Dual Weighted Graph Convolutional Network for POI Recommendation

Zhi Liu, Deju Zhang, Junhui Deng, Guojiang Shen, Xiangjie Kong*
College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China

Abstract

In recent years, with the widespread popularity of location-based social network platforms, the data generated by users on social networks has grown exponentially. There has been a growing focus on the problem of POI (Point-of-Interest) recommendations. Unlike traditional sequence recommendation that primarily considers the temporal dimension, POI recommendation needs to account for the influence of geographical information to a large extent. However, previous works in the graph construction process often only consider the places users have visited, neglecting those they haven't been to. To address this, we propose a Dual Weighted Graph Convolutional Network for POI recommendation called DualPOI. Specifically, we first leverage graph neural networks and attention mechanisms to capture users' local trajectory preferences for visited POIs. A delicately designed spatiotemporal encoder is conducted to model users' local spatiotemporal preferences. Subsequently, using a dual graph convolutional approach, we transfer the user's local preference information to a global scope, thereby modeling novel preferences for unvisited locations. Extensive experiments on four real-world datasets validate the effectiveness of our proposed method in enhancing the accuracy of POI recommendations. Comprehensive ablation studies and parameter analysis further confirm the efficacy of the proposed modules.

Keywords: Social Network; Next POI recommendation; Graph Neural Network

Introduction

The POI recommendation problem has emerged with the flourishing development of location-based social network (LBSN) platforms. Reliable POI recommendations can significantly alleviate information overload in social networks and enhance user experience. The Next POI recommendation aims to predict a user's next interest or preference based on their historical browsing records or interaction sequences, however, in contrast to e-commerce or short video platforms, which primarily model users' personal attribute information (likes, browsing history, etc.), emphasizing the temporal dimension. POI recommendation places greater emphasis on modeling geographical information and explores how geographical locations may influence recommendation results (H. Wang, Shen, Ouyang, & Cheng, 2018).

Considering the simultaneous modeling of the spatiotemporal preferences in user mobility trajectories is challenging. This is because user local movement patterns are not unidirectional, often involving repeated visits and being significantly influenced by both temporal and spatial factors. Existing works have made significant efforts in modeling spatiotemporal features. Early research employed collaborative filtering techniques to perform matrix factorization on geographic location information for POI recommendation (Cheng, Yang, King, & Lyu, 2012). Later, scholars utilized Recurrent Neural Networks (RNNs) or their variants, such as LSTM and GRU models (H. Wang, Shen, & Cheng, 2020), to encode users' historical visiting sequences. Additionally, the utilization of learned embeddings to transform a graph into one or multiple d-dimensional vectors, preserving graph information, has also been employed (Qian, Liu, Nguyen, & Yin, 2019). Nevertheless, these methods merely model the spatial/temporal relationships of POI visit records as a whole, lacking fine-grained modeling of spatiotemporal relationships and different directions.

With the rapid development of deep learning (Ma & Cheng, 2024), Graph Convolutional Networks (GCNs) have become one of the most widely used methods in POI recommendation systems. This is because GCN is particularly well-suited for capturing spatial features in the given data. However, in many practical applications, a user's actual visit history may not cover all locations. For example, in the real world, users exhibit diverse behavioral patterns. Some users tend to visit the same POIs or explore further within familiar regions. On the other hand, some users have a curiosity-driven behavior, preferring to explore locations they haven't visited or seen before, even across different areas. Consequently, modeling geographic preferences becomes challenging as it is hard to cover all locations in the modeling process. The working principle of traditional GCN is to aggregate neighborhood information on the graph, generating similar representations for nodes with similar characteristics. However, in practical applications, user information typically consists only of the visited POI records. The node representations learned by traditional GCN do not distinguish between visited and unvisited POIs. Therefore, traditional GCN methods are not effective for the recommendation task of novel POIs.

To address the aforementioned challenges, this paper proposes a Dual Weighted Graph Convolutional Network for POI recommendation (DualPOI). Specifically, we first construct a spatiotemporal interval graph separately in the temporal and spatial dimensions to model user visit patterns, providing an initial representation of the spatiotemporal features of user trajectories. Additionally, we design a spatiotemporal transition encoder to extract fine-grained bidirectional spatiotemporal weights of user local trajectories. Subsequently, we employ a Dual Graph Convolutional Network (DGCN) to capture the user's novel geographic preferences, enabling the transfer of local transition information to the global set of POIs. Finally, the recommendation results are obtained through joint learning of local transitions and global geographic preferences. The contributions of this study are summarized as follows:

- We design a tailored spatiotemporal transition encoder, which not only models user transition preferences from temporal and spatial perspectives but also explores the di-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society*.

rected weighted problem of visiting POIs at a finer granularity.

- We employ a DGCN to model user geographic preferences, enabling the extension of the user's current local spatiotemporal preferences to the global scope, thereby achieving joint learning of seen and unseen information.
- Experimental results on four real datasets validate the effectiveness of DualPOI, and extensive ablation studies further confirmed the efficacy of the proposed components.

Related Work

Next POI Recommendation

Recurrent neural networks (RNNs) and their variants are well known for their superior sequence-related task-handling capabilities. RNNs have served as the fundamental architecture for several prior POI recommendation models, which map one POI sequence to another (or simply the next POI). ST-RNN (Q. Liu, Wu, Wang, & Tan, 2016) builds upon the RNN model and designs a time and distance matrix to capture temporal cyclical effects and geographic influences. FlashBack (Yang, Fankhauser, Rosso, & Cudre-Mauroux, 2020) leverages RNN with spatiotemporal context and weighted historical states to enhance the model's ability to capture spatiotemporal effects. TMCA (R. Li, Shen, & Zhu, 2018) employs an LSTM-based encoder-decoder network with attention mechanisms to select relevant historical and contextual factors, utilizing embedding to merge heterogeneous context information. LSPL (Wu, Li, Zhao, & Qian, 2019) adopts two embedding learning modules to capture users' sequential behaviors in terms of location and category, and combines them to predict the next POI. LSTPM (K. Sun et al., 2020) uses all trajectories to capture long-term preferences and the last trajectory for short-term preferences.

Graph Neural Networks

The development of deep learning has driven numerous efforts to capture spatiotemporal correlations in graph data (S. Wang, Cao, & Philip, 2020). These methods often employ graph neural networks to model the spatial dependencies of users. GLSP (J. Liu, Chen, Huang, Li, & Min, 2023) proposed a GNN-based model to transform POIs into low-dimensional representations, integrating users' long-term and short-term preferences to comprehensively represent dynamic preferences. STA (B. Liu et al., 2017) views users, POIs, and spatiotemporal pairs as entities and relations in a knowledge graph, employing knowledge graph embeddings to learn users' spatiotemporal dependency information. GSTN (Z. Wang, Zhu, Zhang, et al., 2022) utilizes graph embeddings to explicitly model complex geographical features and capture distance-based and transition-based geographical influences from the designed POI semantic graph. However, current approaches heavily rely on users' visit records, overlooking globally available useful geographical information. The dependence on GNNs to capture user spatial dependencies still has limitations.

Methodology

In this section, we first define the POI recommendation task. Then we elucidate the functionality and operation of the predefined spatiotemporal transition graph. A detailed description of each component of our proposed model is presented subsequently. The overall framework of the model is illustrated in Figure 1.

Problem Definition and Preliminary

Sequential POI Recommendation Similar to classical recommendation problems, given a dataset $\mathcal{D} = \{\mathcal{U}, \mathcal{T}, \mathcal{P}\}$, with POI set $\mathcal{P} = \{p_1, p_2, ..., p_{|\mathcal{P}|}\}$ and user set $\mathcal{U} = \{u_1, u_2, ..., u_{|\mathcal{U}|}\}$, sequential POI recommendation can be defined as follows: Given a user u and his/her check-in history sequence $\mathcal{S} = \{(p_1^u, t_1^u), (p_2^u, t_2^u), ..., (p_n^u, t_n^u)\}$ at corresponding timestamps, the goal is to recommend the top K POIs most likely to be visited by the user at the next timestamp.

Spatiotemporal Transition Graph Considering the local trajectory features of user movements, we need to construct a trajectory graph to define the local visit patterns of the user's historical check-in sequence. Based on a user's historical trajectory $\mathcal{S}$, we construct a transition graph $\mathcal{G}_u = \{\mathcal{V}_u, \mathcal{E}_u\}$, where the vertex set $\mathcal{V}_u$ includes all POIs in this trajectory, and the edge set $\mathcal{E}_u$ represents a continuous check-in trajectory from p_i^u to p_{i+1}^u. This graph characterizes the user's local preferences accordingly. To account for the influence of geographical factors, we construct a geographic preference graph $\mathcal{G}_s = \{\mathcal{P}, \mathcal{E}, \mathcal{A}_s\}$. $\mathcal{A}_s(i, j) = d(p_i, p_j)$, where $d(p_i, p_j)$ represents the geographical distance between any pair of POIs $< p_i, p_j >$, and it is defined as the weight of the edge. $\mathcal{A}_s$ is an $N \times N$ matrix, where N is the number of POIs.

Personalized Spatiotemporal Transition

An intuitive observation is that users exhibit diverse behavioral attributes, with some frequently visiting and recording POIs in sequence, while others only occasionally document their check-in information. Therefore, in modeling temporal intervals for a user's trajectory records, we adopt relative time intervals. Specifically, given a user's historical visit trajectory $\mathcal{S} = \{(p_1^u, t_1^u), (p_2^u, t_2^u), ..., (p_n^u, t_n^u)\}$, the temporal interval between consecutive check-in POIs i and j is denoted as $|t_i^u - t_j^u|$. Suppose the set of all check-in time intervals for user u is represented as $\mathcal{R}_u$, we define the minimum check-in time interval for this user as $t_{min}^u = min(\mathcal{R}_u)$, and the relative transition time interval as $t_{ij}^u = \lfloor \frac{|t_i - t_j|}{t_{min}^u} \rfloor$. Similarly, for spatial transition intervals, the distance between two check-in POIs for a user is denoted as $d(p_i, p_j) = \mathcal{A}_s(i, j)$, where we define the minimum geographical interval within a user's trajectory as d_{min}^u and the relative spatial interval as $s_{ij}^u = \lfloor \frac{\mathcal{A}_s(i,j)}{d_{min}^u} \rfloor$. Consequently, we can obtain the user's spatiotemporal transition matrices, denoted as $M_s, M_t \in \mathbb{R}^{n \times n}$.

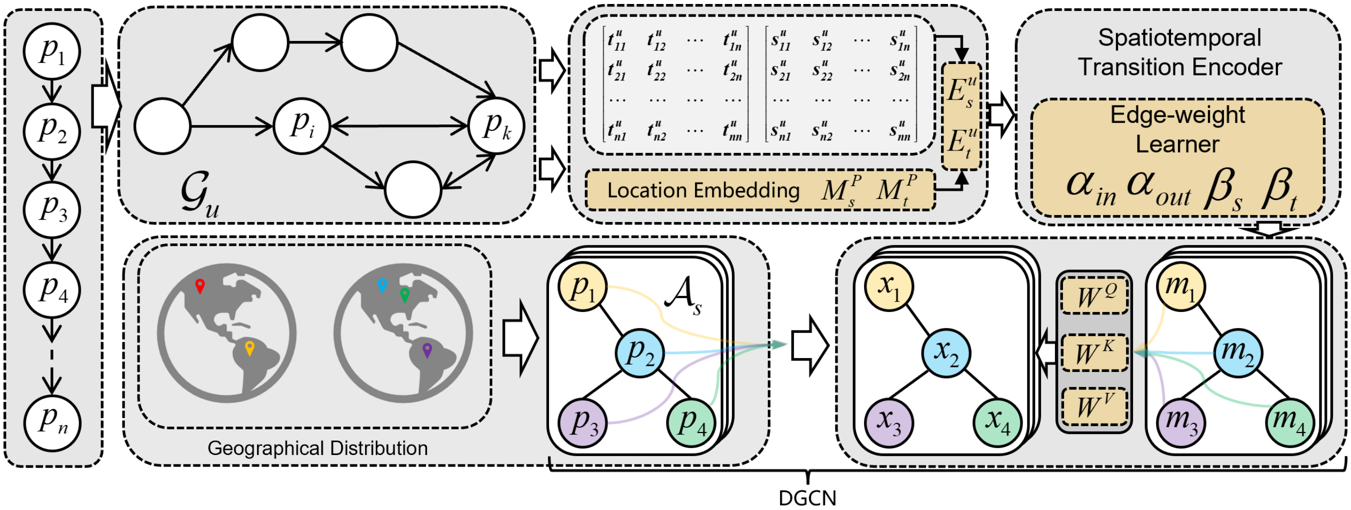

Figure 1: Overview of DualPOI. The DualPOI mainly consists of a spatiotemporal transition encoder and a DGCN for capturing global geographic information.

$$M_t = \begin{bmatrix} t_{11}^u & t_{12}^u & \dots & t_{1n}^u \\ t_{21}^u & t_{22}^u & \dots & t_{2n}^u \\ \dots & \dots & \dots & \dots \\ t_{n1}^u & t_{n2}^u & \dots & t_{nn}^u \end{bmatrix}, M_s = \begin{bmatrix} s_{11}^u & s_{12}^u & \dots & s_{1n}^u \\ s_{21}^u & s_{22}^u & \dots & s_{2n}^u \\ \dots & \dots & \dots & \dots \\ s_{n1}^u & s_{n2}^u & \dots & s_{nn}^u \end{bmatrix} \tag{1}$$

To enhance the generalization capability of the model, allowing it to discern spatiotemporal intervals not encountered during training, and simultaneously avoiding sparse relationship encoding (J. Li, Wang, & McAuley, 2020), we employ a hyperparameter θ for pruning these two matrices. Post the pruning operation, each matrix element is $s_{ij}^u = min(s_{ij}^u, \theta)$, $t_{ij}^u = min(t_{ij}^u, \theta)$, respectively.

We employ two learnable location embedding matrices, $M_s^P \in \mathbb{R}^{n \times d}$ and $M_t^P \in \mathbb{R}^{n \times d}$, to maintain these two interval matrices. This approach is more suitable for self-attention mechanisms, eliminating the need for additional linear transformations (Shaw, Uszkoreit, & Vaswani, 2018). Subsequently, by tensor concatenation of these two location embedding matrices with M_t and M_s, respectively, we obtain two trainable embedding matrices, $E_s^u, E_t^u \in \mathbb{R}^{n \times n \times d}$. In these matrices, any position element $e_{ij}^{s,u}$, $e_{ij}^{t,u}$ represents a d-dimensional embedding vector, where d is a hyperparameter.

Spatiotemporal Transition Encoder

As illustrated in Figure 1, a user's check-in sequence for POIs is not strictly unidirectional. To model these bidirectional spatiotemporal relationships, we design a personalized *spatiotemporal edge-weight learner* for POI embedding e and personalized spatiotemporal transitions to learn such dynamic relationships, which improves the message aggregation mechanism of GGNN (Ruiz, Gama, & Ribeiro, 2020). Specifically, the direction and spatiotemporal information of the user's trajectory can be defined as follows:

$$m_i = \sum_{<i,j> \in \mathcal{E}} \alpha_{ij}^{in} e_j + \sum_{<i,k> \in \mathcal{E}} \alpha_{ik}^{out} e_k \tag{2}$$

$$I_{in} = \beta_{ij}^s e_{ij}^{s,u} + \beta_{ij}^t e_{ij}^{t,u} \tag{3}$$

$$I_{out} = \beta_{ik}^s e_{ik}^{s,u} + \beta_{ik}^t e_{ik}^{t,u} \tag{4}$$

$$\alpha_{ij}^{in} = \delta(I_{in}(W_{in}(e_i \odot e_j))) \tag{5}$$

$$\alpha_{ik}^{out} = \delta(I_{out}(W_{out}(e_i \odot e_k))) \tag{6}$$

where $W_{in}, W_{out} \in \mathbb{R}^d$ are trainable mapping matrices, α_{ij}^{in}, α_{ik}^{out}, β_{ij}^s, β_{ik}^t all are trainable parameters, $\odot$ indicates the product operation, δ is the Softmax activation function, used to map weights to the (0, 1) interval. To effectively model a user's historical preferences, we employ a self-attention mechanism to iteratively update the representation of neighbor influences between each pair of nodes in the trajectory $\mathcal{S}$ for each user. For each input sequence $E_u = [m_1, m_2, ..., m_n]$, we aim to obtain a new sequence representation $X = [x_1, x_2, ..., x_n]$, where $x_i \in \mathbb{R}^d$. Each x_i is obtained by weighted summation, representing the historical check-in preferences of user i:

$$x_i = \sum_{j=1}^{n} \alpha_{ij}(W^V m_j) \tag{7}$$

$$\alpha_{ij} = \frac{exp e_{ij}}{\sum_{k=1}^{n} exp e_{ik}} \tag{8}$$

$$e_{ij} = \frac{W^Q m_i (W^K m_j)^T}{\sqrt{d}} \tag{9}$$

where $W^Q, W^K, W^V \in \mathbb{R}^{d \times d}$ are all trainable mapping matrices. The Spatiotemporal Transition Encoder (STE) is employed to capture the spatiotemporal features of user trajectory transitions, while the self-attention mechanism is capable of modeling the global preferences between long-term visits by the user.

Novel Geographical Preferences

Inspired by (Yue, Liang, Cui, & Bai, 2022), for the POI recommendation task, it is crucial to thoroughly consider the relationship between users and geographic information that has not been visited. Based on this consideration, we propose a Dual Graph Convolutional Neural Network (DGCN) for POI recommendation. In the preceding sections, we initially constructed the transition features X for users, reflecting the preferences for POIs they have encountered. Additionally, POIs seen and unseen by the user need to be represented within the same network. Specifically, we consider a layer of DGCN to adopt a dual graph aggregation mechanism, resulting in $V_u = [v_1, v_2, ..., v_n]$, where $v_i \in \mathbb{R}^d$. The general form can be expressed as follows:

$$V_i = Softmrelu(\hat{M}_s X W^{(1)}) W^{(2)} \hat{\mathcal{A}}_s) \tag{10}$$

where $\hat{M}_s$ is the normalized spatial interval matrix of user trajectories, $\mathcal{A}_s$ is the normalized geographic adjacency matrix defined as the geographical distance between POIs. This can be initially represented as the geographic relationship between POIs, $W^{(1)} \in \mathbb{R}^{d \times d}, W^{(2)} \in \mathbb{R}^{d \times N}$ are trainable mapping matrices. The dimension of V_i is consistent with the total number of POIs, allowing the adjacency representation learned through a unified network to not only reflect the user's preferences for the visited POIs but also integrate unvisited POIs. The traditional GCN aggregates and reflects adjacency information between user trajectories, while DGCN can simultaneously aggregate adjacency information for user trajectories and global POIs. Therefore, DGCN is better suited for recommending the next location a user might visit, even if they have not visited it before. Finally, we use a multi-layer perceptron to encode V_i into $v_i \in \mathbb{R}^d$ for the optimization of the model.

Prediction and Optimization

We previously obtained two corresponding embedding representations, x_u and v_u, representing the user's local trajectory preference and global geographical preference, respectively. Considering the user's preference for the next visited POI, the probability of user u visiting POI i can be expressed as:

$$\hat{y}_i^u = Softmax(\alpha x_u^T + (1-\alpha) v_u^T) \tag{11}$$

where α is a weight coefficient reflecting the importance of the two similarity terms. By default, we set α to 0.5, indicating equal importance for both factors. Given the user's visitation history trajectory $\mathcal{S}$ and the corresponding Ground-Truth target POI y_i^u, the model optimizes using a cross-entropy loss function with regularization L2 loss weighted by a hyperparameter. The overall optimization function is expressed as:

$$\mathcal{L}_{final} = -\frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} y_i^u log(\hat{y}_i^u) + \gamma \|\Theta\|_2^2 \tag{12}$$

Table 1: Statistics of Datasets

Dataset	#Check-ins	#User	#POI	#Avg. Length	#Sparsity
Gowalla	456820	10162	24237	44.95	99.81%
Singapore	194108	2321	5596	83.63	98.51%
NYC	179468	1083	9989	165.71	98.34%
TKY	494807	2293	15177	215.79	98.58%

Experiment

To validate the performance of our proposed DualPOI and evaluate the efficacy of its components, a series of experiments and parameter analyses are conducted on four real-world POI datasets.

Datasets and Parameter Setting

We utilize four real-world datasets to validate the effectiveness of the proposed method, all of which have been widely applied in POI recommendations.

Gowalla[1] is a wildly used benchmark dataset. It comprises user check-in records from 2009 to 2010, with each entry containing user ID, location coordinates, and check-in timestamp.

Singapore dataset is collected from the large-scale check-in platforms Foursquare[2]. The dataset is collected from August 2010 to July 2011.

NYC and Tokyo Check-in Dataset contains check-ins in NYC and Tokyo collected for about 10 months from 12 April 2012 to 16 February 2013, which is also provided by Foursquare.

For a fair comparison, all models use the Adam optimizer to optimize the parameters. The learning rate is set to 0.001 uniformly. For our model, we set the hyperparameters α to 0.5 and γ to 10^{-3} in the loss function. To the dropout rate, we fix the value of it to 0.1. For efficiency, we adopt an early stopping strategy with a patience of 10 during the training process. We evaluate the model using the best epoch determined by this strategy.

Baselines and Metrics

- MF (Mnih & Salakhutdinov, 2007) is a classical collaborative filtering method that decomposes user-poi interaction data to capture patterns to facilitate personalized recommendations.

- BERT4Rec (F. Sun et al., 2019) is a transformer-based model that utilizes a bidirectional attention mechanism to achieve end-to-end historical sequence recommendation.

- LightGCN (He et al., 2020) utilizes graph convolutional neural networks for embedding representations to accomplish POI recommendations.

[1]https://snap.stanford.edu/data/loc-gowalla.html
[2]https://sites.google.com/site/yangdingqi/home/foursquare-dataset

Table 2: Recommendation effectiveness comparisons. Each row highlights the best-performing method in bold and the second-best method with an underline.

Dataset	Metric	MF	LightGCN	SGRec	LSTPM	STAN	BERT4Rec	DRAN	DualPOI
Gowalla	Recall@2	0.0414	0.1102	0.1531	0.1904	0.2195	0.1734	0.2133	**0.2564**
	Recall@5	0.0869	0.1348	0.1987	0.2049	0.2364	0.2042	0.2466	**0.2981**
	Recall@10	0.1473	0.1699	0.2384	0.2618	0.2994	0.2235	0.3056	**0.3332**
	NDCG@2	0.0428	0.0914	0.1402	0.1431	0.1917	0.1683	0.2038	**0.2451**
	NDCG@5	0.0810	0.1162	0.1621	0.1588	0.2152	0.1877	0.2168	**0.2637**
	NDCG@10	0.1078	0.1432	0.1751	0.1742	0.2268	0.2033	0.2367	**0.2750**
Singapore	Recall@2	0.1103	0.1606	0.3058	0.2704	0.2634	0.2566	0.3086	**0.3292**
	Recall@5	0.1766	0.2144	0.3507	0.3253	0.2940	0.2837	0.3554	**0.3793**
	Recall@10	0.2059	0.2419	0.3884	0.3791	0.3280	0.3099	0.3921	**0.4214**
	NDCG@2	0.0744	0.1720	0.2274	0.2610	0.2831	0.2424	0.2972	**0.3161**
	NDCG@5	0.0791	0.1789	0.2697	0.2697	0.2995	0.2584	0.3175	**0.3389**
	NDCG@10	0.0868	0.1892	0.2916	0.2749	0.2892	0.2688	0.3297	**0.3524**
New York	Recall@2	0.2361	0.3789	0.5734	0.5754	0.6027	0.5629	0.5859	**0.6477**
	Recall@5	0.2792	0.4363	0.6175	0.6020	0.6358	0.6034	0.6253	**0.6648**
	Recall@10	0.3046	0.4519	0.6424	0.6312	0.6533	0.6301	0.6478	**0.6776**
	NDCG@2	0.2002	0.3819	0.5490	0.5524	0.5887	0.5660	0.5702	**0.6417**
	NDCG@5	0.2318	0.3867	0.5559	0.5596	0.6092	0.5730	0.5881	**0.6494**
	NDCG@10	0.2426	0.3903	0.5613	0.5681	0.6124	0.5896	0.5956	**0.6535**
Tokyo	Recall@2	0.2896	0.3917	0.5091	0.5029	0.5105	0.5059	0.5225	**0.5850**
	Recall@5	0.3141	0.4473	0.5488	0.5513	0.5489	0.5432	0.5570	**0.6238**
	Recall@10	0.3866	0.4936	0.6173	0.5795	0.6167	0.5745	0.6210	**0.6537**
	NDCG@2	0.1984	0.4207	0.4715	0.4724	0.4990	0.4888	0.5089	**0.5734**
	NDCG@5	0.2071	0.4354	0.4905	0.4881	0.5264	0.5051	0.5390	**0.5910**
	NDCG@10	0.2327	0.4403	0.5112	0.4962	0.5554	0.5203	0.5602	**0.6007**

- LSTPM (K. Sun et al., 2020) is an LSTM-based model that achieves improved recommendation performance by modeling both the long-term and short-term preferences of users.
- SGRec (Y. Li, Chen, Luo, Yin, & Huang, 2021) is a GAT-based method that enhances the representation of nodes by aggregating POIs before and after the target POI in different orders.
- STAN (Luo, Liu, & Liu, 2021) aggregates temporal and spatial information through a dual-layer attention mechanism and performs linear interpolation for spatiotemporal discretization.
- DRAN (Z. Wang, Zhu, Liu, & Wang, 2022) utilizes GCN for feature extraction after decomposing POI into multiple dimensions and employs multi-head attention to process each decomposed dimension.

We employ two evaluation metrics to assess the model, namely Recall@K and NDCG@K (Normalized Discounted Cumulative Gain). K means the consideration of the top K items in a recommendation list, which is a common practice in POI recommendation.

Overall Comparison

We conduct experiments on four real datasets using the open-source code from the aforementioned paper and followed similar procedures as in previous related works. We present the experimental results alongside the best results from similar previous works for comparison (Qin, Wu, Ju, Luo, & Zhang, 2023). The experimental results are presented in Table 2. Here, we can make the following observations.

DualPOI consistently exhibits superior performance across all six metrics compared to all baselines across four datasets. In particular, DualPOI achieved performance improvements

Table 3: Results of ablation experiments with different STE components.

Dataset	Ablation	Recall@2	Recall@5	Recall@10	NDCG@2	NDCG@5	NDCG@10
Gowalla	w MLP	0.2194	0.2621	0.2983	0.2083	0.2275	0.2392
	w GCN	0.2355	0.2775	0.3121	0.2245	0.2433	0.2545
	w GAT	0.2433*	0.2852*	0.3224*	0.2323*	0.2511*	0.2631*
	DualPOI	**0.2564**	**0.2981**	**0.3332**	**0.2451**	**0.2637**	**0.2750**
Singapore	w MLP	0.2746	0.3228	0.3631	0.2621	0.2835	0.2965
	w GCN	0.3089	0.3567	0.3991	0.2972	0.3186	0.3323
	w GAT	0.3188*	0.3671*	0.4074*	0.3052*	0.3268*	0.3399*
	DualPOI	**0.3292**	**0.3793**	**0.4214**	**0.3161**	**0.3389**	**0.3524**
New York	w MLP	0.5999	0.6211	0.6366	0.5943	0.6039	0.6090
	w GCN	0.6241	0.6424	0.6571	0.6180	0.6262	0.6310
	w GAT	0.6316*	0.6491*	0.6623*	0.6262*	0.6341*	0.6384*
	DualPOI	**0.6477**	**0.6648**	**0.6776**	**0.6417**	**0.6494**	**0.6535**
Tokyo	w MLP	0.5174	0.5607	0.5927	0.5040	0.5235	0.5338
	w GCN	0.5544	0.5969	0.6273	0.5414	0.5605	0.5703
	w GAT	0.5790*	0.6189*	0.6480*	0.5660*	0.5839*	0.5934*
	DualPOI	**0.5850**	**0.6238**	**0.6537**	**0.5734**	**0.5910**	**0.6007**

Table 4: Results of ablation experiments using different modules instead of DGCN.

Dataset	Ablation	Recall@2	Recall@5	Recall@10	NDCG@2	NDCG@5	NDCG@10
Gowalla	w MLP	0.2269	0.2675	0.3060	0.2159	0.2341	0.2465
	w GCN	0.2474*	0.2880*	0.3232*	0.2367*	0.2549*	0.2663*
	DualPOI	**0.2564**	**0.2981**	**0.3332**	**0.2451**	**0.2637**	**0.2750**
Singapore	w MLP	0.2906	0.3386	0.3789	0.2784	0.2999	0.3129
	w GCN	0.3106*	0.3593*	0.3986*	0.2980*	0.3199*	0.3326*
	DualPOI	**0.3292**	**0.3793**	**0.4214**	**0.3161**	**0.3389**	**0.3524**
New York	w MLP	0.6188	0.6370	0.6520	0.6130	0.6213	0.6261
	w GCN	0.6333*	0.6517*	0.6679*	0.6281*	0.6363*	0.6416*
	DualPOI	**0.6477**	**0.6648**	**0.6776**	**0.6417**	**0.6494**	**0.6535**
Tokyo	w MLP	0.5595	0.6006	0.6320	0.5470	0.5655	0.5757
	w GCN	0.5721*	0.6111*	0.6430*	0.5600*	0.5775*	0.5878*
	DualPOI	**0.5850**	**0.6238**	**0.6537**	**0.5734**	**0.5910**	**0.6007**

of 9%/16%, 7%/7%, 4%/7%, and 5%/7% on the four datasets, respectively (Recall@10 and NDCG@10) compared to the best baseline model. The user-specific mobility patterns generated by the Spatiotemporal Transition Encoder and the joint modeling by the DGCN reveal special trends in recommending the next visit to POI.

Traditional methods like MF exhibit lower recommendation performance due to the absence of geographical information modeling. LSTPM, STAN, and DRAN, which also leverage geographical information like our approach, outperform other methods. This underscores the importance of geographical information in POI recommendation tasks. Benefiting from the excellent capability of graph neural networks in handling irregular data, graph-based models (LSTPM, STAN, DRAN) exhibit relatively strong performance, including our proposed approach.

Ablation Study

Table 3 presents the effects of replacing STE with other common encoders. We can observe that the performance of the model will be greatly adversely affected when our designed module is replaced by other common components. Therefore, designing a dedicated graph encoder for user trajectory information is crucial. Leveraging GCN for aggregating neighboring node information and GAT for adaptive weighting and global information integration is a best practice for capturing essential details.

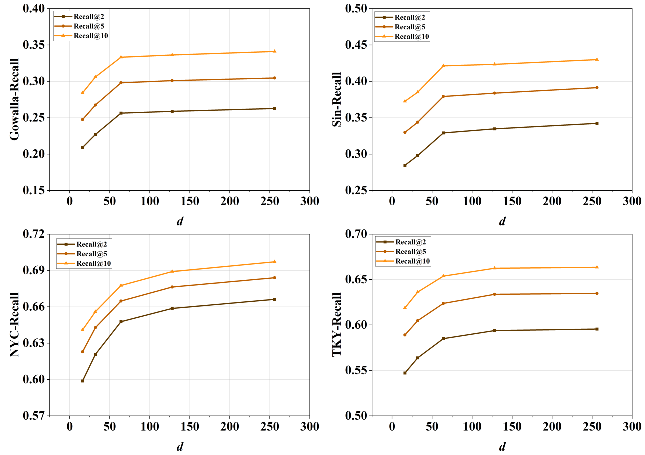

Figure 2: Effect of Hidden Dimensionality d.

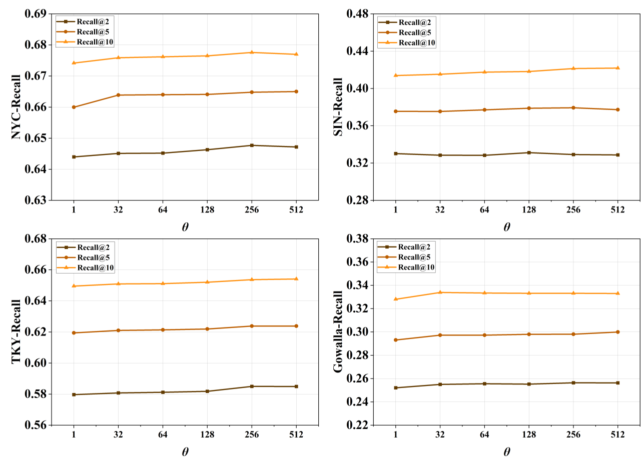

Figure 3: Effect of spatiotemporal interval θ.

Table 4 presents the results after replacing DGCN with different components. We can observe that the model still exhibits excellent performance when using a single layer GCN instead of DGCN, it loses the ability to capture global POI geographical information compared to DGCN. This indicates that even if users have not visited a certain place, our model is still capable of providing constructive recommendations for the next POI.

Parameter Analysis

Figure 2 illustrates the Recall for dimensions d ranging from 16 to 256 (doubling increment) while keeping other parameters constant. In most cases, a larger value of d leads to better model performance. In general, the performance of recommendations tends to improve with the increase in hidden dimensions. We can also observe that as the dimensions increase, the accuracy of recommendations initially increases significantly, then shows signs of slow growth, and finally stabilizes relatively around a hidden dimension of 64. The increase in dimensions implies a larger receptive field for the model, but it also increases parameters and consumption of computational resources. Therefore, choosing appropriate parameters is crucial.

Figure 3 shows the Recall of DualPOI. We choose spatial and temporal intervals $\{1, 32, 64, 128, 256, 512\}$ for comparison. We can see that the Recall of the DualPOI is insensitive to parameter changes across all datasets. It indicates that our model exhibits good robustness and achieves satisfactory results across different ranges of temporal and spatial intervals. Within a limited range, competitive performance can be easily achieved.

Conclusion

In this paper, we studied how to leverage user historical trajectories and global geographical information for the next POI recommendation. Two main modules comprise the model that we proposed. First, we construct a spatiotemporal transition graph and use a spatiotemporal transition encoder to extract the user-specific local preferences. Second, to obtain extra insightful data, we utilize dual graph convolution to transfer information from the user's historical trajectory preference to the global POI environment. Subsequently, we conduct extensive experiments on four real-world datasets, and the experimental results demonstrate that our model significantly improves recommendation performance. Extensive ablation studies and parameter analysis also validate the effectiveness and robustness of the proposed components. Overall, our model is of significant importance for enhancing user experience and reducing issues such as information overload in social networks.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62073295 and Grant 62072409, in part by the Zhejiang Provincial Natural Science Foundation under Grant LR21F020003, and in part by the Zhejiang Province Basic Public Welfare Research Project under Grant LTGG24F020009.

References

Cheng, C., Yang, H., King, I., & Lyu, M. (2012). Fused matrix factorization with geographical and social influence in location-based social networks. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 26, pp. 17–23).

He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., & Wang, M. (2020). 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* (pp. 639–648).

Li, J., Wang, Y., & McAuley, J. (2020). Time interval aware self-attention for sequential recommendation. In *Proceedings of the 13th international conference on web search and data mining* (pp. 322–330).

Li, R., Shen, Y., & Zhu, Y. (2018). Next point-of-interest recommendation with temporal and multi-level context attention. In *2018 ieee international conference on data mining (icdm)* (pp. 1110–1115).

Li, Y., Chen, T., Luo, Y., Yin, H., & Huang, Z. (2021). Discovering collaborative signals for next poi recommendation with iterative seq2graph augmentation. In *Proceedings of the thirtieth international joint conference on artificial intelligence* (pp. 1491–1497).

Liu, B., Qian, T., Liu, B., Hong, L., You, Z., & Li, Y. (2017). Learning spatiotemporal-aware representation for poi recommendation. *arXiv preprint arXiv:1704.08853*.

Liu, J., Chen, Y., Huang, X., Li, J., & Min, G. (2023). Gnn-based long and short term preference modeling for next-location prediction. *Information Sciences*, *629*, 1–14.

Liu, Q., Wu, S., Wang, L., & Tan, T. (2016). Predicting the next location: A recurrent model with spatial and temporal contexts. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 30).

Luo, Y., Liu, Q., & Liu, Z. (2021). Stan: Spatio-temporal attention network for next location recommendation. In *Proceedings of the web conference 2021* (pp. 2177–2185).

Ma, Y., & Cheng. (2024). Lada: Latent dialogue action for zero-shot cross-lingual neural networklanguage modeling. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 45).

Mnih, A., & Salakhutdinov, R. R. (2007). Probabilistic matrix factorization. *Advances in neural information processing systems*, *20*.

Qian, T., Liu, B., Nguyen, Q. V. H., & Yin, H. (2019). Spatiotemporal representation learning for translation-based poi recommendation. *ACM Transactions on Information Systems (TOIS)*, *37*(2), 1–24.

Qin, Y., Wu, H., Ju, W., Luo, X., & Zhang, M. (2023). A diffusion model for poi recommendation. *ACM Trans. Inf. Syst.*, *42*.

Ruiz, L., Gama, F., & Ribeiro, A. (2020). Gated graph recurrent neural networks. *IEEE Transactions on Signal Processing*, *68*, 6303-6318. doi: 10.1109/TSP.2020.3033962

Shaw, P., Uszkoreit, J., & Vaswani, A. (2018). Self-attention with relative position representations. *arXiv preprint arXiv:1803.02155*.

Sun, F., Liu, J., Wu, J., Pei, C., Lin, X., Ou, W., & Jiang, P. (2019). Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th acm international conference on information and knowledge management* (pp. 1441–1450).

Sun, K., Qian, T., Chen, T., Liang, Y., Nguyen, Q. V. H., & Yin, H. (2020). 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* (Vol. 34, pp. 214–221).

Wang, H., Shen, H., & Cheng, X. (2020). Modeling poi-specific spatial-temporal context for point-of-interest recommendation. In *Advances in knowledge discovery and data mining* (pp. 130–141).

Wang, H., Shen, H., Ouyang, W., & Cheng, X. (2018). Exploiting poi-specific geographical influence for point-of-interest recommendation. In *Proceedings of the thirtieth international joint conference on artificial intelligence* (pp. 3877–3883).

Wang, S., Cao, J., & Philip, S. Y. (2020). Deep learning for spatio-temporal data mining: A survey. *IEEE transactions on knowledge and data engineering*, *34*(8), 3681–3700.

Wang, Z., Zhu, Y., Liu, H., & Wang, C. (2022). Learning graph-based disentangled representations for next poi recommendation. In *Proceedings of the 45th international acm sigir conference on research and development in information retrieval* (pp. 1154–1163).

Wang, Z., Zhu, Y., Zhang, Q., Liu, H., Wang, C., & Liu, T. (2022). Graph-enhanced spatial-temporal network for next poi recommendation. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *16*(6), 1–21.

Wu, Y., Li, K., Zhao, G., & Qian, X. (2019). Long-and short-term preference learning for next poi recommendation. In *Proceedings of the 28th acm international conference on information and knowledge management* (pp. 2301–2304).

Yang, D., Fankhauser, B., Rosso, P., & Cudre-Mauroux, P. (2020). Location prediction over sparse user mobility traces using rnns. In *Proceedings of the twenty-ninth international joint conference on artificial intelligence* (pp. 2184–2190).

Yue, Q., Liang, J., Cui, J., & Bai, L. (2022). Dual bidirectional graph convolutional networks for zero-shot node classification. In *Proceedings of the 28th acm sigkdd conference on knowledge discovery and data mining* (p. 2408-2417).