Periodic and Successive Point-of-interest Recommendation under Dual Social Group Influences with Matrix Factorization

Lin Cui\textsuperscript{1,2}, Dechang Pi\textsuperscript{1*}, Hui Li\textsuperscript{1}

\textsuperscript{1}College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, Jiangsu, China;

\textsuperscript{2} School of Information and Engineering, Suzhou University, Suzhou234000, Anhui, China

*Dechang Pi is the corresponding author

Abstract

Successive POI recommendation is a newly emerged research direction in recent years, which tries to recommend new POIs that the user has not visited before under the premise of knowing the current POIs of users. Previous work on successive POI recommendation only made limited improvements from using the geographical context and social information and so on. However, these approaches roughly regarded social influences as neighbor relationships and did not deeply study social influence factors. Hence, in this paper, we propose a unified POI recommendation model, which is called the periodic and successive point-of-interest recommendation based on the dual social influence. We subdivide the social influence into two social groups, namely, the direct social group with link relationship and the indirect social group with the common checked-in POIs. A stochastic gradient descent based algorithm is adopted to learn the matrix factorization and the experimental study on two real-world datasets demonstrate the effectiveness of the proposed approach over the existing state-of-the-art ones.

Keywords: periodic and successive POI recommendation, location-based social network, direct social group, indirect social group

1. INTRODUCTION

1.1 Background

Recent years have witnessed the rapid development of location-based social networks (short for LBSNs), such as taxi driving behavior prediction and other online social media services including Foursquare, Gowalla and Facebook places, etc (Cheng et. al., 2012) (Gavalas et. al., 2014; Zhang and Chow, 2015). More and more researches have focused on POI recommendation for studying the mobile behaviors of users (Zhang et. al., 2015). Through investigating many existing researches on POI recommendation, we find that previous efforts mainly consider the social and geographical influences in terms of social ties under a spatial space, the social-spatial correlation and so on (Cho et. al., 2011), temporal factors have been also taken into account (Gao et. al., 2013). But to the best of our knowledge, successive POI recommendation based on fine granularity of social relations and temporal division has been rarely studied.

In our research, we observe two prominent properties in the check-in sequence: periodic property and social group influence. Hence, we propose a novel matrix factorization method, which is called the Direct social group and Indirect social group based Periodic and Successive POI recommendation using Matrix Factorization (DIDPS-MF for short). To embed the periodic and successive properties based on social groups, our proposed recommendation model not only exploits the periodic and successive properties of user
check-ins, but also takes into account the social group influence in the check-in sequence. To summarize, the main contributions of this paper are three-fold:

We formally describe the user-POI layout information of LBSNs, define some notations related with our model, and then give the graphical representation of DIDPS-MF model.

We introduce the framework of our proposed DIDPS-MF method in detail. Based on this, we then model the DIDPS-MF method step by step.

We conduct extensive experiments to evaluate our proposed models using two real-world datasets from Foursquare and Gowalla. The experimental results demonstrate the superiority of the proposed DIDPS-MF against the existing approaches.

**1.2 Related work**

In this section, we briefly introduce the related work on POI recommendation, temporal influence and social influence analysis. Then we present the connections of our proposed model and prior works.

In recent years, various methods have been proposed for POI recommendation (Dunlavy et. al., 2011; Ye et. al., 2011) (Levandoski et. al., 2012; Lian et. al., 2014). POI recommendation systems depend on the check-in data and the classical user-POI rating matrix in which a rating corresponds to the visit frequency of a user at a given POI. One important research line includes matrix factorization models. Matrix factorization techniques have been widely adopted since the Netflix challenge (Liu and Xiong, 2013). For example, Zheng et al. have proposed a new approach, known as user-centered collaborative location and activity filtering, which pulls many users’ data together and apply collaborative filtering to find like-minded users and like-patterned activities at different locations (Zheng et. al., 2010). A recent study showed that weighted matrix factorization was the most adapted method to CF problems with implicit feedback (Griesner et. al., 2015).

Also, there are studies taking temporal factors into consideration for the purpose of improving the algorithm efficiency. For example, Yuan et al. defined a new style time-aware POI recommendation and developed a collaborative recommendation model, which recommended POIs for a given user at a specified time in a day (Yuan et. al., 2013). However, this methods cannot capture the evolving changes of user preferences. More recently, the importance of the sequential patterns hidden in the historical check-in sequences has been realized for next POI recommendations (Feng et. al., 2015). However, these studies are unable to recommend POI for a specific time period due to the lack of modeling temporal interval information in their methods. Further, the social dimension is another important factor leveraging the accuracy of the model and most of POI recommendation studies exploit social information (Gao et. al., 2012; Gao et. al., 2014; Bao et. al., 2015). Such as Ye et al. put forward a unified POI recommendation framework, which fused the user preference to a POI with social influence and geographical influence to make POI recommendations(Ye et. al., 2011).

Different from the above mentioned researches, our proposed DIDPS-MF model describes the direct social group dependence and indirect social group influence in a unified temporal and social group style. We also adopts the social regularization term, which is based on the assumption that the preference of a user is close to the weighted average preference of his friends.

The rest of the paper is organized as follows: We introduce the user-POI layout information of LBSNs, define some notations, give the graphical representation of DIDPS-MF model, describe the framework of the DIDPS-MF method and list its modeling
process, respectively in Sections 2. The experimental analysis results are presented in Section 3. Finally, we conclude this paper and give the future work in Section 4.

2. MATERIALS AND METHODS

2.1 Definition on the proposed DIDPS-MF Model

In this section, we firstly describe the user-POI layout information of LBSNs, define some notations related with our model, and then give the graphical representation of the proposed DIDPS-MF model.

2.1.1 The user-POI layout information of LBSNs

Location-based social network contains abundant information, Let $u=\{u_1,u_2,\ldots,u_m\} \subset U^m$ be a subset of users and $p=\{p_1,p_2,\ldots,p_n\} \subset P^n$ be a subset of POIs. Then let $R \in R^{m \times n}$ be a user-POI matrix containing $m$ users and $n$ POIs. Value $r_{ij}$ in $R$ refers to the visit frequency of user $u_i$ to the POI $p_j$. Then the problem is converted to predict the unobserved POIs in $R$ and rank the POIs. The rating score $r_{ij}$ indicates the preference of user $u_i$ to POI $p_j$. Hence, predicting the rating $r_{ij}$ plays the central role in POI recommendation.

As Figure 1 shows, in the direct social layer, users keep the direct social links each other and there exist friendships between them. Besides the direct social layer, the implicit indirect social layers are discovered as shown in the ellipses of Figure 1. For example, there are not the direct social links between user $u_2$ and user $u_3$, however, they have a common check-in POI $p_2$, so we define there is an indirect link between user $u_2$ and user $u_3$. The bottom map mainly includes POI information and the corresponding successive check-ins, which shows the check-ins of users are temporal and successive.

![Figure 1. The user-POI check-ins layout of LBSNs](image)

2.1.2 Preliminaries

For ease of reading, we list notations used throughout the paper in Table 1. We will give the following definitions.
Definition 1. (POI) A POI p refers to either an event or venue generated in various LBSNs or EBSNs.

Definition 2. (Check-ins cuboid) POIs check-ins cuboid is a $U \times T \times P$ cuboid, where $U$ is the number of users, $T$ is the number of time intervals and $P$ is the number of POIs. A cell indexed by $r_{u,p}$ stores the number of check-ins that user $u$ assigned to POI $p$ during time interval $t$.

Definition 3. (Time indexing scheme) We split a week into weekday and weekend and a day into the following four sessions: the morning session from 6:00 a.m. To 10:59 a.m., the afternoon and night session from 0:00 a.m. to 2:59 a.m. And 3:00 p.m. to 11:59 p.m., two transitive sessions that ranges from 3:00 a.m. To 5:59 a.m, and 11:00 a.m. To 2:59 p.m.

Definition 4. (latent factor matrix) User latent factor matrix is denoted as $U \in \mathbb{R}^{m \times d}$, which illustrates the individual user check-in preferences determined by a small number latent factors. Let $P \in \mathbb{R}^{n \times d}$ be the POI latent factor matrix, which denotes the role of POI are determined by the latent factors, with $d << \min(m,n)$ being the number of latent preference factors. The basic idea is to embody user $i$ and POI $j$ with the low-dimensional latent factor vectors $U_{i} \in \mathbb{R}^{d}$ and $P_{j} \in \mathbb{R}^{d}$.

Definition 5. (A direct social group) $DN_{u_{i}}^{t}$ denotes a set of direct social friends of user $u_{i}$ at time $t$. In other words, $u_{i}$ only constructs the link relationship with users in social group $DN_{u_{i}}^{t}$, but not consider any POI check-ins. In our work, we use $H \in \mathbb{R}^{n \times n}$ to denote the social relations between users. Formally, a direct social group $DN_{u_{i}}^{t}$ satisfies:

$$H_{ij} = \frac{|G(u_{i}) \cap G(u_{j})|}{|G(u_{i}) \cup G(u_{j})|}$$  \hspace{1cm} (1)

where $G(u)$ denotes the social groups that user $u_{i}$ joins, $G(U)$ denotes the social groups that $u_{j}$ joins, $|X|$ denotes the cardinality of the set $X$. 

### Table 1 List of notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U$</td>
<td>The set of users ${u_{1}, u_{2}, \cdots, u_{m}}$</td>
</tr>
<tr>
<td>$P$</td>
<td>The set of POIs ${p_{1}, p_{2}, \cdots, p_{n}}$</td>
</tr>
<tr>
<td>$T$</td>
<td>The set of time slots ${t_{1}, t_{2}, \cdots, t_{T}}$</td>
</tr>
<tr>
<td>$P_{u}^{t}$</td>
<td>The set of POIs that user $u$ has visited</td>
</tr>
<tr>
<td>$R = [r_{u,p}]$</td>
<td>A $</td>
</tr>
<tr>
<td>$DN_{u_{i}}^{t}$</td>
<td>Set of direct social friends of user $u_{i}$ at time $t$</td>
</tr>
<tr>
<td>$IN_{u_{i}}^{t}$</td>
<td>Set of indirect social friends of user $u_{i}$ with common POIs</td>
</tr>
<tr>
<td>$R_{u_{i},p}^{t}, \hat{R}<em>{u</em>{i},p}^{t}$</td>
<td>Observed and predicted ratings of user $u_{i}$ to POI $p$ at time $t$</td>
</tr>
</tbody>
</table>
Definition 6. (A indirect social group) \( IDN_{u_i}^t \) is a set of the indirect social friends of user \( u_i \) with common POIs. In other words, \( IDN_{u_i}^t \) stores the users who check in the POIs that user \( u_i \) has checked in. In terms of POI \( p_{ij} \) if user \( u_i \) and \( u_j \) co-participate in \( p_{ij} \), they share the similar interests. We use \( Q \in \mathbb{R}^{n \times n} \) to represent the interest similarity between users. Formally, an indirect social group \( IDN_{u_i}^t \) satisfies:

\[
Q_{ij} = \frac{\sum_{o=1}^{\mid T \mid} P_{u_i,o} P_{u_j,o}}{\sqrt{\sum_{o=1}^{\mid T \mid} (P_{u_i,o})^2} \sqrt{\sum_{o=1}^{\mid T \mid} (P_{u_j,o})^2}}
\]

where \( P_{u_i,o} \) denotes the POIs that user \( u_i \) has visited, \( P_{u_j,o} \) represents the POIs that user \( u_j \) has checked-in.

2.2 DIDPS-MF model: the direct social group and indirect social group based periodic and successive POI recommendation using matrix factorization

In this section, we formally introduce the framework of our proposed DIDPS-MF model. Based on this, we then model the proposed DIDPS-MF method step by step.

2.2.1 The framework of the DIDPS-MF model

The proposed DIDPS-MF model is a latent class statistical mixture model and it can be represented by a graphical model. Figure 2 illustrates the working flow of our POI recommendation framework. The whole framework consists of three steps: temporal division, temporal factorization and social group influence. Firstly, the original user-POI matrix \( R \) is divided into \( \mid T \mid \) sub-matrices according to the \( T \) temporal states, with each sub-matrix containing check-in actions that happened at the corresponding temporal state. Secondly, each \( R_t \) is factorized into the user check-in preference \( U_t \) and the POI features \( P \), while \( P \) is shared by all of \( U_t \). Thirdly, each \( C_t \) is factorized into the indirect social group users \( S_{ID} \) and the user check-in preference \( U_t \).

Figure 2. The graphical representation of DIDPS-MF model
Since the temporal division is easy to implement, in the following, we will describe in details the second and third steps, i.e., temporal factorization and social group influence for successive POI recommendation.

### 2.2.2 POI recommendation with matrix factorization

We first introduce a basic location recommendation model based on low-rank matrix factorization without considering temporal effects and social influence. The basic POI recommendation model approximates the check-in preference of user $u_i$ on an unvisited $p_j$ by solving the following optimization problem:

$$\min_{U,p} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - U_iP_j^T)^2 + \gamma (\|U\|^2_F + \|P\|^2_F)$$  

(3)

where $I \in (0,1)^{m \times n}$ is a check-in indicator matrix and is defined as,

$$I_{ij} = \begin{cases} 
1 & \text{if } R_{ij} \neq "\phi" \\
0 & \text{otherwise} 
\end{cases}$$  

(4)

To avoid the over-fitting problem, two smoothness regularization $\gamma (\|U\|^2_F + \|P\|^2_F)$ are added on $U$ and $P$, respectively. Where $\gamma$ is a non-negative parameter to control the capability of $U$ and $P$. $\|\cdot\|_F$ is the Frobenius norm of a matrix.

### 2.2.3 Modeling social group influence for successive POI recommendation

We further formulate the objective function by embedding our proposed direct social group influence $S_D$ and indirect social group influence $S_{ID}$ with the classical matrix factorization as below.

$$\min_{U,p} \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij} (R_{ij} - U_iP_j^T)^2$$

$$+ \alpha \sum_{i=1}^{m} \sum_{k \in S_D} w_{ik} \|U_i - U_k\|^2_F$$

$$+ \beta \sum_{i=1}^{m} \sum_{q=1}^{S_D} w_{iq} (C_{iq} - U_i(S_{ID})^T)^2$$

$$+ \frac{\gamma}{2} (\|U\|^2_F + \|P\|^2_F)$$  

(5)

where $w_{ik} \in (0,1)$ represents the connection weight $u_k$ can exert on node $u_i$ in the range of the direct social friends, $w_{iq} \in (0,1)$ denotes the influence weight $u_q$ can impose on user $u_i$ within the scope of the indirect social group but with the common POIs. Parameters $\alpha$ and $\beta$ aim to balance the mutual effects of social group influence (the second term and the third term) and the traditional collaborative filtering model (the first term).

### 2.2.4 Modeling temporal influence for successive POI recommendation
According to the temporal property of user check-in information, all users exhibit distinct check-in preferences at different hours of the day. This inspires us to consider a user’s check-in behavior as time-dependent check-in preferences. We then define \( U_t \in \mathbb{R}^{m \times d} \) as the time-dependent user check-in preferences under temporal state \( t \). As POI features are inherent properties that do not change much as time goes by. Therefore, we stipulate POI features to be time-independent, denoted as \( P \in \mathbb{R}^{n \times d} \). By approximating the check-in activities at each temporal state \( t \) and minimizing their aggregation, we obtain time-dependent user check-in preferences via the following optimization problem:

\[
\begin{align*}
\min_{U_{t \geq 0}, P_{t \geq 0}} & \frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}(R_{ij} - U_{ij}P_{ij})^2 \\
+ & \frac{\alpha}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{k \in S_{D}} w_{ik} \left\| U_{ij} - U_{ik} \right\|_F^2 \\
+ & \frac{\beta}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{q \in S_{DI}} w_{iq} (C_{iq} - U_{ij}S_{ID})^2 \\
+ & \frac{\gamma}{2} \left\| U_{ij} \right\|_F^2 + \left\| S_{D} \right\|_F^2 + \left\| S_{DI} \right\|_F^2 + \left\| P \right\|_F^2
\end{align*}
\] (6)

where \( R_{ij} \) contains the check-in activities at temporal state \( t \), and \( I_{ij} \) is the corresponding indicator function.

### 2.2.5 Modeling successive POI recommendation with temporal regularization

Based on the discussion of modeling temporal periodicity and successive properties in the above sections, the user temporal check-in preferences can be obtained by solving the following optimization problem. Where \( \lambda \) is a non-negative parameter to control the temporal regularization. Inspired by the temporal successive property, which implies that users on LBSNs tend to have closer check-in preferences on successive temporal state, we propose a temporal regularization to minimize the following terms:

\[
\begin{align*}
\min_{U_{t \geq 0}, P_{t \geq 0}} & \frac{1}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{j=1}^{n} I_{ij}(R_{ij} - U_{ij}P_{ij})^2 \\
+ & \frac{\alpha}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{k \in S_{D}} w_{ik} \left\| U_{ij} - U_{ik} \right\|_F^2 \\
+ & \frac{\beta}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} \sum_{q \in S_{DI}} w_{iq} (C_{iq} - U_{ij}S_{ID})^2 \\
+ & \frac{\lambda}{2} \sum_{t=1}^{T} \sum_{i=1}^{m} S_{t,i-1} \left\| U_{ij} - U_{i,t-1} \right\|_F^2 + S_{t,i} \left\| U_{ij} - U_{i,t+1} \right\|_F^2 \\
+ & \frac{\gamma}{2} \sum_{t=1}^{T} \left\| U_{ij} \right\|_F^2 + \left\| S_{D} \right\|_F^2 + \left\| S_{DI} \right\|_F^2 + \left\| P \right\|_F^2
\end{align*}
\] (7)

where, parameter \( \lambda \) aim to balance the mutual effect of successive influence information model (the fourth term), the collaborative filtering model (the first item), the direct social group influence (the second item) and the indirect social group influence (the third item). \( \gamma \) prevents the over-fitting problem. \( S_{t,i-1} \in [0,1] \) is defined as a
temporal coefficient that measures the temporal closeness of \( u_i \)'s check-in preferences between temporal state \( t \) and \( t-1 \). The larger \( S_{t,t-1} \) is, the closer \( u_i \)'s check-in preferences between \( t \) and \( t-1 \). We use cosine similarity to measure \( S_{t,t-1} \), defined as,

\[
S_{t,t-1} = \frac{C_j(i,;):C_{t-1}(i,:)}{\sqrt{\sum_j C_j^2(i,:)} \cdot \sqrt{\sum_j C_{t-1}^2(i,:)}} \tag{8}
\]

Note that we consider that the temporal state \( t-1 \) as \( T \) when \( t=1 \), i.e. \( U_{t-1} = U_T \) when \( t=1 \).

After some derivations, we can get the matrix form of temporal regularization.

Analogously, \( S_{t,t+1} \) can be computed in the same way.

\[
S_{t,t+1} = \frac{C_j(i,;):C_{t+1}(i,:)}{\sqrt{\sum_j C_j^2(i,:)} \cdot \sqrt{\sum_j C_{t+1}^2(i,:)}} \tag{9}
\]

### 2.3 An optimization method for DIDPS-MF model

In this paper, we adopt SGD (Stochastic Gradient Descent) approach to optimize the objective function. SGD algorithm randomly scans all training data and updates parameters along the gradient descent direction of the objective function for each user-POI entry. Each update is executed by the following formulation:

\[
\Lambda \leftarrow \Lambda - \xi \cdot \frac{\partial F(\Lambda)}{\partial \Lambda} \tag{10}
\]

where \( \xi \) is the learning rate, \( \Lambda \) represents all the involved model parameters, and \( \partial F \) corresponds to the objective function shown in Equation (7).

To get the gradients of Equation (10) w.r.t. \( U^T \), \( S_{t,t}^D \), \( S_{t,t}^ID \) and \( P \), a local minimum of the objective function in Equation (10) can be calculated by performing SGD algorithm and the gradients of all related parameters are calculated as follows:

\[
\frac{\partial F}{\partial U^T} = \sum_{i,j=1}^n I_j^i (R_{ij} - U_j^T P_j^I) P_j^T + \gamma U^T + \alpha \sum_{i,k=1}^n W_k^i (U_j^T - U_k^T) + \beta \sum_{q=1}^{S_{t,t}^D} W_q^j (C_q - U_j^T (S_{t,t}^ID)^T) + \lambda \sum_{t=1}^T (S_{t,t-1}(U^T(i,:)-U^{t-1}(i,:)) + S_{t,t+1}(U^T(i,:)-U^{t+1}(i,:))) \tag{11}
\]

Therefore, \( U^T \) is updated as

\[
U^T \leftarrow U^T - \xi \cdot \frac{\partial F(U^T,S_{t,t}^D,S_{t,t}^ID,P)}{\partial U^T} \tag{12}
\]

the gradient of \( F(t^T,S_{t,t}^D,S_{t,t}^ID,P) \) with respect to \( P \) is given as follows:
\[
\frac{\partial F}{\partial P} = \sum_{i=1}^{m} I_{ij}^t (R_{ij}^t - U_i^t P_j^t) U_i + \gamma P
\]

(13)

And \( P \) is updated as

\[
P \leftarrow P - \xi \cdot \frac{\partial F(U_i^t, S_D^t, S_{ID}^t, P)}{\partial P}
\]

(14)

The gradient of \( F(t^t, S_{ID}, S_{ID}, P) \) with respect to \( S_D^t \) is given as follows:

\[
\frac{\partial F}{\partial S_D^t} = \alpha \sum_{i=1}^{m} k \in S_D^t w_{ik}(U_i^t - U_k^t) + \gamma S_D^t
\]

(15)

And \( S_D^t \) is updated as

\[
S_D^t \leftarrow S_D^t - \xi \cdot \frac{\partial F(U_i^t, S_D^t, S_{ID}^t, P)}{\partial S_D^t}
\]

(16)

The gradient of \( F(t^t, S_{ID}, S_{ID}, P) \) with respect to \( S_{ID}^t \) is given as follows:

\[
\frac{\partial F}{\partial S_{ID}^t} = \beta \sum_{i=1}^{m} q \in S_{ID}^t w_{iq}(C_i^t - U_i(S_{ID}^t)^T) + \gamma S_{ID}^t
\]

(17)

Similarly, \( S_{ID}^t \) is updated as

\[
S_{ID}^t \leftarrow S_{ID}^t - \xi \cdot \frac{\partial F(U_i^t, S_D^t, S_{ID}^t, P)}{\partial S_{ID}^t}
\]

(18)

3. RESULTS AND DISCUSSION

In this section, experiments are conducted to verify the effectiveness of the proposed framework DIDPS-MF. We first list the adopted two real-word datasets, introduce the evaluation metric and all the baseline methods we compare, stipulate the parameter setting in section 3.1. Then, in order to verify whether the proposed model can improve the successive POI recommendation performance by incorporating the direct social group and indirect social group influence, we compare the proposed DIDPS-MF model with the state-of-the-art baselines in section 3.2.

3.1 Experimental setting

Datasets. We use two real-world datasets Foursquare and Gowalla in the experiments. Foursquare dataset includes user-POI check-in information of restaurant venues in NYC collected from Foursquare (https://foursquare.com/) from October 24, 2011 to February 20, 2012 (Yang et. al., 2014). Gowalla dataset was collected from Gowalla (https://gowalla.com/), which includes the user profiles, user friendship, location profiles, and users’ check-in history made before June 1, 2011 (Liu et. al., 2014). We filter out the users with only less than 5 check-ins and remove the POIs with time intervals less than 1 minute and more than 12 hours. Because POIs with time intervals less than 1 minute are not so useful and should be filtered out. In addition, if time intervals between previous check-ins are more than 12h, this means that some POIs are very likely to be
missed, or may deviate the results, so they also be removed. After applying the above filtering process, The details of the datasets are summarized in Table 2.

**Table 2** Statistics information of two experimental datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th># of Users</th>
<th># of POIs</th>
<th># of Check-in Pairs</th>
<th># of Edges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foursquare</td>
<td>2,621</td>
<td>3,005</td>
<td>20,211</td>
<td>37,634</td>
</tr>
<tr>
<td>Gowalla</td>
<td>2,207</td>
<td>5,302</td>
<td>52,632</td>
<td>85,771</td>
</tr>
</tbody>
</table>

Evaluation metrics: In our research problem, the central task is to predict a personalized list of the successive POIs the user wants to visit next, which can be seen as a ranking problem. In order to evaluate the effectiveness of POI recommendation in a real scenario, we generate the list test sets. We adopt the following three metrics to measure the successive POI recommendation performance, namely, Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Precision at position $P@N$. The smaller MAE or RMSE value means better rating prediction accuracy. In the following equations, $N$ is the set of user-POI check-in pairs $(u,p)$ used in the testing set. $P@N$ is mainly used in ranking problems and measures the ratio of the recommended POIs that are actually attended by users.

$$MAE = \frac{1}{N} \sum_{(u,p) \in N} |r_{u,p} - \hat{r}_{u,p}|$$  \hspace{1cm} (19)

$$RMSE = \sqrt{\frac{1}{N} \sum_{(u,p) \in N} (r_{u,p} - \hat{r}_{u,p})^2}$$  \hspace{1cm} (20)

$$P@N = \frac{\#TestSetHits}{\#Recommendations} = \frac{|test \cap topN|}{|topN|}$$  \hspace{1cm} (21)

Baseline methods. We compare the DIDPS-MF model with the following three baseline methods and the three proposed methods in this paper.

1. User Mean: this method uses the mean rating of each user to predict the successive POIs for the corresponding user.

2. POI Mean: this method utilizes the mean rating of each POI to predict the check-in values for the corresponding POI.

3. Biased MF(Takács et al., 2008): this is the MF model with user and POI biases and often widely used as a baseline in recommender systems.

We extend MF method to incorporate influences from multiple factors: direct social group (D), indirect social group (ID), the periodical temporal property (P) and the successive temporal property (S). The proposed methods are denoted using the letters in parentheses to indicate the influences considered in each method.

4. D-MF: this method incorporates direct social group influence.
(5) DID-MF: this method incorporates direct social group and indirect social group influences.

(6) DIDP-MF: this method incorporates direct social group, indirect social group and the periodical temporal influences.

Parameter Setting. The first 70% of ratings are used for training, and the remaining 30% for testing. We perform 5-fold cross-validation on the training set to empirically set the hyper parameters. The number of latent factors \(d=20\). The relative importance of direct social group influence and indirect social group influence are set to \(\alpha=0.8\) and \(\beta=0.6\), respectively. The balance parameter \(\lambda\) is set to 0.5. On the other hand, the regularization parameter is also empirically set to 0.001 for all. The latent factors are learned by SGD algorithm with initial learning rate \(\xi=0.001\), which decreases by a factor of 0.9 after each iteration. The same parameters are used in all methods for fair comparison for all our proposed methods and the baseline methods. For all the methods based on matrix factorization, the reported results are averaged over 5 runs to avoid the impact of initialization in parameter learning.

3.2 Experimental results analysis

We first compare the proposed methods with baseline methods and then evaluate the two schemes for defining the set of the direct social group and the indirect social group for POI. Lastly, we evaluate the proposed methods with cold-start setting.

3.2.1 Methods comparison on MAE and RMSE

The prediction errors measured by MAE and RMSE of all models are reported in Table 3 and Table 4 with best results highlighted in boldface. We make four observations from the results.

<table>
<thead>
<tr>
<th>ID</th>
<th>Method</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>User Mean</td>
<td>0.8305</td>
<td>1.0237</td>
</tr>
<tr>
<td>2</td>
<td>POI Mean</td>
<td>0.8067</td>
<td>1.0127</td>
</tr>
<tr>
<td>3</td>
<td>Biased MF</td>
<td>0.7922</td>
<td>1.0022</td>
</tr>
<tr>
<td>4</td>
<td>D-MF</td>
<td>0.7849</td>
<td>1.0017</td>
</tr>
<tr>
<td>5</td>
<td>DID-MF</td>
<td>0.7563</td>
<td>1.0008</td>
</tr>
<tr>
<td>6</td>
<td>DIDP-MF</td>
<td>0.7341</td>
<td>0.9992</td>
</tr>
<tr>
<td>7</td>
<td>DIDPS-MF</td>
<td>0.7299</td>
<td>0.9883</td>
</tr>
</tbody>
</table>

Table 3 Experimental results of the compared methods w.r.t. MAE and RMSE on Foursquare

<table>
<thead>
<tr>
<th>ID</th>
<th>Method</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>User Mean</td>
<td>0.9321</td>
<td>1.0061</td>
</tr>
<tr>
<td>2</td>
<td>POI Mean</td>
<td>0.8720</td>
<td>1.0003</td>
</tr>
<tr>
<td>3</td>
<td>Biased MF</td>
<td>0.8376</td>
<td>0.9932</td>
</tr>
<tr>
<td>4</td>
<td>D-MF</td>
<td>0.7759</td>
<td>0.9891</td>
</tr>
<tr>
<td>5</td>
<td>DID-MF</td>
<td>0.7294</td>
<td>0.9757</td>
</tr>
<tr>
<td>6</td>
<td>DIDP-MF</td>
<td>0.6836</td>
<td>0.9734</td>
</tr>
<tr>
<td>7</td>
<td>DIDPS-MF</td>
<td>0.6457</td>
<td>0.9692</td>
</tr>
</tbody>
</table>

Table 4 Experimental results of the compared methods w.r.t. MAE and RMSE on Gowalla
From Table 3 and Table 4, we can conclude the following observations. First, D-MF, DID-MF, DIDP-MF and DIDPS-MF obtain better performance than the other methods, which shows that embedding social group influence can improve the performance of POI recommendation. Second, DID-MF obtains better results than D-MF, the main reason is that indirect social group influence further provides the social friends with common POI check-ins and these friends also help the user to check-in more interesting POIs. In addition, DIDP-MF and DIDPS-MF are both outperform D-MF and DID-MF because of the introduction of more fine granular time division. Thus, we further collaborate the social group influence rating and temporal influence for the task of successive POI recommendation. All in all, the proposed DIDPS-MF always obtains the best results with respect to the evaluation metrics MAE and RMSE.

3.2.2 Methods comparison on the precision rate \( P@N \)

Since the methods considering implicit rating always performs better than the pure matrix factorization methods for POI recommendation, we further compare our method DIDPS-MF with D-MF, DID-MF and DIDP-MF w.r.t. \( P@5, P@10, P@15 \) and \( P@20 \). The comparison results are shown in the following Figure 3 and Figure 4, among which, \( N \) in the horizontal coordinates denotes the number of the top recommendation results.

![Figure 3](image-url)

**Figure 3.** Performance comparison in terms of precision rate on Foursquare dataset
From Figure 3 and Figure 4, we can observe that the precision rates of DIDPS-MF are obviously larger than DIDP-MF, DID-MF and D-MF. Moreover, the precision rates of DID-MF are slightly larger than D-MF, which shows that, though the indirect social group influence plays a certain role on POI recommendation, the performance of the algorithm is not greatly improved. However, the introduction of periodic temporal feature and successive temporal feature can obviously improve the performance of POI recommendation at precision rate. Thus, we can conclude that the proposed method can obtain better POI recommendation results. The best prediction accuracy is achieved by DIDPS-MF which considers direct social group influence (D), indirect social group influence (ID), periodic temporal influence (P) and successive temporal influence (S).

4. CONCLUSION

In this paper, we studied a new problem of successive POI recommendation in LBSN. The technical improvement is to jointly model three sources of POI information, i.e., the direct social group influence, the indirect social group influence and periodic temporal influence. We present a new matrix factorization model to integrate these information for accurate POI recommendation. We test the model on two real-world datasets Foursquare and Gowalla, and the results demonstrate the performance of our proposed DIDPS-MF model outperform the other state-of-the-art methods. As a part of our future work, we will include POI categories into the model and to cope with the scalability issues. Other aspects that may improve the prediction quality are content information like attributes of users or POIs, or the content of user reviews, which may be help to provide more precise POI recommendation.

5. ACKNOWLEDGMENTS

This work was supported by the Scientific Research Innovation Program for College Graduates of Jiangsu Province in China (KYLX15_0321), the Key Project of Overseas Visiting for the Outstanding Young Backbone Talents at Universities and Colleges of Anhui Province in China (gxfxZD2016267, gxyqZD2016344), the Key Projects of Anhui Province Colleges and Universities Natural Science Foundation of China (KJ2016A768), and Soft science research project of Anhui science and technology plan in 2016 Center.
6. REFERENCES


