

## Article

# Capturing Dynamic Interests of Similar Users for POI Recommendation Using Self-Attention Mechanism

Xinhua Fan, Yixin Hua \*, Yibing Cao and Xinke Zhao

Institute of Geographic Space Information, PLA Strategic Support Force Information Engineering University, Zhengzhou 450052, China

\* Correspondence: chxyhyxy@163.com

**Abstract:** The integration of location-based social networks and POI recommendation systems has the potential to enhance the urban experience by facilitating the exploration of new and relevant locales. The deployment of graph neural networks (GNNs) drives the development of POI recommendations, but this approach also brings with it the challenge of over-smoothing, where information propagation between nodes in the graph can lead to an excessive homogenization of the data. In prior works that utilized GNNs for POI recommendation, the bipartite graphs constructed from users and POIs as nodes failed to incorporate temporal dynamics, limiting the scope of the analysis to only spatial structure information. To circumvent this issue, the incorporation of a temporal component can be introduced during the aggregation process of graph convolution. In light of these considerations, the present study proposes a novel regionalized temporal GCN (RST-GCN) recommendation model that leverages self-attention mechanism to capture various levels of temporal information to better reflect the dynamic changes of time. By combining the graph's spatial structure with geospatial features, similar users are distributed into distinct regional subgraphs, effectively avoiding the influence of non-similar users. The efficacy of the proposed model has been demonstrated through empirical evaluations conducted on two real-world datasets.

**Keywords:** POI recommender system; graph convolutional neural network; self-attention; regional subgraph



**Citation:** Fan, X.; Hua, Y.; Cao, Y.; Zhao, X. Capturing Dynamic Interests of Similar Users for POI Recommendation Using Self-Attention Mechanism. *Sustainability* **2023**, *15*, 5034. <https://doi.org/10.3390/su15065034>

Academic Editors: Haoran Wei, Zhendong Wang, Yuchao Chang and Zhenghua Huang

Received: 13 February 2023

Revised: 9 March 2023

Accepted: 11 March 2023

Published: 13 March 2023



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

## 1. Introduction

With the proliferation of social media, location data, and mobile Internet, urban life is enhanced by the presence of various points of interest (POIs) such as restaurants, stores, parks, and cultural venues. POI recommendation systems play a key role in helping city dwellers discover new and exciting places based on their preferences and past experiences. In recent years, POI recommendation has become an increasingly important task.

In early POI recommendation, algorithms mainly achieved mining POIs of interest to users by analyzing the interaction between users and POIs and using rating information to construct user preferences for POIs, and most of the work used collaborative filtering [1] techniques to recommend POIs to users. In recent years, the emergence of graph neural networks (GNNs) has driven the development of POI recommendation. The graph model can be used in LBSN [2] to construct the interaction between POI and user as a bipartite graph network. The utilization of graph neural networks (GNNs) has proven to be an effective approach to capturing the high-order connectivity [3] inherent in POI recommendation systems. By incorporating multi-hop neighbors in the graph [4], the collaborative filtering effect can be effectively represented and integrated into the learning process through embedding propagation and aggregation. The utilization of GNNs in graph data learning allows for the acquisition of deeper insights, owing to the unique information propagation mechanism inherent in this approach. Zhong et al. [5] proposed a hybrid graph convolutional network model to construct a spatial graph using the geographic distance between

pairs of interest points, Zhang et al. [6] proposed a GNN-POI algorithm to use graph neural network to construct a social network graph that uses a bi-directional long and short-term memory model to simulate users' sequential check-in behavior.

Because of the unique information propagation mechanism of GNN, as the number of convolutional layers is added, the larger the number of neighboring nodes that can be taken into account when each node embedding is updated. The model is better able to represent the global features of the graph. However, not all information from neighbors can be beneficial for embedding learning, and users with non-similar interests often bring negative information propagation [7]. How to target information propagation becomes a problem. In previous work based on GNN recommendations, the bipartite graphs [8] constructed with users and POIs as nodes do not reflect the temporal factor [9], and their topologies only involve information about the graph spatial structure, while a user's POI recommendation needs to be closely integrated with the user's historical check-in information to capture the user's dynamic interest [10]. Therefore, the dynamic temporal factor needs to be involved in the aggregation process of graph convolution.

In view of the above limitations, in this paper, we propose a new Regional Spatio-Temporal GCN (RST-GCN) recommendation model. The model applies scaled dot product attention and captures different levels of temporal information using a multi-head attention mechanism. The acquired temporal information can better reflect the dynamic change of time, and its participation in information propagation as neighboring temporal information in graph convolution can achieve the capture of users' dynamic interests. To prevent the impact of non-similar users, we categorize users and the POIs they visit into separate regional subgraphs, and the information propagation can only take place in the regional subgraphs. To demonstrate the validity of our proposed model, we conducted extensive experiments on both Foursquare and Gowalla datasets.

In summary, the main contributions of this paper are summarized as follows:

1. We propose a model called Regional Spatio-Temporal GCN (RST-GCN), which combines graph structural features with geospatial features to divide regional subgraphs and reduce the negative impact brought by non-similar users in information propagation;
2. We use a self-attention mechanism to capture dynamic temporal information as temporal features of users and POIs in the GCN to participate in neighbor information aggregation;
3. We conduct extensive experiments on relevant datasets to demonstrate the effectiveness of our proposed RST-GCN.

The remainder of this paper is arranged in the following manner. In Section 2, we revisit the relevant literature pertaining to POI recommendation. In Section 3, we conceptualize the problem at hand. The specifics of the proposed RST-GCN model are expounded upon in Section 4. In Section 5, we subject our proposed model to an evaluation through comparison with existing POI recommendation models, as well as through the examination of ablation experiments and the analysis of sensitivity parameters. Finally, in Section 6, we summarize the paper and chart a course for future work.

## 2. Related Work

### 2.1. GNN-Based POI Recommendation

Graph-based approaches have been experimentally shown to be dominant in recommendation problems, and the following are relevant recent works. STP-UDGAT [11] designs global spatial-temporal preference (STP) neighborhoods to find high-order POI neighbors by random wandering; GARG [12] treats POIs in sequences equally and adaptively divides importance, and automatically identifies POI relevance; GPR [13] estimates user preferences by capturing the nonlinear geographic influence of the user-POI network and using geographically latent representations of input and output influences; LightGCN [14] retains the original graph model structure simplifying the GCN model by learning embeddings through linear propagation in the user-item interaction graph making it more applicable to recommendation tasks; IA-GCN [15] emphasizes relevant information in the user's neigh-

borhood and assigns higher attention weights to similar neighborhoods; PinSage [16] uses random walking sampling in combination with graph convolution to learn embeddings on the item-item graph of image recommendation tasks. The unequal relevance of nodes in graph convolution requires targeted weight assignment for a selective aggregation process. The above models take into account the variability of realistic geographic locations by designing POI-POI graphs, which shows the need to combine graph structural features as well as geospatial features to achieve more accurate recommendations.

## 2.2. Sequence-Based Recommendation

Sequence-based recommendations are widely used because they can capture the order of users' check-in POIs to obtain their dynamic preferences. Markov chains [17] were first applied to sequential recommendations, using the most recently clicked items to infer the change process of items from the next item. Deep learning-based algorithms have further improved the efficiency of sequential-based recommendations, and Chen et al. [18] proposed a supervised learning prediction model based on recurrent neural networks (RNNs) considering location interests and contextual information of similar users; ATST-LSTM [19] uses an attention mechanism to pick relevant historical check-in behaviors in the input sequence using spatio-temporal contextual information; ATCA-GRU [20] proposes a GRU model that uses an attention mechanism to perceive POI categories to predict the most likely types of POIs to be visited at future moments; SASRec [21] uses self-attention techniques to model sequential patterns and capture long-term semantics from them to make effective recommendations; BERT4Rec [22] uses a deep bidirectional sequential model optimization on top of the base model to make sequential recommendations; SAE-NAD [23] uses a multi-head attention mechanism to distinguish user preferences and uses inner product and radial basis functions for similar neighbor perception.

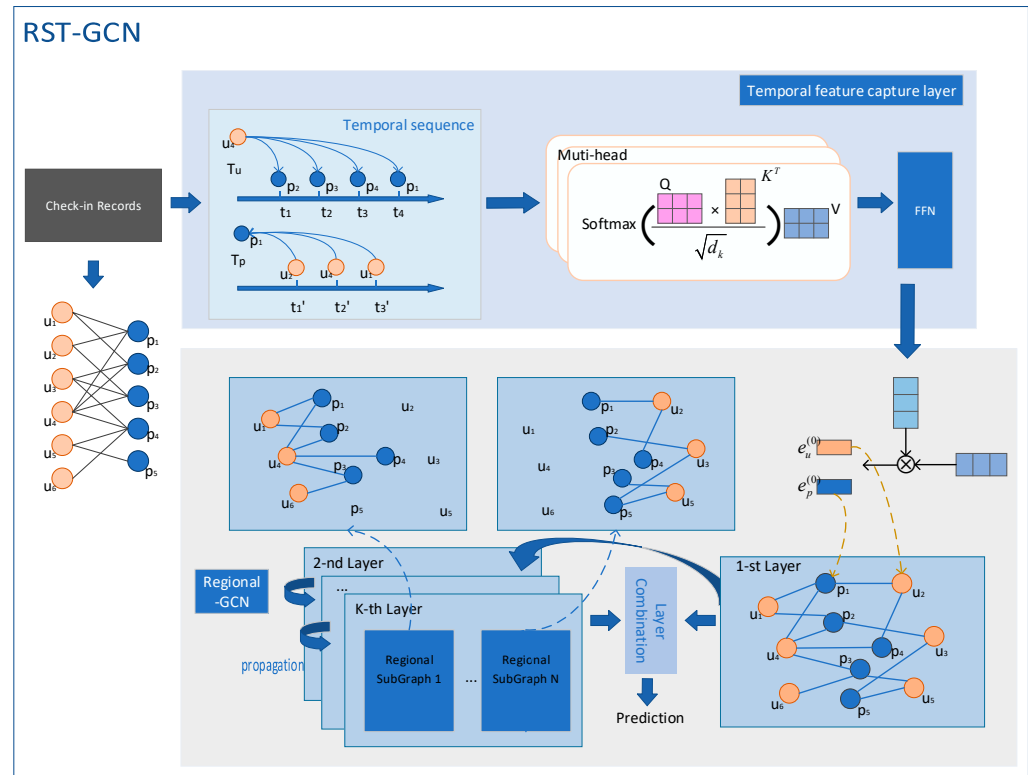
## 3. Problem Definition

This paper introduces a set of notations to describe a POI recommendation system. The set of users,  $U$ , is represented as  $\{u_1, u_2, \dots, u_M\}$ , where  $M$  is the number of users. Each user is characterized by their ID and region information and is denoted by  $u_m = \langle ID_M, Region \rangle$ . The set of POIs is represented by  $P = \{p_1, p_2, \dots, p_N\}$ , where  $N$  is the number of POIs. Each POI is described by its ID, longitude, and latitude information, represented by  $l_N = \langle ID_N, lon_N, lat_N \rangle$ . The user-POI check-in matrix is denoted as  $A \in \mathbb{R}^{M \times N}$ , which is used to construct the user-POI bipartite graph,  $\mathcal{G} = (\mathcal{W}, \mathcal{E})$ . The node set  $\mathcal{W}$  consists of user nodes and POI nodes, and  $\mathcal{E}$  denotes the edge set. The check-in behavior of users at different POIs is used to collect temporal information. The user's check-in temporal sequence and POI visited temporal sequence are represented by  $S_u = \{p_1, p_2, \dots, p_S\}$  and  $S_p = \{u_1, u_2, \dots, u_S\}$ , respectively, where  $S$  represents the length of the sequence.

Based on the above information, the problem of POI recommendation is to make recommendations of POIs in  $P$  to users, taking into account the constructed user-POI bipartite graph  $\mathcal{G} = (\mathcal{W}, \mathcal{E})$  and incorporating the temporal information from the user's check-in temporal sequence and POI visited temporal sequence.

## 4. Methodology

In this section, we present our model in detail. RST-GCN improves GNN-based POI recommendation algorithm by setting regional subgraphs to obtain similarity users' high-order interests and combining Transformer to obtain dynamic temporal features. Figure 1 shows the general architecture of RST-GCN.



**Figure 1.** The general architecture of RST-GCN.

#### 4.1. Temporal Feature Capture Layer

In this layer, we generate the user's check-in temporal sequence  $S_u = \{p_1, p_2, \dots, p_S\}$  and POI visited temporal sequence  $S_p = \{u_1, u_2, \dots, u_S\}$ , where  $S$  is the length of the temporal sequence. We designed the location encoding to efficiently capture the temporal location embedding  $V_u = \{v_u^1, v_u^2, \dots, v_u^S\}$  and  $V_p = \{v_p^1, v_p^2, \dots, v_p^S\}$ . As in most of the work, we add the position encoding to the embedding in the form of a sum, and the final temporal user input embedding and final temporal user input embedding are denoted as  $T_u = \{e_p^1 + v_p^1, e_p^2 + v_p^2, \dots, e_p^S + v_p^S\}$ ,  $T_p = \{e_u^1 + v_u^1, e_u^2 + v_u^2, \dots, e_u^S + v_u^S\}$ , respectively.

For the above temporal user input embedding  $T_u = \{e_p^1 + v_p^1, e_p^2 + v_p^2, \dots, e_p^S + v_p^S\}$ , the query, key, and value vectors are computed as follows and the self-attention function can be formulated as:

$$h^{T_u} = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

$$Q = T_u W_Q, K = T_u W_K, V = T_u W_V, \quad (2)$$

where  $h^{T_u}$  denotes the attention output embedding matrix;  $W_Q$ ,  $W_K$ , and  $W_V$  are the weight matrices of query, key, and value; softmax function is used to normalize these scores into attention weights;  $d_k$  is the dimension of the  $K$  vector.  $T_p$  is also calculated in the same way.

We use multi-head attention to capture temporal information from different latent perspectives and are fed into a feed-forward network (FFN). The final output is calculated as follows:

$$h_u = \text{FFN}(h_1^{T_u} \dots || h_i^{T_u} \dots || h_k^{T_u}) \quad (3)$$

$$h_p = \text{FFN}(h_1^{T_p} \dots || h_i^{T_p} \dots || h_k^{T_p}) \quad (4)$$

where  $k$  denotes the number of attention functions.

## 4.2. Regional-GCN

### 4.2.1. Embedding Layer

In the embedding layer, let  $E_U \in \mathbb{R}^{M \times d}$  be the user embedding matrix,  $E_P \in \mathbb{R}^{N \times d}$  be the POI embedding matrix, which is the projection in the lower dimensional representation, where  $d$  is the embedding size. The user-POI check-in matrix is denoted as  $A \in \mathbb{R}^{M \times N}$ . The user-POI bipartite graph  $\mathcal{G} = (\mathcal{W}, \mathcal{E})$  is constructed by the user-POI check-in matrix. The set of nodes,  $\mathcal{W}$ , comprises both user nodes and POI nodes, while  $\mathcal{E}$  represents the set of edges connecting these nodes. Direct interaction between the user and the POI as the most reliable information will be updated by GCN aggregation and propagation.

### 4.2.2. Embedding Propagation Layer

Our model adapts the widely recognized and efficacious framework of LightGCN, a GCN-based recommendation model, in its underlying architecture. By leveraging the simplicity of the network structure demonstrated in LightGCN, we aim to develop a recommendation model that is both performant and computationally tractable.

By constructing a regional subgraph, users with similar interests are classified using graph spatial structure and geospatial features, and users of the same class are assigned to the same regional subgraph, and one-hop neighbor POIs connected to the users are also grouped into that regional subgraph. Therefore, the same POI may exist in multiple region subgraphs, and users belong to only the regional subgraph in which they are located.  $R = \{r_1, r_2, \dots, r_i\}$  denotes the regional subgraph, where  $i$  denotes the number of regional subgraphs. The way of constructing regional subgraphs will be elaborated in Section 4.2.3. The initial embedding of all users and POIs are incorporated into the first-order graph convolution operation to obtain the check-in relationship between every user and POI. This is described as follows:

$$e_u^{(1)} = \sum_{i \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_p|}} e_p^{(0)} \quad (5)$$

$$e_p^{(1)} = \sum_{u \in \mathcal{N}_p} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_p|}} e_u^{(0)} \quad (6)$$

where  $e_u^{(1)}$  and  $e_p^{(1)}$  represent the user and POI embeddings after the first convolution;  $e_u^{(0)}$  and  $e_p^{(0)}$  represent the user and POI initial embedding matrices, respectively;  $\mathcal{N}_u$  represents the set of POIs visited by the user;  $\mathcal{N}_p$  represents the set of users who have visited POIs;  $\frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_p|}}$  is used to achieve symmetric normalization.

Each convolutional layer processes only one order neighborhood information, incorporating a self-degree matrix and a normalization operation on the adjacency matrix. The propagation formula is shown as follows:

$$H^{(l+1)} = \sigma(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}) \quad (7)$$

In the high-order graph convolution calculation, through the division of regional subgraphs, the user node belongs to only one of the regional subgraphs, but the POI distribution lies in the subgraphs where all the user nodes associated with it are located, and the POI embedding is the sum of the POI embeddings in all regional subgraphs, the propagation through the  $k - 1$  layer graph convolution is defined as:

$$e_u^{(k)} = \sum_{pr \in \mathcal{N}_u} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_p|}} e_{pi}^{(k-1)} \quad (8)$$

$$e_{pi}^{(k)} = \sum_{u \in \mathcal{N}_p^i} \frac{1}{\sqrt{|\mathcal{N}_u|} \sqrt{|\mathcal{N}_p|}} e_u^{(k-1)} \quad (9)$$

$$e_p^{(k)} = \sum_{s \in R} e_{pi}^{(k)} \quad (10)$$

where  $R$  denotes the set of each regional subgraph in which the POI is located;  $e_{pi}$  denotes the embedding representation of the POI in the regional subgraph  $r_i$ .

#### 4.2.3. Regional Subgraph Construction

The regional subgraph setup module is designed to filter out the negative impact of non-similar users and group users with similar interests into the same subgraph. The feature vector representing each user is a combination of two parts: the user's ID embedding after first-order graph convolution, which provides information about the user's relationship with neighboring POIs (graph spatial features), and the user's most frequently visited location (geospatial features). To determine the most active geographic location of a user, we use the latitude and longitude data of the POI that the user visits most frequently, which is normalized before being used as an input feature.

$$Feature_u = \sigma(W_1(e_u^{(1)} + e_{up}) + b_1) \quad (11)$$

where  $Feature_u$  denotes the user feature vector used for classification;  $e_u^{(1)}$  is the user embedding after first layer of graph convolution, representing the graph spatial structure obtained by aggregating first-order neighbor POIs;  $e_{up}$  is the representation of the geographic location where the user most frequently visits the POI;  $\sigma(\cdot)$  we use LeakyReLU [24] as the activation function;  $W_1, b_1$  denote the weight matrix and bias vector, respectively. After obtaining the user feature vector, we use a two-layer neural network to project the obtained user feature.

$$U = W_3(W_2 Feature_u + b_2) + b_3 \quad (12)$$

where  $U$  denotes the classification prediction vector obtained by projection, and users with similar representations will be classified into the same regional subgraph. After generating the specified number of regional subgraphs, users only aggregate neighbor information in the regional subgraph where they are located and filter the neighbor relationships in the non-regional subgraphs.

#### 4.3. Using RST-GCN for Recommendation

In this paper, we construct a user-POI bipartite graph  $\mathcal{G} = (\mathcal{W}, \mathcal{E})$  using the user-POI check-in matrix. We collect temporal information based on the check-in behavior of users at different POIs and construct user check-in temporal sequence  $S_u = \{p_1, p_2, \dots, p_s\}$  and POI visited temporal sequence  $S_p = \{u_1, u_2, \dots, u_s\}$ . We use Transformer to obtain the temporal features of each user node and POI node as part of the initial embedding to participate in GCN. Then, we divide the regional subgraphs to shield the negative information propagation of high-order non-similar users, and get the final embedding of users and POIs by multi-layer GCN propagation. Finally, we rank the POIs and keep the top  $k$  items as the recommended POIs.

#### 4.4. Optimization

In this paper, we use the widely adopted BPR [25] algorithm to optimize the parameters of the model.

$$\mathcal{L}_{pred} = - \sum_{\substack{(u,p) \in \mathcal{R}, \\ (u,z) \in \mathcal{R}^-}} \ln \sigma(y_{pred}(u,p) - y_{pred}(u,z)) + \lambda \|\Theta\|_2^2 \quad (13)$$

where  $\sigma(\cdot)$  is the sigmoid function;  $\Theta$  is trainable parameters;  $\lambda$  controls the  $L_2$  regularization strength;  $\mathcal{R}$  is the existing user check-in behavior in the training dataset; and  $\mathcal{R}^-$  represents the negative sampling strategy.

## 5. Experiments and Results

In this section, we first present the experimental setup of this paper, which includes the dataset, evaluation metrics, and baseline. Then, the experimental parameter settings of the proposed method are analyzed. Finally, we describe the comparison of the experimental results with other related methods and the results of the ablation experiments.

### 5.1. Experimental Setup

#### 5.1.1. The Datasets

In order to validate the efficacy of our proposed approach, we have selected two well-known, publicly accessible datasets in the domain of location-based social networks (LBSN), namely Foursquare and Gowalla, for our experiments. These datasets have been extensively used and validated in previous studies, and therefore serve as a reliable benchmark for the evaluation of POI recommendation algorithms. The statistics of the dataset are presented in Table 1.

**Table 1.** Statistics of datasets.

| Dataset    | # of Users | # of POIs | # of Check-ins |
|------------|------------|-----------|----------------|
| Foursquare | 24,941     | 28,593    | 1,196,248      |
| Gowalla    | 18,737     | 32,510    | 1,278,274      |

In each of these datasets, we collected five key pieces of information pertaining to the check-in behavior of users, specifically, the user ID, POI ID, latitude and longitude of the POI, and the corresponding check-in timestamp.

In the preprocessing phase of the dataset, we mitigated the impact of users and POIs with insufficient interactions by removing inactive users who had checked into fewer than 10 POIs and POIs that had been visited by fewer than 10 users. We then partitioned the processed data into two segments, with an 8:2 ratio, for use as the training set and the test set, respectively.

#### 5.1.2. Evaluation Metrics

The proposed approach is evaluated using two commonly adopted metrics in the field of POI recommendation, NDCG@K and Recall@K. The metrics are calculated based on the average value obtained after multiple experiments, with K representing the number of POIs recommended to each user, the metrics are calculated as follows:

$$\text{Recall@K}(u) = \frac{1}{|N|} \sum_{u=1}^N \frac{|R^K(u) \cap T(u)|}{|T(u)|} \quad (14)$$

$$\text{NDCG@K} = \frac{1}{|U|} \sum_{u \in U} \frac{\sum_{k=1}^K \frac{I(R_k^K(u) \in T(u))}{\log(k+1)}}{\sum_{k=1}^K \frac{1}{\log(k+1)}} \quad (15)$$

### 5.2. Baselines

To verify the validity of our proposed RST-GCN method, we selected relevant baseline methods for comparison, where the influencing factors in baselines are shown in Table 2. These methods are described below.

**Geo Teaser [26]:** Geo Teaser uses geo-influence capture in user check-in temporal sequence to learn the POI representation of certain specific moments, using geo-influence to represent user preferences.

**LightGCN [14]:** LightGCN is a graph convolutional network model that removes feature transformations and non-linear activation designs, which simplifies GCN to be more suitable for collaborative filtering for recommendation tasks.

**GeoMF [27]:** GeoMF is a matrix factorization model that uses two-dimensional kernel density estimation to explain the aggregation of geospatial regions.

GPR [13]: GPR is a graph neural network-based geographic potential representation model that uses graphical autoencoders to train incoming and outgoing influences, and uses trained geographic potential representations to estimate users' preferences.

GNN-POI [6]: GNN-POI is a model that uses graph neural networks to learn graph spatial structure and combines bi-directional long short-term memory to simulate users' continuous check-in behavior to obtain geographic and temporal features.

Table 2. Influential factors in baselines.

|           | Graph Spatial Structure | Sequential Effect | Geographical Influence | Temporal Influence |
|-----------|-------------------------|-------------------|------------------------|--------------------|
| GeoMF     | ✗                       | ✗                 | ✓                      | ✗                  |
| GeoTeaser | ✗                       | ✓                 | ✓                      | ✓                  |
| LightGCN  | ✓                       | ✗                 | ✗                      | ✗                  |
| GPR       | ✓                       | ✗                 | ✓                      | ✗                  |
| GNN-POI   | ✓                       | ✓                 | ✓                      | ✓                  |
| RST-GCN   | ✓                       | ✓                 | ✓                      | ✓                  |

‘✗’ indicates that the model does not involve the corresponding influence factor; ‘✓’ indicates that the model involves the corresponding influence factor.

### 5.3. Performance Comparison

In this section, we evaluate the experimental results of all methods on the datasets Foursquare and Gowalla using the evaluation metrics of Section 5.1.2. Figures 2–5 depict the results obtained from the experiment. The performance comparison results are summarized as follows:

The data presented in the figure support the conclusion that our novel RST-GCN approach is superior to all other baseline methods across both Foursquare and Gowalla datasets, across all evaluation metrics. This observation underscores the strength of our proposed technique. To unpack these results, comparison of each method and its corresponding performance is shown below.

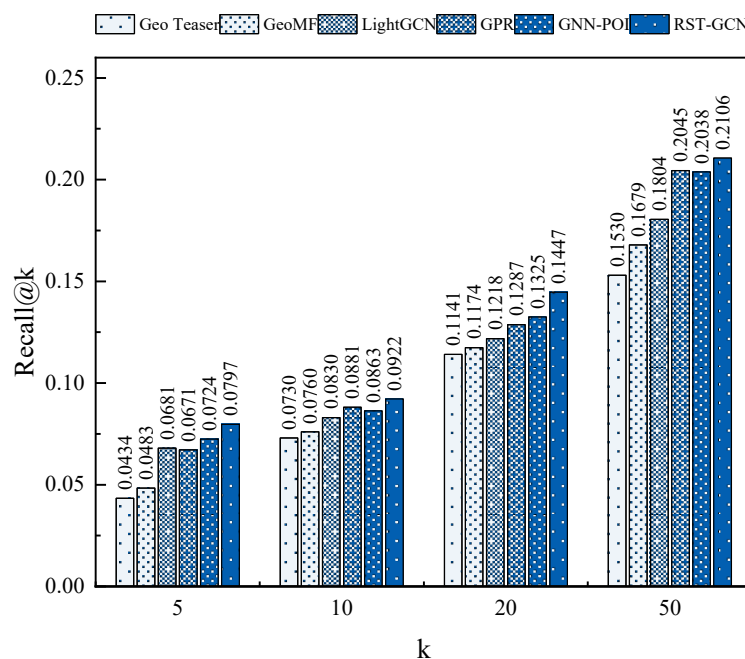
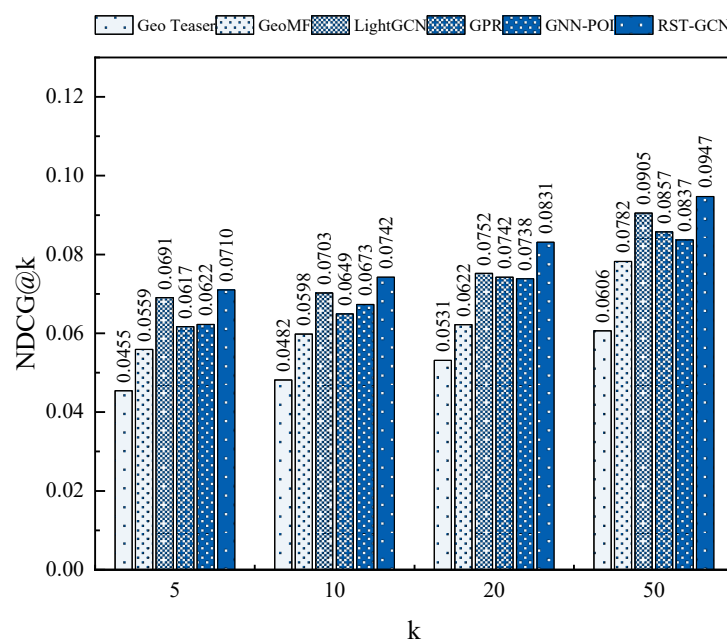
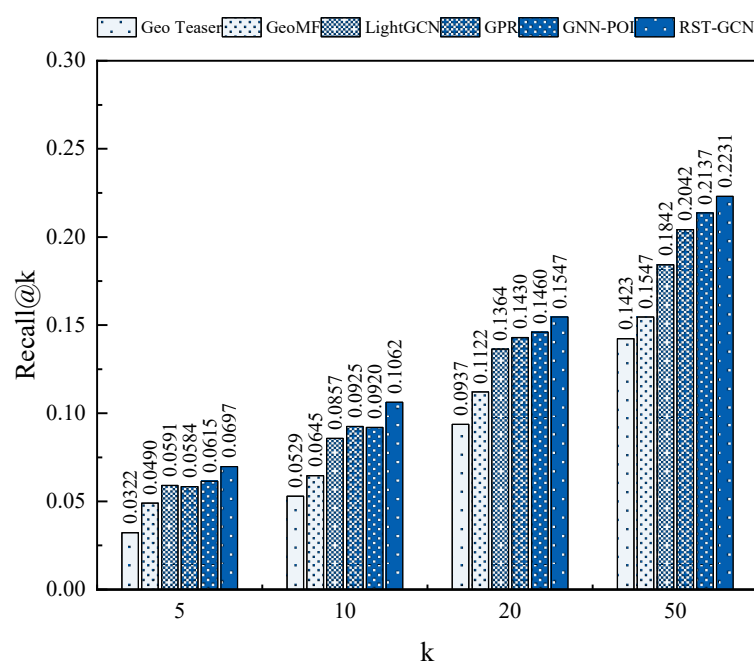


Figure 2. Comparison of Recall performance of RST-GCN and baselines on Foursquare dataset at top-k.



**Figure 3.** Comparison of NDCG performance of RST-GCN and baselines on Foursquare dataset at top-k.

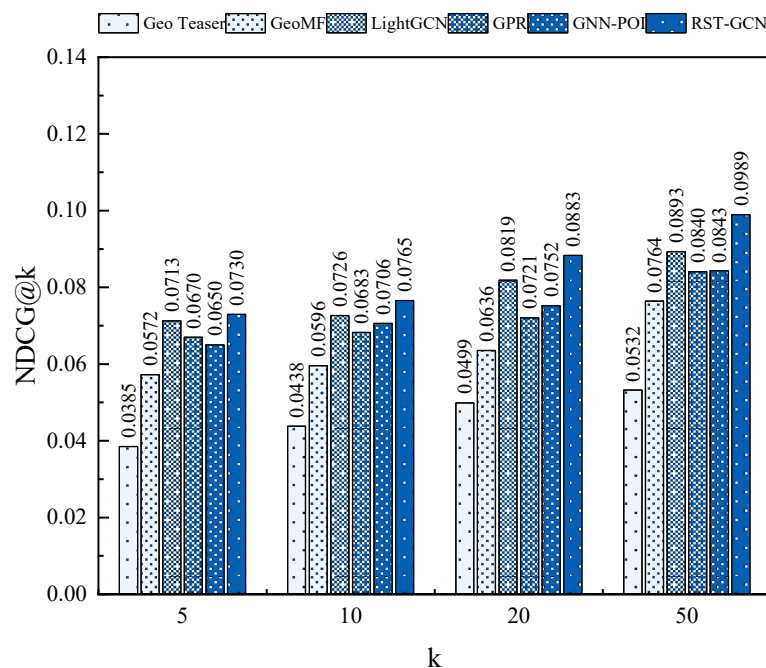


**Figure 4.** Comparison of Recall performance of RST-GCN and baselines on Gowalla dataset at top-k.

The Geo-Teaser and GeoMF techniques in the baseline category are conventional embedding-based methods that display lackluster performance across all metrics. It is evident that utilizing the spatial topology of the graph and recursively updating the node embeddings can more effectively express the user's preferences. In comparison, GeoMF outperforms Geo-Teaser due to its use of matrix decomposition to vectorize the user and POI, while Geo-Teaser relies on the intersection of POI text information and geolocation information for similarity calculation.

Excluding the aforementioned techniques, the RST-GCN approach outperforms all other GNN-based methods across both metrics and datasets. (1) Compared with GPR, GPR incorporates the outgoing geographical influence of the POI-POI directed graph as a user part feature during information propagation, while representing the POI with the

incoming geographical influence. However, GPR only models the relationship between POIs in successive order in the temporal sequence. (2) Compared with GNN-POI, GNN-POI leverages the BiLSTM method from the check-in sequence to acquire temporal features as part of the user representation, thereby appending information regarding temporal and geospatial features to the user and POI's respective characteristics. Nevertheless, the MLP model, as a classical feedforward neural network, disregards the interconnection between nodes.



**Figure 5.** Comparison of NDCG performance of RST-GCN and baseline on Gowalla dataset at top-k.

The comparisons outlined above clearly illustrate the strength of our proposed framework design: (1) we employ the structural features of the graph in combination with geospatial features to segment the regional subgraphs, thereby mitigating the negative impact of non-similar users in the information dissemination process; (2) we extract temporal information from the temporal sequence to participate in the aggregation of neighboring information; (3) we utilize the self-attention mechanism and multi-headed attention mechanism to further enhance performance.

As depicted in Figures 2–5, our method achieves significant improvement over the strongest baseline method, with a 9.21% and 10.5% increase in the recall@20 and ndcg@20 metrics, respectively, for the Foursquare dataset. Similarly, for the Gowalla dataset, it achieves a 5.96% and 7.81% increase in the recall@20 and ndcg@20 metrics, respectively. These results highlight the efficacy and superiority of our proposed method over other baseline and GNN-based methods.

#### 5.4. Ablation Experiments

In this section, we set up ablation experiments based on the proposed approach, removing the regional subgraph as well as the temporal feature capture layer, respectively, to investigate the impact of the core content in the proposed model on the performance and present the results in Table 3 with -R -T for each of the two variants. In terms of the results, it is observed that both variants are weaker than RST-GCN, respectively, marking that both designed components are to some extent able to facilitate the acquisition of user preferences.

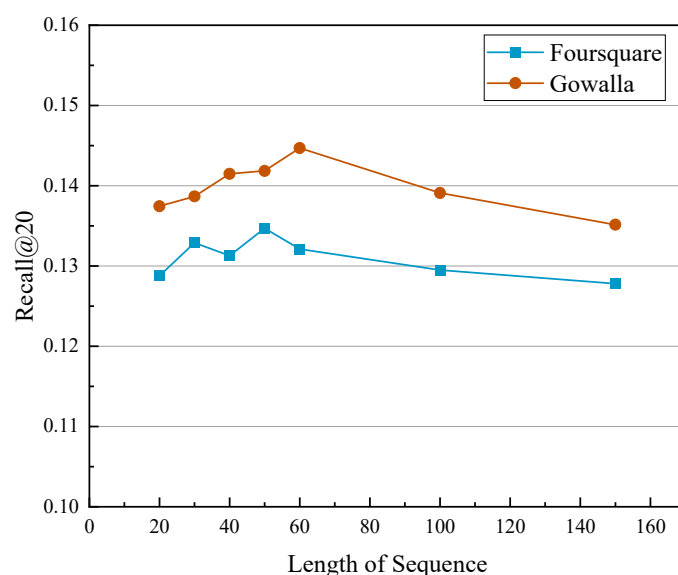
**Table 3.** Results of ablation experiments.

| Method     | Foursquare |         |           |         | Gowalla   |         |           |         |
|------------|------------|---------|-----------|---------|-----------|---------|-----------|---------|
|            | Recall@10  | NDCG@10 | Recall@20 | NDCG@20 | Recall@10 | NDCG@10 | Recall@20 | NDCG@20 |
| RST-GCN -R | 0.0839     | 0.0696  | 0.1288    | 0.0756  | 0.0884    | 0.0698  | 0.1375    | 0.0821  |
| RST-GCN -T | 0.0856     | 0.0703  | 0.1313    | 0.0775  | 0.0913    | 0.0715  | 0.1404    | 0.0832  |
| RST-GCN    | 0.0922     | 0.0742  | 0.1447    | 0.0831  | 0.1062    | 0.0765  | 0.1547    | 0.0883  |

### 5.5. Parameter Sensitivity Analysis

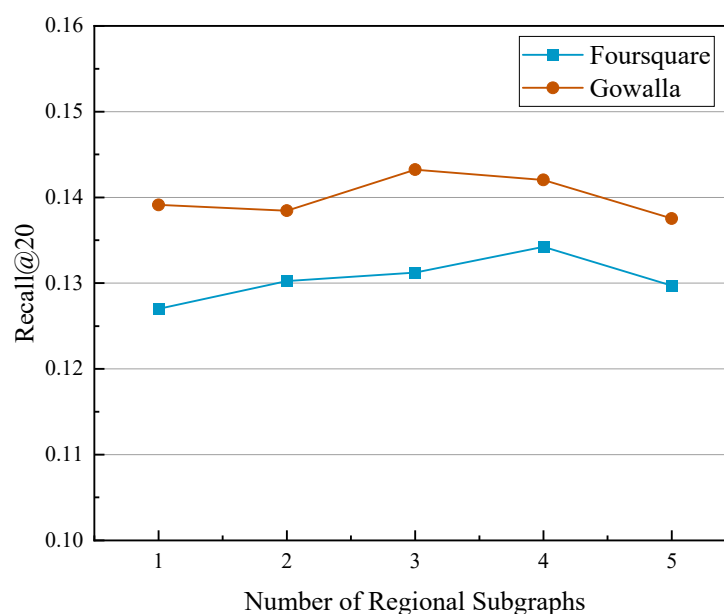
#### 5.5.1. Temporal Sequence Length

In this paper, we use temporal sequence to capture the dynamic interests of users. It is difficult to obtain accurate preferences with short sequence length, and the computational complexity of the model's multi-head self-attention mechanism increases when the sequence length is too long. We need to set the optimal temporal sequence length according to the properties of the dataset. We selected sequence lengths in the range {20, 30, 40, 50, 60, 100, 150} for comparison experiments and obtained the best settings for different datasets, setting the temporal sequence length to 50 and 60 for the Foursquare and Gowalla datasets, respectively, to obtain the best performance. As seen in Figure 6, the model performance fluctuates and grows with increasing sequence length and decreases after achieving the peak. The optimal temporal sequence length was found to be correlated with the average length of user check-in sequences for different datasets (Foursquare: 47.96, Gowalla: 68.22), so the sequence length can be selected for the new dataset in a targeted manner.

**Figure 6.** Comparison of Recall performance with different length of sequence.

#### 5.5.2. Number of Regional Subgraphs

For the number of area subgraphs, we selected the values 1–5 for the corresponding tests, where the value of 1 is used to set no regional subgraphs. The experimental results are shown in the Figure 7 below, with the increase of the number of subgraphs, the similar user grouping strategy becomes fine from coarse, users can get better information about higher-order neighbors, and the reduction of non-similar users makes the experimental performance gradually improves, and the best performance is obtained when the values of 3 and 4 are taken respectively. The performance does not increase significantly when the value continues to grow, probably because the best number of similar users grouped in the dataset has been obtained, and too many regional subgraphs are set to reduce the number of similar users in the subgraph where the users are located, which reduces the number of learnable neighbor embeddings and thus weakens the performance.



**Figure 7.** Comparison of Recall performance with different number of regional subgraphs.

## 6. Conclusions

In this study, we present a POI recommendation approach based on graph neural networks (GNNs). Our method leverages the multi-head attention mechanism to capture the temporal features of both the user's check-in sequence and the POI visited sequence. By using graph spatial structure and geospatial features to classify users into regional categories, we construct regional subgraphs that allow users with similar preferences to better learn from the node embeddings of their peers, thereby reducing the propagation of information from non-similar nodes. Experiments conducted on real-world datasets, such as Foursquare and Gowalla, have demonstrated the efficacy of our proposed approach.

In future work, we aim to further exploit user social relationship and behavioral data to achieve precise user classification, which will enhance the embedding learning of similar users within the regional subgraphs. Additionally, we will investigate more fine-grained temporal sequence partitioning and optimization of temporal feature capturing by incorporating aspects such as time intervals and periodicity.

**Author Contributions:** Conceptualization, X.F. and Y.H.; methodology, X.F.; software, X.F. and X.Z.; validation, X.F., Y.C. and X.Z.; formal analysis, Y.H. and Y.C.; investigation, X.F. and X.Z.; resources, Y.H.; data curation, X.Z.; writing—original draft preparation, X.F.; writing—review and editing, Y.H. and Y.C.; visualization, X.Z. and X.F.; supervision, Y.H. and Y.C.; project administration, Y.H.; funding acquisition, Y.H. and Y.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Key R&D Program of China, grant number 2021YFB3-900900.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on the website: <https://sites.google.com/site/yangdingqi/home/foursquare-dataset>, <http://snap.stanford.edu/data/loc-gowalla.html> (accessed on 13 February 2023).

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

## References

- Si, Y.; Zhang, F.; Liu, W. An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features. *Knowl.-Based Syst.* **2019**, *163*, 267–282. [\[CrossRef\]](#)
- Islam, M.A.; Mohammad, M.M.; Das, S.S.S.; Ali, M.E. A survey on deep learning based Point-of-Interest (POI) recommendations. *Neurocomputing* **2022**, *472*, 306–325. [\[CrossRef\]](#)
- Wang, X.; He, X.; Wang, M.; Feng, F.; Chua, T.-S. Neural graph collaborative filtering. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Paris, France, 21–25 July 2019; pp. 165–174.
- Gao, C.; Zheng, Y.; Li, N.; Li, Y.; Qin, Y.; Piao, J.; Quan, Y.; Chang, J.; Jin, D.; He, X. A Survey of Graph Neural Networks for Recommender Systems: Challenges, Methods, and Directions. *ACM Trans. Recomm. Syst.* **2022**, arXiv:2109.12843. [\[CrossRef\]](#)
- Zhong, T.; Zhang, S.; Zhou, F.; Zhang, K.; Trajcevski, G.; Wu, J. Hybrid graph convolutional networks with multi-head attention for location recommendation. *World Wide Web* **2020**, *23*, 3125–3151. [\[CrossRef\]](#)
- Zhang, J.Y.; Liu, X.; Zhou, X.F.; Chu, X.W. Leveraging graph neural networks for point-of-interest recommendations. *Neurocomputing* **2021**, *462*, 1–13. [\[CrossRef\]](#)
- Liu, F.; Chen, Z.Y.; Zhu, L.; Gao, Z.; Nie, L.Q. Interest-aware Message-Passing GCN for Recommendation. In Proceedings of the 30th World Wide Web Conference (WWW), Ljubljana, Slovenia, 12–23 April 2021; pp. 1296–1305.
- Wu, S.W.; Sun, F.; Zhang, W.T.; Xie, X.; Cui, B. Graph Neural Networks in Recommender Systems: A Survey. *ACM Comput. Surv.* **2023**, *55*, 1–37. [\[CrossRef\]](#)
- Yuan, Q.; Cong, G.; Sun, A. Graph-based point-of-interest recommendation with geographical and temporal influences. In Proceedings of the 23rd ACM International Conference on Information and Knowledge Management, Shanghai, China, 3–7 November 2014; pp. 659–668.
- Yang, Z.; Ding, M.; Xu, B.; Yang, H.X.; Tang, J.; Acm. STAM: A Spatiotemporal Aggregation Method for Graph Neural Network-based Recommendation. In Proceedings of the 31st ACM Web Conference (WWW), Lyon, France, 25–29 April 2022; pp. 3217–3228.
- Lim, N.; Hooi, B.; Ng, S.K.; Wang, X.O.; Goh, Y.L.; Weng, R.R.; Varadarajan, J.; Assoc Comp, M. STP-UDGAT: Spatial-Temporal-Preference User Dimensional Graph Attention Network for Next POI Recommendation. In Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM), Galway, Ireland, 19–23 October 2020; pp. 845–854.
- Wu, S.W.; Zhang, Y.X.; Gao, C.L.; Bian, K.G.; Cui, B. GARG: Anonymous Recommendation of Point-of-Interest in Mobile Networks by Graph Convolution Network. *Data Sci. Eng.* **2020**, *5*, 433–447. [\[CrossRef\]](#)
- Chang, B.R.; Jang, G.; Kim, S.; Kang, J.; Assoc Comp, M. Learning Graph-Based Geographical Latent Representation for Point-of-Interest Recommendation. In Proceedings of the 29th ACM International Conference on Information and Knowledge Management (CIKM), Galway, Ireland, 19–23 October 2020; pp. 135–144.
- He, X.N.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.D.; Wang, M.; Acm. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR), Xi'an, China, 25–30 July 2020; pp. 639–648.
- Kazi, A.; Farghadani, S.; Navab, N. IA-GCN: Interpretable Attention based Graph Convolutional Network for Disease prediction. *arXiv* **2021**, arXiv:2103.15587.
- Ying, R.; He, R.N.; Chen, K.F.; Eksombatchai, P.; Hamilton, W.L.; Leskovec, J.; Acm. Graph Convolutional Neural Networks for Web-Scale Recommender Systems. In Proceedings of the 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), London, UK, 19–23 August 2018; pp. 974–983.
- He, R.; McAuley, J. Fusing similarity models with markov chains for sparse sequential recommendation. In Proceedings of the 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, Spain, 12–15 December 2016; pp. 191–200.
- Chen, M.; Li, W.Z.; Qian, L.; Lu, S.L.; Chen, D.X. Next POI Recommendation Based on Location Interest Mining with Recurrent Neural Networks. *J. Comput. Sci. Technol.* **2020**, *35*, 603–616. [\[CrossRef\]](#)
- Huang, L.W.; Ma, Y.T.; Wang, S.B.; Liu, Y.B. An Attention-Based Spatiotemporal LSTM Network for Next POI Recommendation. *IEEE Trans. Serv. Comput.* **2021**, *14*, 1585–1597. [\[CrossRef\]](#)
- Liu, Y.W.; Pei, A.X.; Wang, F.; Yang, Y.H.; Zhang, X.Y.; Wang, H.; Dai, H.N.; Qi, L.Y.; Ma, R. An attention-based category-aware GRU model for the next POI recommendation. *Int. J. Intell. Syst.* **2021**, *36*, 3174–3189. [\[CrossRef\]](#)
- Kang, W.C.; McAuley, J.; Ieee. Self-Attentive Sequential Recommendation. In Proceedings of the 18th IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018; pp. 197–206.
- Sun, F.; Liu, J.; Wu, J.; Pei, C.H.; Lin, X.; Ou, W.W.; Jiang, P.; Acm. BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management (CIKM), Beijing, China, 3–7 November 2019; pp. 1441–1450.
- Ma, C.; Zhang, Y.X.; Wang, Q.L.; Liu, X. Point-of-Interest Recommendation: Exploiting Self-Attentive Autoencoders with Neighbor-Aware Influence. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM), Torino, Italy, 22–26 October 2018; pp. 697–706.
- Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. In Proceedings of the 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; p. 3.
- Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L. BPR: Bayesian Personalized Ranking from Implicit Feedback. In Proceedings of the Uncertainty in Artificial Intelligence, Montreal, QC, Canada, 18–21 June 2009; pp. 452–461.

26. Zhao, S.L.; Zhao, T.; King, I.; Lyu, M.R.; Assoc Comp, M. Geo-Teaser: Geo-Temporal Sequential Embedding Rank for Point-of-interest Recommendation. In Proceedings of the 26th International Conference on World Wide Web (WWW), Perth, Australia, 3–7 May 2017; pp. 153–162.
27. Lian, D.F.; Zhao, C.; Xie, X.; Sun, G.Z.; Chen, E.H.; Rui, Y.; Acm. GeoMF: Joint Geographical Modeling and Matrix Factorization for Point-of-Interest Recommendation. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), New York, NY, USA, 24–27 August 2014; pp. 831–840.

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