



Personalized POI Recommendation: Spatio-Temporal Representation Learning with Social Tie

Shaojie Dai, Yanwei Yu^(✉), Hao Fan, and Junyu Dong

Department of Computer Science and Technology,
Ocean University of China, Qingdao, China

daishaojie@stu.ouc.edu.cn, {yuyanwei,fanhao,dongjunyu}@ouc.edu.cn

Abstract. Recommending a limited number of Point-of-Interests (POIs) a user will visit next has become increasingly important to both users and POI holders for Location-Based Social Networks (LBSNs). However, POI recommendation is a challenging task since complex sequential patterns and rich contexts are contained in extremely sparse user check-in data. Recent studies show that embedding techniques effectively incorporate POI contextual information to alleviate the data sparsity issue, and Recurrent Neural Network (RNN) has been successfully employed for sequential prediction. Nevertheless, existing POI recommendation approaches are still limited in capturing user personalized preference due to separate embedding learning or network modeling. To this end, we propose a novel unified spatio-temporal neural network framework, named PPR, which leverages users' check-in records and social ties to recommend personalized POIs for querying users by joint embedding and sequential modeling. Specifically, PPR first learns user and POI representations by joint modeling User-POI relation, sequential patterns, geographical influence, and social ties in a heterogeneous graph, and then models user personalized sequential patterns using the designed spatio-temporal neural network based on LSTM model for the personalized POI recommendation. Extensive experiments on three real-world datasets demonstrate that our model significantly outperforms state-of-the-art baselines for successive POI recommendation in terms of Accuracy, Precision, Recall and NDCG. The source code is available at: <https://github.com/dsj96/PPR-master>.

Keywords: POI recommendation · Location-based social network · Spatio-temporal neural network · Heterogeneous graph

1 Introduction

Newly emerging LBSNs has become an important mean for people to share their experience, write comments, or even interact with friends. With the prosperity of LBSNs, many users check in at various POIs via mobile devices in real time. Therefore, a large amount of check-in data is being generated, which is crucial to understand the users' preferences and behaviors. POI recommendation not only helps

users explore attractive and interesting places, but also gives guidance to location-based service providers, where to launch advertisements to target customers for marketing. Due to the great significance to both of users and businesses, how to use spatio-temporal information effectively, and recommend a limited number of POIs users more likely visit next have been attracting increasing attention in both industry and academia.

In particular, several studies [2, 10, 14, 20, 24] have been conducted to recommend successive POIs for users based on users' spatio-temporal check-in sequence in LBSNs. Based on Markov chain model, LORE [24] and NLPMM [2] explore users' successive check-in patterns by considering temporal and spatial information. ST-RNN [10] employs RNN to capture the users' sequential check-in behaviors. In a follow-up work, STGN [28] carefully designs the time gates and distance gates in LSTM to model users' sequential visiting behaviors by enhancing long short term memory. Additionally, some models [1, 4] based on Word2Vec [15] framework to capture the preference and mobility pattern of users and the relationship among POIs also achieved decent performance. GE [20] uses graph embedding to combine the sequential effect, geographical influence, temporal effect and semantic effect in a unified way for location-based recommendation. Recently, SAE-NAD [14] utilizes a self-attentive encoder to differentiate the user preference and a neighbor-aware decoder to incorporate the geographical context information for POI recommendation.

However, location-based POI recommendation still faces three major challenges. First, *data sparsity*, unlike the general e-commerce, music and movie recommendation, which can be collected and verified just online, location-based POI recommendation systems usually associate with the POI-entities. Only when a user visits a POI-entity, a check-in record is generated. Therefore, the check-in records in the POI recommendation task is much sparser. This issue has plagued many POI recommendation models based on the collaborative filtering. Furthermore, data sparsity problem in check-in data makes it difficult to capture user's sequential pattern, because the check-in sequence is very short or is not continuous in time. Second, *contextual factors*, POI recommendation may be affected by various contextual factors, including social tie influence, geographical influence, temporal context, and so on. In fact, social ties are often available in LBSNs, and recently studies show that social networks associated with users are important in POI recommendation task since users are more likely to be influenced by their close friends (Who keeps company with the wolf will learn to howl). In this work, we incorporate social ties, check-in time interval, sequential and geographical effect into user-POI interaction graph to joint learn user and POI representations. Lastly, *dynamic and personalized preferences*, users' preferences are changing dynamically over time. At different time and circumstances, users may prefer different POIs. For example, some users prefer to visit gourmet restaurants in the local area, but when they go to a new city, some prefer to visit the cultural landscapes, while some prefer the natural landscapes. Dynamically and accurately capturing this trend has been proved to be essential for personalized POI recommendation task. However, effectively modeling the personalized sequential transitions from the sparse check-in data is challenging.

To address the aforementioned challenges, in this work, we stand on advances in embedding technique and RNN network, and propose our model, named **PPR**, which is a *spatial-temporal representation learning* framework for personalized and successive POI recommendation. First, we jointly model the user-POI relation, sequential effect, geographical influence and social ties by constructing a heterogeneous graph, and then develop a densifying trick by adding second-order neighbors to nodes with low in/out-degrees to alleviate the data sparsity issue. Then, we learn user and POI representations by embedding the densified heterogeneous graph into a shared low-dimensional space. Furthermore, to better capture the user dynamic and personalized preference, we also design a spatio-temporal neural network by concatenating user embedding, POI embedding and POI category as personalized sequence input to feed the network.

The main contributions of this paper are summarized as follows:

- We propose a novel PPR model for personalized POI recommendation, which incorporates users’ check-in records and social ties. We construct a heterogeneous graph by jointly taking user-POI relation, sequential pattern, geographical effect and social ties into consideration to learn the representations of users and POIs.
- We propose a spatio-temporal neural network to model users’ dynamic and personalized preference by concatenating user, POI embedding and POI category to generate personalized behavior sequence.
- We conduct extensive experiments to compare our method with state-of-the-art baselines, and our method significantly outperforms state-of-the-art baselines for successive POI recommendation task.

2 Related Work

General POI Recommendation. The most well-known approaches of personalized recommendation are collaborative filtering (CF) and Matrix Factorization (MF). The conventional CF techniques have been widely studied for POI recommendation. LARS [6] employs item-based CF to make POI recommendation with the consideration of travel penalty. FCF [22] is a friend-based CF model based on the common visited POIs among friends, which considers the social influence. UTE [23] is a collaborative recommendation model that incorporates with temporal and geographical information. However, such methods suffer the data sparsity problem, leading them difficult to identify similar users.

Recommendation models based on MF and embedding learning [7, 11, 12] have been intensively studied. Rank-GeoFM [9] fits the users’ preference rankings for POIs to learn the latent embeddings. By incorporating the geographical context, it utilizes a geographical factorization method for calculating the recommendation score. TSG-MF [26] models the multi-tag influences via extracting a user-tag matrix and the social influences via social regularization, and uses a normalized function to model geographical influences.

Next POI Recommendation. In the literature, next POI recommendation issues have been studied in [19, 28], in which the main objective is to exploit the user’s check-in sequence between different POIs and dynamic preference.

Markov Chains (MC) Based Methods. MC based models aim to predict the next behavior according the historical sequential behaviors. FPMC-LR [3] considers first-order Markov chain for POI transitions and distance constraints. HMM [21] exploits check-in category information to capture the latent user movement pattern by using a mixed hidden Markov chain. LORE [24] incrementally mines sequential patterns and represents it as a dynamic location-location transition graph. By utilize an additive Markov chain, LORE fuses the sequential, geographical and social influence in a unified way.

Graph-Based Methods. Graph-based approaches are exploring in the literature of next POI recommendation. GE [20] jointly captures the latent relations among the POI, region, time slot and words related to the POIs by constructing four bipartite graphs. HME [5] projects the entities into a hyperbolic space after study multiple contextual subgraphs. Although the above approaches achieve promising performance, they can not model the sequential patterns effectively.

RNN-Based Methods. Recently, RNNs such as LSTM or GRU have demonstrated groundbreaking performance on predicting sequential problem. ST-RNN [10] utilizes RNN structure to model the temporal contexts by carefully designing the time-specific and distance-specific transition matrices. NEXT [25] encodes the sequential relations within the pre-trained POI embeddings by adopting DeepWalk [16] technique. Time-LSTM [30] employs LSTM with time gates to capture time interval among users’ behaviors. CAPE [1] first uses a check-in context layer to capture the geographical influence of POIs and a text content layer to model the characteristics of POIs from text content. Then, CAPE employs RNN as recommendation component to predict successive POIs. PEU-RNN [13] proposes a LSTM based model that combines the user and POI embeddings, which are learned from Word2Vec. ASPPA [27] proposes to identify the semantic subsequences of POIs and discover their sequential patterns. Recently, STGN [28] extends the LSTM gating mechanism with the spatial and temporal gates to capture the user’s space and time preference. However, these approaches fail to capture users’ personalized preferences.

3 Problem Definition

In this section, we first give the key concepts used in this paper. Then, the problem definition for personalized POI recommendation is formulated.

Definition 1 (POI). *A POI is a uniquely identified venue in the form of $\langle p, \ell, cat \rangle$, where p is the POI identifier, cat denotes the category of the POI, and ℓ represents the geographical coordinates of the POI (i.e., longitude and latitude).*

Definition 2 (Check-in record). *A check-in record is a triple $c = \langle u, v, t \rangle$ that represents user u visiting POI v at timestamp t .*

The collection of all users is denoted as U , and the collection of all POIs is denoted as V .

Definition 3 (Trajectory). *The trajectory of a user u is a sequence of all check-in records $(\langle u, v_1, t_1 \rangle, \langle u, v_2, t_2 \rangle, \dots, \langle u, v_n, t_n \rangle)$ made by user u in chronological order. We denote it as T_u .*

Definition 4 (Social Ties). *Social ties among users is defined as a graph $\mathcal{G}_u = (U, \mathcal{E}_u)$, where U is the set of users, and \mathcal{E}_u is the set of edges between the users. Each edge $e_{ij} \in \mathcal{E}_u$ represents users u_i and u_j being friends in LBSNs and is associated with a weight $w_{ij} > 0$, which indicates their tie strength.*

Problem 1 (Successive POI Recommendation). *Given users' check-in records and their social ties, and a querying user u with his/her current check-in $\langle u, v, t \rangle$, our goal is to recommend top- k POIs that user u would be interested in in the next τ time period.*

4 Methodology

In this section, we first present the details of the proposed framework PPR. Then we introduce our model how to utilize PPR model to make personalized POI recommendation.

4.1 Heterogeneous Graph Construction

We first introduce the heterogeneous User-POI graph to model users' sequential check-ins and social relationships. Specifically, we employ a heterogeneous graph $\mathcal{G} = (V, U, \mathcal{E}, W)$ to jointly model the multiple relations between users and POIs. U and V are the user collection and POI collection respectively, and \mathcal{E} is the set of all edges between nodes in \mathcal{G} , which are categorized into three edge types, *i.e.*, \mathcal{E}_u , \mathcal{E}_v , and $\mathcal{E}_{u,v}$. As mentioned in Definition 4, each edge $e_{i,j} \in \mathcal{E}_u$ represents that user u_i and u_j are friends. Each edge $e_{i,j} \in \mathcal{E}_v$ denotes that there exists at least one user visits POI v_j after visiting POI v_i , and each edge $e_{i,j} \in \mathcal{E}_{u,v}$ indicates that user u_i visits POI v_j at least one time. Notice that each edge $e \in \mathcal{E}_u \cup \mathcal{E}_{u,v}$ is a bi-directed edge and each edge $e \in \mathcal{E}_v$ is a directed edge, and each edge is associated with a weight $w \in W (w > 0)$, which indicates the strength of the relation.

Modeling User-POI Relation. Intuitively, we consider that if user u_i visit POI v_j more frequent, u_i and v_j have a stronger relation than with other POIs. Therefore, we formulate the weight between user u_i and POI v_j as:

$$w_{i,j} = \text{freq}(u_i, v_j), \quad (1)$$

where $\text{freq}(\cdot)$ denotes check-in frequency of user u_i visiting POI v_j . Since we aim to build a directed graph to accommodate the following work, we define $w_{i,j} = w_{j,i}$ for the bi-directed edge $e_{i,j} \in \mathcal{E}_{u,v}$ between user u_i and POI v_j .

Modeling Sequential and Geographical Effect. Compared with general POI recommendation, successive POI recommendation pays more attention to sequential pattern. The impact of user’s recent check-in behaviors are greater than those of a long time ago when making POI recommendations [20]. To further model the sequential effect, we carefully design a weighting strategy for the edges in \mathcal{E}_v .

Let $\Delta t_{k,k+1}^u$ be the time interval between two consecutive check-in records in the trajectory T_u of user u . $l_{k,k+1}^u$ is the flag that indicates the status of a pair of consecutive check-in records in the trajectory T_u , which is defined as:

$$l_{k,k+1}^u = \begin{cases} 1 & \text{if } \Delta t_{k,k+1}^u < \theta \\ 0 & \text{else} \end{cases}, \quad (2)$$

where θ is a predefined time threshold.

Given an edge $e_{i,j} \in \mathcal{E}_v$ from POI v_i to POI v_j , the sequential weight $w_{i,j}^{(seq)}$ for the edge $e_{i,j}$ is defined as:

$$w_{i,j}^{(seq)} = \sum_{u \in U} \sum_{k=1}^{|T_u|-1} l_{k,k+1}^u, \text{ if } v_k = v_i \text{ and } v_{k+1} = v_j. \quad (3)$$

Namely, the weight $w_{i,j}^{(seq)}$ for the edge from POI v_i to POI v_j is the total number of times that all users visit v_i first and then v_j in their trajectories.

Furthermore, geographical influence indicates the impact of geographical distance to the users’ spatial behaviors. According to [5, 8], the distribution of the geographical distance between two successive POIs follows the power-law distribution, which means users are more willing to visit POIs close to the current location. Therefore, we incorporate the geographical distance into our model as follows:

$$w_{i,j}^{(geo)} = \frac{d_{i,j}^\kappa}{\sum_{v_k \in N(v_i)} d_{i,k}^\kappa}, \quad (4)$$

where $N(v_i)$ represents the set of out-neighbor POIs of POI v_i in \mathcal{E}_v , $d_{i,j}$ denotes the Euclidean distance between POIs v_i and v_j , and κ is the negative exponent (*i.e.*, $\kappa < 0$). Finally, we combine the sequential and geographical influence as follows:

$$w_{i,j} = w_{i,j}^{(seq)} \cdot w_{i,j}^{(geo)}. \quad (5)$$

In such way, the sequential, time interval and geographical information are all reflected in graph \mathcal{G} .

Modeling Social Tie Strength. Users in an LBSN have multiple types of relations with other users, such as friends, family and colleagues. The preference of a user in social network are easily affected by his/her close friends or other users which has some kind of relations with them. Recently, these social ties are incorporated into the POI recommendation system [25] to improve the

recommendation performance. In this work, we propose to assign the weight between the users based on their historical check-in interactions. Specifically, for two socially connected users u_i and u_j , we assign the edge weight $w_{i,j}$ as:

$$w_{i,j} = \frac{\varepsilon + \sum_{v \in V} \min(f_{u_i,v}, f_{u_j,v})}{|T_{u_i} \cap T_{u_j}| + 1}, \tag{6}$$

where ε is a very small float number to avoid two users have connection but no common visited POIs, $f_{u_i,v}$ denotes the frequency of user u_i visiting at POI v , and $|T_{u_i} \cap T_{u_j}|$ represents the number of the common visited POIs for user u_i and u_j . Therefore, the common preferences between socially connected users are also taken into account in the User-POI graph \mathcal{G} .

Densifying Graph. Most recommendation models need to take the data sparsity into consideration, but the check-in data in POI recommendation area is much sparser. To address the data sparsity issue, we propose to construct a dense graph based on the graph \mathcal{G} . Specifically, we regard each user and POI as a node, and expand the neighbors of those nodes with low in/out degrees by adding higher order neighbors. In this work, we only consider expanding second-order neighbors to every node. If the out-degree of a node in \mathcal{G} is less than a predefined threshold ρ , we create an edge from node v_i to its second-order out-neighbor node v_j and assign the weight as follows:

$$w_{i,j} = \sum_{v_k \in N(v_i)} w_{i,k} \frac{w_{k,j}}{d_k^{(o)}}, \tag{7}$$

where $N(v_i)$ is the set of out-neighbors of node v_i , and $d_k^{(o)}$ is the out-degree of the node v_k . The densifying method for nodes with a low in-degree less than ρ is same. After densifying the User-POI graph, we can get the a more dense network, denoted by \mathcal{G}_{dense} . Then we use \mathcal{G}_{dense} instead of \mathcal{G} and exploit embedding technique to learn the nodes' representation vectors.

4.2 Learning Latent Representation

Inspired by LINE [17], which learns the first- and second-order relations representations of homogeneous networks. We develop it to learn heterogeneous node representations on our constructed heterogeneous graph \mathcal{G}_{dense} .

Specifically, we regard each user or POI as a node v and ignore their node type. In graph \mathcal{G}_{dense} , each node plays two roles: the node itself and a specific "context" of other nodes. We use \vec{v}_i to denote the embedding vector of node v_i when it is treated as a node, and \vec{v}_i^T to denote the embedding vector of v_i when it is treated as a specific "context". In particular, we use a binary cross-entropy loss to encourage nodes and their "context" connected with an edge, to have similar embeddings. Therefore, we minimize the following objective function:

$$\mathcal{O} = - \sum_{e_{i,j} \in \mathcal{E}} \left(\log(\sigma(\vec{v}_j'^T \cdot \vec{v}_i)) + w_n \sum_{v_n \in Neg(v_i)} \log(1 - \sigma(\vec{v}_n'^T \cdot \vec{v}_i)) \right), \tag{8}$$

where $\sigma()$ is the sigmoid function, \vec{v}_j^T denotes vector transpose, $Neg(v_i)$ is a negative edge sampling *w.r.t.* node v_i in \mathcal{G}_{dense} , and w_n denotes the negative sampling ratio, which is a tunable hyper-parameter to balance the positive and negative samples.

By minimizing the objective function \mathcal{O} with ASGD (asynchronous stochastic gradient) optimization and edge sampling technique, we can learn a d -dimensional embedding vector for each user and POI in \mathcal{G}_{dense} . Additionally, the representation learning is highly efficient and is able to scale to very large graphs because of the use of edge sampling technique.

4.3 Modeling User Dynamic and Personalized Preference

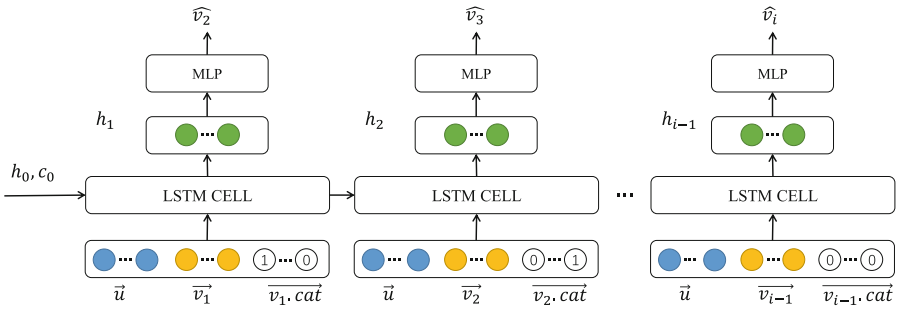


Fig. 1. Architecture of the proposed model

After representation learning, all users and POIs are mapped into a low dimensional space. However, the latent representations only capture the users’ preferences or POIs’ characteristics in a general way. Although it can model sequence transition patterns and geographical influence, some personalized preference may not be preserved in the node representations.

Furthermore, the categories of POIs are very useful to make a better representation of venues and improve the recommendation performance. In order to model user dynamic and personalized preference, we propose to concatenate user embedding, POI embedding and POI category to generate a new and more personalized embedding to represent a check-in record. More concretely, we use one-hot encoding to represent the POI category information.

Additionally, to better model user dynamic preference and sequential behavior patterns, we utilize LSTM model to construct a spatio-temporal neural network. As illustrated in Fig. 1, h_t and c_t denote the hidden state and cell state of LSTM at time t respectively. Given a user u and his/her trajectory sequence T_u , first, we concatenate the user embedding, POI embeddings with POI categories that he/she visited, and we can get a new embedding sequence. Second, we feed LSTM network with these new embedding sequences of all users. Specifically, we utilize the first $i - 1$ POIs as input to train the network, and predict the

$(i + 1)$ -th POI as the recommended POI based on the current i -th POI. At the output layer, we also connect a multi-layer perceptron (MLP). Therefore, we use the following objective function to train the model:

$$\mathcal{O}_{lstm} = \sum_{t=1}^{i-1} MSE(MLP(h_t), \overrightarrow{v_{t+1}}), \quad (9)$$

where h_t is hidden representation at time step t , $MSE(\cdot, \cdot)$ is a criterion that measures the mean squared error (*e.g.*, squared L2 norm) between each element.

4.4 Personalized POI Recommendation

As described in Sect. 4.3, the user embedding and the first i POI embedding sequence are used to train the spatio-temporal neural network. For the querying user u , the embedding vector of the $(i + 1)$ -th POI can be predicted by the current POI v_i as:

$$\widehat{v_{i+1}} = MLP(h_i). \quad (10)$$

Therefore, for each POI v , we calculate its recommendation score as follows:

$$Score(v|\widehat{v_{i+1}}, u, T_u) = 1 - MSE(\widehat{v_{i+1}}, \overrightarrow{v}). \quad (11)$$

Finally, we rank all POIs by their recommendation scores and select top- k POIs as the candidate that user u is more likely to visit in the next τ time period.

5 Experiments

5.1 Datasets

We conduct extensive experiments on three public real-world large-scale datasets: Foursquare¹, Gowalla² and Brightkite³. The basic statistics of these three datasets are summarized in Table 1. Notice that we preprocess these datasets utilizing the same method of [29] by filtering the POIs visited by less than five users and the users with less than ten check-in records.

- **Foursquare:** This dataset contains 483,813 check-in records generated by 4,163 users who live in California from December 2009 to July 2013.
- **Gowalla:** Gowalla is a location-based social networking website where users share their locations by checking-in. We choose data from Asian area for our experiments. It includes 251,378 check-in records generated by 6,846 users over the period of February 2009 to October 2010.
- **Brightkite:** Brightkite is also a location-based social networking service provider. We use the same selection strategy to obtain the check-in records generated by Asian users, which contains 572,739 records of 5,677 users.

Notice that there are 35 POI categories in Foursquare, and no category information is attached to Gowalla and Brightkite datasets.

¹ <https://sites.google.com/site/dbhongzhi/>.

² <http://snap.stanford.edu/data/loc-Gowalla.html>.

³ <http://snap.stanford.edu/data/loc-Brightkite.html>.

Table 1. Basic statistics of three datasets

Dataset	Foursquare	Gowalla	Brightkite
# of users	4,163	6,846	5,677
# of POIs	121,142	74,856	128,799
# of check-ins	483,813	251,378	572,739
# of categories	35	/	/
Time span	Dec. 2009–Jul. 2013	Feb. 2009–Oct. 2010	Apr. 2008 - Oct. 2010

5.2 Evaluation Metrics

To evaluate the recommendation model performance, we use four widely-used metrics, *i.e.*, Accuracy ($Acc@k$), Precision ($Pre@k$), Recall ($Rec@k$) and Normalized Discounted Cumulative Gain ($NDCG@k$), which are also used to evaluate top- k POI recommendation in [1, 18, 27, 28].

Let $\#hit@k$ denote the number of hits in the test set, and $|D_{Test}|$ is the number of all test records. $Acc@k$ is defined as:

$$Acc@k = \frac{\#hit@k}{|D_{Test}|}. \quad (12)$$

Let R_k denote the top- k POIs with the highest recommendation score, and T_k be the ground truth of the corresponding record, respectively. $Pre@k$ and $Rec@k$ are defined as:

$$Pre@k = \frac{1}{|D_{Test}|} \sum \frac{|R_k \cap T_k|}{|R_k|}, \quad (13)$$

$$Rec@k = \frac{1}{|D_{Test}|} \sum \frac{|R_k \cap T_k|}{|T_k|}. \quad (14)$$

To better measure the ranking quality, we further utilize $NDCG@k$, which assigns higher scores to POIs at top position ranks, to evaluate the model. The $NDCG@k$ for each test case is defined as:

$$NDCG@k = \frac{DCG@k}{IDCG@k}, \quad (15)$$

where $DCG@k = \sum_{i=1}^k \frac{2^{rel_i} - 1}{\log_2(i+1)}$, $IDCG@k = \sum_{i=1}^k \frac{1}{\log_2(i+1)}$ and $rel_i = 1$ refers to the graded relevance of result ranked at position i . We use the binary relevance in our experiments, *i.e.*, $rel_i = 1$ if the recommended POI is in the ground truth, otherwise, $rel_i = 0$.

5.3 Baselines

We compare our model against the following baselines for successive POI recommendation:

- **Rank-GeoFM** [9]: It is a ranking based geographical factorization model, which earns the embeddings of users and POIs by combining geographical and temporal influence in a weighting scheme.
- **ST-RNN** [10]: ST-RNN is a RNN-based model with spatial and temporal contexts for next POI recommendation.
- **GE** [20]: GE jointly learns the embedding of POIs, regions, time slots and word into a shared low dimensional space by constructing four bipartite graphs.
- **PEU-RNN** [13]: It is a LSTM based model that combines the user and POI embeddings, which are learned from Word2Vec, for modeling the dynamic user preference and successive transition influence.
- **SAE-NAD** [14]: SAE-NAD exploits the self-attentive encoder to differentiate the user preference and the neighbor-aware decoder to incorporate the geographical context information for POI recommendation.

Notice that STGN [28] and ASPPA [27] are not compared in our experiment due to no publicly available source code. However, our PPR consistently outperforms ASPPA and STGN in terms of $Acc@k$ on both Foursquare and Gowalla datasets according to the experimental results reported in [27] (e.g., PPR vs. STGN vs. ASPPA: 0.3008: 0.2: 0.2796 in $Acc@5$, 0.3935: 0.2592: 0.3371 in $Acc@10$ on Foursquare; PPR vs. STGN vs. ASPPA: 0.3835: 0.1947: 0.2363 in $Acc@5$, 0.4905: 0.2367: 0.2947 in $Acc@10$ on Gowalla).

To further validate the effectiveness of each component in our model, we design four variations of PPR:

- **PPR-RL**: This is a simplified version of PPR, which do not use LSTM network for personalized preference modeling. After representation learning on \mathcal{G}_{dense} , we use $Score(v|v_c, u) = \vec{u} \cdot \vec{v} + \vec{v}_c \cdot \vec{v}$ to calculate the recommendation score, where v_c is the current location of the querying user u .
- **PPR-Seq**: This variation do not model the sequential and geographical effect (i.e., ignore POI-POI edges) in graph \mathcal{G}_{dense} , and the other components remain the same.
- **PPR-Den**: This variation directly learns representations for users and POIs on graph \mathcal{G} , which do not densify the graph. And the other components remain the same.
- **PPR-GRU**: In this variation, we use GRU to replace LSTM in user personalized preference modeling, and the other components remain the same.

5.4 Parameter Setting

Following [24, 28, 29], we utilize the first 80% chronological check-ins of each user as the training set, the remaining 20% as the test data.

We use the source code released by their authors for baselines. We set learning rate to 0.0025 in graph embedding, embedding dimension d to 128, the number of negative samples to 5, threshold θ to 24 h, κ to -2 , ε to 0.5 and in/out-degree threshold ρ to 400. Following [5], we uniformly set the next time period as

$\tau = 6$ h for all methods unless stated otherwise, and other parameters of all baselines are tuned to be optimal. In the experiment, we use a two-layer stacked LSTM, the hidden state size is 128. The learning rate of LSTM is set to 0.001 with epoch decay, which makes the learning rate becomes 1/10 of the original value when the number of training rounds reaches 75%.

Table 2. Performance comparison on Foursquare dataset

Methods	<i>Acc@5</i>	<i>Acc@10</i>	<i>Pre@5</i>	<i>Pre@10</i>	<i>Rec@5</i>	<i>Rec@10</i>	<i>NDCG@5</i>	<i>NDCG@10</i>
Rank-GeoFM	0.2456	0.2983	0.0618	0.0413	0.0509	0.0669	0.0683	0.0468
ST-RNN	0.1642	0.2150	0.0167	0.0118	0.1207	0.1790	0.0175	0.0152
GE	0.1357	0.3100	0.0378	0.0342	0.1579	0.1919	0.0431	0.0362
PEU-RNN	0.2021	0.2775	0.0495	0.0276	0.1888	0.2848	0.0494	0.0375
SAD-NAE	0.2429	0.3221	0.0588	0.0442	0.0333	0.0505	0.0672	0.0542
PPR	0.3008	0.3935	0.0698	0.0501	0.2471	0.3387	0.0802	0.0628

Table 3. Performance comparison on Gowalla dataset

Methods	<i>Acc@5</i>	<i>Acc@10</i>	<i>Pre@5</i>	<i>Pre@10</i>	<i>Rec@5</i>	<i>Rec@10</i>	<i>NDCG@5</i>	<i>NDCG@10</i>
Rank-GeoFM	0.2162	0.2643	0.0647	0.0453	0.0887	0.1180	0.0696	0.0499
ST-RNN	0.1865	0.2246	0.0278	0.0217	0.0817	0.1075	0.0606	0.0574
GE	0.1763	0.4060	0.0391	0.0203	0.1363	0.3135	0.0813	0.0157
PEU-RNN	0.3329	0.3766	0.0663	0.0390	0.2504	0.3613	0.0919	0.0627
SAD-NAE	0.3273	0.4300	0.0849	0.0645	0.1102	0.1600	0.0956	0.0777
PPR	0.3835	0.4905	0.0936	0.0687	0.2573	0.3430	0.1055	0.0840

5.5 Performance Comparison

Table 4. Performance comparison on Brightkite dataset

Methods	<i>Acc@5</i>	<i>Acc@10</i>	<i>Pre@5</i>	<i>Pre@10</i>	<i>Rec@5</i>	<i>Rec@10</i>	<i>NDCG@5</i>	<i>NDCG@10</i>
Rank-GeoFM	0.3681	0.4270	0.0968	0.0618	0.2497	0.2983	0.1058	0.0700
ST-RNN	0.2396	0.3540	0.0389	0.0394	0.2279	0.3400	0.1166	0.1074
GE	0.1903	0.4259	0.0869	0.0483	0.1303	0.4119	0.1313	0.1217
PEU-RNN	0.7187	0.7383	0.1437	0.0720	0.6944	0.7204	0.2348	0.1538
SAD-NAE	0.2578	0.3383	0.0645	0.0499	0.0703	0.1047	0.0708	0.0584
PPR	0.8717	0.8966	0.1788	0.0927	0.8485	0.8741	0.2875	0.1889

First, we evaluate the overall performance of our model PPR compared with five baselines on three real-world datasets. We repeat 10 runs for all methods on each dataset and report average *Acc@k*, *Pre@k*, *Rec@k* and *NDCG@k* in Table 2, Table 3 and Table 4, respectively.

From Table 2, we observe that PPR is significantly better than all baselines in terms of four evaluation metrics on Foursquare dataset. Specifically, PPR achieves 0.3008 in $Acc@5$ and 0.3935 in $Acc@10$, improving 22.5% and 22.2% over second-best baseline Rank-GeoFM and SAD-NAE, respectively. Additionally, our PPR slightly outperforms the strong baselines (*e.g.*, SAD-NAE) in $Pre@k$, but it is significantly better than the strong baselines in $Rec@k$.

As depicted in Table 3, our PPR also significantly outperforms all baselines in terms of $Acc@k$, $Pre@k$, $Rec@k$ and $NDCG@k$ on Gowalla dataset. In particular, PPR performs better than the second-best baseline by 14.6% in $Acc@k$ and 9.2% in $NDCG@k$ on average. PPR shows slightly poor performance compared to PEU-RNN in terms of $Rec@10$. This phenomenon can be explained that PEU-RNN uses a distance constraint, which may significantly reduce the potential POIs as k increases.

As we can see in Table 4, PPR consistently significantly outperforms all baselines in terms of all evaluation metrics on Brightkite dataset. PPR achieves the state-of-the-art performance, *e.g.*, 0.8717 in $Acc@5$ and 0.8485 in $Rec@5$. More specifically, our PPR achieves about 21.3%, 24.4%, 22.2% and 22.4% improvement compared to state-of-the-art RNN-based method PEU-RNN in terms of $Acc@5$, $Pre@5$, $Rec@5$ and $NDCG@5$, respectively. Furthermore, all methods achieve better performance on Brightkite than the other datasets. This is because users in Brightkite have more check-in records than users in Foursquare and Gowalla on average, which may enable all methods to model users’ behavior and preference more accurately.

5.6 Ablation Study

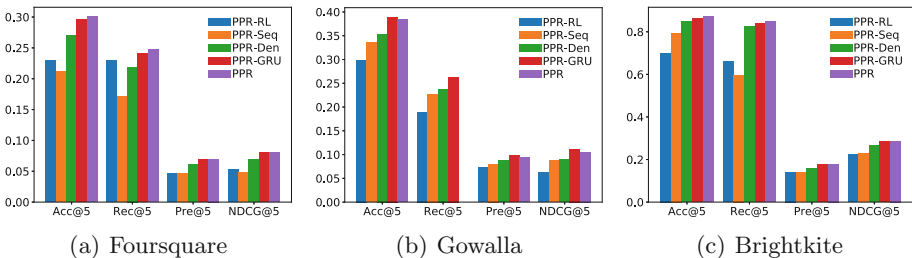


Fig. 2. Performance comparison of variations

To explore the benefits of incorporating the sequential and geographical effect, densifying technique and modeling personalized preference into PPR respectively, we compare our model with four carefully designed variations, *i.e.*, PPR-RL, PPR-Seq, PPR-Den and PPR-GRU. We show the results in terms of $Acc@5$, $Pre@5$, $Rec@5$, and $NDCG@5$ on three datasets in Fig. 2.

Based on the results, we have the following observations: First, PPR achieves the best performance in most cases on three datasets, indicating that PPR

benefits from simultaneously considering the various contextual factors and personalized preference in a joint way. Second, the contributions of different components to recommendation performance are different. Sequential and geographical effect and modeling personalized preference have comparable importance, specifically, the later contributes more on Gowalla, and the former contributes more on Foursquare. And both of them are necessary for improving performance. Furthermore, through the comparison of PPR and PPR-Den, it is obvious that the densifying trick works for alleviating the data sparse issue. Third, our PPR and PPR-GRU exhibit a decent performance compared to other variations, which indicates that sequential pattern and users' dynamic and personalized preference play an important role in location-based recommendation.

5.7 Sensitivity of Hyper-parameters

We now investigate the sensitivity of our model compared against three strong baselines (*i.e.*, Rank-GeoFM, PEU-RNN, and SAE-NAD) with respect to the important parameters, including embedding dimension d , the number of recommended POIs k , and next time period τ . To clearly show the influence of these parameters, we report $Acc@5$ with different parameter settings on Foursquare and Gowalla datasets. Figure 3 and Fig. 4 show the experimental results.

As shown in Figs. 3(a) and 4(a), PPR achieves best performance compared to the three strong baselines with the increasing number of dimension d . Meanwhile, PPR achieves the best result when $d = 128$, and then begins to decline as d further increases. From the results in Figs. 3(b) and 4(b), we can see that the recommendation accuracy of all methods increases as k increases. This is expected, because the more results are recommended, the easier they are to fall into the ground truth. However, we also observe that our PPR exhibits an increasing performance improvement compared to all baselines, as k increases. In Figs. 3(c) and 4(c), as τ increases, our PPR is also consistently better than the strong baselines. More specifically, PPR improves the recommendation accuracy more significantly for near future prediction (*e.g.*, $\tau = 2$ vs. $\tau = 12$), indicating that our PPR can effectively capture users' personalized preferences, especially short-term preferences.

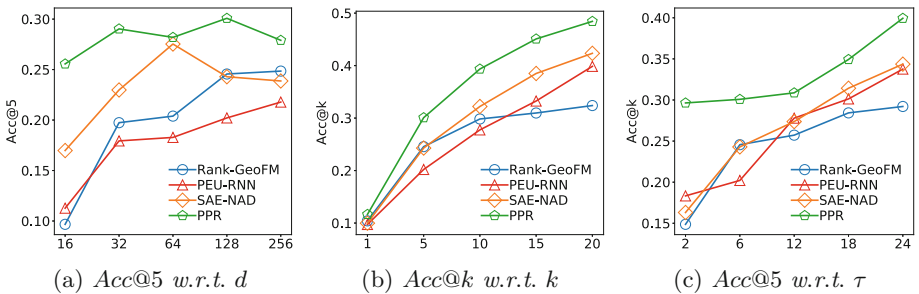


Fig. 3. Parameter sensitivity *w.r.t.* parameter d , k and τ on Foursquare

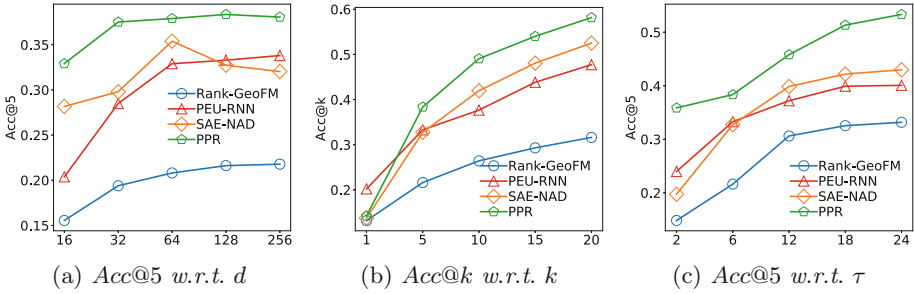


Fig. 4. Parameter sensitivity *w.r.t.* parameter d , k and τ on Gowalla

6 Conclusion

In this work, we propose a novel spatio-temporal representation learning model for personalized POI recommendation. By incorporating the user-POI relation, sequential effect, geographical effect and social ties, we construct a heterogeneous network. Afterwards, we exploit the embedding technique to learn the latent representation of users and POIs. In light of recent success of RNN on sequential prediction problem, we feed the spatio-temporal network with concatenated user and POI embedding sequences for capturing the users' dynamic and personalized preference. The results on three real-world datasets demonstrate the superiority of our proposal over state-of-the-art baselines. Furthermore, we explore the importance of each factor in improving recommendation performance. We observe that sequential effect, geographical effect, and users' dynamic and personalized preference play a vital role in POI recommendation task.

Acknowledgments. This work is partially supported by the National Natural Science Foundation of China under grant Nos. 61773331, U1706218 and 41927805, the National Key Research and Development Program of China under grant No. 2018AAA0100602, and the Natural Science Foundation of Shandong Province under grant No. ZR2020QF030.

References

1. Chang, B., Park, Y., Park, D., Kim, S., Kang, J.: Content-aware hierarchical point-of-interest embedding model for successive poi recommendation. In: IJCAI, pp. 3301–3307 (2018)
2. Chen, M., Liu, Y., Yu, X.: NLPMM: a next location predictor with Markov modeling. In: Tseng, V.S., Ho, T.B., Zhou, Z.H., Chen, A.L.P., Kao, H.Y. (eds.) PAKDD'2014. LNCS, vol. 8444, pp. 186–197. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-06605-9_16
3. Cheng, C., Yang, H., Lyu, M.R., King, I.: Where you like to go next: successive point-of-interest recommendation. In: IJCAI (2013)

4. Feng, S., Cong, G., An, B., Chee, Y.M.: POI2VEC: geographical latent representation for predicting future visitors. In: AAAI, pp. 102–108 (2017)
5. Feng, S., Tran, L.V., Cong, G., Chen, L., Li, J., Li, F.: HME: a hyperbolic metric embedding approach for next-poi recommendation. In: SIGIR, pp. 1429–1438 (2020)
6. Levandoski, J.J., Sarwat, M., Eldawy, A., Mokbel, M.F.: LARS: a location-aware recommender system. In: ICDE, pp. 450–461. IEEE (2012)
7. Li, K., Lu, G., Luo, G., Cai, Z.: Seed-free graph de-anonymization with adversarial learning. In: CIKM, pp. 745–754 (2020)
8. Li, X., Han, D., He, J., Liao, L., Wang, M.: Next and next new POI recommendation via latent behavior pattern inference. *ACM Trans. Inf. Syst. (TOIS)* **37**(4), 1–28 (2019)
9. Li, X., Cong, G., Li, X.L., Pham, T.A.N., Krishnaswamy, S.: Rank-GeoFM: a ranking based geographical factorization method for point of interest recommendation. In: SIGIR, pp. 433–442 (2015)
10. Liu, Q., Wu, S., Wang, L., Tan, T.: Predicting the next location: a recurrent model with spatial and temporal contexts. In: AAAI (2016)
11. Liu, Z., Huang, C., Yu, Y., Fan, B., Dong, J.: Fast attributed multiplex heterogeneous network embedding. In: CIKM, pp. 995–1004 (2020)
12. Liu, Z., Huang, C., Yu, Y., Song, P., Fan, B., Dong, J.: Dynamic representation learning for large-scale attributed networks. In: CIKM, pp. 1005–1014 (2020)
13. Lu, Y.-S., Shih, W.-Y., Gau, H.-Y., Chung, K.-C., Huang, J.-L.: On successive point-of-interest recommendation. *World Wide Web* **22**(3), 1151–1173 (2018). <https://doi.org/10.1007/s11280-018-0599-5>
14. Ma, C., Zhang, Y., Wang, Q., Liu, X.: Point-of-interest recommendation: exploiting self-attentive autoencoders with neighbor-aware influence. In: CIKM, pp. 697–706 (2018)
15. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint [arXiv:1301.3781](https://arxiv.org/abs/1301.3781) (2013)
16. Perozzi, B., Al-Rfou, R., Skiena, S.: DeepWalk: Online learning of social representations. In: KDD, pp. 701–710 (2014)
17. Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q.: LINE: large-scale information network embedding. In: WWW, pp. 1067–1077 (2015)
18. Wang, Q., Yin, H., Chen, T., Huang, Z., Wang, H., Zhao, Y., Viet Hung, N.Q.: Next point-of-interest recommendation on resource-constrained mobile devices. In: WWW, pp. 906–916 (2020)
19. Wu, Y., Li, K., Zhao, G., Xueming, Q.: Personalized long-and short-term preference learning for next POI recommendation. *IEEE Trans. Knowl. Data Eng.* (2020)
20. Xie, M., Yin, H., Wang, H., Xu, F., Chen, W., Wang, S.: Learning graph-based poi embedding for location-based recommendation. In: CIKM, pp. 15–24 (2016)
21. Ye, J., Zhu, Z., Cheng, H.: What’s your next move: user activity prediction in location-based social networks. In: SDM, pp. 171–179. SIAM (2013)
22. Ye, M., Yin, P., Lee, W.C.: Location recommendation for location-based social networks. In: SIGSPATIAL, pp. 458–461 (2010)
23. Yuan, Q., Cong, G., Ma, Z., Sun, A., Thalmann, N.M.: Time-aware point-of-interest recommendation. In: SIGIR, pp. 363–372 (2013)
24. Zhang, J.D., Chow, C.Y., Li, Y.: LORE: exploiting sequential influence for location recommendations. In: SIGSPATIAL, pp. 103–112 (2014)
25. Zhang, Z., Li, C., Wu, Z., Sun, A., Ye, D., Luo, X.: Next: a neural network framework for next POI recommendation. *Front. Comput. Sci.* **14**(2), 314–333 (2020)

26. Zhang, Z., Liu, Y., Zhang, Z., Shen, B.: Fused matrix factorization with multi-tag, social and geographical influences for POI recommendation. *World Wide Web* **22**(3), 1135–1150 (2019)
27. Zhao, K., Zhang, Y., Yin, H., Wang, J., Zheng, K., Zhou, X., Xing, C.: Discovering subsequence patterns for next POI recommendation. In: *IJCAI*, pp. 3216–3222 (2020)
28. Zhao, P., et al.: Where to go next: a spatio-temporal gated network for next POI recommendation. *IEEE Trans. Knowl. Data Eng.* (2020)
29. Zhao, S., Zhao, T., King, I., Lyu, M.R.: Geo-Teaser: geo-temporal sequential embedding rank for point-of-interest recommendation. In: *WWW*, pp. 153–162 (2017)
30. Zhu, Y., et al.: What to do next: modeling user behaviors by time-LSTM. In: *IJCAI*, pp. 3602–3608 (2017)