mathbb{R}^{n \times r_2}$, and $\mathbf{T} \in \mathbb{R}^{k \times r_3}$ are the low-rank eigenmatrices in the U, L, and T dimensions, respectively. \mathbb{R} refers to the low-dimensional feature space formed by the factorization of each dimensional feature of the original tensor. The core tensor $\mathbf{G} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$ represents the interaction between different eigenmatrices, retains the main information of the original tensor, and is stable. The core tensor \mathbf{G} has dimensions $r_1 \times r_2 \times r_3$ and is much smaller than the original vector. After the least-squares iterative solutions of the eigenmatrix and core tensor are found, the approximate tensor element values are calculated according to Formula (15).

$$\hat{z}_{u, l, t} = \sum_{\tilde{u}} \sum_{\tilde{l}} \sum_{\tilde{t}} \hat{g}_{\tilde{u}, \tilde{l}, \tilde{t}} \cdot \hat{u}_{u, \tilde{u}} \cdot \hat{l}_{l, \tilde{l}} \cdot \hat{t}_{t, \tilde{t}} \quad (15)$$

where $\hat{z}_{u, l, t}$ is the approximate tensor element value, $\hat{g}_{\tilde{u}, \tilde{l}, \tilde{t}}$ is the core tensor, $\hat{u}_{u, \tilde{u}}$ is the element value of the low-rank eigenmatrix in the U dimension, $\hat{l}_{l, \tilde{l}}$ is the element value of the low-rank eigenmatrix in the L dimension, $\hat{t}_{t, \tilde{t}}$ is the element value of the low-rank eigenmatrix in the T dimension, “ \sim ” is the label for an index of the feature dimension, and “ $\hat{\cdot}$ ” is the label for the elements in the eigenmatrix.

After obtaining the approximate tensor element values of the tensor model \mathbf{Z} by the tensor factorization method, a recommended list of Top-N POIs is generated according to Formula (16). It denotes a list of locations l , which are the N POIs at which user u is most likely to check in when the time is t .

$$Top - N = \underset{t \in T}{\operatorname{argmax}} \hat{z}_{u, l, t} \quad (16)$$

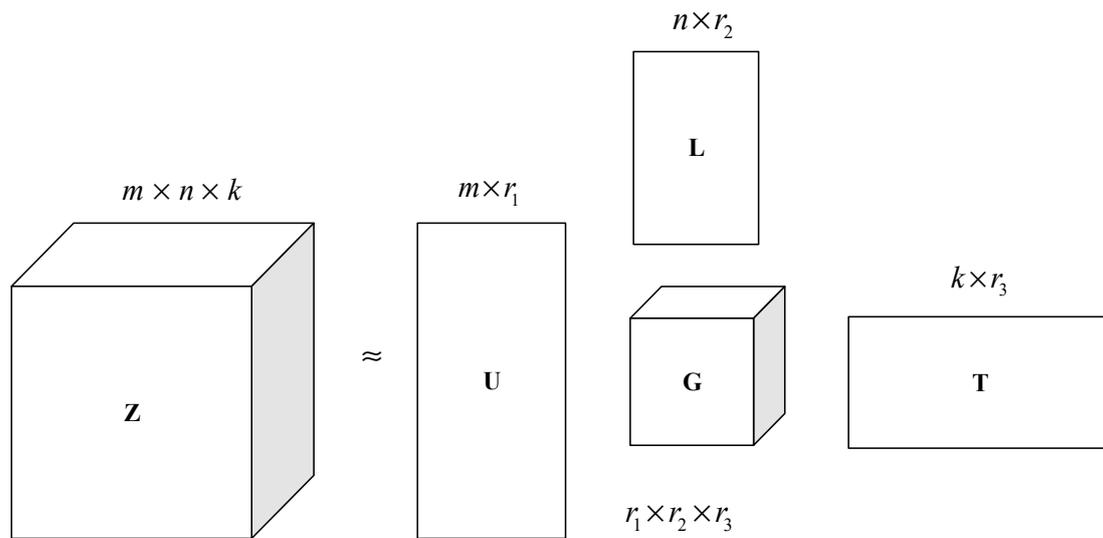


Figure 5. Tucker factorization form of a third-order tensor.

5. Experiments

5.1. Datasets

The experiments use publicly available datasets from Brightkite, a location-based social networking service provider. The Brightkite datasets include a social dataset and check-in dataset in the US region (<http://snap.stanford.edu/data/loc-brightkite.html>, accessed on 1 March 2023). The experimental data comprise the data from the Brightkite datasets from 1 October 2009 to 30 September 2010, which are processed before the experiment. For example, false information is deleted. The experimental dataset profile is shown in Table 2. The social dataset contains 4844 users and 186,071 relations. The check-in dataset contains 388,148 check-in records, and each record consists of a check-in time, a longitude and a latitude coordinate, and a location id of POI corresponding to the check-in coordinates. The check-in dataset contains 7685 POIs in total. To clearly show the geographic distribution of user check-ins, we use the map to display a large number of user check-ins based on POI information. Each red dot indicates a POI that users have checked in, as shown in Figure 6.

Table 2. Brightkite datasets.

Description	Statistics
Time Range	1 October 2009–30 September 2010
Geographic Range	19.27° N–71.29° N 67.84° W–159.67° W
Number of Users	4844
Number of POIs	7685
Number of Social Relations	186,071
Number of Check-ins	388,148

5.2. Evaluation Metrics

The two most common metrics used to evaluate POI recommendation methods are *Precision@N* and *Recall@N* [42]. *Precision@N*, as defined in Formula (17), is the ratio of correctly predicted POIs to the total number of recommended POIs. *Recall@N*, as defined in Formula (18), is the ratio of correctly predicted POIs to the total number of POIs where a check-in occurred.

$$Precision@N = \frac{1}{|U|} \sum_{u \in U} \frac{|R(u) \cap T(u)|}{N} \quad (17)$$

$$Recall@N = \frac{1}{|U|} \sum_{u \in U} \frac{|R(u) \cap T(u)|}{|T(u)|} \quad (18)$$

where N is the number of POIs recommended to the user, $R(u)$ denotes the recommended list of *Top-N* POIs that user u would like to visit, and $T(u)$ is the list of POIs that user u has visited.

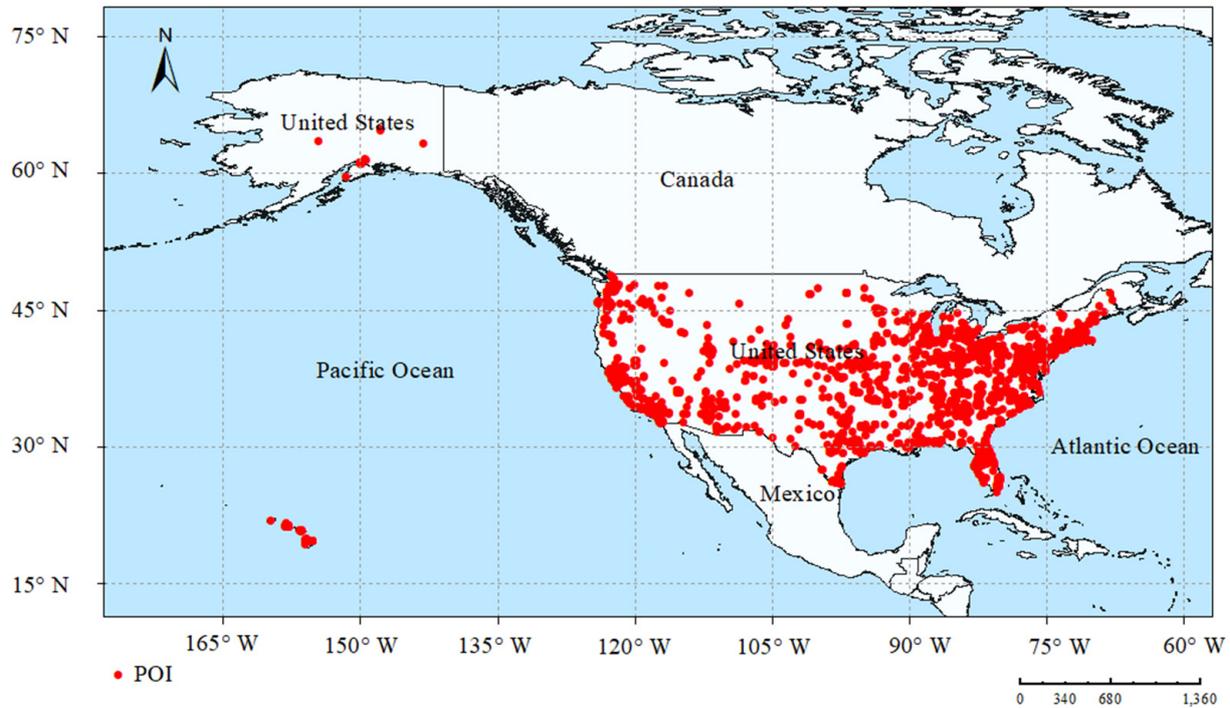


Figure 6. Distribution of check-in data.

5.3. Clustering Parameter Analysis

As described in Section 4.1, the proposed method requires setting the user clustering center parameter C and the ROI clustering parameter K . The activity of each user can be calculated based on the Brightkite dataset. The normalized user activities of the ten most active users are shown in Table 3. Theoretically, the larger the value of C is, the larger the number of user clusters, the higher the similarity between users within a cluster, and accordingly, the higher the accuracy of POI recommendation based on these similar users. However, too large a value of C can cause excessive partitioning of user clusters, resulting in more sparse check-in data of user clusters; too small a value of C may decrease the similarity between users within a cluster, which makes it difficult to ensure the accuracy of POI recommendation. Therefore, this section analyzes the effect of parameter C on the standard deviation of users and average user similarity to set a suitable value of parameter C .

Table 3. Top 10 user activities.

Ranking	User ID	User Activity	Ranking	User ID	User Activity
1	1863	1.000000	6	620	0.608652
2	1864	0.822594	7	0	0.483306
3	143	0.808769	8	1302	0.468147
4	2149	0.732460	9	212	0.414451
5	35	0.623615	10	208	0.385934

The standard deviation of users reflects the difference between the number of users in each cluster and the average number of users; a smaller value indicates that the number of users in each cluster is closer to the average number of users. The average number of users is the ratio of the total number of users to the number of clusters C . Figure 7 shows the average number of users and the standard deviation of users for different values of parameter C . When C is 3, the standard deviation of users decreases substantially, and the average number of users is close to 1615. However, a large number of users will affect the performance of tensor modeling and tensor calculation. We set parameter C to 7 in this paper because the standard deviation of users at this time is close to that when C is equal to 3, and the average number of users is only 692.

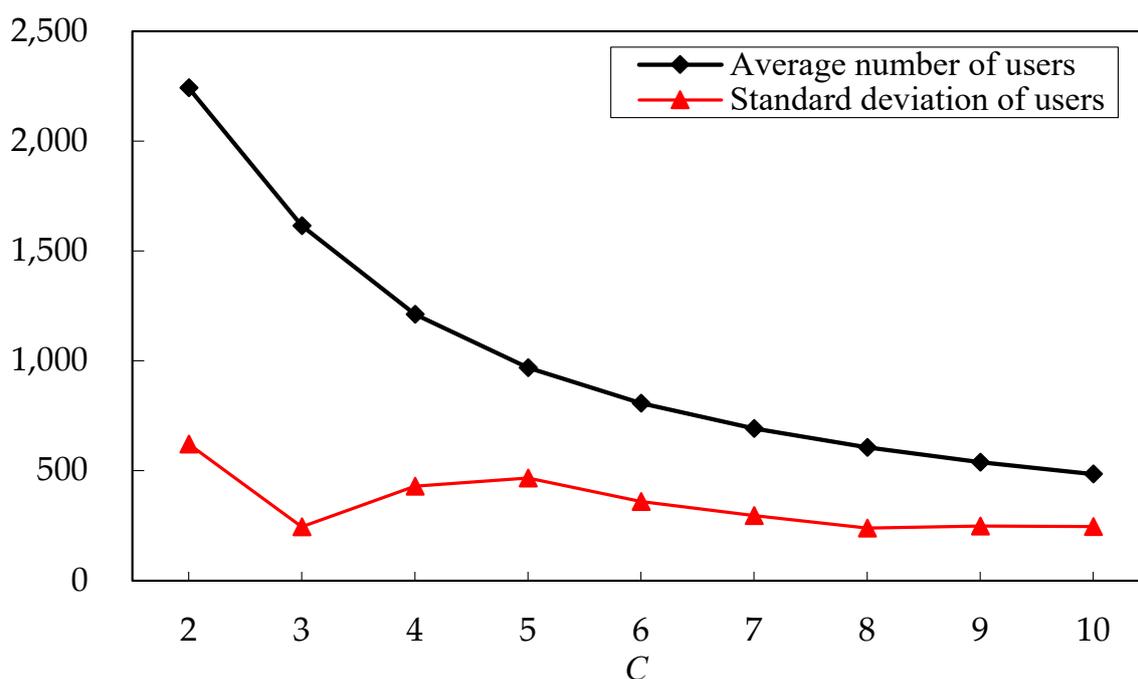


Figure 7. Average number of users and standard deviation of users.

Figure 8 shows the maximum proportion of users and the average user similarity with different values of parameter C . The maximum proportion of users is the ratio of the number of users in the largest user cluster to the total number of all users, and the larger the value is, the more users are concentrated in the same cluster. The user similarity of the cluster is defined as the average user similarity between each user and the central user in the same cluster, and a higher value indicates that the check-in behavior of the users in the cluster is more similar. The average user similarity is the average of the user similarity of C clusters, and a higher value indicates that each cluster has a higher user similarity. As shown in Figure 8, as parameter C increases, the maximum proportion of users tends to decrease, and the average user similarity tends to increase; the curve of the maximum proportion of users starts to level off when C is 7, and the curve of the average user similarity is also relatively high at this time. When C is 7, Table 4 gives the user ID of the cluster center and the user similarity of the cluster for each user cluster. In Table 4, the cluster's minimum value of user similarity reaches 0.72074 among the seven user clusters, indicating that all user clusters have high user similarity. All the above experiments show that it is reasonable to set parameter C to 7.

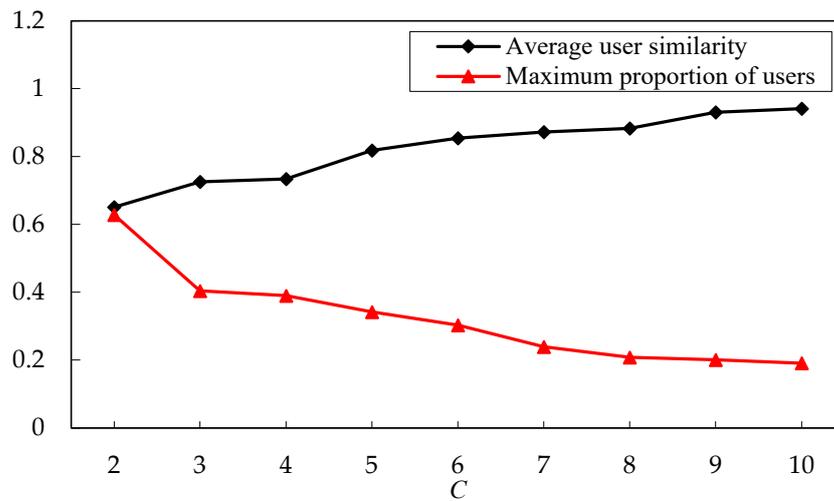


Figure 8. Maximum proportion of users and average user similarity.

Table 4. User cluster information (C = 7).

Cluster No.	User ID of Cluster Center	User Similarity of Cluster
1	1863	1.00592
2	1864	0.72074
3	143	0.90315
4	2149	0.86348
5	35	0.81491
6	620	0.90564
7	0	0.89310

To calculate the similarity of check-in locations between users, we cluster POIs into K ROI clusters, as described in Section 4.1.2. The results of ROI clustering directly affect the similarity of check-in locations. Therefore, to set a reasonable ROI clustering parameter K , the silhouette coefficient is used to evaluate the results of ROI clustering in this experiment. The range of the silhouette coefficient is $[-1, 1]$, and the closer to 1 the value is, the better the clustering result. The silhouette coefficients of ROI clustering with different values of K are shown in Figure 9. Figure 9 shows that the silhouette coefficient is the highest when parameter K is 12 and the number of ROI clusters is moderate; therefore, the ROI clustering parameter K is set to 12 in this paper.

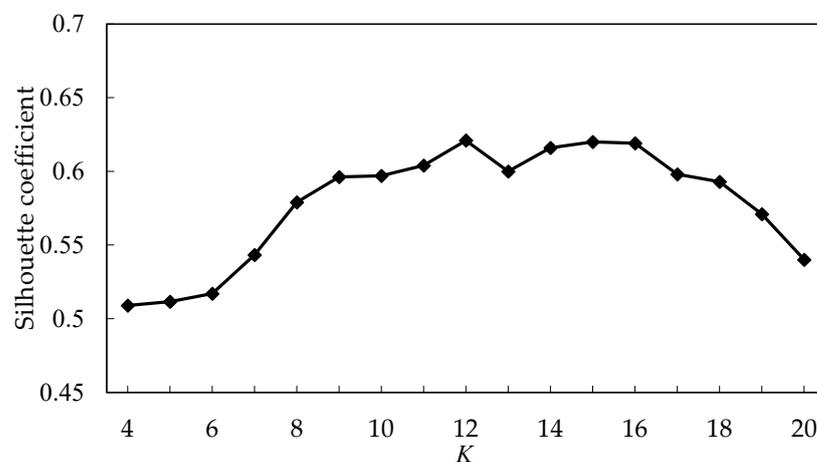


Figure 9. Silhouette coefficient of ROI clustering.

5.4. Time Slot Analysis

In this paper, we encode a large number of users' check-in timestamps to a limited number of time slots. Theoretically, if the number of time slots is too small, then each time slot will have a large number of user check-ins, making the time periodicity of user check-in behavior more difficult to see. If the number of time slots is too large, the time periodicity of user check-in behavior can be found more accurately, but the more time slots there are, the greater the cost of tensor modeling. To set the number of time slots reasonably, we consider low (32 time slots), medium (96 time slots), and high (192 time slots) values and analyze the number of user check-ins for different time slots to find the temporal distribution characteristic of user check-in behavior. Compared with the 96 time slots described in Section 4.1.1, the 32 time slots use "Quarter" instead of "Month," where one quarter consists of three months (i.e., there are four quarters in one year), and the "Day" and "Hour" time slots remain unchanged; for 192 time slots "Hour" is partitioned into eight time slots, where every three hours is a time slot, and the "Month" and "Day" time slots remain unchanged.

The quantity distributions of user check-ins when the number of time slots is set to 32, 96, and 192 are shown in Figures 10, 11, and 12, respectively. When the number of time slots is set to 32, the check-in distribution regularity is not obvious; when the number of time slots is set to 96 or 192, the distribution of users' check-ins shows obvious periodic characteristics.

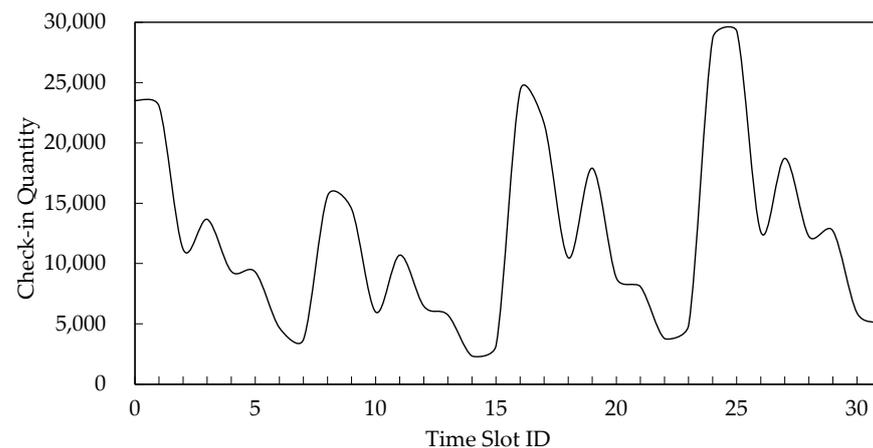


Figure 10. The distribution of check-ins for 32 time slots.

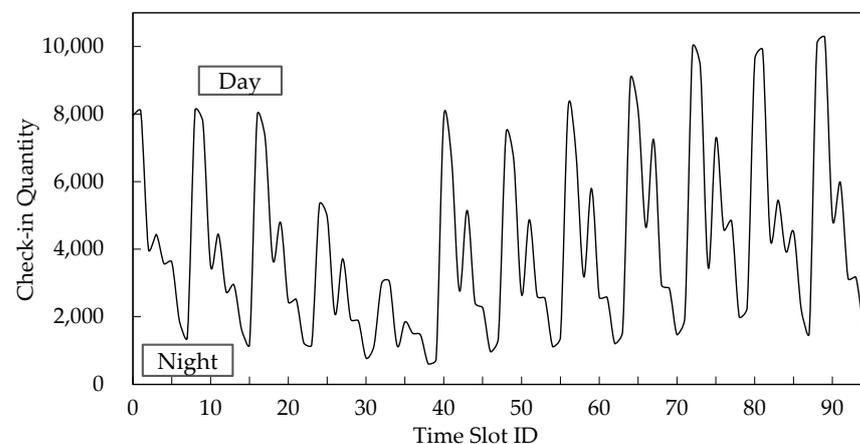


Figure 11. The distribution of check-ins for 96 time slots.

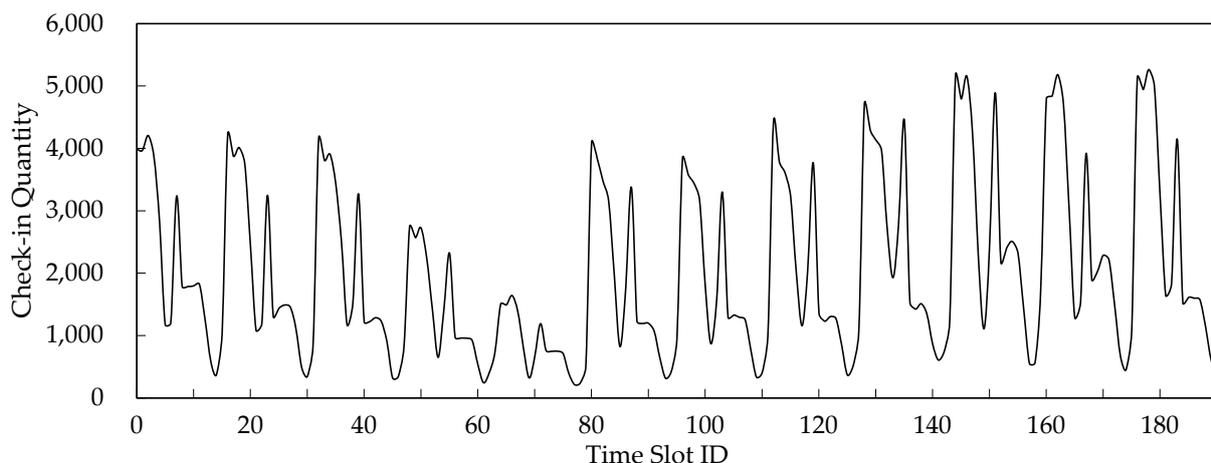


Figure 12. The distribution of check-ins for 192 time slots.

As shown in Figure 11, between time slots 30 and 40 (March to May), the number of user check-ins is at a low level, which indicates that user willingness to travel is lower from March to May than in other months. The number of users' check-ins gradually decreases from day to night during a day, and the distribution of check-ins in different months also varies. These results indicate that the temporal distribution characteristics of user check-in behavior can be more clearly shown through time slots.

As shown in Figure 12, when the number of time slots is 192, although more time slots can show more detailed check-in distribution regularity, excessive time slots in a day do not bring significant benefits; for example, most users' check-in behavior and intention are very low during the late night to early morning hours (00:00–05:59), and dividing this time slot into two time slots (00:00–02:59) and (03:00–05:59) has no significant impact on mining users' check-in behavior regularity. In addition, too many time slots will significantly increase the cost of tensor modeling; therefore, the number of time slots is set to 96 in our experiments, which can reflect the temporal periodic characteristic of users' check-in behavior while controlling the cost of tensor modeling, and the number of time slots is moderate.

5.5. Experimental Results and Analysis

To analyze the effectiveness of the CULT-TF recommendation method, four POI recommendation methods integrating different types of contextual information are chosen for comparison as follows:

- USG [15]: USG is a collaborative recommendation method based on the geographical influence that models the geographical clustering phenomenon by means of a naive Bayesian approach and integrates social influence;
- MGMPFM [17]: MGMPFM is a POI recommendation method based on matrix factorization that models the geographical influence of users' check-in behavior based on a multicenter Gaussian model (MGM) and fuses social and geographical influence into a matrix factorization framework;
- GeoSoCa [36]: GeoSoCa is a POI recommendation method that exploits geographical correlations, social correlations, and category correlations among users and POIs;
- LORE [28]: LORE is a location recommendation method with sequential influence based on an additive Markov chain (AMC), which integrates sequential influence with geographical influence and social influence into a unified location recommendation framework.

In these experiments, the Brightkite datasets are used; 20% of the data are randomly selected as the test set, and 80% of the data are used as the training set. The parameter N of $Top-N$ is set to 5, 10, 15, and 20 to compare the precision and recall of the CULT-TF method with those of the above four methods, and the experimental results are shown in Figure 13. The trends in the precision and recall of all evaluated methods are intuitive. All five recommendation methods show a trend of decreasing precision and increasing recall with increasing N , which means that the more POIs are recommended to users, the more likely users are to find POIs that they are willing to visit; however, some of the recommended POIs are less likely to be visited by users.

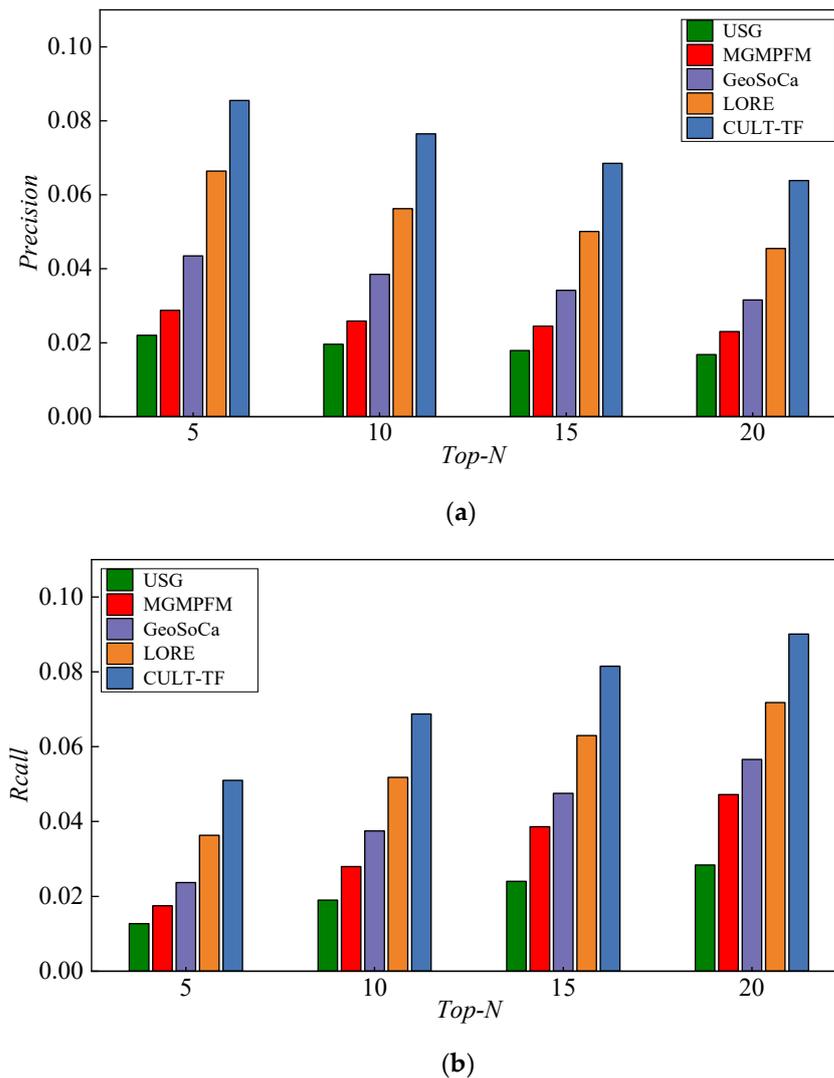


Figure 13. Recommendation accuracy of five recommendation methods. (a) $Top-N$ Precision; (b) $Top-N$ Recall.

The absolute accuracy of POI recommendation methods is usually not high due to the sparsity of check-in data. However, POI recommendation will perform better as more check-in data are collected. This phenomenon has been observed repeatedly in previous works [28,36]. So, we focus on contrasting the relative accuracy of all evaluated methods.

USG linearly integrates social influence and geographical influence, and it considers only the impact of distance on user check-in behavior in terms of geographical influence. As shown in Figure 13, USG gives the worst recommendation result with respect to the $Top-N$ values. MGMPFM adopts MGM to model the geographical influence and integrates social and geographical influence into matrix factorization. The geographical influence can

be captured more accurately by MGMPFM than USG. Thus, the recommendation accuracy of MGMPFM is a little better than that of USG. However, it does not consider the popularity of POI categories. As a result, it reports the second-lowest recommendation accuracy.

From Figure 13, GeoSoCa achieves better performance than MGMPFM and USG. According to Table 5, for example, when the value of *Top-N* is 10, the precision increases from 0.0259 for MGMPFM to 0.0385 for GeoSoCa, and the recall increases from 0.0280 for MGMPFM to 0.0375 for GeoSoCa. This improvement in performance is mainly due to GeoSoCa not only considering the geographical influence and social influence but also taking the popularity of POI categories into account. Nonetheless, its improvement is limited in comparison to MGMPFM and USG because GeoSoCa ignores the temporal influence of POI recommendations.

Table 5. Precision and recall of different recommendation methods.

Metrics	USG	MGMPFM	GeoSoCa	LORE	CULT-TF
<i>Precision@5</i>	0.0220	0.0288	0.0435	0.0664	0.0855
<i>Recall@5</i>	0.0127	0.0175	0.0237	0.0363	0.0510
<i>Precision@10</i>	0.0196	0.0259	0.0385	0.0563	0.0765
<i>Recall@10</i>	0.0190	0.0280	0.0375	0.0518	0.0687
<i>Precision@15</i>	0.0179	0.0245	0.0342	0.0501	0.0685
<i>Recall@15</i>	0.0240	0.0386	0.0475	0.0630	0.0815
<i>Precision@20</i>	0.0168	0.0230	0.0316	0.0455	0.0639
<i>Recall@20</i>	0.0284	0.0472	0.0566	0.0718	0.0901

As shown in Table 5, CULT-TF and LORE significantly outperform GeoSoCa, MGMPFM, and USG in all metrics. For example, when *Precision@15*, the precision of CULT-TF attains 0.0685 and LORE attains 0.0501, while the precision of GeoSoCa, MGMPFM, and USG achieves 0.0342, 0.0245, and 0.0179, respectively. When *Recall@15*, the recall of CULT-TF attains 0.0815 and LORE attains 0.0630, while the recall of GeoSoCa, MGMPFM, and USG achieves 0.0475, 0.0386, and 0.0240, respectively. This implies that temporal influence plays a significant role in POI recommendation. By integrating temporal influence, we can provide a much better performance of POI recommendations.

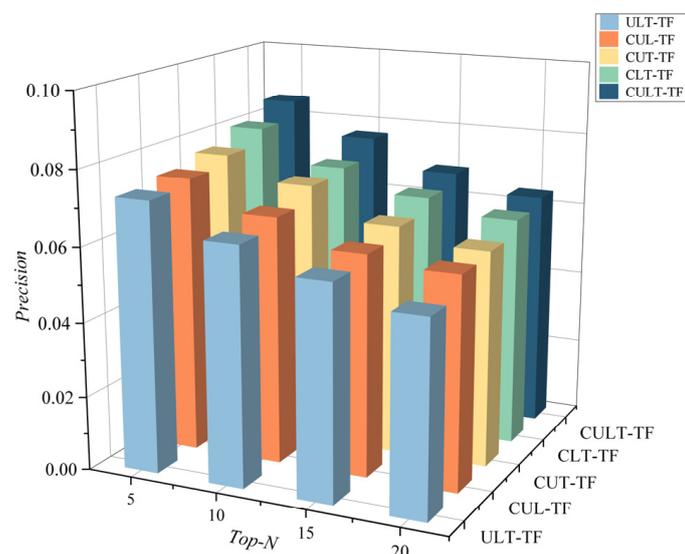
CULT-TF always exhibits the best recommendation performance in terms of precision and recall. In particular, it achieves a significant improvement compared to the second-best recommendation method LORE, mainly because it takes good advantage of contextual information from similar users and incorporates user activity, POI popularity, and time slot popularity into the recommendation matrix, achieving a significant improvement compared to other recommendation methods.

To analyze the effects of similar user clustering, user activity, POI popularity, and time slot popularity on POI recommendation, the following four baseline methods are designed for comparison with our proposed method.

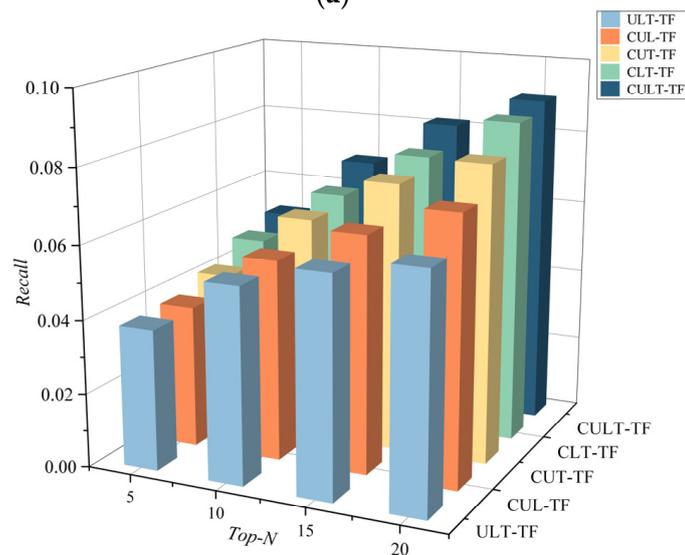
- ULT-TF: This is a recommended method based on CULT-TF that does not conduct similar user clustering, i.e., it does not consider the influence of similar users;
- CLT-TF: This is a simplified version of CULT-TF in terms of the U dimension, which does not introduce the U dimensional eigenvalue “user activity”;
- CUT-TF: This is a simplified version of CULT-TF in terms of the L dimension, which does not introduce the L dimensional eigenvalue “POI popularity”;
- CUL-TF: CUL-TF is a simplified version of CULT-TF in terms of the T dimension, which does not introduce the T dimensional eigenvalue “time slot popularity.”

The experimental results of CULT-TF and the four baseline methods are shown in Figure 14 and Table 6, which indicates that the precision and recall of CULT-TF are always superior to those of the four baseline methods. As shown, ULT-TF exhibits the lowest recommendation precision and recall. That suggests that similar user clustering has the most important impact on recommendation accuracy. The recommendation performance of CUL-TF is slightly better than that of ULT-TF; in fact, they are close. As shown in Table 6,

for example, the precision and recall of ULT-TF are 0.0726 and 0.0382 when the value of *Top-N* is 5, and those of CUL-TF are 0.742 and 0.387, respectively. CUL-TF gives the second worst recommendation result. The result implies that temporal influence plays the second most important role in improving recommendation quality. CUT-TF is better than CUL-TF. This means that the impact of time slot popularity on the decrease in recommendation accuracy is greater than that of POI popularity. From Table 6, CLT-TF outperforms CUT-TF, CUL-TF, and ULT-TF; that is, according to the impacts of the four baseline methods on the recommendation results, the importance of the four factors is ranked as follows: similar user clustering > time slot popularity > POI popularity > user activity. As shown in Figure 14 and Table 6, CULT-TF always gives the best performance of POI recommendation, and the results suggest that a single factor alone cannot accurately reflect user preferences for POIs. CULT-TF shows the strength of combining all four factors of similar user clustering, user activity, POI popularity, and time slot popularity.



(a)



(b)

Figure 14. Recommendation accuracy of CULT-TF and four baseline methods. (a) *Top-N* Precision; (b) *Top-N* Recall.

Table 6. Precision and recall of different baseline methods.

Metrics	ULT-TF	CUL-TF	CUT-TF	CLT-TF	CULT-TF
<i>Precision@5</i>	0.0726	0.0742	0.0768	0.0807	0.0855
<i>Recall@5</i>	0.0382	0.0387	0.0424	0.0478	0.0510
<i>Precision@10</i>	0.0641	0.0665	0.0706	0.0717	0.0765
<i>Recall@10</i>	0.0530	0.0546	0.0609	0.0634	0.0687
<i>Precision@15</i>	0.0576	0.0597	0.0622	0.0656	0.0685
<i>Recall@15</i>	0.0593	0.0641	0.0730	0.0763	0.0815
<i>Precision@20</i>	0.0522	0.0573	0.0584	0.0620	0.0639
<i>Recall@20</i>	0.0634	0.0723	0.0804	0.0875	0.0901

6. Discussion

CULT-TF always exhibits the best recommendation quality in terms of both precision and recall. It achieves a significant improvement compared to USG, MGMPFM, GeoSoCa, and LORE. This is because USG and MGMPFM consider only geographical and social influence, and GeoSoCa further integrates POI popularity, but all these three methods ignore temporal influence for POI recommendation. LORE fuses sequential influence with geographical and social influence into a recommendation framework, and LORE implicitly expresses temporal influence via the sequential influence on users' check-in behavior. CULT-TF is distinct from the abovementioned methods. (1) CULT-TF exploits contextual information from similar users and integrates geographical, temporal, and social influence into a unified recommendation framework. (2) CULT-TF explicitly models temporal influence as time slot popularity. Social and geographical influence are modeled as user activity and POI popularity, respectively. (3) CULT-TF constructs a U-L-T tensor matrix based on a similar user cluster that can capture user preferences more accurately by integrating user activity, POI popularity, and time slot popularity into the tensor. (4) CULT-TF reduces the complexity of tensor modeling by clustering similar users, clustering ROIs, and encoding time slots, significantly improving recommendation quality.

CULT-TF is significantly superior to each baseline method, i.e., ULT-TF, CUL-TF, CUT-TF, and CLT-TF. The reason is that, in reality, users are affected by varying degrees of geographical, temporal, and social influences. It is unable to model users' check-in behavior by considering only one influence. Similar user clustering has the greatest impact on recommendation accuracy, which can significantly improve the recommendation effect. Temporal influence also plays an important role in improving the quality of recommendations. That is mainly because time slot popularity can reduce the sparsity of the T dimension by mapping discrete check-in timestamps to time slots to capture the temporal preference of user check-in behavior more accurately. The impact of POI popularity on the recommendation results is greater than that of user activity, partly because users within a cluster have greater similarity after user clustering, resulting in user activity differences that are not obvious, and partly because POI popularity fusing the contextual information of the user, location, and time levels, can capture the location preference of user check-in behavior more accurately.

Our proposed CULT-TF method can be extended to other heterogeneous information networks containing rich semantic information. Nonetheless, there are still some potential limitations to our study. One limitation of our method is that the values of clustering parameters C and K affect the recommendation results, so it is necessary to set reasonable parameter values according to the datasets. In addition, we encode check-in timestamps to time slots by the month, day, and period time granularities and obtain the check-in pattern of users over the course of a year. The number of time slots should be set appropriately for datasets with longer time information, which affects the performance of tensor modeling.

7. Conclusions and Future Work

This paper proposes a POI recommendation method (CULT-TF) that integrates the contextual information of similar users to capture user preferences more accurately. A user activity model and user similarity model are presented to find active users as clustering centers and then cluster similar users. CULT-TF integrates social, geographical, and temporal influence into a unified location recommendation framework, in which social, geographical, and temporal influences are modeled as user activity, POI popularity, and time slot popularity, respectively. In CULT-TF, a U-L-T tensor matrix is constructed based on the clusters of similar users, and the *Top-N* list of POI recommendations is obtained through tensor factorization, which not only fuses the context of similar users but also reduces the complexity of tensor modeling. Finally, we conducted extensive experiments to evaluate the performance of CULT-TF on the Brightkite dataset. The experimental results show that the precision and recall of CULT-TF are always superior to those of the other recommendation methods evaluated in our experiments. This indicates that integrating the contexts of similar users can significantly improve recommendation accuracy.

In the future, we plan to integrate the semantics of POIs and the textual information derived from user comments, opinions, and views on check-in locations into a unified recommendation framework to further improve the recommendation quality of CULT-TF. In addition, we aim to perform more efficient and accurate clustering for users by considering other advanced clustering algorithms and addressing the cold start issue.

Author Contributions: Conceptualization, Yan Zhou; Methodology, Yan Zhou, Kaixuan Zhou, and Shuaixian Chen; Formal analysis, Yan Zhou; Investigation, Kaixuan Zhou and Shuaixian Chen; Validation, Kaixuan Zhou and Shuaixian Chen; Supervision, Yan Zhou; Funding acquisition, Yan Zhou; Writing—Original draft, Kaixuan Zhou; Writing—Review and editing, Yan Zhou. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 41871321), the National Key Research and Development Program of China (2022YFC3005702), and the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources (KF-2021-06-033).

Data Availability Statement: The datasets analyzed in this study are available on the website <http://snap.stanford.edu/data/loc-brightkite.html> (accessed on 1 March 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liao, G.; Jiang, S.; Zhou, Z.H.; Wan, C.X. POI Recommendation of Location-Based Social Networks Using Tensor Factorization. In Proceedings of the 19th IEEE International Conference on Mobile Data Management (MDM), Aalborg, Denmark, 25–28 June 2018; pp. 116–124.
2. Li, Z.; Huang, X.; Yuan, K. Survey of research on point-of-interest recommendation methods based on location-based social networks. *Appl. Res. Comput.* **2022**, *39*, 3211–3219.
3. Liu, Y.; Pham, T.; Cong, G. An experimental evaluation of point-of-interest recommendation in location-based social networks. *Proc. VLDB Endow.* **2017**, *10*, 1010–1021. [[CrossRef](#)]
4. Xu, S.; Fu, X.; Cao, J.; Liu, B.; Wang, Z. Survey on user location prediction based on geo-social networking data. *World Wide Web.* **2020**, *23*, 1621–1664. [[CrossRef](#)]
5. Lian, D.; Zhao, C.; Xie, X.; Sun, G.Z.; Chen, E.H.; Rui, Y. GeoMF: 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, New York, NY, USA, 24–27 August 2014; pp. 831–840.
6. Huang, J.; Tong, Z.; Feng, Z. Geographical POI recommendation for Internet of Things: A federated learning approach using matrix factorization. *Int. J. Commun. Syst.* **2022**, *2022*, e5161. [[CrossRef](#)]
7. Rahmani, H.A.; Aliannejadi, M.; Ahmadian, S.; Baratchi, M.; Afsharchi, M.; Crestani, F. LGLMF: Local geographical based logistic matrix factorization model for POI recommendation. In Proceedings of the 15th Asia Information Retrieval Societies Conference, Hong Kong, China, 7–9 November 2019; pp. 66–78.
8. Ying, Y.; Chen, L.; Chen, G. A temporal-aware POI recommendation system using context-aware tensor decomposition and weighted HITS. *Neurocomputing* **2017**, *242*, 195–205. [[CrossRef](#)]
9. Ren, X.Y.; Song, M.N.; Song, J.D. Point-of-Interest recommendation based on the user check-in behavior. *Chin. J. Comput.* **2017**, *40*, 28–51.

10. Li, X.; Jiang, M.; Hong, H.; Liao, L. A time-aware personalized point-of-interest recommendation via high-order tensor factorization. *ACM T. Inform. Syst.* **2017**, *35*, 1–23. [[CrossRef](#)]
11. Yao, L.; Sheng, Q.Z.; Qin, Y.; Wang, X.; Shemshadi, A.; He, Q. Context-aware point-of-interest recommendation using tensor factorization with social regularization. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, 9–13 August 2015; pp. 1007–1010.
12. Luan, W.; Liu, G.; Jiang, C. Collaborative tensor factorization and its application in POI recommendation. In Proceedings of the IEEE 13th International Conference on Networking, Sensing, and Control, Mexico City, Mexico, 28–30 April 2016; pp. 1–6.
13. Xing, S.; Liu, F.; Zhao, X.; Li, T. Points-of-interest recommendation based on convolution matrix factorization. *Appl. Intell.* **2018**, *48*, 2458–2469. [[CrossRef](#)]
14. Zhou, X.; Tian, J.; Peng, J.; Su, M. A Smart Tourism Recommendation Algorithm Based on Cellular Geospatial Clustering and Multivariate Weighted Collaborative Filtering. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 628. [[CrossRef](#)]
15. Ye, M.; Yin, P.; Lee, W.C.; Lee, D.L. 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, Beijing, China, 24–28 July 2011; pp. 325–334.
16. Maroulis, S.; Boutsis, I.; Kalogeraki, V. Context-aware point of interest recommendation using tensor factorization. In Proceedings of the IEEE International Conference on Big Data, Washington, DC, USA, 5–8 December 2016; pp. 963–968.
17. Cheng, C.; Yang, H.; King, I.; Lyu, M. Fused matrix factorization with geographical and social influence in location-based social networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Toronto, Canada, 22–26 July 2012; pp. 17–23.
18. Zhai, X.; Zheng, W.; Xiao, Y.; Liu, K. Point-of-interest recommendation system based on deepwalk and tensor decomposition. In Proceedings of the IEEE 25th International Conference on Computer Supported Cooperative Work in Design, Hangzhou, China, 4–6 May 2022; pp. 867–872.
19. Karatzoglou, A.; Amatriain, X.; Baltrunas, L.; Oliver, N. Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering. In Proceedings of the Fourth ACM Conference on Recommender Systems, Barcelona, Spain, 26–30 September 2010; pp. 79–86.
20. Jean-Benoit, G.; Talel, A.; Hubert, N. POI Recommendation: Towards Fused Matrix Factorization with Geographical and Temporal Influences. In Proceedings of the 9th ACM Conference on Recommender Systems, Vienna, Austria, 16–20 September 2015; pp. 301–304.
21. Lu, J.; Wu, D.; Mao, M. Recommender system application developments: A survey. *Decis. Support Syst.* **2015**, *74*, 12–32. [[CrossRef](#)]
22. Tobler, W.R. A computer movie simulating urban growth in the detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [[CrossRef](#)]
23. Rahmani, H.; Aliannejadi, M.; Baratchi, M. Joint geographical and temporal modeling based on matrix factorization for point-of-interest recommendation. In Proceedings of the 42th European Conference on IR Research, Lisbon, Portugal, 14–17 April 2020; pp. 205–219.
24. Luan, W.; Liu, G.; Jiang, C.; Qi, L. Partition-based collaborative tensor factorization for POI recommendation. *IEEE/CAA J. Autom. Sin.* **2017**, *4*, 437–446. [[CrossRef](#)]
25. Jiao, X.; Xiao, Y.; Zheng, W.; Wang, H.; Jin, Y. R2SIGTP: A novel real-time recommendation system with integration of geography and temporal preference for next point-of-interest. In Proceedings of the World Wide Web Conference, New York, NY, USA, 13–17 May 2019; pp. 3560–3563.
26. Ren, X.Y.; Song, M.; Haihong, E.; Song, J. Context-aware probabilistic matrix factorization modeling for point-of-interest recommendation. *Neurocomputing* **2017**, *241*, 38–55. [[CrossRef](#)]
27. Liu, Y.; Pei, A.; Wang, F.; Yang, Y.; Zhang, X.; Wang, H.; Dai, H.; Qi, L.; Ma, R. An attention-based category aware GRU model for the next POI recommendation. *Int. J. Intell. Syst.* **2021**, *36*, 3174–3189. [[CrossRef](#)]
28. Zhang, J.D.; Chow, C.Y. Spatiotemporal sequential influence modeling for location recommendations: A gravity-based approach. *ACM Trans. Intel. Syst. Tec.* **2015**, *7*, 11–25. [[CrossRef](#)]
29. Zhao, S.; King, I.; Lyu, M.R. Aggregated temporal tensor factorization model for point-of-interest recommendation. *Neural Process. Lett.* **2018**, *47*, 975–992. [[CrossRef](#)]
30. Zhang, J.D.; Chow, C.Y.; Li, Z. iGeoRec: A personalized and efficient geographical location recommendation framework. *IEEE T. Serv. Comput.* **2015**, *8*, 701–714. [[CrossRef](#)]
31. Zhu, J.; Wang, C.; Guo, X.; Ming, Q.; Li, J.; Liu, Y. Friend and POI recommendation based on social trust cluster in location-based social networks. *EURASIP J. Wirel. Comm.* **2019**, *2019*, 89. [[CrossRef](#)]
32. Gao, H.; Tang, J.; Hu, X.; Liu, H. Exploring temporal effects for location recommendation on location-based social networks. In Proceedings of the Seventh ACM Conference on Recommender Systems, Hong Kong, China, 12–16 October 2013; pp. 93–100.
33. Ye, M.; Shou, D.; Lee, W.C.; Yin, P.; Janowicz, K. On the semantic annotation of places in location-based social networks. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego, CA, USA, 21–24 August 2011; pp. 520–528.
34. Pennacchiotti, M.; Gurusurthy, S. Investigating topic models for social media user recommendation. In Proceedings of the 20th International Conference on World Wide Web, Hyderabad, India, 28 March–1 April 2011; pp. 101–102.
35. Li, H.; Hong, R.; Wu, Z. A spatial-temporal probabilistic matrix factorization model for point-of-interest recommendation. In Proceedings of the 2016 SIAM International Conference on Data Mining, Miami, FL, USA, 5–7 May 2016; pp. 117–125.

36. Zhang, J.; Chow, C.Y. 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, Santiago, Chile, 9–13 August 2015; pp. 443–452.
37. Fang, J.F.; Meng, X.F. POI recommendation based on LBSN and multi-graph fusion. *Act. Geo. Cart. Sinica*. **2022**, *51*, 739–749.
38. Tucker, L. Some mathematical notes on three-mode factor analysis. *Psychometrika* **1966**, *31*, 279–311. [[CrossRef](#)]
39. Zhao, S.; Zhao, T.; Yang, H.; Lyu, M.; King, I. STELLAR: Spatial-temporal latent ranking for successive point-of-interest recommendation. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 315–321.
40. Liu., B.; Xiong, H. Point-of-Interest recommendation in location based social networks with topic and location awareness. In Proceedings of the 2013 SIAM International Conference on Data Mining, Austin, TX, USA, 2–4 May 2013; pp. 396–404.
41. Macqueen, J. Classification and Analysis of Multivariate Observations. In *5th Berkeley Symposium on Mathematical Statistics and Probability*, Los Angeles, CA, USA; University of California Press: Berkeley, CA, USA; pp. 281–297.
42. Wang, X.; Liu, Y.; Zhou, X.; Wang, X.; Leng, Z. A Point-of-Interest Recommendation Method Exploiting Sequential, Category and Geographical Influence. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 80. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.