

Article

Next Location Recommendation with Bi-Directional Inference and Similarity-Aware Enhancement

Xixi Li ¹ and Kaijun Miao ^{2,*} ¹ School of Computer Science and Technology/School of Artificial Intelligence, China University of Mining and Technology, Xuzhou 221116, China; xixi@cumt.edu.cn² School of Mines, China University of Mining and Technology, Xuzhou 221116, China

* Correspondence: miaokj@cumt.edu.cn

Abstract

The greatest challenge of location recommendation is data sparsity. Some works introduce auxiliary information such as social links and semantics to enrich the representation. However, there are still two main limitations. Firstly, existing models only investigate how a person chooses POIs (Point-of-Interests), highlighting the user interest in locations from a human perspective, limited by users' sparse trajectory. Secondly, they recommend POIs by one's trajectory independently. However, it is hard to uncover specific behavior patterns or daily routines individually for each user. To address these problems, we develop a novel method for the next location recommendation via bi-directional inference and similarity-aware enhancement (SimBIN). Specifically, it considers the location recommendation task from both the human and location view, conducting the next POI inference and the next visitor inference synchronously. In particular, we employ the adaptive learning strategy to refine the model based on the conformity of two inference parts. Moreover, for the POI inference, we take the highly similar users to approximate a user's behavior for enriching the next POI candidates. It is similar for the user inference. Experiments show that the proposed SimBIN significantly outperforms the state-of-the-art approaches. Furthermore, each key component can be refined and generally used in other frameworks. By establishing a bidirectional probabilistic collaboration framework, this study not only advances the theoretical understanding of mutual regularization in spatial-temporal modeling, but also provides a possible solution for personalized location-aware services and urban trajectory planning.

Keywords: location recommendation; human mobility modeling; data sparsity; bi-directional inference; similarity-aware



Academic Editor: Andrea Prati

Received: 7 May 2026

Revised: 1 July 2026

Accepted: 9 July 2026

Published: 13 July 2026

Copyright: © 2026 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Nowadays, large amounts of behavioral data have been produced by users on location-based social networks (LBSNs), fueling many meaningful location-based applications [1]. It is of significant importance for human behavior understanding and the construction of the smart city. One of the most significant applications is the next Point-of-Interest (POI) recommendation. Its main purpose is to recommend the next POI for a user based on his/her trajectory history.

Despite the successful development of the next location recommendation on check-in data in recent years, data sparsity is still one of the most challenging problem. Many works still attend to alleviating this issue [2–5]. Specifically, data sparsity mainly lies in three

aspects: (a) Users' mobility traces are sparse. The spatial distance and time interval between successive check-ins are too large. (b) POIs for each user are scarce in contrast to the whole POI set. (c) Visitors for each POI are scarce in contrast to the whole visitor set. To illustrate the sparsity issue, we do some preliminary data analysis on Gowalla.

Distance between successive check-ins. We first investigate the distance distribution between check-ins, which are significant factors for transition learning. Figures 1a,b plot the probability of the temporal interval and geographical distance between check-in points. For every two successive check-ins, we calculate the time interval by hour and the spatial distance by km. As shown in Figure 1a, in most cases, the time interval is longer than one day, indicating the sparse distribution on time interval. Figure 1b shows that users may like to move in a region but not check in consecutively. Hence it is challenging when the distance is too far to provide effective transition patterns for the sequence modeling.

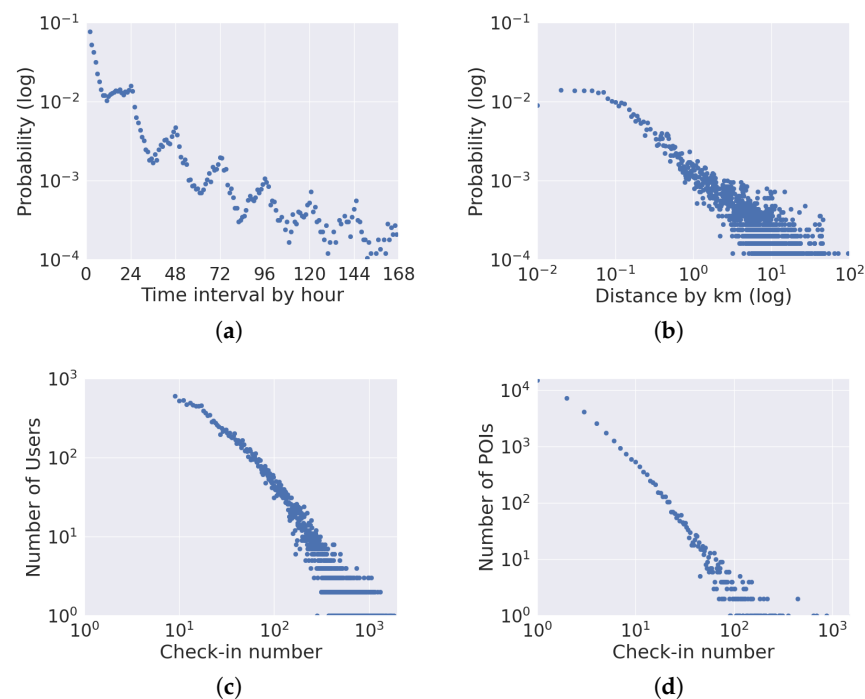


Figure 1. Some preliminary results for illustrating the sparsity issue. (a) Check-in probability vs. the temporal interval. (b) Check-in probability vs. the geographical distance. (c) The count of users vs. each user's check-in numbers. (d) The count of POIs vs. the count of each POI's check-in numbers.

Check-in frequency. Then we investigate the check-in frequency of users and POIs. For each user/POI, the check-in numbers are counted to exploit the frequency distribution from both the view of the location and human. Figure 1c shows that most users check in less than 1000 times. Similarly, from Figure 1d we can find that most POIs appeal to a small group of people, and only a few POIs have more than 100 visitors. These observations demonstrate that the number of samples from the history may be too low to mine efficient patterns for the recommendation task.

Over the past years, lots of approaches have been developed for the next location recommendation [6]. However, as far as we know, existing models only investigate how a person chooses a POI, i.e., highlighting their interests in POIs from the human view. Regrettably, they are still limited by the data sparsity despite the success of recent studies. On one hand, these works for mining users' interests from the check-in history are intrinsically limited since most users share few check-ins online. On the other hand, existing methods do the recommendations according to users' trajectory data independently. However, it is

difficult to uncover specific behavior patterns or daily routines individually for each user. Therefore, the next location recommendation remains challenging.

To this end, we develop a similarity-aware framework for the next location recommendation via adaptive bi-directional inference. The characteristics of our model lie in: (a) From the view of the location: we assume that the POI appeals to people consciously. Hence we take the POI as the subject and intend to design a model for predicting the next visitor for this POI. (b) Similarity-aware enhancement: we believe that similar targets with similar behavior can help enhance the recommendation performance. Thus a target's behavior can be speculated from the most similar groups. The following are elaborations on our motivation.

From the view of the location. Firstly, we assume that POIs appeal to people automatically. For a POI, we find that some users visit persistently and some visit rarely. This phenomenon is common in real life. For example, a library will attract people who like to read but not those who have no related hobbies. Intuitively, it is understandable to consider improving the location recommendation problem from the view of the location, i.e., predicting the POI's next visitor to help speculate whether the target user's next POI is possible. To further carry out the next visitor prediction, we assume that POIs show different temporal patterns, inspired by users' temporal patterns. For instance, a pleasure ground may be busier than usual on weekends, whereas the patterns in offices are just the opposite. In addition, a museum only collects people in the morning, but a bar is busy at night normally. The phenomena suggest that there exist some behavior patterns for a POI, and a place can attract specific people. Thus we can learn POIs' temporal patterns from the historical sequences to predict the next visitor. On this basis, we develop the next visitor prediction model to do the bi-directional inference.

Similarity-aware enhancement. Secondly, we assume that users may show similar behavior if they have similar interests. In like manner, similar POIs may have similar appeal to visitors. For the sparse check-in data, it inspires us to take a close approximation for a subject by its highly similar ones for enriching its behavior mining. When we cannot recommend accurate POIs for a user due to the incomplete and sparse trace, it might be a reasonable choice to consider the decisions of those similar to this user. Likewise, visitor candidates from similar places may be useful for predicting the future visitors.

Some preliminaries for the statistical analysis are done to validate our motivation above. Detailed analysis can be found in Section 3.1. Inspired by the assumptions and observations above, we propose a similarity-aware framework for the next location recommendation via adaptive bi-directional inference (SimBIN). It not only speculates when and where the user is likely to go from the view of the human but also predicts when and who will visit this place from the view of the location. To be specific, we first conduct the POI inference and user inference by employing the U-Net and P-Net to mine users' interests and POIs' appeal simultaneously. Then we take the similarity-aware enhancement to further enrich the candidates. For the POI inference, we take the highly similar users to approximate a user's behavior to mine more potential POI candidates. It is similar for the user inference part. For further optimizing the model, we propose the adaptive learning strategy to refine the model based on the conformity of outputs from two inference parts. Experimental results on real datasets show the significant improvements of SimBIN over the state-of-the-art approaches.

The main contributions are listed below:

- To tackle the sparsity issue, we propose a novel framework for the next location recommendation via bi-directional inference and similarity-aware enhancement. It explores users' interests for the POI inference and POIs' appeals for the visitor inference syn-

chronously. To our best knowledge, we are the first to consider the recommendation problem from both the human view and location view.

- We propose the similarity-aware enhancement to further alleviate the sparsity by mining highly similar users or POIs to enrich the candidates. We think that someone is similar to certain people rather than all users, thus we can take a close approximation by the similar ones for enriching this user's behavior mining. Similarly, we can approximate a POI's potential visitors according to POIs similar to it.
- In particular, we employ the adaptive learning strategy to refine the model based on the conformity of the bi-directional inference, making the model synchronously optimize the recommendation performance. Experimental results on public check-in datasets indicate that our SimBIN significantly outperforms the state-of-the-art models.

2. Related Work

2.1. Methods for Location Recommendation

There are three main types of location recommendation researches: Matrix Factorization (MF) [7], Markov Chain (MC) [8], and neural network (NN) based methods [9]. MF-based models can capture users' general preferences but fail to mine the sequential information. MC-based models are capable of learning the sequential transitions but they fail to learn users' long-term preferences through first-order MCs. Some hybrid models [10] use high-order MCs to mine more latent patterns. Wang combined the MF and MC for long- and short-term preferences [11]. Nevertheless, these approaches mainly aim to mine users' preferences by modeling the user-POI interactions or the transitions, suffering from the data sparsity heavily. This is because most users share few records online, making user training difficult with few samples.

In the past few years, neural networks have become the popular alternative for the recommendation task [12,13]. Recurrent neural network (RNN) including LSTM and GRU, have shown huge advantages in sequential modeling. STGN [14] incorporates long- and short-term preferences together. It also considers geographical relations between the nonconsecutive POIs. LSTPM [15] uses a nonlocal network and a geo-dilated RNN to learn the long- and short-term interests separately. Other deep models including memory network, attention mechanism and graph networks are also widely employed. Deep-Move [16] captures the multi-level periodicity by introducing two attention modules into the RNN to select highly correlated check-ins for the current one. Stp-udgat [17] takes a graph attention network to exploit users' preferences. It also mines new POIs from others. GeoSAN [3] employs the self-attention for the recommendation. It uses location gridding to model the geographical influence. However, it fails to mine the explicit information of spatial-temporal distances.

Some studies [18–26] introduce attention mechanism to learn the contribution weights of different factors or to capture sequence periodicity. STAN [27] considers correlations between both adjacent and non-adjacent check-ins. It first aggregates spatio-temporal correlations within user trajectories and then combines POI frequency for recommendation with self-attention models. Other studies further optimize recommendation models through multi-task learning or collaborative analysis [28]. GETNext [29] captures global transition patterns, users' general preferences, spatio-temporal contextual information, and time-aware category embeddings.

Some studies perform location recommendation based on graph models [30–37]. RELINE [38] captures social influence, geographical influence, temporal influence, and user preference dynamics by embedding eight relational graphs into a shared latent space. Some studies incorporate POI category or region information to enrich modeling [21]. ROTAN [39] further explores rotation-based time-specific modeling to handle finer-grained

temporal contexts. DCHL [40] learns collaborative, transition, and geographical influences separately by constructing disentangled hypergraphs. Some works also address the data sparsity issue [41,42].

Difference. To sum up, the above models mainly focus on modeling users' mobility patterns, i.e., mining the latent representation of one's interests in POIs. Nevertheless, they only consider the recommendation task from the human view, which intrinsically cannot escape the sparsity limitation of each individual's data.

This work is a comprehensive extension of BSDA [43]. In this work, we proposed a new similarity-aware enhancement module and an adaptive learning strategy, so that the model can effectively enrich more appropriate candidates for a target and synchronously optimize the recommendation performance. Detailed theoretical research and explanations are included in Section 4.

2.2. Data for Location Recommendation

In addition to the check-in data, cellular data and GPS log are also used. Compared with the check-in data from LBSNs, cellular data and GPS data are relatively dense because they can be collected by second while users check in online sparsely. However, we can only access the raw GPS coordinates from cellular data and GPS data.

For real application scenarios, cellular data and GPS data are not easy to get. Here are some reasons. Firstly, we have no public access to cellular data. Since the next location recommendation usually benefits users and business, publicly available data is necessary. However, the cellular data and GPS data may involve the privacy-preserving problem [44]. Secondly, cellular data and GPS data are usually collected by second. However, the dense sampling usually means high cost. In addition, it is impossible to obtain users' private location information anytime. Thirdly, the computational complexity is a new problem if we take all fine-grained data for training to model users' behavior patterns. Otherwise, we need to pre-process the raw data through stay points detection before we apply these data for the location recommendation, which brings extra work and errors.

Different from cellular data and GPS data, check-in data is easily available from LBSNs, where users share their lives as they wish. Moreover, besides GPS coordinates, the check-in records usually contain the POI ID, semantics, and social relationships, which are helpful for understanding users' complex mobility patterns. Therefore, the next location recommendation on check-in data becomes popular.

2.3. Similarity-Based Methods

Data sparsity has always been the most difficult challenge due to the distinct characteristic of the check-in data. Some approaches leverage multiple contexts to enrich the representation [2]. Some works introduce the social links [45]. For instance, Xu et al. introduced the location propagation for the location prediction based on the social ties, textual contents, and users' residence locations [46]. ARNN [4] exploits the transitions of neighbors through the heterogeneous contexts to tackle the sparsity problem. However, existing methods that employ social links for recommendations do not always work because the social relationship may not mean the friends' similar behavior patterns.

Some other works involve the embeddings to solve the sparsity issue [3–5]. For instance, for modeling the implicit relevance of neighbors, NGAT4Rec [47] adopts neighbor-aware graph attention to determine weights to each neighbor. Then it produces the next layer embedding for every node via fusing the neighbor embeddings according to the weights.

Difference. Our motivation for similarity-aware enhancement is different from existing works that involve embedding or social links. Here three points need to be stressed. (a)

Our similarity-aware enhancement is to reorder the recommended candidate list. Existing methods usually share parameters or use embedding during the sequential training process, but the reranking task proposed in this paper occurs after the sequence training. (b) The similarity matrices in this paper are to exploit the potential candidates for the recommendation list. Our similarity is more reasonable because the possible candidates are reordered according to similar patterns, rather than just sharing data in the input. (c) Our similarity-aware enhancement focuses on a group of users that are highly similar in behavior patterns, not only socially connected. We declare that similar behavior patterns can make more appropriate recommendations, while the social relationships or intersections don't exactly correspond to similar behavior.

2.4. User Prediction

Feng et al. considers a new research problem of predicting visitors [48]. They introduced POI2Vec that can predict the future users by learning user interest from POI sequence. Thus POI2Vec is more close to our U-Net that models users' patterns. In other words, it is the same as the traditional location recommendation framework from the view of the human.

Difference. The differences between our SimBIN and POI2Vec are mainly reflected in the following aspects. (a) We adopt the bi-directional inference with two sequential modeling modules, where users and POIs are treated as subjects respectively. (b) We utilize the U-Net to mine users' patterns and POIs' sequential transitions. (c) In the P-Net, we aim to learn the visitor transition and the locations' patterns. Besides, we further employ the similarity-aware enhancement to explore more potential candidates.

3. Preliminary

3.1. POI Behavior Analysis

We conduct preliminary experiments for the POI behavior analysis. It serves two purposes: (a) To verify our assumptions about the POI behavior. (b) To choose indispensable features for modeling POIs' appeal to users and then predicting the next visitor. All preliminary statistics and conclusions are based on the whole Gowalla rather than a deliberately chosen subset.

3.1.1. Personalized Appeal

In the literature, we know that users have personalized interests in various POIs. Inspired by this, we assume that POIs have personalized appeals to users as well. Here we investigate whether POIs have similar behavior patterns by counting visitors with different visiting frequencies for each POI.

Figure 2a is to verify whether a POI has distinct appeals to the users. Only when this condition is met can our user inference for a POI be meaningful. For each POI, we divide its visitors into three sectors according to the visiting frequency. Moreover, a POI may usually attract users nearby, leaving the visitors within a small range. Here we pay attention to the number of visitors who visits this POI less than two times or more than eight times. We randomly select a subset for visualization several times and the results are always similar. Hence, to visualize the differences among POIs more directly, we only show one random subset. For each POI, we find that some visitors visit persistently while some visit rarely, demonstrating POIs' personalized appeal to users, confirming our hypothesis. Therefore, it is reasonable to do the bi-directional inference from the view of the human and location.

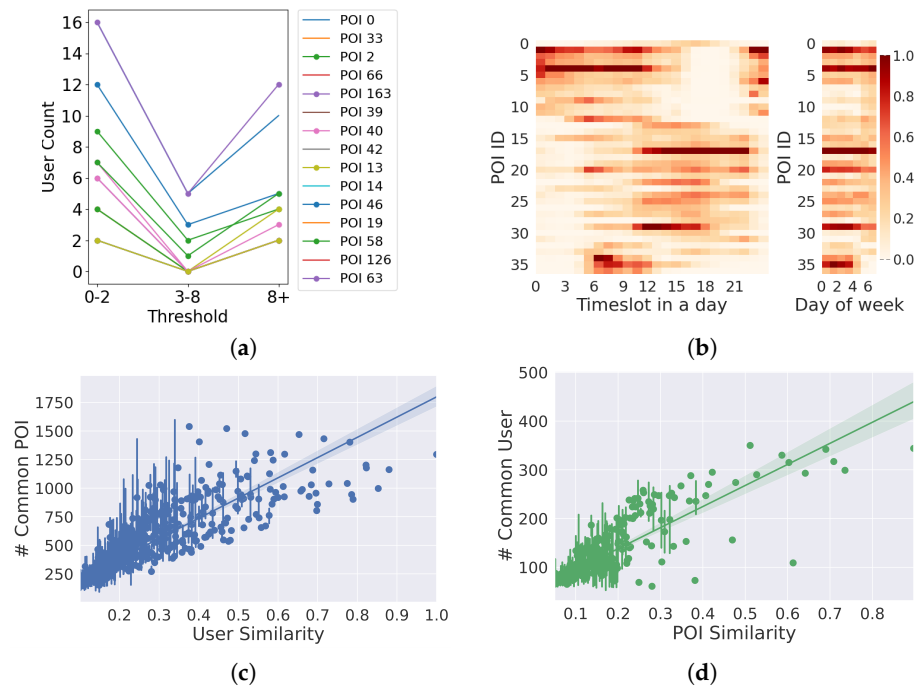


Figure 2. The preliminary analysis. (a) is to verify whether a POI has distinct appeals to the users. The horizontal axis is the threshold of the visiting times of a user. (b) is to investigate the POIs' temporal patterns on different scales. (c,d) are scatter plots to explore the similarities among users and POIs separately.

3.1.2. Temporal Patterns

Next, we intend to verify the feasibility of the idea of next visitor inference. In current location recommendation models that are from the view of the human, users' historical check-ins are usually represented as sequences, and the temporal patterns are of great importance for users' behavior modeling. We investigate whether the POIs have similar temporal patterns. Only if this condition is satisfied can our user inference for a POI be realized by modeling sequential patterns.

Figure 2b is to validate whether different POIs have distinct temporal patterns on two scales. Here we only select a random subset for visualization due to the space limitation. As depicted in Figure 2b, we can find that densities of visitors at two scales show a distinct distribution, demonstrating that POIs' behavior exhibits temporal patterns. This is consistent with our assumption. Considering that this phenomenon in POI behavior is similar to that in the user behavior, we use recurrent networks to model the POIs' sequential information and mining latent temporal patterns. Moreover, due to the difference between the daily pattern and weekly pattern, we consider both the date feature and time slot feature in the sequence modeling, i.e., day of the week and hour of the day.

3.2. User Similarity

We further conduct preliminaries with two purposes: (a) To verify our assumption that targets with similar interests may show similar behavior. (b) To explore significant factors for the similarity-aware enhancement.

In this paper, we declare that users may have similar mobility patterns if they have common interests. Additionally, exploring the appropriate similarities to group similar users or POIs is important for the similarity-aware enhancement. In order to satisfy both, we should analyze the proposed user similarity and users' common interests together. If two users share more common visiting POIs, they will show more similar mobility patterns.

Only in this way can we use similar users to mining the potential POIs for the target user. Therefore, our proposed user similarity should be proportional with the common interests.

The user similarity is determined based on whether they appear in a same POI in the same day. Figure 2c is the scatter plot, illustrating the relationship between the similarity of each user pair and the number of locations they all visited. We depicted all the data on the figure to measure the similarity. As shown in the figure, the two have a proportional relation. This result inspires us that similar users have similar behaviors. Hence, similar users' decisions can be considered when the next location recommendation is hard due to a user's sparse trace. Meanwhile, this indicates that our proposed user similarity definition would be valid for enhancing the location recommendation.

3.3. POI Similarity

The process of selecting the POI similarity definition is similar to the above. We declare that POI may have similar appeal to visitors if they have highly common visitors. Similar to Section 3.2, we should analyze our proposed POI similarity and POIs' common interests together.

We judge the POI similarity according to whether they attract the same users on the same day. If two POIs have more common visitors, they will show more similar attraction to people. Figure 2d demonstrates the scatter plot of POI similarity versus the number of common visitors of each POI pair. We can find that the two are highly related. This enlightens us that future visitors from the most similar POIs are possible choices of the focused POI. In the meantime, this demonstrates that our proposed POI similarity would be useful for the user prediction performance.

Based on the discussion above, we give detailed definitions of similarities in the next section.

3.4. Definitions and Problem Formulation

User similarity. The user similarity is developed to measure how similar a user pair's behavior patterns are. It is defined according to whether the two users come to the same POI over a similar period. We calculate the overlap of the user pair's POIs: $\tau_{m,n} = \frac{|L_m \cap L_n|}{|L_m|}$, where L_m denotes the set of all locations that the user m visits.

POI similarity. The POI similarity is to measure how similar a POI pair's attraction is. The similarity $\delta_{p,q}$ between the POI l_p and the POI l_q is the number of days POIs l_p and l_q attract the same visitor. We calculate the overlap of days when they have the same visitor: $U_{p,d} \cap U_{q,d} \neq \emptyset$, where $U_{p,d}$ denotes the visitors of POI p on day d .

Check-in point of U-Net. For a user u , when he/she checks in a POI l at timestamp t , the check-in point is recorded as $r_t^u = \langle u, t, lon, lat, l \rangle$. lat and lon represent the geographical coordinates of l .

Visiting point of P-Net. Since the target in the P-Net is a POI, the check-in point is denoted as $f_t^l = \langle l, t, lon, lat, u \rangle$.

POI's visiting sequence of P-Net. Given a POI l , the visiting sequence is represented as a chronologically collection of visiting points $\mathcal{F}_l = \{f_{t_1}^l, f_{t_2}^l, \dots, f_{t_{T'}}^l\}$, where T' is the length of sequence.

The next user prediction. Formally, given a POI's visiting history in the last $T' - 1$ steps, the main purpose of the next user prediction task is to predict the top- k most possible visitors $\{\mathcal{U}_k\}$ that will visit this POI at the next time $t_{T'}$.

User's trajectory of U-Net. Given a user u , his/her trajectory is represented as a chronologically collection of check-in points $\mathcal{R}_u = \{r_{t_1}^u, r_{t_2}^u, \dots, r_{t_T}^u\}$, where T is the length of trajectory.

Problem formulation. For the next location recommendation task, the input are the user's historical check-ins in the last $T - 1$ moments $\{r_{t_1}^u, r_{t_2}^u, \dots, r_{t_{T-1}}^u\}$. The output is the POIs' score list at the next time t_T . The main goal is to predict the top- k POI candidates $\{\mathcal{L}_k\}$ by ranking the scores.

4. Proposed Methodology

4.1. Overview

The framework of our SimBIN is illustrated in Figure 3. It mainly consists of (a) the U-Net for modeling users' interests, (b) the P-Net for modeling POIs' appeals, (c) the similarity-aware enhancement, and (d) the recommendation layer for the bi-directional inference.

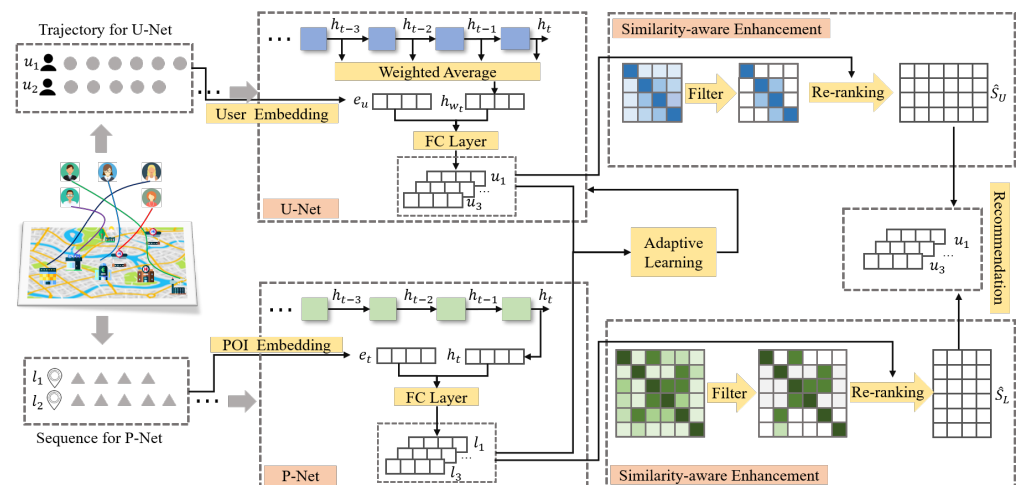


Figure 3. The framework of SimBIN. It mainly consists of bi-directional inference and similarity-aware enhancement. The bi-directional inference is mainly implemented by the U-Net and P-Net, which infers the next POI and the next visitor separately. The adaptive learning is to refine the model based on the conformity of the two inference parts. The similarity-aware enhancement is to select the highly similar users/POIs to enrich the candidates. Finally, it fuses the two parts to do the final speculation for the next location recommendation.

To be specific, we design the U-Net to explore user interest to infer where a user will visit. Meanwhile, in the P-Net, we aim to mine POIs' appeal to users to infer who might visit the POI. To further improve the performance, we propose the adaptive learning strategy to refine the model based on the conformity of the two inference parts. After that, we develop the similarity-aware enhancement approach to reorder the inference list by exploiting the potential candidates. On one hand, it provides possible POI candidates with highly similar users for improving the POI inference from the U-Net. On the other hand, it provides possible user candidates with highly similar POIs for improving the user inference from P-Net. Finally, we can recommend the next POI via bi-directional inference.

4.2. U-Net for POI Inference

The U-Net is to infer where a user will visit, from the view of the human. Its primary goal is to explore users' interests in POIs, thus it is the traditional location recommendation task. In the U-Net, the input are users' check-in sequences, and the output are the POIs' score lists. We assume that there exist sequential patterns in users' behavior and the decision on the next POI is influenced by recent trajectory. Hence we need the U-Net module to capture sequential patterns and the influence of the historical check-ins. Benefited by Flashback [6], we take it as baseline to capture the sequential transitions for inferring the next POI. In Flashback, it chooses the RNN, GRU, LSTM as the basic unit to implement the task and the RNN unit performs best. Thus we choose the RNN as the basic unit

for U-Net in this work. Flashback has great advantages on the sequential modeling than the vanilla RNN. It aggregates the historical hidden states by computing their weighted average through a decay weight $\omega(\Delta T, \Delta D)$, where the output is denoted as h_{w_t} . The decay weight $\omega(\Delta T, \Delta D)$ is adopted to measure the relevance between the recent hidden state h_j and the current one $h_t, j < t$: $\omega(\Delta T_{i,j}, \Delta D_{i,j}) = \text{hvc}(2\pi\Delta T_{i,j})e^{-\alpha\Delta T_{i,j}}e^{-\beta\Delta D_{i,j}}$, where ΔT and ΔD are the temporal interval and the geographical distance respectively. The $\text{hvc}(\cdot)$ is the Havercosine to capture the temporal periodicity. Thus the U-Net can explore the sequential transitions, temporal patterns, and the relevance of recent check-ins.

Firstly, for each user, we split his/her historical trajectory $\mathcal{R}_u$ into segments and feed them into the RNN block. The hidden state at time t is represented as h_t . Next, the spatial and temporal distances are leveraged to produce the decay weight $\omega(\Delta T, \Delta D)$. Then we can calculate the aggregated output $h_{w_t} = \frac{\sum_{j=0}^t w_j * h_j}{\sum_{j=0}^t w_j}$. Moreover, for the personalized user modeling, we merge the user embedding e_u and the aggregation h_{w_t} to generate the possible POI score list s_U^u .

4.3. P-Net for User Inference

On the basis of the discoveries from Figures 2a,b, we believe that the POI could attract people autonomously, and they appear personalized patterns. Thus we expect the P-Net to mine POIs' appeals from the sequences and infer the next visitor. Moreover, as discussed in Section 3.1, we find that POIs' behavior shows temporal patterns on two time scales. Hence, we intend to explore POIs' dynamic attraction by the RNNs. In particular, we consider different slots of the day and days of the week when mining temporal patterns. Different from the U-Net, we don't consider the relationships among the POI's visitors, hence we remove the flashbacks operation on the hidden states in RNNs.

To be specific, given the historical records, we denote them as POIs' visiting sequences. For the POI l , its visiting point is represented as $f_t^l = \langle l, t, lon, lat, u \rangle$, where f_t^l denotes the raw data. We further process the visiting point into a new form denoted as $a_t^l = \langle l, u, s, d \rangle$. Here we make use of time t on two scales, i.e., the slot of the day and the day of the week, denoted as s and d separately. Then the whole visiting sequence is denoted as $A_l = \{a_{t_1}^l, a_{t_2}^l, \dots, a_{t_{T'}}^l\}$, where T' is the sequence length. In the P-Net, the input are POIs' visiting sequences, and the output are visitors' score lists. We employ the RNN unit to learn POIs' appeal. The current hidden state h_t is taken to infer the next user. Additionally, for the personalized appeal modeling, we merge the POI embedding e_l and the current hidden state h_t . Finally, it outputs the possible user score list s_l^l for the POI l .

4.4. Similarity-Aware Enhancement

4.4.1. Similar Users for Potential POIs

Due to the data sparsity, the inference performance of the U-Net will not be satisfactory if there is too few records for a user. To resolve this issue, we employ the similarity-aware enhancement to explore potential POIs for the recommendation through a user's similar ones.

Specifically, to measure how similar a user pair's patterns are, we generate the user similarity matrix Corr_U , where $\text{Corr}_U \in \mathbb{R}^{M \times M}$. Each item $\tau_{m,n}, m, n \leq M$ in Corr_U is the similarity between user u_m and user u_n . Corr_U is normalized to $[0, 1]$. The detailed definition of user similarity is in Section 3.2.

For further refining the similarity, we select the highly similar users and then reorder the POI candidate list according to the selected users. Usually, users have personalized mobility patterns, showing differences in their preferences on POIs. Although we assume that users with similar interests may visit the same places, it does not mean that we need to consider all users. This is easy to understand since not all users are relevant enough

to provide useful candidates for a user. In BSDA [43], for each user, it treats all users as similar. For the cross-user association, it normalizes the similarity vector and does the linear combination with all other users. On one hand, some users who have no social relationship or similar behavior patterns may become noisy associations. On the other hand, due to the large quantities of users, the weights of most similar users may be submerged, making the association less effective. Therefore, it is more essential and appropriate to do the cross-user representation for u with the most similar users rather than all users. To tackle this problem, we develop an improved approach, aiming to take a close approximation for a user by the highly similar users for enriching the behavior mining.

In detail, for each item in the user similarity matrix $\mathbf{Corr}_U$, we empirically set a threshold θ_u to eliminate the irrelevant users and reserve the distinct ones. We adopt the filtering and get a new user similarity matrix $\hat{\mathbf{C}}_U$. The equation is as follows:

$$\hat{\mathbf{C}}_U = \begin{cases} \tau_{m,n}, & \text{if } \tau_{m,n} \geq \theta_u \\ 0, & \text{if } \tau_{m,n} < \theta_u, \end{cases} \quad (1)$$

where $\tau_{m,n} \in \mathbf{Corr}_U$.

Meanwhile, we generate a POI candidate score matrix $\mathbf{S}_U = [\mathbf{s}_U^{u_1}; \mathbf{s}_U^{u_2}; \dots; \mathbf{s}_U^{u_M}]$, $\mathbf{S}_U \in \mathbb{R}^{M \times N}$. $\mathbf{s}_U^u$ is the possible POI score list for user u at time t_T , $u \in \mathcal{U}$. We declare that users with similar behavior are more likely to share common interests. When the POI inference is poor due to the user's sparse trace, the considerations of his/her similar ones could be reasonable choices. Therefore, we integrate the POI score matrix $\mathbf{S}_U$ and the filtered similarity matrix $\hat{\mathbf{C}}_U$ to achieve the reordered one $\hat{\mathbf{S}}_U$ for the reranking:

$$\hat{\mathbf{S}}_U = \hat{\mathbf{C}}_U * \mathbf{S}_U, \quad (2)$$

where $\hat{\mathbf{S}}_U \in \mathbb{R}^{M \times N}$. Lastly, we can get the reordered POI score list $\hat{\mathbf{s}}_U^u$ of the user u from $\hat{\mathbf{S}}_U$.

4.4.2. Similar POIs for Potential Visitors

The visitor inferring performance of a POI is unsatisfactory when few visitors come to this POI. To tackle the issue above, similar to the approach for the U-Net, we adopt the similarity-aware enhancement to explore potential visitors through a POI's similar ones.

Specifically, to measure how similar a POI pair's attraction is, we generate the POI similarity matrix $\mathbf{Corr}_L$, where $\mathbf{Corr}_L \in \mathbb{R}^{N \times N}$. Each item $\delta_{p,q}$, $p, q \leq N$ in $\mathbf{Corr}_L$ is the similarity between two POIs. $\mathbf{Corr}_L$ is normalized to $[0, 1]$. The detailed definition of POI similarity is in Section 3.2.

We further select the highly similar POIs to refine the similarity. Specifically, for each item in the POI similarity matrix $\mathbf{Corr}_L$, we empirically set a threshold θ_l to eliminate the irrelevant POIs and reserve the distinct ones. Then we get a new POI similarity matrix $\hat{\mathbf{C}}_L$. The equation is as follows:

$$\hat{\mathbf{C}}_L = \begin{cases} \delta_{p,q}, & \text{if } \delta_{p,q} \geq \theta_l \\ 0, & \text{if } \delta_{p,q} < \theta_l, \end{cases} \quad (3)$$

where $\delta_{p,q} \in \mathbf{Corr}_L$.

Meanwhile, we generate a user candidate score matrix $\mathbf{S}_L$, where $\mathbf{S}_L = [\mathbf{s}_L^1; \mathbf{s}_L^2; \dots; \mathbf{s}_L^N]$, $\mathbf{S}_L \in \mathbb{R}^{N \times M}$. $\mathbf{s}_L^l$ is the user score list for POI l at time t_T , $l \in \mathcal{L}$. We assume that similar POIs appeal to common users. The user candidates of the most similar POIs may benefit the user inference. Therefore, we integrate the user candidate score matrix $\mathbf{S}_L$ and the POI similarity matrix $\mathbf{Corr}_L$ to generate the reordered score matrix $\hat{\mathbf{S}}_L^l$:

$$\hat{\mathbf{S}}_L = \hat{\mathbf{C}}_L * \mathbf{S}_L, \quad (4)$$

where $\hat{\mathbf{S}}_L \in \mathbb{R}^{N \times M}$.

4.5. Bi-Directional Inference

The U-Net and the P-Net are integrated for the bi-directional inference to improve the performance of the recommendation. Thus we can get more possible and appropriate POI candidates because we consider the recommendation problem from both the view of the human and the location. In detail, we adopt a fusion layer to integrate the reordered POI score matrix $\hat{\mathbf{S}}_U$ and the user score matrix $\hat{\mathbf{S}}_L$.

Fusion strategy. We discuss some alternative fusion strategies in detail. The purpose of the fusion layer is to assign each candidate a new probability by combining the original POI and user candidate score matrices. The element in the POI candidate score matrix represents the probability of this POI being visited. The element in the user candidate score matrix represents the probability of this user visiting a POI. The original POI candidate score from the U-Net may have some deficiency and inconsistency with the results from the P-Net. For example, a POI may be very important for a user and he/she is very likely to visit this POI. However, the probability of this POI is very low limited by the number of samples or the sequential transition. Moreover, the probability of a POI is relatively high in the original POI candidate score list, but this POI is unlikely to be visited by this user according to the user candidate score list. In this case, we hope to consider the probability of this POI being visited and the probability of this user visiting a POI simultaneously. Then we use the pooling mechanisms to represent the information from two modules and the function is as below:

$$\mathbf{S}_F = \text{POOL}([\hat{\mathbf{S}}_U, \hat{\mathbf{S}}_L^T]), \quad (5)$$

where the *POOL* is used to calculate the aggregated probability of the outputs from the U-Net and the P-Net.

We want to find the determining factor for the next POI from the human side or the location side. Hence we choose the max pooling operation for fusion, which takes the larger item at each point of $\hat{\mathbf{S}}_U$ and $\hat{\mathbf{S}}_L^T$. The output is denoted as $\mathbf{S}_F \in \mathbb{R}^{M \times N}$. The key formula of the hybrid representation $\mathbf{S}_F$ is as follows:

$$\mathbf{S}_F = \text{maxpooling}([\hat{\mathbf{S}}_U, \hat{\mathbf{S}}_L^T]), \quad (6)$$

For the user u , the final POI score list $\mathbf{s}_F^u$ can be found from $\mathbf{S}_F$. Then, the top- k POIs can be selected for the location recommendation by ranking the score list $\mathbf{s}_F^u$.

The pooling mechanism in our work is to fuse the results from the U-Net and P-Net to conduct the bi-directional inference. It is inspired by the relevant works that use the pooling mechanism for data aggregation or feature fusion [49]. For instance, Wang et al. take different aggregation operations to combine IME and OME module [50], including the max pooling, concatenation operation, and the average pooling. The results indicates the poor performance of max pooling. In the work [51], researchers integrate different pooling strategies and find the mean pooling performs the best. Inspired by these, we investigate different fusion strategies in our experiments and find the max pooling is the most suitable approach for aggregation.

4.6. Model Learning

The adaptive learning strategy based on conformity. To further improve the performance, we propose the adaptive learning strategy to refine the model based on the conformity of outputs from two inference parts. The loss is calculated according to the outputs from two modules and then feedback for adaptive learning. As a result, we can find the inconsistency among the two modules, i.e., a POI l is one of the top- k candidates in the

output of U-Net for the user u but the user u visits the POI l with very low probability in the output of P-Net for the POI l . If the loss is below a certain threshold $\tau = 0.05$, we think there is consistency between the U-Net and P-Net. Here the loss serves as a monitoring signal rather than the ultimate optimization goal. When the loss remains below τ for 5 epochs, we cease updating to prevent unnecessary computational overhead.

The loss is defined as below:

$$loss_a = \sum |(\alpha_{i,j} - \beta_{i,j})|, \alpha_{i,j} \in S_U; \beta_{i,j} \in S_L^T. \quad (7)$$

Parameter learning. Our SimBIN mainly needs the following key parameters to learn: (a) RNN parameter matrices and RNN bias vectors in the U-Net and P-Net, (b) the decay weight $\omega(\Delta T_{i,j}, \Delta D_{i,j})$ in the U-Net. Algorithm 1 outlines the training procedure of SimBIN.

Algorithm 1 SimBIN

Input: Users' check-in trajectory: $R^U = \{R_{u_1}, R_{u_2}, \dots, R_{u_{|U|}}\}$; POIs' visiting sequences: $\mathcal{F}^L = \{\mathcal{F}_{l_1}, \mathcal{F}_{l_2}, \dots, \mathcal{F}_{l_{|L|}}\}$;
Output: Probability Pr for location recommendation.
1: // **U-Net**
2: Initialize the parameters;
3: **repeat**
4: Get recurrent hidden state: $h_i = RNN(l^i, h_{i-1})$.
5: Compute the weight decay:
6: $\omega(\Delta T_{i,j}, \Delta D_{i,j}) = \text{hvc}(2\pi\Delta T_{i,j})e^{-\alpha\Delta T_{i,j}}e^{-\beta\Delta D_{i,j}}$.
7: Compute the aggregated hidden state: h_{w_t} , t is the last step.
8: $s_U \leftarrow \text{softmax}(W_f h_{w_t} + b_f)$ // **Inference**
9: **until** (convergence)
10: // **P-Net**
11: Initialize the parameters;
12: **repeat**
13: Get recurrent hidden state: $h_i = RNN(u^i, h_{i-1})$.
14: Take the hidden state at the last time step: h_t, t .
15: $s_L \leftarrow \text{softmax}(W_f h_t + b_f)$ // **Inference**
16: **until** (convergence)
17: // **Similarity-aware enhancement**
18: Compute the adjusted score matrix $\hat{S}_U = \hat{C}_U * S_U$;
19: Compute the adjusted score matrix $\hat{S}_L = \hat{C}_L * S_L$;
20: // **Bi-directional inference**
21: Get the final score matrix $S_F = \text{maxpooling}([\hat{S}_U, \hat{S}_L^T])$
22: **return** Pr

5. Experimental Setup

5.1. Research Questions

We intend to evaluate the effectiveness of SimBIN and to answer the questions below:

RQ1: What is the performance of SimBIN for the next location recommendation? Does it outperform state-of-the-art approaches?

RQ2: Do P-Net and U-Net contribute to the performance of SimBIN?

RQ3: How well does the SimBIN perform by introducing the similarity-aware enhancement?

RQ4: How do the parameters affect the performance of SimBIN?

RQ5: How about the sensitivity of SimBIN to data sparsity?

5.2. Data Preprocessing

We perform experiments on the following public check-in datasets.

- Gowalla is obtained from the Gowalla Website. The period is from Feb. 2009 to Oct. 2010.
- Foursquare dataset [52] is collected from the other platform. We take the data from Apr. 2012 to Jan. 2014 in New York.

We list the reasons why the two datasets are appropriate. (a) They are widely used in the location recommendation task, commonly acknowledged as the data for evaluation. (b) They have distinct characteristics including the data sparsity, time span, user groups, and platforms, which can help verify the robustness and generalization of the recommendation method. The records in two datasets include the information of userID, POIID, timestamp, and GPS coordinate. The first 80% check-ins are training data, and the others are testing data. In this work, we pre-process the data following our baseline method [6]. The statistical description is shown in Table 1.

Table 1. Data statistics. The density refers to the average records per user per POI.

Dataset	Gowalla	Foursquare
#user	52,979	46,065
#check-in	3,300,986	945,032
#POI	121,851	69,005
Avg. #check-in per user	62.31	20.51
Avg. #POI per user	2.29	1.50
Min #check-in per user	10	100
Min #user per POI	1	1
density	0.00051	0.00030

5.3. Compared Methods

We compare our SimBIN with the following models:

RNN [53] is a vanilla RNN architecture that simply learns the sequential information for the recommendation task.

PRME [8] takes sequential information, individual preference, and spatial effect to improve performance. It models the sequential transitions by the personalized ranking metric embedding.

STRNN [54] extends RNN and introduces the temporal and spatial contexts. It utilizes the transition matrices for different temporal and spatial distances separately.

DeepMove [16] predicts the mobility by an attention-based RNN. It captures sequential transitions with a multi-modal embedding recurrent network and captures the multi-level periodicity with a historical attention model on sequence-level.

LBSN2Vec [30] is a hypergraph embedding approach for automatic feature learning, which uses the categorical information and social relationships. It first uses the random-walk-with-stay to represent user check-ins and friendships. Then it learns node embeddings for feature learning.

STGN [14] is a spatiotemporal gated network to learn the spatial-temporal relations between successive check-ins. In detail, it controls the short- and long-term interest updates respectively by introducing new time gates and distance gates.

Flashback [6] learns the sequential patterns based on the RNN. In detail, it calculates the predictive power of the past hidden states and then uses it to aggregate them to predict the next location.

STAN [27] exploits relative spatiotemporal information of all historical trajectories by the self-attention. It captures the interaction between non-consecutive check-ins and non-adjacent locations.

MobTCast [28] is a Transformer-based context-aware network. It explores the influence of temporal, semantic, social, and geographical contexts.

SGRec [55] is a GNN-based method that designs a Seq2Graph augmentation to mine collaborative signals among POIs.

GSTN [31] captures complex geographical influences by POI semantic graphs and learns temporal dependencies using Time-LSTM, after which the spatial and temporal dependencies are aggregated via the attention.

ContraRec [56] devises context-target and context-context contrastive learning tasks to promote representational similarity among trajectories that have the same target.

Diff-POI [57] utilizes two distinct graphs for spatiotemporal representation learning and introduces a diffusion-based sampling approach to explore the spatial preferences of users.

AGRAN [58] leverages adaptive graph learning for POI representation learning, with temporal and spatial information being extended to facilitate the self-attention mechanism.

GUGEN [59] is a graph-enhanced model that fuses information from both global and user-specific perspectives by constructing a global POI relation graph and individual user historical graphs.

DASGRAN [60] enhances recommendation performance through simpler adaptive graph learning and an augmented self-attention mechanism, while reducing domain knowledge dependency and mitigating over-smoothing and over-correlation issues.

We summarize the baselines and their characteristics in Table 2. We compare our SimBIN with the state-of-the-arts on POI category, the view of targets, similarity, and embedding. The reasons are as follows: (a) Our SimBIN considers the next POI problem from both the view of the human and the location. (b) To some degree, considering from the view of the location is related to the method that integrates the geographical semantics. (c) Our proposed similarity-aware enhancement is related to the method that infers the concept of similarity to enrich the information.

Table 2. The characteristics of compared models. × indicates that the model does not involve the corresponding characteristics, while ✓ indicates that it does.

Methods	Category	View	POI Category	Similarity	Embedding
RNN	RNN	×	×	×	✓
PRME	MC + ME	×	×	×	✓
STRNN	RNN	×	×	×	✓
DeepMove	RNN	×	×	×	✓
LBSN2Vec	Graph	×	✓	×	✓
STGN	RNN	×	×	×	✓
Flashback	RNN	×	×	×	✓
STAN	Self-attention	×	×	✓	✓
MobTCast	Transformer	×	✓	✓	✓
SGRec	Graph + attention	×	✓	×	✓
GSTN	Graph + LSTM	×	×	×	✓
ContraRec	Transformer	×	×	×	✓
Diff-POI	Graph + Self-attention	×	×	×	✓
AGRAN	Graph + Self-attention	×	×	×	✓
GUGEN	Graph + Transformer	×	✓	×	✓
DASGRAN	Graph + Self-attention	×	×	×	✓
SimBIN	RNN	✓	×	✓	✓

5.4. Evaluation Metrics

We take the average Accuracy@K (Acc@K for short) and Mean Reciprocal Rank (MRR) to evaluate the performance of the recommendation. Acc@K is the percentage of true recommendations in the recommendation list. K is the list length. MRR is a position-aware metric.

5.5. Parameters

In our experiments for this paper, the learning rate is initialized as 0.01. The hidden size in the RNN is set to be 10 and 20 respectively. The minimum sequence length is 20.

In Flashback, the recurrent network can be RNN, GRU, or LSTM. We choose the basic RNN unit in both U-Net and P-Net. The default batch size is 50. The embedding size is empirically set to be 10. Parameters α and β in the U-Net are the same as that in [6]. We repeat the experiment 15 times and average the results to obtain more persuasive results.

6. Results and Discussions

6.1. Recommendation Effectiveness (RQ1)

The results of compared methods and our SimBIN on Gowalla and Foursquare are illustrated in Table 3. From the table we can get the following observations:

Table 3. Overall performance comparison on Gowalla and Foursquare. The best results are in bold. Improved (%) denotes the performance improvement of SimBIN over the best baseline. The second to last line indicates the standard deviation and significance. ** indicates $p < 0.01$ compared to the best baseline using the paired Wilcoxon signed-rank test over 15 independent runs.

Methods	Foursquare				Gowalla			
	Acc@1	Acc@5	Acc@10	MRR	Acc@1	Acc@5	Acc@10	MRR
PRME	0.0982	0.3167	0.4064	0.2040	0.0740	0.2146	0.2899	0.1503
LBSN2Vec	0.2190	0.3955	0.4621	0.2781	0.0864	0.1186	0.1390	0.1032
RNN	0.1824	0.4334	0.5237	0.2984	0.0881	0.2140	0.2717	0.1507
STRNN	0.2290	0.4310	0.5050	0.3248	0.0900	0.2120	0.2730	0.1508
GUGEN	0.2068	0.4696	0.5412	0.3271	0.0952	0.2328	0.2808	0.1583
DeepMove	0.2400	0.4319	0.4742	0.3270	0.0625	0.1304	0.1594	0.0982
STGN	0.2094	0.4734	0.5470	0.3283	0.0624	0.1586	0.2104	0.1125
Contrarec	0.2176	0.4720	0.5554	0.3424	0.1133	0.2909	0.3701	0.1965
SGRec	0.2352	0.5230	0.6139	0.3786	0.1145	0.2680	0.3501	0.1921
STAN	0.2400	0.5380	0.6214	0.3739	0.1165	0.2733	0.3467	0.1917
Flashback	0.2496	0.5399	0.6236	0.3805	0.1158	0.2754	0.3479	0.1925
GSTN	0.2510	0.5499	0.7112	0.4227	0.1220	0.2910	0.3813	0.2014
MobTCast	0.2563	0.5688	0.6399	0.3905	0.1243	0.2959	0.3702	0.2007
Diff-POI	0.2556	0.5916	0.6910	0.4005	0.1298	0.3110	0.3808	0.2132
AGRAN	0.2692	0.5846	0.6602	0.4231	0.1306	0.3206	0.3801	0.2172
DASGRAN	0.2922	0.6331	0.6870	0.4333	0.1398	0.3398	0.4070	0.2316
SimBIN	0.3199	0.6899	0.7417	0.4697	0.1488	0.3698	0.4361	0.2514
(mean $\pm$ std)	± 0.041 **	± 0.103 **	± 0.104 **	± 0.070 **	± 0.019 **	± 0.051 **	± 0.065 **	± 0.036 **
Improved (%)	9.464	8.979	7.968	8.402	6.463	8.836	7.145	8.560

OBS1. Our proposed SimBIN consistently outperforms all compared models, demonstrating its advantages for the next location recommendation. PRME gets the worst performance as expected. Although PRME integrates the sequential information and introduces a pair-wise ranking metric embedding to exploit the latent relationship among items by incorporating multiple factors, it only uses the first-order MC, failing to capture the relationships among items. Table 3 shows mean $\pm$ standard deviation and p value. ** indicates $p < 0.01$ compared to the best baseline using the paired Wilcoxon signed-rank test over 15 independent runs.

OBS2. LBSN2Vec achieves better performance than PRME. The result is as expected. LBSN2Vec uses the graph to represent the data and integrates the social relationship and semantic information of locations, which can take advantage of the intrinsic structure of the check-in data. Moreover, it considers both the co-location and social proximity when finding potential friendship. However, it performs worse than other RNN-based methods. This may be because it fails to capture some important information such as fine-grained temporal and spatial distance. In addition, the graph might not work if the data is too sparse to make the connections between nodes invalid.

OBS3. All RNN-based approaches get better performance than PRME because the RNN can model high-order sequential transitions. Overall, the results of RNN and STRNN are similar. They both integrate the temporal and spatial factors into the RNN unit, showing RNN's superiority of modeling the high-order sequential transitions. Specifically, STGN performs better than RNN and STRNN on Foursquare but worse on Gowalla. This is because STGN introduces new gates to capture the spatial and temporal relations between check-in points, indicating the advantage of the gates mechanism in STGN and the importance of transitions. Nevertheless, the difference in results on the two datasets indicates that the effect of gates is not remarkable on Gowalla. The performance of DeepMove on two datasets is similar to STGN. It uses attention to capture the multi-periodicity, but the result seems unsatisfactory on Gowalla. We explain this phenomenon as data dependence of the model.

OBS4. Flashback performs better than other RNN-based compared methods. It outperforms PRME significantly, because it models users' mobility by applying flashbacks in RNN. In this case the long- and short-term transitions are both considered. Moreover, it takes the transitions among several check-ins not only the successive pair. The excellent performance indicates the importance of the relationship among nonadjacent check-ins. STAN gets similar performance with Flashback, because it also considers the relationship among non-adjacent check-in points. MobTCast performs better than RNN-based methods on two datasets. It utilizes the transformer-based architecture to capture the complex periodicity of human mobility patterns. Moreover, it considers the social links and the semantic information of locations, benefiting its superior performance.

OBS5. The ContraRec method outperforms most RNN-based approaches, likely because it considers the similarity between trajectories. Methods such as GSTN, diff-POI, AGRAN, GUGEN, and DASGRAN all integrate self-attention or transformers with graph structures. Overall, these methods perform better than RNN-based ones because they incorporate graph knowledge and better exploit the structural information inherent in the data. AGRAN achieves superior performance, possibly due to its design of an adaptive graph learning to represent the spatiotemporal information of POIs. DASGRAN performs the best, likely because it further refines the adaptive graph learning to reduce information redundancy and mitigate the issue of excessive associations.

OBS6. SimBIN outperforms all compared approaches on Gowalla and Foursquare, indicating its effectiveness for the next location recommendation. SimBIN improves 8.0% and 7.1%, at Acc@10 on Foursquare and Gowalla separately compared with DASGRAN. The advantage of SimBIN is summarized as follows. (a) It is an RNN-based network that can capture sequential transitions effectively. (b) The bi-directional inference considers the candidates from both the view of the location and the human, which is of great importance for the improvement. (c) The proposed similarity-aware enhancement alleviates the sparsity problem to some extent by providing more possible candidates from similar ones.

OBS7. All methods obtain better performance on Foursquare than on Gowalla, in terms of Acc@K and MRR. This phenomenon may be caused by the differences between two datasets in check-in number, check-ins per user, and density. The data in Gowalla is more sparse so that modeling the appropriate information is more challenging (as analyzed in Section 1). Specifically, on one hand, the number of check-ins for each user may be less on Gowalla so that the model could not get enough samples for learning. On the other hand, the temporal interval and geographical distance may also influence the effectiveness of sequential learning between successive check-ins.

6.2. Analysis on Key Components (RQ2 & RQ3)

We further intend to verify the efficacy of each key component in SimBIN, including the similarity-aware enhancement, the P-Net for the user prediction, and the adaptive learning strategy. We implement the following variants: (a) SimBIN *w/o* user prediction. It contains the U-Net and the user-similarity enhancement. (b) SimBIN *w/o* POI-similarity. We remove the POI-similarity enhancement. (c) SimBIN *w/o* user-similarity. We remove the user-similarity enhancement. (d) SimBIN *w/o* adaptive learning. We remove the adaptive learning.

Figure 4 depicts the result of the ablation experiments. It shows that each key component is effective for the performance of SimBIN. The following are detailed discussions.

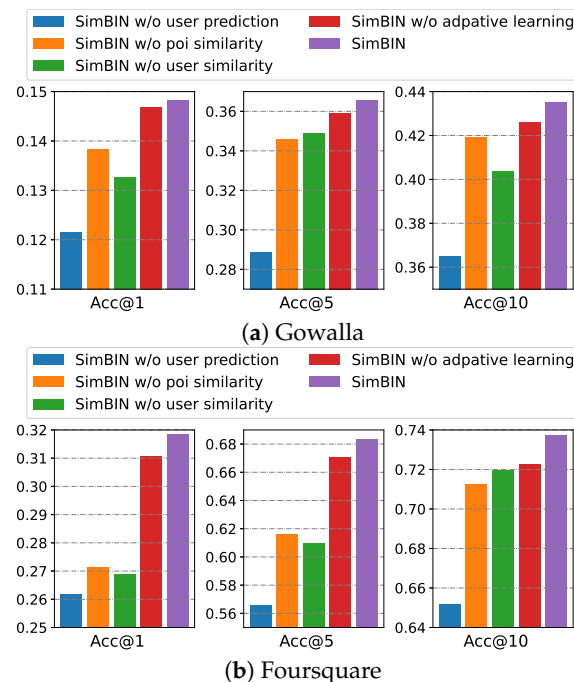


Figure 4. Comparison of variants on Gowalla and Foursquare. The results are about the effectiveness of key components in SimBIN.

The adaptive learning strategy. Figure 4 shows that SimBIN *w/o* adaptive learning gets worse performance than SimBIN, indicating the effective contribution of the adaptive strategy.

Similarity-aware enhancement. SimBIN *w/o* POI similarity and SimBIN *w/o* user similarity both perform slightly worse than SimBIN, indicating the effectiveness of the similarity-aware enhancement approach. We select the highly similar groups of a target user/POI and take a approximation of the behavior by the similar subset. The result demonstrate that we can predict the user's or POI's behavior through the similarity to alleviate the sparsity issue.

Bi-directional inference. SimBIN *w/o* user prediction performs significantly worse than SimBIN. It is because SimBIN *w/o* user prediction only considers the problem from the view of the human. This result indicates the importance of bi-directional inference.

User inference. We further discuss the performance of the U-Net and P-Net. In the P-Net, the input data is from part of the input of the U-Net. However, as illustrated in Table 4, the user inference is unsatisfactory. The explanation is that the data preprocessing for this paper is user-specific rather than POI-specific. Thus the user inference will perform poorly if there are no enough visiting records for most predicted POIs. In spite of this, the user prediction module indeed benefits the performance of SimBIN. We give a rational

explanation as follows. The user prediction module lists users' coming probability for a given POI. A POI may attract a specific group, limiting visitors to a relatively small set. In this way, this probability list can distinguish the users. Therefore, this collaborative analysis is meaningful because the user prediction can help find abnormal values and make the location recommendation list more appropriate through reranking.

Table 4. Performance of U-Net and P-Net on Gowalla and Foursquare.

Methods	Foursquare				Gowalla			
	Acc@1	Acc@5	Acc@10	MRR	Acc@1	Acc@5	Acc@10	MRR
U-Net for POI inference	0.2496	0.5399	0.6236	0.3805	0.1158	0.2754	0.3479	0.1925
P-Net for user inference	0.2015	0.4308	0.4597	0.2963	0.1013	0.1827	0.2187	0.1395

POI inference in the U-Net. In the U-Net, we mainly use the Flashback block, and the analysis of its performance has been discussed in Section 6.1. It is worth noting that our key components including the U-Net, P-Net, similarity-aware enhancement can be replaced separately, i.e., our model is easy to refine and can be generally improved.

6.3. Impact of Parameters (RQ4)

6.3.1. Training Parameter

We investigate the effect of key parameters on SimBIN including the hidden size and the learning rate. Because of the space restriction, we only illustrate experimental results on Gowalla, at Acc@10, as presented in Figure 5.

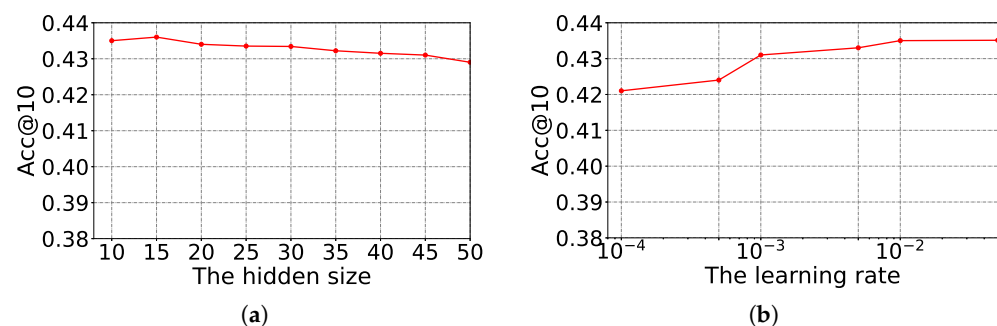


Figure 5. Impact of parameters. (a) is about the impact of the hidden size. (b) is about the impact of the learning rate.

The hidden size. We first study the effect of the hidden size of RNN in the U-Net and find the best setting for our model. From Figure 5a, we can find that the performance of SimBIN at Acc@10 first increases and then decreases as the hidden size varies from 10 to 50. This may be because the larger hidden size can involve more latent information but there will be overfitting with large size. For the trade-off between efficiency and accuracy, we set the hidden size as 10 in the U-Net.

The learning rate. Then we examine how the learning rate influences the performance of SimBIN. Figure 5b shows the result with different values of learning rate. We can find that the performance of SimBIN first increases and then changes little, indicating that the performance is not sensitive to the learning rate when it varies from 0.005 to 0.05. For the trade-off between computational complexity and prediction accuracy, the learning rate is set as 0.01.

6.3.2. The Filtering Threshold

We conduct some experiments to explore the influence of the filtering threshold in the similarity matrices. Here we implement three variants of SimBIN with different filtering

strategies for the similarity-aware enhancement. The variants are below: (a) SimBIN_v1: we replace the similarity-aware enhancement in SimBIN with Cross-User and Cross-POI Associations, i.e., we leave all items unchanged in matrix $Corr_U$ and $Corr_L$. (b) SimBIN_v2: in the similarity-aware enhancement module, the threshold is calculated from the average of all item values in matrix $Corr_U$ or $Corr_L$, except 0 and 1. (c) SimBIN_v3: we rank all similarity items of each user or POI in matrix $Corr_U$ or $Corr_L$ to select the most similar 50 ones for the similarity-aware enhancement. SimBIN_v3 denotes the final SimBIN model. The final results of SimBIN are from this variant. The results are shown in Table 5.

Table 5. Variants of SimBIN with different filtering strategies for the similarity-aware enhancement. The best results are in bold.

Variant	Foursquare				Gowalla			
	Acc@1	Acc@5	Acc@10	MRR	Acc@1	Acc@5	Acc@10	MRR
v1	0.3094	0.6675	0.7201	0.4539	0.1465	0.3562	0.4233	0.2434
v2	0.3182	0.6836	0.7372	0.4610	0.1482	0.3655	0.4350	0.2481
v3	0.3199	0.6899	0.7417	0.4697	0.1488	0.3698	0.4361	0.2514

Our initial intention was to set a uniform threshold for each to select similar groups. However, from the table, we find that the performance of SimBIN_v2 is unsatisfactory. We find that the number of similar users for each target user is different after filtering by the threshold. To be specific, for a target user, most of his/her similarity items become 0. But for another target user, most of his/her similarity items don't change. This is probably why this filtering approach does not work well. SimBIN_v3 performs slightly better than SimBIN_v2, indicating the effectiveness of the modified strategy. This indicates that not all targets are similar and a target's behavior can be a close approximation with the most similar ones'. Here the number 50 is empirically set and the sensitivity is not obvious when it changes from 30 to 100, shown in Figure 6. Although the similarity-aware enhancement brings a slight improvement, more efficient strategies need further research in the future.

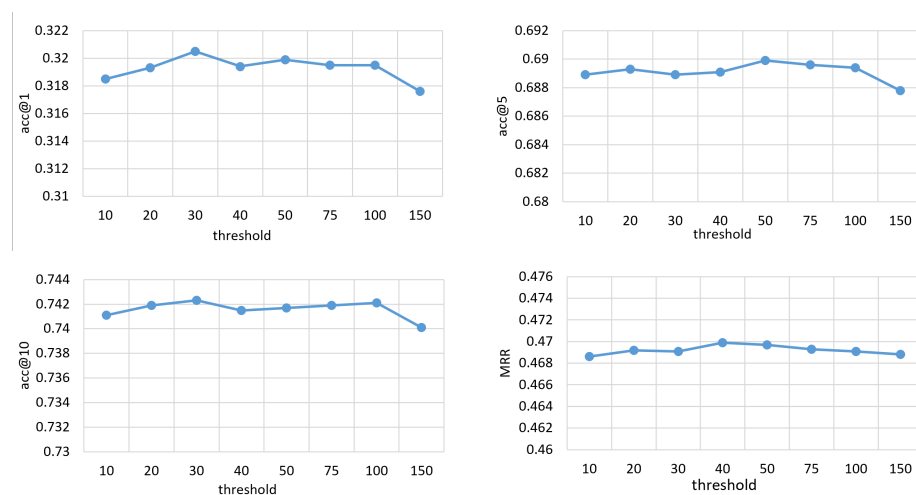


Figure 6. Sensitivity to filtering threshold on Foursquare.

6.4. Sensitivity to Data Sparsity (RQ5)

To study the sensitivity of SimBIN and the baselines to the sparse traces, we extensively perform experiments with different ratios of training data. To be specific, we train the model by taking 20%, 30%, 40%, 50%, 60%, 70%, and 80% of the Gowalla data, respectively. We take the last 20% for the testing evaluation. Table 6 shows the performance of SimBIN under different ratios. Figure 7 depicts the drop under different ratios in terms of that

under 0.8. The drop value is calculated by the ratio of the decrease to SimBIN under 0.8. We compare the SimBIN with Flashback, omitting the traditional compared methods and other RNN-based ones. The following are our observations from Table 6 and Figure 7.

Table 6. Performance of SimBIN on Gowalla under different training ratios. The best results are in bold.

Ratio	Flashback				SIMBIN			
	Acc@1	Acc@5	Acc@10	MRR	Acc@1	Acc@5	Acc@10	MRR
0.2	0.0773	0.1803	0.2210	0.1261	0.1112	0.2611	0.3111	0.1758
0.3	0.0861	0.2059	0.2558	0.1430	0.1200	0.2818	0.3346	0.1907
0.4	0.0953	0.2283	0.2859	0.1589	0.1308	0.3204	0.3794	0.2129
0.5	0.1002	0.2411	0.3027	0.1676	0.1366	0.3407	0.3961	0.2279
0.6	0.1064	0.2547	0.3201	0.1772	0.1415	0.3532	0.4165	0.2373
0.7	0.1095	0.2647	0.3344	0.1838	0.1438	0.3582	0.4227	0.2408
0.8	0.1158	0.2754	0.3479	0.1925	0.1488	0.3698	0.4361	0.2514

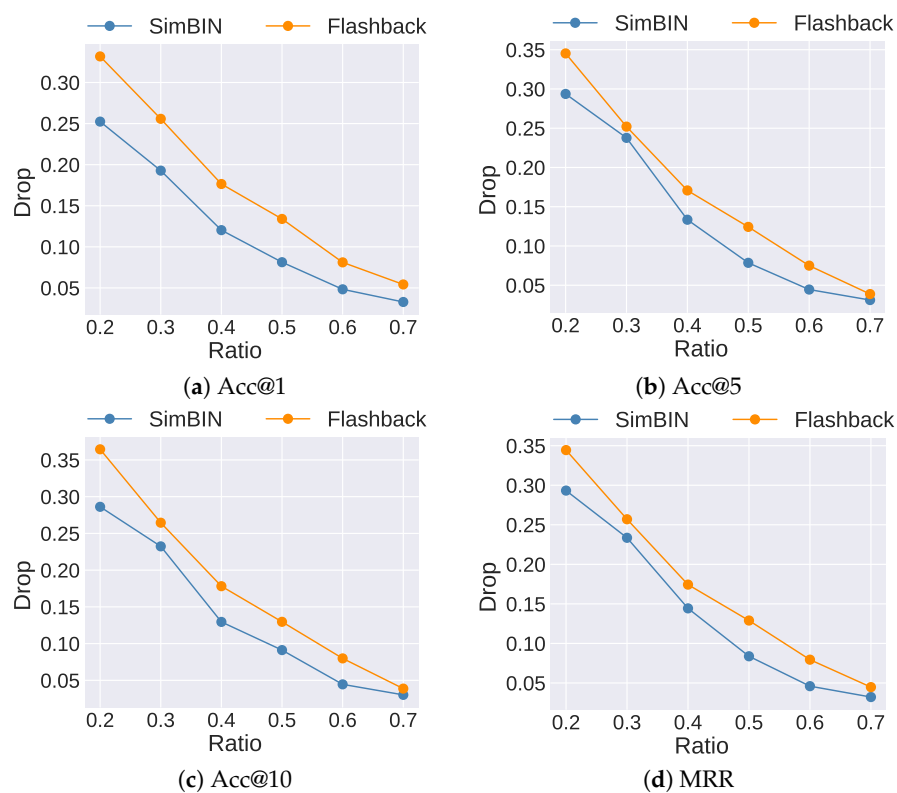


Figure 7. Drop of SimBIN under different training ratios in terms of that under 0.8. The value is calculated by the ratio of the decrease to SimBIN under 0.8.

OBS1. SimBIN outperforms Flashback consistently under all training ratios. The performances of SimBIN and Flashback get worse with the decreasing of training data. However, for each metric, Flashback drops more than SimBIN. In particular, SimBIN drops more slowly in terms of Acc@1, but Flashback drops quickly. These findings above demonstrate that our SimBIN is more robust to the sparsity.

OBS2. As the training ratio changing from 0.8 to 0.4, the performances of SimBIN and Flashback drop steadily, but they decline more sharply when the ratio changes from 0.4 to 0.2.

Quantitatively, SimBIN bridges the performance gap between sparse and dense users more effectively than ST-based methods. Qualitatively, unlike existing spatio-temporal studies that primarily treat space and time as independent contextual constraints or simply

concatenate their embeddings, SimBIN's similarity-awareness captures semantic peer-group behaviors, making it more resilient to temporal perturbations.

7. Conclusions

In this paper, we investigate how a POI attracts a person and develop a novel approach for the next location recommendation. We conduct the adaptive bi-directional inference which captures users' interests and POIs' attraction simultaneously. We further employ the adaptive learning strategy to refine the model and use the similarity-aware enhancement to rerank the recommendations. Extensive experiments demonstrate the advantage of our SimBIN for the location recommendation task.

We consider the location recommendation problem from both the view of the location and the human. As the direct implication, this idea can be transferred to many similar tasks, i.e., consider the problem from the bi-directional views. Moreover, our model can be well generalized and extended. Each key component in our SimBIN can be refined and generally used in any other recommendation framework, including the bi-directional inference, adaptive learning strategy, and the similarity-enhancement.

This study focuses on the next location recommendation task by modeling the spatiotemporal behavior patterns. Besides, the similarity-aware enhancement only involves spatiotemporal information. However, maybe social relationships and categorical semantics can enrich the mobility modeling and similarity measurement. We explicitly acknowledge: (1) the current model does not integrate rich semantic cues (e.g., text reviews) or social network graphs; (2) maintaining the bi-direction architecture increases memory overhead. We outline that exploring lightweight bi-directional distillation and fusing social graphs are promising future directions. Moreover, we may try more effective similarity measurement through various contexts for more convincing similarity-aware enhancement.

While our model demonstrates superior performance, we acknowledge the temporal limitation of the evaluated datasets. The Gowalla and Foursquare benchmarks, though indispensable for fair comparisons with existing arts, were collected over a decade ago. We argue, however, that the core algorithmic challenge—learning robust representations from power-law distributed and sparse trajectories—reflects a universal and time-invariant structural reality of human mobility, extending beyond the specific POI identities of that era. In future work, we aim to validate the generalizability of our dual-perspective collaborative framework on more recent and geographically diverse mobility datasets, further adapting the model to evolving urban dynamics and contemporary spatial-temporal patterns.

A notable limitation of our current experimental design is the asymmetry in data preprocessing. Because our framework requires strict consistency with baseline methods to ensure a fair, controlled ablation study in Table 4, we adopted the conventional user-specific preprocessing pipeline. While this guarantees an apples-to-apples architectural comparison, it inherently disadvantages the POI-centric branch (P-Net), as its input features are not optimally tailored to capture POI-side semantics. Consequently, the performance of P-Net reported in Table 4 likely represents a lower bound of its true capability. This asymmetry slightly tempers our core bi-directional argument.

Another limitation of our work is the memory overhead associated with large-scale similarity matrices. Our experiments were conducted on a Linux server equipped with A100 GPU. On the Foursquare dataset, Flashback completes training in approximately 6 h, whereas our proposed model requires approximately 25 h. Future work could explore approximate nearest neighbor techniques to further reduce memory consumption.

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

Funding: This research was funded by Fundamental Research Funds for the Central Universities (2024QN11052).

Data Availability Statement: The Foursquare dataset used in this study can be accessed at <https://sites.google.com/site/yangdingqi/home>.

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

References

1. Zhang, Q.; Yang, P.; Yu, J.; Wang, H.; He, X.; Yiu, S.M.; Yin, H. A survey on point-of-interest recommendation: Models, architectures, and security. *IEEE Trans. Knowl. Data Eng.* **2025**, *37*, 3153–3172.
2. Yu, F.; Cui, L.; Guo, W.; Lu, X.; Li, Q.; Lu, H. A category-aware deep model for successive POI recommendation on sparse check-in data. In Proceedings of the Web Conference, Taiwan, China, 20–24 April 2020; pp. 1264–1274.
3. Lian, D.; Wu, Y.; Ge, Y.; Xie, X.; Chen, E. Geography-aware sequential location recommendation. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Virtual Event, 23–27 August 2020; pp. 2009–2019.
4. Guo, Q.; Sun, Z.; Zhang, J.; Theng, Y.L. An attentional recurrent neural network for personalized next location recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 April 2020; pp. 83–90.
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 Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, Xian, China, 25–30 July 2020; pp. 1429–1438.
6. Yang, D.; Fankhauser, B.; Rosso, P.; Cudre-Mauroux, P. Location prediction over sparse user mobility traces using rnns. In Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, Yokohama, Japan, 7–15 January 2020; pp. 2184–2190.
7. Liu, B.; Fu, Y.; Yao, Z.; Xiong, H. Learning geographical preferences for point-of-interest recommendation. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Chicago, IL, USA, 11–14 August 2013; pp. 1043–1051.
8. Feng, S.; Li, X.; Zeng, Y.; Cong, G.; Chee, Y.M. Personalized ranking metric embedding for next new poi recommendation. In Proceedings of the 24th International Conference on Artificial Intelligence, Buenos Aires, Argentina, 25–31 July 2015; pp. 2069–2075.
9. Karatzoglou, A.; Jablonski, A.; Beigl, M. A Seq2Seq learning approach for modeling semantic trajectories and predicting the next location. In Proceedings of the 26th Acm Sigspatial International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 6–9 November 2018; pp. 528–531.
10. 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, Baltimore, MY, USA, 18–21 December 2016; pp. 191–200.
11. Wang, P.; Guo, J.; Lan, Y.; Xu, J.; Wan, S.; Cheng, X. Learning hierarchical representation model for nextbasket recommendation. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, Santiago, Chile, 9–13 August 2015; pp. 403–412.
12. Zhu, Y.; Li, H.; Liao, Y.; Wang, B.; Guan, Z.; Liu, H.; Cai, D. What To Do Next Model. User Behav. By Time-LSTM. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI 2017), Melbourne, Australia, 19–25 August 2017; pp. 3602–3608.
13. Wang, C.; Yuan, M.; Yang, Y.; Peng, K.; Jiang, H. Revisiting long-and short-term preference learning for next POI recommendation with hierarchical LSTM. *IEEE Trans. Mob. Comput.* **2024**, *23*, 12693–12705.
14. Zhao, P.; Luo, A.; Liu, Y.; Xu, J.; Li, Z.; Zhuang, F.; Zhou, X. Where to go next: A spatio-temporal gated network for next poi recommendation. *IEEE Trans. Knowl. Data Eng.* **2020**, *34*, 2512–2524.
15. Sun, K.; Qian, T.; Chen, T.; Liang, Y.; Nguyen, Q.V.H.; Yin, H. Where to go next: Modeling long-and short-term user preferences for point-of-interest recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 April 2020; pp. 214–221.
16. Feng, J.; Li, Y.; Zhang, C.; Sun, F.; Meng, F.; Guo, A.; Jin, D. Deepmove: Predicting human mobility with attentional recurrent networks. In Proceedings of the 2018 World Wide Web Conference, Lyon, France, 23–27 April 2018; pp. 1459–1468.

17. Lim, N.; Hooi, B.; Ng, S.K.; Wang, X.; Goh, Y.L.; Weng, R.; Varadarajan, J. 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, Virtual Event, 19–23 October 2020; pp. 845–854.
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 Proceedings of the Web Conference, Taiwan, China, 20–24 April 2020; pp. 906–916.
19. Li, X.; Hu, R.; Wang, Z. Next-point-of-interest recommendation based on joint mining of regularity and randomness. *Knowl. Based Syst.* **2022**, *258*, 110052.
20. Sun, G.; Qi, H.; Shen, Y.; Yin, B. TCSA-Net: A temporal-context-based self-attention network for next location prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 20735–20745.
21. Wang, X.; Sun, G.; Fang, X.; Yang, J.; Wang, S. Modeling Spatio-temporal Neighbourhood for Personalized Point-of-interest Recommendation. In Proceedings of the 31st International Joint Conference on Artificial Intelligence, Shenzhen, China, 23–29 July 2022; pp. 3530–3536.
22. Li, S.; Chen, W.; Wang, B.; Huang, C.; Yu, Y.; Dong, J. Mcn4rec: Multi-level collaborative neural network for next location recommendation. *ACM Trans. Inf. Syst.* **2024**, *42*, 105.
23. Li, P.; de Rijke, M.; Xue, H.; Ao, S.; Song, Y.; Salim, F.D. Large language models for next point-of-interest recommendation. In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, Washington, DC, USA, 14–18 July 2024; pp. 1463–1472.
24. He, X.; He, W.; Liu, Y.; Lu, X.; Xiao, Y.; Liu, Y. ImNext: Irregular interval attention and multi-task learning for next POI recommendation. *Knowl. Based Syst.* **2024**, *293*, 111674.
25. Wu, Y.; Bu, Y. Temporal-preference-aware attention in latent diffusion for personalized POI recommendation. *Inf. Sci.* **2026**, 123434.
26. Yu, Y.; Xu, Y.; Zhu, L.; Ma, J.; Li, F. GANPOI: A Multi-factor Generative Adversarial Framework for POI Recommendation. *Expert Syst. Appl.* **2025**, *304*, 130765.
27. Luo, Y.; Liu, Q.; Liu, Z. Stan: Spatio-temporal attention network for next location recommendation. In Proceedings of the Web Conference, New York, NY, USA, 19–23 April 2021; pp. 2177–2185.
28. Xue, H.; Salim, F.; Ren, Y.; Oliver, N. MobTCast: Leveraging auxiliary trajectory forecasting for human mobility prediction. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 30380–30391.
29. Yang, S.; Liu, J.; Zhao, K. Getnext: Trajectory flow map enhanced transformer for next poi recommendation. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, Madrid, Spain, 11–15 July 2022; pp. 1144–1153.
30. Yang, D.; Qu, B.; Yang, J.; Cudre-Mauroux, P. Revisiting user mobility and social relationships in lbsns: A hypergraph embedding approach. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–14 May 2019; pp. 2147–2157.
31. Wang, Z.; Zhu, Y.; Zhang, Q.; Liu, H.; Wang, C.; Liu, T. Graph-enhanced spatial-temporal network for next POI recommendation. *ACM Trans. Knowl. Discov. Data* **2022**, *16*, 104.
32. Zhang, J.; Li, Y.; Zou, R.; Zhang, J.; Jiang, R.; Fan, Z.; Song, X. Hyper-relational knowledge graph neural network for next POI recommendation. *World Wide Web* **2024**, *27*, 46.
33. Zeng, J.; Tao, H.; Tang, H.; Wen, J.; Gao, M. Global and local hypergraph learning method with semantic enhancement for POI recommendation. *Inf. Process. Manag.* **2025**, *62*, 103868.
34. Rao, X.; Jiang, R.; Shang, S.; Chen, L.; Han, P.; Yao, B.; Kalnis, P. Next point-of-interest recommendation with adaptive graph contrastive learning. *IEEE Trans. Knowl. Data Eng.* **2024**, *37*, 1366–1379.
35. Liu, Y.; Qi, L.; Mao, X.; Liu, W.; Fan, X.; Ni, Q.; Beheshti, A. Hyperbolic-Enhanced Mixture-of-Experts Mamba for Sequential Recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence, Singapore, 20–27 March 2026; pp. 15403–15411.
36. Wang, J.; Wang, Z.; Lu, T.; Guan, Y. Task-Aware Meta-Learning on Heterogeneous Knowledge Graph for POI Recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence, Singapore, 20–27 March 2026; pp. 15833–15840.
37. Pan, Q.; Zheng, H.; Zhao, Z.; Kong, X.; Shen, G.; Zhao, X. CTKGRec: A context-aware temporal knowledge graph reasoning model for next POI recommendation. *Expert Syst. Appl.* **2025**, *297*, 129224.
38. Christoforidis, G.; Kefalas, P.; Papadopoulos, A.N.; Manolopoulos, Y. RELINE: Point-of-interest recommendations using multiple network embeddings. *Knowl. Inf. Syst.* **2021**, *63*, 791–817.
39. Feng, S.; Meng, F.; Chen, L.; Shang, S.; Ong, Y.S. Rotan: A rotation-based temporal attention network for time-specific next poi recommendation. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Barcelona, Spain, 25–29 August 2024; pp. 759–770.
40. Lai, Y.; Su, Y.; Wei, L.; He, T.; Wang, H.; Chen, G.; Wang, X. Disentangled contrastive hypergraph learning for next POI recommendation. In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval, Washington, DC, USA, 14–18 July 2024; pp. 1452–1462.

41. Pan, X.; Cai, X.; Song, K.; Baker, T.; Gadekallu, T.R.; Yuan, X. Location recommendation based on mobility graph with individual and group influences. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 8409–8420.
42. Dang, W.; Wang, H.; Pan, S.; Zhang, P.; Zhou, C.; Chen, X.; Wang, J. Predicting human mobility via graph convolutional dual-attentive networks. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, Tempe, AZ, USA, 21–25 February 2022; pp. 192–200.
43. Hu, R.; Wang, Z.; Yamasaki, T.; X, L. Location predicts you: Location prediction via bi-direction speculation and dual-level association. In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, Virtual Event, 19–26 August 2021; pp. 529–536.
44. Xu, F.; Tu, Z.; Li, Y.; Zhang, P.; Fu, X.; Jin, D. Trajectory recovery from ash: User privacy is not preserved in aggregated mobility data. In Proceedings of the 26th International Conference on World Wide Web, Perth, Australia, 3–7 April 2017; pp. 1241–1250.
45. Li, H.; Ge, Y.; Hong, R.; Zhu, H. Point-of-interest recommendations: Learning potential check-ins from friends. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 975–984.
46. Xu, D.; Cui, P.; Zhu, W.; Yang, S. Find you from your friends: Graph-based residence location prediction for users in social media. In Proceedings of the 2014 IEEE International Conference on Multimedia and Expo, Shanghai, China, 12–19 July 2014; pp. 1–6.
47. Song, J.; Chang, C.; Sun, F.; Song, X.; Jiang, P. Ngat4rec: Neighbor-aware graph attention network for recommendation. *arXiv* **2020**, arXiv:2010.12256.
48. Feng, S.; Cong, G.; An, B.; Chee, Y.M. Poi2vec: Geographical latent representation for predicting future visitors. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017; pp. 102–108.
49. Zhu, N.; Cao, J.; Liu, Y.; Yang, Y.; Ying, H.; Xiong, H. Sequential modeling of hierarchical user intention and preference for next-item recommendation. In Proceedings of the 13th International Conference on Web Search and Data Mining, Houston, TX, USA, 3–7 January 2020; pp. 807–815.
50. Wang, M.; Ren, P.; Mei, L.; Chen, Z.; Ma, J.; De Rijke, M. A collaborative session-based recommendation approach with parallel memory modules. In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Paris, France, 21–25 July 2019; pp. 345–354.
51. Chen, Z.; Wang, X.; Xie, X.; Wu, T.; Bu, G.; Wang, Y.; Chen, E. Co-attentive multi-task learning for explainable recommendation. In Proceedings of the 28th International Joint Conference on Artificial Intelligence, Macau, China, 10–16 August 2019; pp. 2137–2143.
52. Yang, D.; Zhang, D.; Zheng, V.W.; Yu, Z. Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Trans. Syst. Man. Cybern. Syst.* **2014**, *45*, 129–142.
53. Zhang, Y.; Dai, H.; Xu, C.; Feng, J.; Wang, T.; Bian, J.; Liu, T.Y. Sequential click prediction for sponsored search with recurrent neural networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Quebec, QC, Canada, 1 June 2014; pp. 1369–1375.
54. Liu, Q.; Wu, S.; Wang, L.; Tan, T. Predicting the next location: A recurrent model with spatial and temporal contexts. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; pp. 194–200.
55. Li, Y.; Chen, T.; Yin, H.; Huang, Z. Discovering collaborative signals for next POI recommendation with iterative Seq2Graph augmentation. In Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, Virtual Event, 19–26 August 2021; pp. 1491–1497.
56. Wang, C.; Ma, W.; Chen, C.; Zhang, M.; Liu, Y.; Ma, S. Sequential recommendation with multiple contrast signals. *ACM Trans. Inf. Syst.* **2023**, *41*, 11.
57. Qin, Y.; Wu, H.; Ju, W.; Luo, X.; Zhang, M. A diffusion model for poi recommendation. *ACM Trans. Inf. Syst.* **2023**, *42*, 54.
58. Wang, Z.; Zhu, Y.; Wang, C.; Ma, W.; Li, B.; Yu, J. Adaptive graph representation learning for next POI recommendation. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Taiwan, China, 23–27 July 2023; pp. 393–402.
59. Zuo, C.; Zhang, X.; Yan, L.; Zhang, Z. GUGEN: Global user graph enhanced network for next POI recommendation. *IEEE Trans. Mob. Comput.* **2024**, *23*, 14975–14986.
60. Li, Y.; Wang, S.; Li, X.; Li, Z.; Wu, L.; Tian, W.; Li, G. Decorrelated Adaptive Simple Graph Representation Learning for Next POI Recommendation. *Knowl. Based Syst.* **2025**, *330*, 114585.

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.