Graph-Enhanced Spatio-Temporal Interval Aware Network for Next POI Recommendation in Mobile Environment

Xu Zhang[1,2,3], Deao Liu[1,3], Liang Yan[1,2], Zhiqing Zhang[2], Yan Li[4]*

[1] *Department of Computer Science and Technology, Chongqing University of Posts and Telecommunications, China*
[2] *Key Laboratory of Monitoring, Evaluation and Early Warning of Territorial Spatial Planning Implementation, Ministry of Natural Resources, China*
[3] *Key Laboratory of Tourism Multisource Data Perception and Decision, Ministry of Culture and Tourism, China*
[4] *Department of Electrical and Computer Engineering, Inha University, South Korea*
zhangx@cqupt.edu.cn, S210231115@stu.cqupt.edu.cn, d210201030@stu.cqupt.edu.cn, 425707067@qq.com, leeyeon@inha.ac.kr

Abstract

With the rapid spread of mobile device, technologies in mobile cloud increased quick and introduce huge volume of mobile data and computation. Human movement between POIs are recorded in mobile cloud, which indicate personalized behaviors. Most POI recommendation method in mobile cloud proposed to utilize recurrent neural networks and self-attention mechanism to discover users' potential movement behaviors. In this paper, we propose a graph-enhanced spatio-temporal interval aware network (GESTIAN) to recommend the next POI. In GESTIAN, we propose a graph-based general pattern learning module to learn common behavior patterns based on a global trajectory flow graph to address the challenges caused by cold start. Furthermore, we propose a heterogeneous network with spatio-temporal interval aware with self-attention and gate recurrent unit to extract users' long-term and short-term spatio-temporal dependencies, respectively. In addition, we leverage the scale between positive and negative samples by randomly sampling negative samples. We conduct extensive experiments based on two real-world check-in datasets. The experimental evaluations demonstrate that the proposed GESTIAN outperforms most challenging baselines by approximately 2%-10%, and achieves better performance over few-check-in history users.

Keywords: Mobile cloud, POI Recommendation, Spatio-temporal interval aware, Global trajectory flow graph

1 Introduction

The Point-of-Interest (POI) recommendation technology is a personalized recommendation task, which discovery potential patterns based on historical check-ins and other multimodal information to predict the next set of POIs suitable for certain user [1]. In recent years, the increasing amount of mobile devices bring a significant development in mobile cloud technologies including mobile storage, computation and LBS-based AI services. The mobile cloud service providers such as Foursquare [2] and Gowalla [3] allow users to leverage the advantages of mobile cloud technology to share multiple forms of check-in data, including location (Lat, Lng), timestamp and rich semantic information. The massive check-in data generated in mobile cloud preserve rich spatial and temporal contextual information is of great significance for understanding user travel preferences and predicting the next location [4]. Meanwhile, the POI recommendation tasks help service providers analyze user behavior patterns and optimize the personalized recommendation systems [5].

Early POI recommendation researches mainly focus pattern-based methods to extract predefined movement patterns from users' historical trajectories and predict future POI sequences [6]. The Markov chains [7] and matrix factorization [8] methods are widely studied with the idea of modeling historical data from the perspective of sequence prediction and achieve excellent results. With the development of deep learning, various studies have attempted to discover human mobility with convolutional neural networks (CNN) [9-10], long short-term memory (LSTM) [11], gate recurrent unit (GRU) [12] and self-attention [13-14] mechanism, which significantly improve the performance of the POI recommendation tasks. The deep learning based methods achieve remarkable results in academia and industries and has become a strong and promising tool for POI recommendation. Recent work propose to incorporate the influence of spatio-temporal correlations into networks [15-16], which further improves the accuracy of predictions and become increasingly attractive.

However, it is highly challenging due to the complexity of the urban geospatial structure and the highly nonlinear spatial and temporal dependence on human mobility and heterogeneity in mobile cloud platform [37]. 1) Cold start, which is the most common and unavoidable problem in the field of recommendation systems. Generally, existing POI recommendation models need to learn the potential users' movement patterns based on certain historical trajectory sequences. However, user may have large historical

*Corresponding Author: Yan Li; E-mail: leeyeon@inha.ac.kr
DOI: https://doi.org/10.70003/160792642024072504012

preferences in mobile cloud provider A and few in B, therefore cannot provide effective recommendations for new registers or users with few historical trajectories. 2) The difficulty of capturing the spatio-temporal relationship between non-adjacent check-ins. Most models only focus on the transition patterns between adjacent check-ins in the sequence while ignoring the spatio-temporal dependencies between non-adjacent check-ins. 3) The user's travel preferences are complex with real-time and dynamic features, even distinguished behavior across different mobile cloud platforms. The POI that users will access may dynamically change over time, location, and user characteristics. As shown in Figure 1, the travel preferences of different users have entirely different characteristics. Furthermore, even after accessing the same POI, the destination of the next POI may also different. Therefore, modeling user travel preferences from multiple scales and extracting personalized travel preferences is still challenging.

Figure 1. Example of trajectories and check-in sequence

To address the abovementioned issues, we propose a ***G***raph-***E***nhanced ***S***patio-***T***emporal ***I***nterval ***A***ware ***N***etwork GESTIAN for the POI recommendation task. Specifically, we design a novel user general behavior pattern extraction module to solve the cold start problem. We first construct a graph based on the trajectory from the same batch of users, then use message passing mechanism to aggregate and update node features, and finally obtain a universal behavior pattern representation of the current user set through global average pooling at the graph level. In addition, we also design two spatio-temporal interval aware modules to enhance the learning effect of user-personalized preferences. More specifically, we divide the check-in sequences into historical and current trajectories and capture user travel preferences at different time scales through self-attention modules and GRU modules, with the spatio-temporal perception structures respectively. Finally, the extracted general patterns are fused with the personalized preferences to predict the next POI that matches user characteristics while minimizing the impact of the cold start.

In summary, our contributions can be listed as follows:

- We propose a sequence prediction model in mobile cloud platform, GESTIAN integrating spatio-temporal attention, spatio-temporal recurrent networks and graph convolution to explore user-personalized travel preferences and general behavior patterns. The transfer rules between non-adjacent check-in are well captured and provide recommendation services in the case of insufficient historical data.
- We propose to transform the user check-in sequence gathered through mobile network into a trajectory flow graph and learn global universal features on it. The insufficient features of users with few check-in records are well processed, thereby improving the model performance in sparse data situations.
- We propose a heterogeneous spatio-temporal aware module to extract the long-term stable preferences and short-term dynamic preferences, which capture the travel patterns of users in different time periods and provide multi-scale feature basis for personalized destination prediction.
- We conduct extensive experiments using two real-world datasets. Experimental evaluations demonstrate that the proposed model significantly outperforms state-of-the-art methods.

2 Related Work

2.1 Next POI Recommendation

The early work of POI recommendation mainly focuses on feature engineering and traditional machine learning methods, most of which come from the exploration of related fields of recommendation systems. Methods such as Markov chain [17] and matrix factorization [8, 18] have been successfully applied to POI recommendation after success in other fields. After that, other traditional methods such as support vector machine [19], collaborative filtering [20], Gaussian modeling [21], and Bayesian personalized ranking [22] methods are also used for personalized POI recommendation. However, a significant drawback of all these methods is the high dependence on expertise in related fields. Moreover, with the increasing of the multi-modal data, such as images, text and remote sensing, the prediction tasks become increasingly challenging due to great effort of manually processing multi-modal data [1].

Recenty, the deep learning methods have gained considerable research interest for the POI recommendation. The recurrent neural network (RNN) is the first deep learning model that achieve excellent performance in the field of POI recommendation. The RNN can memorize the previously input information and apply it to the current output results, which overcomes the shortcomings of the hidden Markov model that only relying on the previous hidden state [6]. However, the vanishing gradient problem of RNN makes it difficult for the model to capture long-term dependencies. To overcome this shortcoming, researchers proposed LSTM and GRU. The GRU additionally solves the problem that most RNN models rely on the last hidden layer and limit user information learning. In recent years, the self-attention mechanism has gradually attract researchers interest, which can imitate the human attention-focusing mechanism to give greater attention [23] to critical parts and help to decompose complex input into simple expressions. The graph neural network (GNN) [24] is introduced to process the feature transfer in the POI recommendation tasks.

2.2 Potential Features in Next POI Selection

Previous studies have shown that incorporating all available influencing features into predictions does not always have a positive contribution to the results [25]. Therefore, selecting and utilizing relevant influencing features to improve prediction accuracy is essential and necessary. Most researchers generally believe that spatial location, time periodicity, semantic information and social relation can significantly improve the performance of POI recommendations [26].

Spatial Location: Integrating spatial information into the prediction process is one of the most significant differences compared with general recommendation tasks. The first law of geography states that everything is related to other things, but neighboring things are more related than distant things. Users are usually more inclined to choose points of interest that are close to their own location for access [27], so it is necessary to discovery contextual information with the consideration of current location.

Time Periodicity: Users' access to POI has significant periodicity, and their travel preferences may change over time. The periodicity of temporal features can reflect users' long-term stable travel intentions, while the real-time feature can clearly depict users' current preferences. Therefore, considering both periodicity and real-time features in POI recommendations can predict users' interest preferences more accurately in the next moment [28].

Semantic Information: The POI semantic information mainly includes category description information and user comments information [29]. The POI category information in location-based social networks can be depicted as a hierarchical category tree, and POIs with consistent or similar categories usually have similar attributes and features. The frequency of users' access to different categories of POIs can reflect their preferences, and the continuous access between different POIs in the historical trajectory can also reflect the way users' preferences transition between POIs. Considering POI semantic information in the prediction model are generally adopted recently.

Social Relation: Social relationships may influence users' decision when choosing the next POI to access [30]. For example, when choosing between multiple similar POIs, users usually tend to choose POIs recommended by acquaintances or POIs with higher social popularity [27]. This indicates that extracting features from social networks or general social travel preferences is critical important to improve the poor performance caused by the cold start.

3 Preliminaries

This section briefly provides some basic concepts, symbols, and formulas used in POI recommendation. We assume that $U = \{u_1, u_2, \ldots, u_M\}$ denotes the set of users, $P = \{p_1, p_2, \ldots, p_N\}$ represents the set of POIs, $T = \{t_1, t_2, \ldots, t_T\}$ denotes the set of time stamps.

3.1 Basic Definition

Definition 1 (Check-in): A check-in consists of a tuple $c = (u, p, s, t)$, indicating the user u visited the location p with category s at time t.

Definition 2 (Check-in Sequence): A user's check-in sequence consists of $C_u = \{c_u^1, c_u^2, \ldots, c_u^n\}$ where c_u^i denotes the i-th check-in tuple of user u. The check-in sequences for all users consist of $C_U = \{C_{U1}, C_{U2}, \ldots, C_{UM}\}$.

Definition 3 (Spatio-Temporal Interval Distance): A spatio-temporal interval distance Δ_{ij} comprises the time difference Δ_{tij} and spatial difference Δ_{sij} between adjacent i-th and j-th check-in.

Definition 4 (Spatio-Temporal Interval Matrix): A spatio-temporal interval matrix Δ comprises the time differences Δt and spatial differences Δs between any two check-ins in a check-in sequence C_u.

Definition 5 (Trajectory Flow Graph): A trajectory flow graph $G = (V, E, R)$ is constructed by given user trajectory sequences C_U. $V = \{v_1, v_2, \ldots, v_m\}$ denotes the set of points in the graph, where v_i belongs to the POI set P. $R = \{r_1, r_2, \ldots, r_m\}$ denotes the set of node attributes, where r_i denotes the attribute of node v_i. It is mainly composed of the embedded representation of the location and category information of the POI. $E = \{e_1, e_2, \ldots, e_n\}$ denotes the set of edges in the graph, and if there is an edge between v_i and v_j, it means a user is traveling from v_i to v_j.

3.2 POI Recommendation Task

A POI recommendation task involves learning the user's historical check-in sequence C_u, constructing a spatio-temporal interval matrix, modeling and analyzing the user's travel preferences, and ultimately predict the user's next possible POI from the POI candidate set.

4 Methods

4.1 Framework Overview

The architecture of the proposed graph-enhanced spatio-temporal interval aware network (GESTIAN) is shown in Figure 2. It consists of five main components: 1) an **embedding module** that transforms the check-in tuple into an embedding in the latent vector representation; 2) a **general pattern learning module** that constructs a user trajectory flow graph based on the check-in sequences of users C_U and performs graph convolutional operations to learn general behavioral patterns of users; 3) a **spatio-temporal aware self-attention module** that learns the user's long-term stable travel preferences; 4) a **spatio-temporal aware**

GRU module that learns the user's short-term dynamic travel preferences; 5) a **candidate POI matching module** that integrates universal behavioral patterns with the user's personalized travel preferences and calculates the similarity with a set of candidate POIs.

During the training process, GESTIAN considers the trajectory excluding the last $c_u^1 \rightarrow c_u^2 \rightarrow \ldots \rightarrow c_u^{n-1}$ as the historical trajectory and truncates it at a specific ratio to get $c_u^m \rightarrow c_u^{m+1} \rightarrow \ldots \rightarrow c_u^{n-1}$ as the current trajectory. GESTIAN takes the historical and current trajectories as inputs and finally obtains the probability of each candidate POI being the user's next POI.

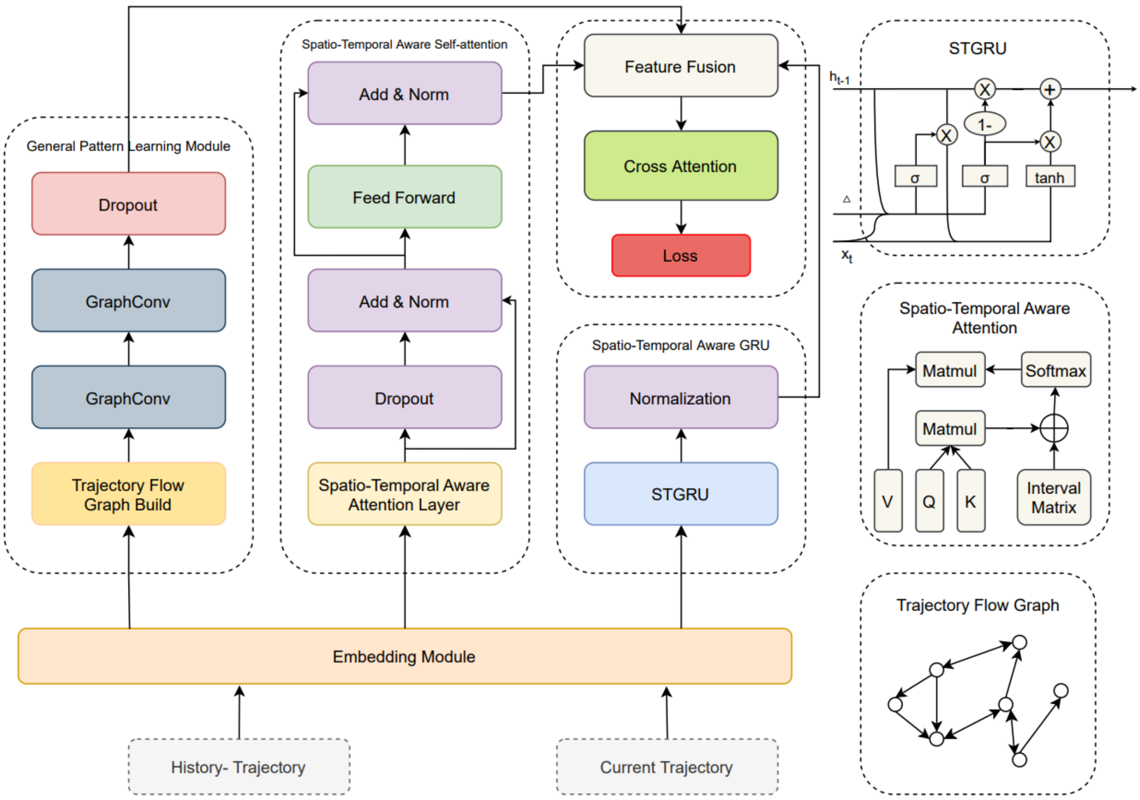


Figure 2. The architecture of the proposed GESTIAN model

4.2 Embedding Module

The primary function of the embedding module is to transform the discrete attributes in the check-in data and spatio-temporal interval matrix into continuous vector representations. It mainly consists of the trajectory embedding layer and the spatio-temporal interval embedding layer.

4.2.1 Trajectory Embedding Layer

For a given check-in tuple $c = (u, p, s, t)$, the trajectory embedding layer transforms the user, location, POI category and timestamp into latent representation, denoted as $e^u \in \mathbf{R}^{du}$, $e^p \in \mathbf{R}^{dp}$, $e^s \in \mathbf{R}^{ds}$, and $e^t \in \mathbf{R}^{dt}$ respectively. Then, the embedding representations of the four attributes are merged into a single representation $e \in \mathbf{R}^d$. For a given check-in sequence $C_u = \{c_u^1, c_u^2, \ldots, c_u^n\}$, the trajectory embedding layer will produce the embedding sequence $E_u = \{e_u^1, e_u^2, \ldots, e_u^n\} \in \mathbf{R}^{n \times d}$.

4.2.2 Spatio-Temporal Interval Embedding Layer

Users' choices of POIs are often influenced by spatial distance and temporal periodicity. It is essential to model the spatio-temporal information between POIs. The spatio-temporal interval aware layer constructs the spatio-temporal interval matrix $\Delta \in \mathbf{R}^{n \times n \times 2}$ based on the historical trajectory, which is represented as follows:

$$\Delta = \begin{bmatrix} \Delta_{11} & \Delta_{12} & \cdots & \Delta_{1n} \\ \Delta_{21} & \Delta_{22} & \cdots & \Delta_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta_{n1} & \Delta_{n2} & \cdots & \Delta_{nn} \end{bmatrix}, \tag{1}$$

where Δ_{ij} is composed of $[\Delta_{ij}^t, \Delta_{ij}^s]$.

To reduce the computational complexity of embedding, the values of the spatio-temporal interval matrix are normalized.

$$\Delta_{ij}^{t} = \frac{|t_i - t_j|}{t_{\max} - t_{\min}}. \quad (2)$$

$$\Delta_{ij}^{s} = \frac{|s_i - s_j|}{s_{\max} - s_{\min}}. \quad (3)$$

In equation (2), Δ_{ij}^{t} denotes the time difference between the i -th and j -th check-in, $t_{\max}$ and $t_{\min}$ denote the maximum and minimum time values in a batch of user data, respectively. The symbol definitions for equation (3) can be inferred from equation (2).

Afterward, Δ is transformed into an embedding matrix $M \in \mathbf{R}^{n \times n \times d}$:

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix}. \quad (4)$$

4.3 General Pattern Learning Module

The general pattern learning module consists of two components: trajectory flow graph and graph convolution module.

4.3.1 Trajectory Flow Graph

To generate the trajectory flow graph, all POIs that appear in the historical trajectories of all users in a batch are added as nodes to the trajectory flow graph G. The location embedding representation and category embedding representation of the POI are combined as the initial features of the corresponding graph nodes. For each user's check-in sequence, a transition between v_i and v_j is considered a directed edge in graph G.

4.3.2 Graph Convolution Layer

Given the constructed trajectory flow graph G, the next step is to aggregate the information of neighboring nodes through graph convolutional networks and use it to update the feature of the current node. The specific calculation formula for GCN is as follows:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(l)}W^{(l)}\right), \quad (5)$$

where $H^{(l)}$ denotes the node features of the l -th layer, $W^{(l)}$ denotes the weight matrix of the l -th layer, $\tilde{A} = A + I$ denotes the adjacency matrix A with self-loops added, $\tilde{D}$ denotes the degree matrix of $\tilde{A}$, and σ denotes the activation function.

Each execution of the above process allows each node in the graph to aggregate information from its neighboring nodes once. By stacking multiple GCN layers, the model's performance can be enhanced. And we can get the graph-level general behavioral pattern by averaging node features across the node dimension:

$$r = \frac{1}{N}\sum_{n=1}^{N} x_n, \quad (6)$$

where x_n denotes the node feature, N denotes the number of nodes, and r denotes the general behavioral pattern.

4.4 Spatio-Temporal Aware Self-attention Module

Due to the excellent performance of self-attention in capturing long-term dependencies, we consider using self-attention to capture long-term stable travel preferences hidden in user historical trajectories. However, self-attention only learns the order between check-in sequences and cannot perceive specific spatio-temporal interval information between check-ins. Therefore, we propose a spatio-temporal aware self-attention module that extracts information from the user's historical trajectory and the corresponding spatio-temporal interval matrix.

Firstly, the module will map the embedding representation of the user's historical trajectory E_u into query and key-value pairs:

$$Q = E_u W_Q, K = E_u W_K, V = E_u W_V, \quad (7)$$

where Q, K, $V \in \mathbf{R}^{n \times d}$ and W_Q, W_K, $W_V \in \mathbf{R}^{d \times d}$. Then, a new sequence $L \in \mathbf{R}^{n \times d}$ containing the user's long-term travel preferences is obtained by calculating according to the following equation:

$$\begin{aligned} &Attention\ (Q, K, V, M) = \\ &soft\max\ (\frac{QK^T}{\sqrt{d}} + M)V, \end{aligned} \quad (8)$$

where $QK^T / \sqrt{d} \in \mathbf{R}^{n \times n}$ denotes the attention score matrix, and M denotes the embedded spatio-temporal interval matrix.

Then the output of the self-attention layer is processed using residual connections and layer norm to accelerate convergence. In order to introduce non-linearity into the network, we finally use a feed-forward neural network layer with activation function ReLU to process the results:

$$L = LayerNorm(L + Attention(L)). \quad (9)$$

$$L = \max(0, LW_1 + b_1). \quad (10)$$

4.5 Spatio-Temporal Aware GRU Module

GRU uses gate mechanisms to control the flow of information. However, it also faces the problem of being unable to perceive specific spatio-temporal intervals mentioned earlier. Therefore, we propose a STGRU to capture the spatio-temporal correlations in the user's current trajectory. We introduce the spatio-temporal interval information into the update gate and reset gate of GRU to control the influence brought by the hidden features of historical time steps in the calculation. This way, we can better capture the user's short-term dynamic travel preferences.

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + O_z m_{ij}). \quad \textbf{(11)}$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + O_r m_{ij}). \quad \textbf{(12)}$$

Here, $z_t \in \mathbf{R}^d$ denotes the update gate, $r_t \in \mathbf{R}^d$ denotes the reset gate, $x_t \in \mathbf{R}^d$ denotes the embedded representation of the user's check-in at time t, $h_{t-1} \in \mathbf{R}^d$ denotes the previous hidden state, and $m_{ij} \in \mathbf{R}^d$ denotes the spatio-temporal interval when the user transitions from the i -th check-in to the j -th check-in.

Next, we use the reset gate and update gate that incorporate spatio-temporal interval information to calculate the candidate hidden state and hidden state at the current time step, respectively, and form a new sequence $S \in \mathbf{R}^{n \times d}$ that contains the user's short-term travel preferences:

$$\tilde{h}_t = \tanh(W x_t + U(r_t \circ h_{t-1})), \quad \textbf{(13)}$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \tilde{h}_t, \quad \textbf{(14)}$$

where $\tilde{h}_t \in \mathbf{R}^d$ denotes the current candidate hidden state, $h_t \in \mathbf{R}^d$ denotes the current hidden state.

4.6 Candidate POI Matching Module

In order to obtain the representation of user travel characteristics, this module will integrate the learned general behavior patterns, personalized long-term stability preferences, and personalized short-term change preferences through the following equation:

$$F = \max(0, \text{concat}[L + S), r]W + b), \quad \textbf{(15)}$$

where $F \in \mathbf{R}^{n \times d}$.

Then, the match score is calculated using the user's travel characteristic F and the embedded candidate POI matrix D. To achieve better performance; we introduced the spatio-temporal interval matrix $N \in \mathbf{R}^{N \times n \times d}$ between the candidate POI and the user trajectory in the calculation process, which is calculated in the same way as the spatio-temporal interval matrix mentioned above:

$$Q = DW_Q, K = FW_K, V = FW_V. \quad \textbf{(16)}$$

$$Match(Q, K, V, N) = soft\max(\frac{QK^T}{\sqrt{d}} + N)V. \quad \textbf{(17)}$$

$$Output = \max(0, Match(Q, K, V, N)W_m + b_m). \quad \textbf{(18)}$$

Here, the role of the feed-forward neural network is to compress the last dimension of the feature matrix and obtain a feature vector that represents the probability of visiting each candidate POI.

We choose cross-entropy as the loss function of the model and perform random sampling on negative samples during each calculation to address the issue of imbalanced positive and negative sample quantities. The specific calculation formula of cross-entropy is as follows:

$$Loss = -\sum_{S_u \in S} \sum_{i=1}^{n} \log \sigma(y_i, o_i) + \sum_{S_u \in S} \sum_{i=1}^{n} \sum_{j=1}^{N} \log(1 - \sigma(y_i, j)), \quad \textbf{(19)}$$

where S_u denotes the check-in sequence of user u, o_i, denotes the ground truth at step i, y_i denotes the prediction result the model gives, and N denotes the negative samples randomly sampled.

5 Experiments

5.1 Datasets

We validated the proposed model GESTIAN on two real-world datasets: NYC and TKY. Both datasets include user ID, POI ID, POI category ID, check-in time, longitude, and latitude, as shown in Table 1. To make the datasets more representative of real-world scenarios, we referred the previous works to filter out all POIs with visit frequencies less than 10 in NYC and TKY and their corresponding check-in records. At the same time, to verify the model's performance during cold start, we retained all users in the original datasets, including those who have few check-in numbers and are filtered out according to pre-processing rules of previous works.

Table 1. Basic dataset statistics

	NYC	TKY
#user	1083	2293
#POIs	5135	7873
#categories	321	292
#check-ins	147938	447570
#min longitude	-74.270751	139.47177
#max longitude	-73.692021	139.906286
#min latitude	40.557295	35.513059
#max latitude	40.987646	35.866030

5.2 Baseline Models

We compare our model with the following baselines:

- **DeepMove** [12]: A model for mobility prediction from long-term sparse trajectories using a recurrent neural network based on attention mechanism.
- **ARNN** [31]: A model that utilizes the attention mechanism to capture neighbor transfer rules on a knowledge graph.
- **TiSASRec** [32]: A model incorporating absolute position encoding and relative time intervals in the attention mechanism for sequence prediction.

- **GeoSAN** [36]: A model for location recommendation using a self-attention network-based geographic perceptual sequential recommendation algorithm.
- **STAN** [14]: A model that stacks two-layer attention mechanisms and introduces direct spatio-temporal differences to enhance performance.
- **CyGNet** [33]: A new temporal knowledge graph representation learning model based on time aware replication generation mechanism.
- **STGCAN** [16]: A model for modeling multiple relationships through multi-graph convolutional networks also involves coupling-based RNNs to improve performance.
- **CARAN** [34]: A model that combines weather conditions with spatio-temporal context and uses attention mechanisms to focus on the nearest visited location.
- **MSTHN** [35]: A spatio-temporal enhancement network model based on user POI interaction that aggregates and propagates spatio-temporal correlations asymmetrically.

Table 2. Recommendation performance comparison with baselines

		NYC			TKY	
	Recall@1	Recall@5	Recall@10	Recall@1	Recall@5	Recall@10
DeepMove	0.134	0.327	0.401	0.124	0.268	0.351
ARNN	-	0.197	0.348	-	0.185	0.270
TiSASRec	0.119	0.366	0.502	0.098	0.227	0.369
GeoSAN	0.167	0.401	0.527	0.125	0.296	0.374
STAN	0.181	0.467	0.596	0.133	0.346	0.426
CyGNet	0.207	0.413	0.586	0.109	0.294	0.406
STGCAN	**0.257**	0.544	0.629	**0.171**	0.357	0.457
CARAN	-	0.534	0.649	-	0.391	0.461
MSTHN	-	0.408	0.440	-	0.338	0.393
GESTIAN	0.219	**0.560**	**0.687**	0.148	0.392	0.507

Table 3. Ablation study

		NYC			TKY	
	Recall@1	Recall@5	Recall@10	Recall@1	Recall@5	Recall@10
GESTIAN	**0.219**	**0.560**	**0.687**	**0.148**	**0.392**	**0.507**
-GPL	0.189	0.524	0.670	0.132	0.364	0.484
-STSA	0.196	0.512	0.615	0.149	0.378	0.492
-STGRU	0.196	0.536	0.678	0.140	0.364	0.476

5.3 Metrics

To verify the effectiveness of our model, we selected the top-k recall rate (Recall@1, Recall@5, Recall@10) as the evaluation metric. Recall rate counts the rate of true positive samples in all positive samples. The larger the value of the evaluation indicator, the better the performance of the model.

5.4 Settings

The specific hyper-parameter settings of the model are shown as follows. For the NYC and TKY datasets, we set the embedding dimension of user ID to 10, POI ID to 60, POI category ID to 20, timestamp to 10, and the final hidden layer dimension to 100. We used the Adam optimizer to train the model and set the learning rate to 0.0003 and dropout to 0.5. To balance positive and negative samples, 10 negative samples were randomly selected each time for NYC, while for TKY, 20 negative samples were randomly selected each time. We trained GESTIAN for 100 epochs on both datasets. Our model is implemented based on Pytorch 1.9.1.

5.5 Results

The experimental results of the proposed GESTIAN and other baselines on two real-world datasets are shown in Table 2. The performance of the model on the NYC dataset is generally better than that on the TKY dataset due to the smaller data volume in the NYC dataset and faster learning and convergence of the model. It can be seen that GESTIAN outperforms existing baseline models in most evaluation indicators. For example, in the TKY dataset, we achieved a top 10 recall rate of 50.7%. At the same time, the optimal result for this metric in the baseline model was 46.1%, resulting in a performance improvement of approximately 10%. We also achieved approximately 3% and 6% improvements on the NYC dataset in Recall@5 and Recall@10, respectively. It is worth noting that we still have a higher recall rate than the baseline model without filtering out users with very few check-in records, which indicates that GESTIAN has overcome the cold start to some extent.

Meanwhile, we also found that we still have a certain gap compared to STGCAN in the top 1 recall rate, and for the top k recommendation problem, as the k increases, the superiority of our model's performance continues to increase. After analyzing the model, we believe that the main reason for this result may be the general pattern learning (GPL) module we proposed. GPL learns the general transfer rules on the graph by constructing a trajectory flow graph, which may alleviate cold start while also weakening the role of user-personalized behavior patterns in the prediction process. This compromise is acceptable in real-world recommendation scenarios, we often do not only focus on the top 1 recommended item. Compared to the accuracy of recommendations, the usability of recommendations is more easily perceived by users. As shown in Table 2, we also found that the performance of baselines such as STAN and STGCAN with spatio-temporal information is significantly better than traditional deep learning models such as DeepMove and ARNN. This further validates the main significance of spatio-temporal relationship extraction for POI Recommendation tasks.

5.6 Ablation Study

We conducted a set of ablation experiments to evaluate the impact of each module we proposed on the performance of the model on two datasets. Specifically, we achieve experimental results by removing different modules and retaining other experimental settings. The specific modules are as follows:

- General Pattern Learning Module (GPL): This indicates using the global average pooling representation of the graph as a general pattern.
- Spatio-Temporal Self-Attention Module (STSA): This indicates the use of self-attention mining for spatio-temporal dependencies as long-term travel preferences.
- Spatio-Temporal GRU Module (STGRU): This indicates the use of gate recurrent unit to mine spatio-temporal dependencies as short-term travel preferences.

After removing the above modules separately, we compared the experimental results with the complete model. The experimental results are shown in Table 3. We found that the complete model achieved optimal performance on all evaluation indicators. For existing components, experimental results show that the self-attention module based on spatio-temporal aware plays a greater role than other components. On the NYC dataset, after removing the STSA module, the evaluation metric Recall@10 from 68.7% to 61.5%, decreasing 10.5%. This indicates that self-attention based on spatio-temporal perception is better at extracting useful features from historical user trajectories. And other components also have positive implications for improving model performance.

6 Conclusion

In this paper, we propose a graph-enhanced spatio-temporal interval aware network (GESTIAN) for POI recommendation in mobile cloud. We proposed a spatio-temporal perception architecture that captures users' travel preferences at different time periods for the transfer patterns between non-adjacent check-ins. At the same time, in order to alleviate the poor effect caused by cold start, the general transfer law is aggregated on the trajectory flow diagram. In two real-world datasets, we demonstrated that the proposed model outperformed the state-of-the-art methods in most evaluation metrics and analyzed the impact of each components. In the future, we will further consider the impact of social networks on general behavior patterns and establish a trajectory flow diagram cross mobile cloud platform.

Acknowledgements

This research is supported in part by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-M202400603), and in part by the Key Laboratory of Monitoring, Evaluation and Early Warning of Territorial Spatial Planning Implementation, Ministry of Natural Resources, China (No. LMEE-KF2023003), and in part by the Project of Key Laboratory of Tourism Multisource Data Perception and Decision, Ministry of Culture and Tourism, China (No. H2023009). This work was supported by INHA UNIVERSITY Research Grant.

References

[1] M. A. Islam, M. M. Mohammad, S. S. S. Das, M. E. Ali, A survey on deep learning based point-of-interest (POI) recommendations, Neurocomputing, Vol. 472, pp. 306-325, February, 2022. https://doi.org/10.1016/j.neucom.2021.05.114

[2] D.-Q. Yang, D.-Q. Zhang, V.-W. Zheng, Z.-Y. Yu, Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 45, No. 1, pp. 129–142, January, 2015. https://doi.org/10.1109/TSMC.2014.2327053

[3] E. Cho, S. A. Myers, J. Leskovec, Friendship and mobility: user movement in location-based social networks, *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, USA, August, 2011, pp. 1082–1090. https://doi.org/10.1145/2020408.2020579

[4] P. Han, Z.-X. Li, Y. Liu, P.-L. Zhao, J. Li, H. Wang, S. Shang, Contextualized point-of-interest recommendation, *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, Yokohama, Japan, 2021, pp. 2484-2490. https://dl.acm.org/doi/abs/10.5555/3491440.3491784

[5] S.-H. Jiang, X.-M. Qian, J.-L. Shen, Y. Fu, T. Mei, Author topic model-based collaborative filtering for personalized poi recommendations, *IEEE Transactions on Multimedia*, Vol. 17, No. 6, pp. 907–918, June, 2015. https://doi.org/10.1109/TMM.2015.2417506

[6] Q. Liu, S. Wu, L. Wang, T. Tan, Predicting the next location: A recurrent model with spatial and temporal contexts, *Proceedings of the Thirtieth AAAI*

Conference on Artificial Intelligence (AAAI), Phoenix, Arizona, USA, 2016, pp. 194–200. https://dl.acm.org/doi/10.5555/3015812.3015841

[7] S. Rendle, Factorization machines, *The 10th IEEE International Conference on Data Mining (ICDM)*, Sydney, Australia, 2010, pp. 995–1000. https://doi.org/10.1109/ICDM.2010.127

[8] Y.-J. Su, X. Li, B.-P. Liu, D.-R. Zha, J. Xiang, W. Tang, N. Gao, Fgcrec: Fine-grained geographical characteristics modeling for point-of-interest recommendation, *IEEE International Conference on Communications (ICC)*, Dublin, Ireland, 2020, pp. 1–6. https://doi.org/10.1109/ICC40277.2020.9148797

[9] C. Yin, S. Ding, J. Wang, Mobile marketing recommendation method based on user location feedback, *Human-centric Computing and Information Sciences*, Vol. 9, Article No. 14, May, 2019. https://doi.org/10.1186/s13673-019-0177-6

[10] X. Lu, H. Zhang, A Content-Aware POI Recommendation Method in Location-Based Social Networks Based on Deep CNN and Multi-Objective Immune Optimization, *Journal of Internet Technology*, Vol. 21, No. 6, pp. 1761-1772, November, 2020. https://doi.org/10.3966/160792642020112106017

[11] F.-Q. Yu, L.-Z. Cui, W. Guo, X.-D. Lu, Q.-Z. Li, H. Lu, A category-aware deep model for successive POI recommendation on sparse check-in data, *The Web Conference 2020 (WWW)*, Taipei, Taiwan, 2020, pp. 1264–1274. https://doi.org/10.1145/3366423.3380202

[12] J. Feng, Y. Li, C. Zhang, F.-N. Sun, F.-C. Meng, A. Guo, D.-P. Jin, Deepmove: Predicting human mobility with attentional recurrent networks, *Proceedings of the 2018 World Wide Web Conference on World Wide Web (WWW)*, Lyon, France, 2018, pp. 1459–1468. https://doi.org/10.1145/3178876.3186058

[13] Q.-Y. Guo, J.-Z. Qi, *SANST: A self-attentive network for next point-of-interest recommendation*, January, 2020. https://arxiv.org/abs/2001.10379

[14] Y.-T. Luo, Q. Liu, Z.-C. Liu, STAN: spatio-temporal attention network for next location recommendation, *Proceedings of the web conference 2021*, Virtual Event / Ljubljana, Slovenia, 2021, pp. 2177–2185. https://doi.org/10.1145/3442381.3449998

[15] E. Wang, Y.-H. Jiang, Y.-B. Xu, L. Wang, Y.-J. Yang, Spatial-temporal interval aware sequential POI recommendation, *IEEE 38th International Conference on Data Engineering (ICDE)*, Kuala Lumpur, Malaysia, 2022, pp. 2086–2098. https://doi.org/10.1109/ICDE53745.2022.00202

[16] X.-L. Wang, G.-H. Sun, X. Fang, J. Yang, S.-J. Wang, Modeling spatio-temporal neighbourhood for personalized point-of-interest recommendation, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI)*, Vienna, Austria, 2022, pp. 3530–3536.

[17] M. Chen, Y. Liu, X.-H. Yu, *NLPMM: a next location predictor with markov modeling*, March, 2020. https://arxiv.org/abs/2003.07037

[18] H. Wang, H.-W. Shen, W.-T. Ouyang, X.-Q. Cheng, Exploiting poi-specific geographical influence for point-of-interest recommendation, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI)*, Stockholm, Sweden, 2018, pp. 3877–3883.

[19] R. Gao, J. Li, X.-F. Li, C.-F. Song, Y.-F. Zhou, A personalized point-of-interest recommendation model via fusion of geo-social information, *Neurocomputing*, Vol. 273, pp. 159–170, January, 2018. https://doi.org/10.1016/j.neucom.2017.08.020

[20] J. Li, G.-J. Liu, C.-G. Yan, C.-J. Jiang, LORI: A learning-to-rank-based integration method of location recommendation, *IEEE Transactions on Computational Social Systems*, Vol. 6, No. 3, pp. 430–440, June, 2019. https://doi.org/10.1109/TCSS.2019.2907563

[21] W. Wang, J.-Y. Chen, J.-Z. Wang, J.-X. Chen, Z.-G. Gong, Geography-aware inductive matrix completion for personalized point-of-interest recommendation in smart cities, *IEEE Internet of Things Journal*, Vol. 7, No. 5, pp. 4361–4370, May, 2020. https://doi.org/10.1109/JIOT.2019.2950418

[22] S. Rendle, C. Freudenthaler, Z. Gantner, L. Schmidt-Thieme, *BPR: bayesian personalized ranking from implicit feedback*, May, 2012. https://arxiv.org/abs/1205.2618

[23] S.-J. Wen, X. Zhang, R.-X. Cao, B.-M. Li, Y. Li, MSSRM: A Multi-Embedding Based Self-Attention Spatio-temporal Recurrent Model for Human Mobility Prediction, *Human-centric Computing and Information Sciences*, Vol. 11, pp. 1-16, September, 2021. https://doi.org/10.22967/HCIS.2021.11.037

[24] S. Yang, J.-M. Liu, K.-Q. Zhao, Getnext: Trajectory flow map enhanced transformer for next POI recommendation, *Proceedings of the 45th International ACM SIGIR Conference on research and development in information retrieval*, Madrid, Spain, 2022, pp. 1144-1153. https://doi.org/10.1145/3477495.3531983

[25] H. A. Rahmani, M. Aliannejadi, M. Baratchi, F. Crestani, A systematic analysis on the impact of contextual information on point-of-interest recommendation, *ACM Transactions on Information Systems*, Vol. 40, No. 4, Article No. 88, pp. 1–35, October, 2022. https://doi.org/10.1145/3508478

[26] P. Sanchez, A. Bellogín, Point-of-interest recommender systems based on location-based social networks: A survey from an experimental perspective, *ACM Computing Surveys*, Vol. 54, No. 11s, Article No. 223, pp. 1–37, January, 2022. https://doi.org/10.1145/3510409

[27] B. Chang, Y. Park, D. Park, S. Kim, J. Kang, Content-aware hierarchical point-of-interest embedding model for successive POI recommendation, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI)*, Stockholm, Sweden, 2018, pp. 3301–3307.

[28] M. Xie, H.-Z. Yin, H. Wang, F.-J. Xu, W.-T. Chen, S. Wang, Learning graph-based POI embedding for location-based recommendation, *Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM)*, Indianapolis, IN, USA, 2016, pp. 15–24. https://doi.

org/10.1145/2983323.2983711

[29] M. Ye, D. Shou, W.-C. Lee, P.-F. Yin, K. Janowicz, On the semantic annotation of places in location-based social networks, *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, USA, 2011, pp. 520–528. https://doi.org/10.1145/2020408.2020491

[30] G. Christoforidis, P. Kefalas, A. N. Papadopoulos, Y. Manolopoulos, RELINE: point-of-interest recommendations using multiple network embeddings, *Knowledge and Information Systems*, Vol. 63, No. 4, pp. 791-817, April, 2021. https://doi.org/10.1007/s10115-020-01541-5

[31] Q. Guo, Z. Sun, J. Zhang, Y.-L. Theng, An attentional recurrent neural network for personalized next location recommendation, *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*, New York, NY, USA, 2020, pp. 83–90.

[32] J.-C. Li, Y.-J. Wang, J. McAuley, Time interval aware self-attention for sequential recommendation, *The Thirteenth ACM International Conference on Web Search and Data Mining (WSDM)*, Houston, TX, USA, 2020, pp. 322–330. https://doi.org/10.1145/3336191.3371786

[33] C.-C. Zhu, M.-H. Chen, C.-J. Fan, G.-Q. Cheng, Y. Zhang, Learning from history: Modeling temporal knowledge graphs with sequential copy-generation networks, *Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI)*, Virtual Event, 2021, pp. 4732–4740.

[34] M. B. Hossain, M. S. Arefin, I. H. Sarker, M. Kowsher, P. K. Dhar, T. Koshiba, CARAN: A context-aware recency-based attention network for point-of-interest recommendation, *IEEE Access*, Vol. 10, pp. 36299–36310, April, 2022. https://doi.org/10.1109/ACCESS.2022.3161941

[35] Y.-T. Lai, Y.-J. Su, L.-W. Wei, G.-D. Chen, T.-C. Wang, D.-R Zha, Multi-view spatial-temporal enhanced hypergraph network for next POI recommendation, *Database Systems for Advanced Applications - 28th International Conference (DASFAA)*, Tianjin, China, 2023, pp. 237–252. https://doi.org/10.1007/978-3-031-30672-3_16

[36] D.-F. Lian, Y.-J. Wu, Y. Ge, X. Xie, E.-H. Chen, Geography-aware sequential location recommendation, *The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, Virtual Event, CA, USA, 2020, pp. 2009–2019. https://doi.org/10.1145/3394486.3403252

[37] X. Zhang, S.-J. Wen, L. Yan, J.-F. Feng, Y. Xia, A hybrid-convolution spatial-temporal recurrent network for traffic flow prediction, *The Computer Journal*, Vol. 67, No. 1, pp. 236–252, January, 2024. https://doi.org/10.1093/comjnl/bxac171

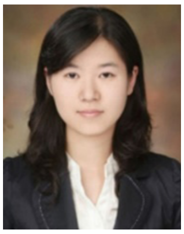
Biographies

Xu Zhang received his Ph.D. in Computer and Information Engineering at Inha University, Incheon, Korea, in 2013. He is currently a professor at Chongqing University of Posts and Telecommunications. His research interests include urban computing, intelligent transportation system, and few-shot learning.

Deao Liu received the B.S. degree and is currently pursuing a M.S. degree in Computer Science and Technology from Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include next POI recommendation based on historical location, smart city based on spatio-temporal sequences predicting.

Liang Yan received the M.S. degree and is currently pursuing the Ph.D. degree in Computer Science and Technology from Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include urban computing, trajectory mining and privacy preservation.

Zhiqing Zhang serves as the Chairman of the Chongqing Surveying and Geographic Information Society, the manager of the Technology Innovation Center for Territorial & Spatial Big Data, MNR, and a senior engineer at the Key Laboratory of Monitoring, Evaluation and Early Warning of Territorial Spatial Planning Implementation, Ministry of Natural Resources (LMEE).



Yan Li received the B.S. degree in computer science and engineering from Chongqing University of Posts and Telecommunications, Chongqing, China, and the M.S. and Ph.D. degrees from Inha University, Incheon, South Korea, in 2017. Her research interests include cloud computing, crowd sensing, and big data analytics.