Local and Global Exploration for Next New POI Recommendation

KE SUN, School of Computer Science and Technology, Wuhan University of Science and Technology, China
LIYU ZHOU, Naval University of Engineering, China
MAYI XU, School of Computer Science, Wuhan University, China
TIEYUN QIAN*, School of Computer Science, Wuhan University, China

The next point-of-interest (POI) recommendation is a hot-spot for both industry and academia, which helps users better experience the physical world. However, existing methods suffer from a severe bias towards recommending repeat POIs that have been visited by the target user before, and perform inefficiently when recommending new POIs that have not been visited by the target user yet. To overcome this issue, we delve into the next new POI recommendation and uncover the coexistence of local and global exploration patterns in users' visits to new POIs, showing their willingness to explore not only nearby new POIs but also those distant ones. Subsequently, we develop a novel **L**ocal and **G**lobal **E**xploration framework (**LGE**) for the next new POI recommendation. In particular, LGE involves three key modules: 1) a Zone-Aware Local Exploration (ZLE) module, which encourages users to explore POIs in the local area by learning zone-aware POI representations and regularizing POI prediction with zone information; 2) an Intention-Aware Global Exploration (IGE) module, which recommends POIs that meet user intentions without distance constraints by extracting static and dynamic intentions from category information; 3) a fusion module, which contains a Mean Pooling (MP) strategy and a Weighted Pooling (WP) strategy to aggregate the outputs of local and global exploration modules for the final recommendation. Experiments carried out on real-world datasets have shown the effectiveness of LGE in recommending new POIs.

CCS Concepts: • **Information systems → Recommender systems**.

Additional Key Words and Phrases: Next New POI recommendation, Local and Global Exploration.

1 Introduction

The next point-of-interest recommendation (NPR) involves predicting users' subsequent preferred POIs, given their historical trajectories [40]. Precise POI recommendations not only improve users' experience, but also contribute to business success, thus gaining popularity in research.

When recommending to users, the system has two generic options. The first option helps users reexperience their history and suggests POIs that have already been visited by the target users, namely *repeat POIs*. The second option, by contrast, is more radical and encourages users to explore the world-recommending *new POIs* unvisited by the target users. In the literature, existing methods have been shown to achieve remarkable overall performance on real-world datasets such as Gowalla, Foursquare, and Yelp. However, they usually confuse above two types of recommendations and do not pay much attention to the challenge of recommending new POIs.

*Corresponding authors.

Authors' addresses: Ke Sun, sunke1995@whu.edu.cn, School of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, Hubei, China; Liyu Zhou, 2420242032@nue.edu.cn, Naval University of Engineering, Wuhan, Hubei, China; Mayi Xu, xumayi@whu.edu.cn, School of Computer Science, Wuhan University, Wuhan, Hubei, China; Tieyun Qian, qty@whu.edu.cn, School of Computer Science, Wuhan University, Wuhan, Hubei, China.

ACM 1556-472X/2026/4-ART
https://doi.org/10.1145/3807950

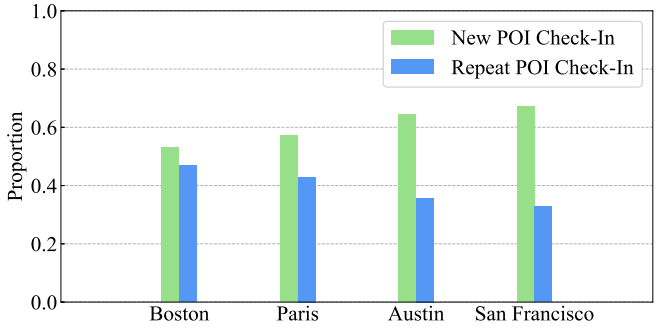


Fig. 1. The proportion of new and repeat POI check-ins within different datasets. Boston and Paris are from the Foursquare platform, while Austin and San Francisco are from the Gowalla platform.

At the early stage, pioneering researchers [3] have pointed out that users want to explore new POIs, and recommending new POIs is much harder but more meaningful than repeat POIs. According to the statistics of four cities shown in Fig. 1, new POI check-ins make up a substantial proportion of the data. This rate even surpasses 60% in some city datasets, e.g., 64.27% for Austin and 67.15% for San Francisco. In light of this, we focus on solving the next new POI recommendation (N^2PR) in this paper.

As an extension to NPR, N^2PR aims to recommend new POIs to be visited next given a user's historical check-ins [7]. Several preliminary efforts have been made to address this problem. Their general idea is to explore the relationships between users and unvisited new POIs by leveraging self-supervised signals or side information. In particular, PRME-G [7] and JTLL [20] use unobserved visits to create negative samples, and PPDM [16] learns contextual features (e.g., POI category). Despite their unique designs, these methods are still deficient in recommending new POIs.

We consider two reasons for the deficiency of previous methods in recommending new POIs. First, the recommendation of new POIs is more difficult than that of repeat POIs, as noted in pioneering work [3], but previous methods treat them equally during training, causing a strong bias towards the easier task of recommending repeat POIs. Second, users' visits to new POIs are more complex than those to repeat POIs. Frequent visits to a certain POI imply higher round-trip transportation costs. Therefore, due to cost considerations, users generally tend to repeatedly visit a geographically proximate POI instead of a remote POI. In contrast, for new POIs, users are willing to explore not only those near their locations, but also those located at considerable distance. For example, a Foursquare user in Boston likes bars and has been actively searching for five different bars with a mean distance of 1.87*km*. This behavior can be termed *local exploration*. Meanwhile, users might be attracted to a famous but distant unvisited POI, and pay transportation costs for an initial visit, e.g., a food enthusiast may choose to experience a highly acclaimed restaurant far from his/her home. We term this behavior as *global exploration*. The coexistence of local and global exploration poses a challenge for modeling users' visits to new POIs, which, however, has been overlooked by previous work.

To take a closer look at the characteristics of visit patterns for repeat and new POIs, we further analyze the geographical distances between the POIs visited and the main active areas of users[1]. As shown in Fig. 2, users in Boston and Paris are willing to visit nearby POIs, whether they are repeat or new POIs. However, the left-skewness is much sharper and the mean value is much lower for repeat POIs than for new POIs, reflecting a

[1]For ease of computation, we simplify a user's main active area to the POI most frequently visited by him/her.

stronger impact of global exploration in users' visits to new POIs. Based on these observations, it is imperative to develop a customized method for N^{2}PR.

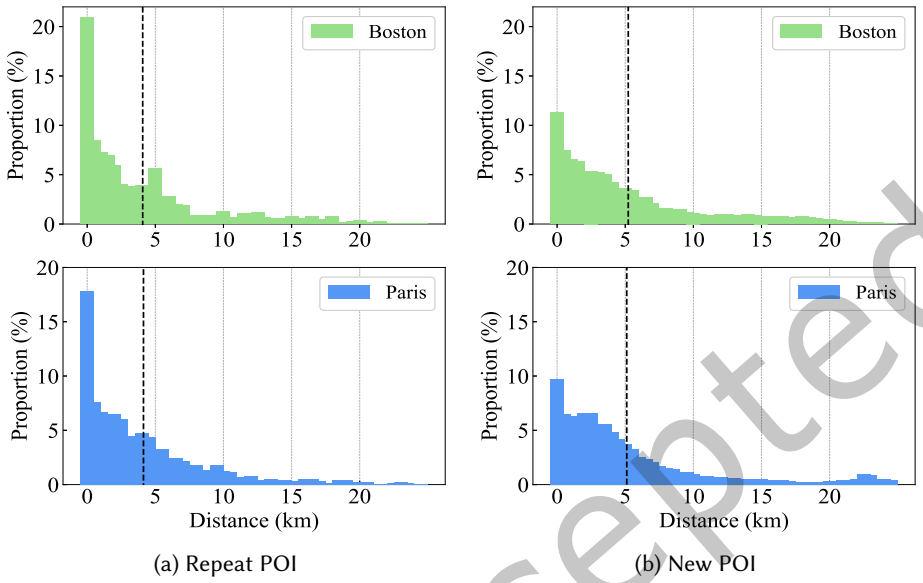


Fig. 2. Visit distance distributions for repeat POIs and new POIs in Boston and Paris city datasets collected by Foursquare. The bold dashed line indicates the mean value of the distribution.

To this end, we introduce **LGE** (Local and Global Exploration), a customized method for the next new POI recommendation. The fundamental flow of LGE is to first recommend POIs from the perspectives of local exploration and global exploration separately and then combine the results to generate the final recommendation. Specifically, LGE consists of the following three modules:

- A Zone-Aware Local Exploration (ZLE) module, which encourages users to explore new POIs in the neighborhood using zone information. ZLE first conducts contrastive learning to learn zone-aware POI representations and then introduces a zone prediction regularizer to supervise the POI prediction process. This approach gives LGE the ability to recommend not only the target POI but also other POIs in the same zone, ultimately increasing the exposure rate of nearby unexplored POIs to users.
- An Intention-Aware Global Exploration (IGE) module, which recommends POIs that meet user intentions without distance constraints. IGE encodes the sequential and distributional category features of the user visit history to extract dynamic and static intentions. The dynamic intention varies over time, while the static one remains stable. Both intentions are spatially agnostic and can be utilized to search for unexplored POIs globally.
- A fusion module that aggregates the results of the ZLE and IGE modules for the final recommendation list. This module involves two fusion strategies: The Mean Pooling (MP) strategy and the Weighted Pooling (WP) strategy. The MP strategy simply aggregates the outputs of two exploration modules by averaging, while the WP strategy first calculates users' future exploration trend and then combines the outputs by weights.

To prove the effectiveness of LGE on the N^{2}PR task, we conduct extensive experiments on four real-world city datasets collected from two platforms, and the results show that our approach outperforms the existing NPR and N^{2}PR methods.

2 Related Work

2.1 Next POI Recommendation

Next POI recommendation task is currently a research hotspot, with the aim of predicting the next POIs that align with users' demands based on their check-in history. In this subsection, we first discuss methods based on conventional techniques, and then introduce methods grounded in language models.

Pioneering Markov Chain (MC) based methods [3, 7, 9, 45, 49] hold the naive Markov assumption and model transition relation between two consecutively visited POIs with geographical/temporal constraints. To learn more complex patterns, subsequent Recurrent Neural Network (RNN) based models [11, 22, 30, 48] are proposed, of which a typical practice is utilizing time differentials or geographical distances to control the influence of historical records on the current state. Although performing better than MC based methods, RNN based methods still suffer from the long-term dependency issue. To overcome this problem, Self-Attention Mechanism (SAM) are further employed to learn the dependency between any pair of POIs within the user's historical check-in sequence [5, 6, 15, 18, 23, 46, 47]. In the last few years, there is a growing trend to adopt Graph Neural Network (GNN) in next POI recommendation [12, 17, 19, 21, 26, 27, 36, 37, 39, 40, 42, 43, 51], which possesses the capability to learn from global views. By performing convolution on constructed graphs, e.g., geographical graph, transitional graph, or collaborative graph, GNN is able to enhance the expressiveness of POI representations. Recently, researchers have gradually become interested in disentangling diverse factors that drive user movements based on graph learning [12, 26, 29, 52]. For example, Sun et al. [29] disentangle transferable and non-transferable knowledge by learning different graphs to achieve cross-city knowledge transfer, while Zhou et al. [52] disentangle causal and bias knowledge in the constructed graphs to mitigate the bias issue.

Inspired by the success of language models, researchers have started to investigate their potential for next POI recommendations. One line of studies leverages the textual knowledge of language models to enhance representations [1, 4]. For example, STKG-PLM [1] and POI-Enhancer [4] describe user behaviors and POI characteristics with texts and use language models to encode them into vectors as auxiliary features. The other line of studies relies on the reasoning capabilities of Large Language Models (LLMs) and directly employs them as recommender systems. Typically, these methods contain two stages [2, 8, 14, 33]. The tokenization stage converts POIs into discrete IDs that can be processed by LLMs. The generation stage formulates the recommendation task in text and prompts LLMs to generate recommendations. For example, LLM4POI [14] and LLMMove [8] use random IDs to represent POIs and design sophisticated textual prompts that contains various features. Furthermore, the most recent GNPR-SID [33] and QT-Mob [2] tokenize POIs into semantic IDs which reveal the semantic relationships among POIs.

These efforts have been continuously improving the overall performance in recommending the next POI. However, they do not take into account the characteristics of user behavior to new POIs, resulting in deficient new POI recommendation.

2.2 Next New POI Recommendation

In this subsection, we first introduce related work from the cold-start perspective, then discuss methods tailored for new POI recommendation.

New POIs that have not been visited by the target user can be regarded as user-specific cold-start POIs, i.e., POIs are cold-start only to the target users. Therefore, next new POI recommendation is actually a special case of the cold-start sequential recommendation. In the literature, one straight solution for the cold-start problem is using side information. For instance, CMCLRec [41] and MML [24] utilize the multimodal information of users or items to enhance representations. Pembek et al. [25] investigate the impact of content information on the cold-start problem. However, in the real scenario, side information is not always accessible due to the privacy policy. To this end, meta-learning become a more promising solution for the cold-start problem, which transfers

knowledge from well-learned users or items to cold-start ones. For example, the metric-based method Mecos [50] collects evidence from constructed support sets for cold-start item recommendation. MetaCSR [10], MetaTL [35], and CTM [38] are optimization-based methods which learn good model initialization from constructed few-shot learning tasks for incoming cold-start tasks. Overall, existing methods have achieved considerable success in dealing with the cold-start problems, but they are mainly targeted at data-specific cold-start users or items with few interactions across the whole dataset. In contrast, our setting concerns user-specific cold-start POIs that are unseen to the target users, which falls outside the scope of existing methods.

Next new POI recommendation is a long-standing and challenging problem, which is harder than repeat POI recommendation. Early work [3] studies the dynamics of user visits to new POIs, showing that users are willing to check-in new POIs in the near future, especially those closer in proximity. Then, to address this issue, the community has proposed several methods [7, 16, 20] with a general idea of learning the relationships between users and unvisited POIs. PRME-G [7] uses a pair-wise ranking loss to learn unobserved POI transition patterns. Li *et al.* [16] incorporate side information, e.g., contextual features, to build connections between unvisited POIs and visited POIs. The most recent work JTLL [20] develops a triplet loss to learn additional unvisited relations between users and POIs. Nevertheless, these methods overlook the coexistence of local and global exploration in users' visits to new POIs, and their designed supervision signals are too weak to handle the complex exploration behavior. In contrast, our method thoroughly learns the exploration nature of users from both local and global views, increasing the likelihood that users encounter correct new POIs.

3 Problem Statement

Let $\mathcal{U} = \{u_1, u_2, ..., u_{|\mathcal{U}|}\}$ be a set of users, $\mathcal{V} = \{v_1, v_2, ..., v_{|\mathcal{V}|}\}$ be a set of POIs, $\mathcal{C} = \{c_1, c_2, ..., c_{|\mathcal{C}|}\}$ be a set of categories. Each POI $v \in \mathcal{V}$ has a unique GPS coordinate (lat, lon) of latitude and longitude. Based on GPS coordinates, we apply the K-Means algorithm to group geographically proximate POIs into zones and formulate the result as $\mathcal{Z} = \{z_1, z_2, ..., z_{|\mathcal{Z}|}\}$. Then each POI v_i is associated with a category c_i and a zone z_i. For each user $u \in \mathcal{U}$, his/her historical check-in sequence is denoted by $S^u = [s_1^u, s_2^u, ..., s_k^u]$, where each check-in tuple $s_i^u = (v_i, c_i, z_i, t_i)$ indicates that user u visits POI v_i characterized by category c_i and zone z_i at timestamp t_i. Subsequently, we denote the historical POI, category, and zone sequences of the user u as $S_v^u = [v_1, v_2, ..., v_k]$, $S_c^u = [c_1, c_2, ..., c_k]$, and $S_z^u = [z_1, z_2, ..., z_k]$, respectively.

Problem 1 (Next New POI Recommendation). Give a user u and his/her historical check-in sequence S^u, the goal is to recommend u a suitable new POI $v_{k+1} \notin S_v^u$, one that u is inclined to visit next and has **not** been visited by u before.

4 The Proposed Model

In this section, we present the details of our proposed method LGE, as shown in Fig. 3, which consists of three main modules: 1) a Zone-Aware Local Exploration (ZLE) module to broaden users' exploration range in the nearby area with a zone-aware contrastive loss and a zone prediction regularizer; 2) an Intention-Aware Global Exploration (IGE) module, which recommends POIs that meet user intentions without distance constraints by modeling the sequential and distributional category features; 3) a fusion module with a Mean Pooling (MP) strategy and a Weighted Pooling (WP) strategy to combine the outputs of local and global exploration modules for the final recommendation.

4.1 Zone-Aware Local Exploration

Modern civilization creates different zones in cities [44], and POIs located in the same zone tend to show similarity in multiple aspects, such as function and popularity. When users visit a POI in a particular zone, they probably have an interest in exploring other POIs in the same zone. Inspired by this fact, we propose a Zone-Aware

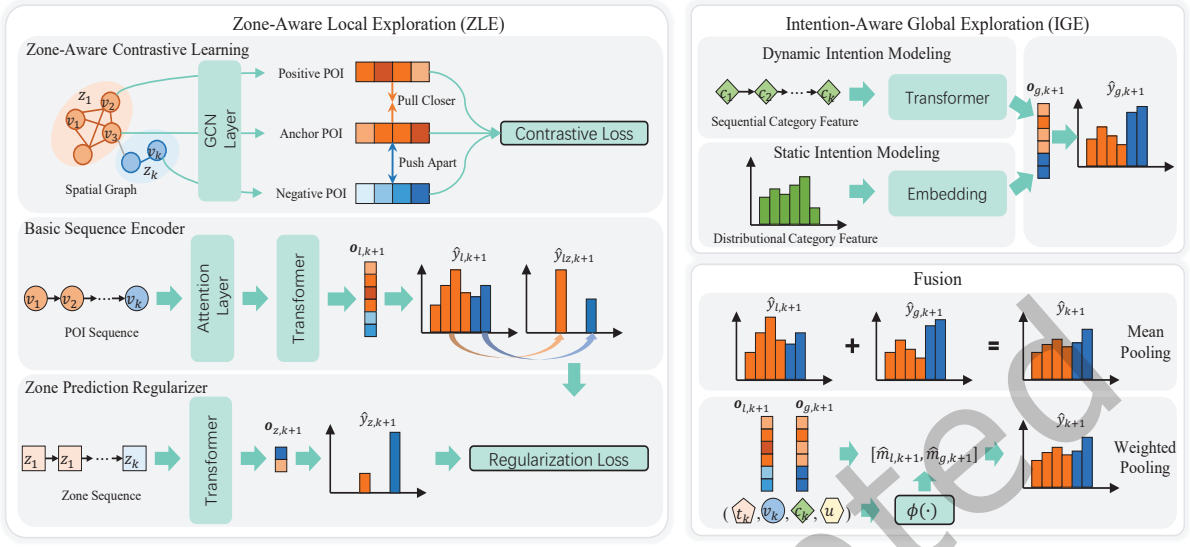


Fig. 3. The overall architecture of LGE.

Local Exploration (ZLE) module to help users explore nearby POIs using the zone information. The ZLE module is built upon a basic sequence encoder for user behavior modeling. Furthermore, ZLE learns zone-aware POI representations through contrastive learning and supervises the POI prediction process with a zone prediction regularizer.

4.1.1 Basic Sequence Encoder. This encoder models the POI sequence $S_v^u = [v_1, v_2, ..., v_k]$ and predicts the next POI v_{k+1}. Following previous work [34, 43], we adopt the widely used transformer [32] as our encoder due to its outstanding performance in sequential data processing.

Given a POI sequence S_v^u, we extract its initialized sequence representation $[\mathrm{v}_{l,1}; \mathrm{v}_{l,2}; ...; \mathrm{v}_{l,k}]$ from a POI embedding matrix $\mathrm{V}_l \in \mathbb{R}^{|\mathcal{V}| \times d}$. Here, we use l to denote the parameters of the **l**ocal exploration module ZLE, distinguishing them from their counterparts in the subsequent global exploration module IGE[2]. d is the embedding dimension. Before sending the sequence representation to a transformer, we first enrich $\mathrm{v}_{l,i}$ with its located zone z_i. Specifically, we introduce an attention layer and aggregate all POIs in the zone z_i to compute an enhanced representation $\mathrm{h}_{l,i}$ for v_i. The calculation is defined as:

$$\mathrm{h}_{l,i} = \sum\nolimits_{v_j \in \mathcal{N}_{z_i}} a_{ij} \mathrm{v}_{l,j}, \tag{1}$$

$$a_{ij} = \mathrm{Softmax}\left(\frac{\mathrm{v}_{l,i}^{\top} \mathrm{v}_{l,j}}{\sum_{v_n \in \mathcal{N}_{z_i}} \mathrm{v}_{l,i}^{\top} \mathrm{v}_{l,n}}\right), \tag{2}$$

where $a_{i,j}$ is the similarity score between $\mathrm{v}_{l,i}$ and $\mathrm{v}_{l,j}$, $\mathcal{N}_{z_i}$ is a POI set contains all POIs located in the zone z_i. Through the attention layer, the enhanced representation $\mathrm{h}_{l,i}$ integrates the zone information, which implicitly strengthens the module's ability to recommend local POIs.

[2]In the **g**lobal exploration module IGE, we adopt g for notation.

Then we stack the enhanced representations to form an input tensor $[\mathrm{h}_{l,1}^{(0)}; \mathrm{h}_{l,2}^{(0)}; ...; \mathrm{h}_{l,k}^{(0)}]$ and send it to a L-layer transformer. For brevity, we choose not to present the computation process of the transformer, as it is already known to the community. After L layers, we obtain an output vector $\mathrm{h}_{l,k}^{(L)} \in \mathbb{R}^d$ from the last layer and use it for the next new POI prediction on a local scale. The predicted score $\mathrm{o}_{l,k+1} \in \mathbb{R}^{|\mathcal{V}|}$ and the probability distribution $\hat{\mathrm{y}}_{l,k+1}$ for local exploration are defined as:

$$\mathrm{o}_{l,k+1} = \mathrm{V}_l \mathrm{h}_{l,k}^{(L)}, \tag{3}$$

$$\hat{\mathrm{y}}_{l,k+1} = \mathrm{Softmax}(\mathrm{o}_{l,k+1}). \tag{4}$$

For training the encoder, we employ the Cross-Entropy loss as:

$$\mathcal{L}_l = -\mathrm{y}_{l,k+1}^{\top} \log(\hat{\mathrm{y}}_{l,k+1}), \tag{5}$$

where $\mathrm{y}_{l,k+1}$ is the one-hot label of the next ground-truth new POI v_{k+1}.

4.1.2 Zone-Aware Contrastive Learning. Although the representation $\mathrm{h}_{l,i}$ has integrated the zone information, the similarity score $a_{i,j}$ to combine neighbor POIs in the same zone is solely supervised by the next visit label, which is often sparse and weak in the data. To this end, we perform contrastive learning to explicitly learn zone-aware POI representations. The main idea is to pull the embeddings of POIs in the same zone closer in the latent space and to push those of POIs in different zones apart. These embeddings allow LGE to sense unexplored POIs around the target POI, increasing their chances of being recommended.

However, directly employing contrastive loss on V_l is inappropriate because two nearby POIs could be assigned to different zones. The assignment might cause a big gap between the embeddings of adjacent POIs. To alleviate this problem, we introduce a spatial graph $\mathcal{G}$ and use the spectral GCN to smooth POI embeddings, ensuring that each POI embedding can perceive its geographical neighborhoods. Formally, we define the spatial graph $\mathcal{G}$ as follows:

Definition 1 (Spatial Graph). The spatial graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathrm{A}\}$ is an undirected graph, where $\mathcal{V}$ and $\mathcal{E}$ are the sets of POI nodes and edges, respectively. $\mathrm{A} \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}$ denotes the adjacent matrix. $\mathrm{A}(i, j) = 1$ if the geographical distance between POI v_i and v_j is less than $0.1km$ or POI v_i and v_j belong to the same zone, otherwise $\mathrm{A}(i, j) = 0$.

Then, we apply the spectral GCN over V_l to get a smoothed POI embedding matrix $\widetilde{\mathrm{V}}_l \in \mathbb{R}^{|\mathcal{V}| \times d}$:

$$\widetilde{\mathrm{V}}_l = \mathrm{D}^{-\frac{1}{2}} \mathrm{A} \mathrm{D}^{-\frac{1}{2}} \mathrm{V}_l, \tag{6}$$

where D is the degree matrix of the spatial graph $\mathcal{G}$.

To perform contrastive learning, we randomly sample POI triplets from the POI set $\mathcal{V}$. Each triplet (v_a, v_p, v_n) consists of an anchor POI v_a, a positive POI v_p, and a negative POI v_n. The positive POI v_p is selected from the same zone as the anchor POI v_a, while the negative POI v_n is selected from a different zone. Mention that the sampling process is constrained by only one condition, that is, the negative and target POIs should be from different zones. Geographic distance is not explicitly considered in our sampling strategy because it has been implicitly embedded within the zone information. Given (v_a, v_p, v_n) and $\widetilde{\mathrm{V}}_l$, the contrastive loss is formulated as a margin-based loss:

$$\mathcal{L}_{con}(v_a, v_p, v_n) = max(d_{a,p} - d_{a,n} + \alpha, 0), \tag{7}$$

$$d_{i,j} = ||\tilde{\mathrm{v}}_{l,i} - \tilde{\mathrm{v}}_{l,j}||_2, \tag{8}$$

where $\tilde{\mathrm{v}}_{l,i}, \tilde{\mathrm{v}}_{l,j} \in \widetilde{\mathrm{V}}_l$ are smoothed embeddings for v_i, v_j, respectively. α is a predefined margin, which is equal to 0.5 in our settings.

4.1.3 Zone Prediction Regularizer. After learning zone-aware POI representations, we move one step forward to the encoder and instruct it to recommend not only the target POI but also the target zone. In the real world, there exist sequential transition patterns among zones. For example, office workers usually commute from the office POI zones to the nearby restaurant POI zones for lunch in the afternoon. These patterns can provide hints to the basic encoder on which zone should be recommended to users for local exploration. In light of this, we propose a zone prediction regularizer to assist the encoder in recommending new POIs located in the correct zone.

Specifically, we introduce another L-layer transformer as the sequence encoder to capture zone transition patterns in the zone sequences. For a zone sequence S_z^u, we extract its initialized sequence representation $[\mathrm{z}_1; \mathrm{z}_2; ...; \mathrm{z}_k]$, where $\mathrm{z}_i \in \mathrm{Z}$ and $\mathrm{Z} \in \mathbb{R}^{|\mathcal{Z}| \times d}$ denotes the initial zone embedding matrix. Taking as input the sequence representation, the L-layer transformer outputs an encoded vector $\mathrm{h}_{z,k}^{(L)}$, which is adopted for the next zone prediction. The final predicted score and probability distribution are formulated as:

$$\mathrm{o}_{z,k+1} = \mathrm{Z}\mathrm{h}_{z,k}^{(L)}, \tag{9}$$

$$\hat{\mathrm{y}}_{z,k+1} = \mathrm{Softmax}(\mathrm{o}_{z,k+1}). \tag{10}$$

For optimization, we employ the Cross-Entropy loss again to get the objective $\mathcal{L}_z$.

With the goal of teaching the basic encoder to recommend the correct zones, we implement the regularizer by minimizing the KL divergence between $\hat{\mathrm{y}}_{z,k+1}$ for the next zone prediction and $\hat{\mathrm{y}}_{l,k+1}$ for the next POI prediction. Since they are in different dimensions, we transform $\hat{\mathrm{y}}_{l,k+1}$ into a $|\mathcal{Z}|$-dimensional distribution $\hat{\mathrm{y}}_{lz,k+1}$. The i-th value in $\hat{\mathrm{y}}_{lz,k+1}$ is equal to the sum of the probability scores in $\hat{\mathrm{y}}_{l,k+1}$ that correspond to the POIs located in zone z_i, denoting the predicted probability of the basic encoder to visit z_i. Finally, the regularization loss can be defined as:

$$\mathcal{L}_{reg} = KL(\hat{\mathrm{y}}_{lz,k+1} || \hat{\mathrm{y}}_{z,k+1}), \tag{11}$$

where $KL(\cdot||\cdot)$ denotes the Kullback Leibler (KL) divergence between two discrete distributions. Minimizing $\mathcal{L}_{reg}$ enhances the potential of the basic encoder to recommend the target zone z_{k+1}, directly increasing the exposure rate of unvisited POIs within z_{k+1}.

4.2 Intention-Aware Global Exploration

When exploring distant POIs, user intentions play a more critical role in decision-making than geographic factors. Generally, user intentions demonstrate the type of services that users want, reflecting the subjective reasons why users visit POIs. For example, a food enthusiast could be motivated by a dining intention to check-in a distant but highly rated restaurant, even though the transportation cost is high. Therefore, we design an Intention-Aware Global Exploration (IGE) module to accurately model user intentions and correspondingly recommend POIs without geographical constraints.

The IGE module models two types of user intentions: (1) dynamic intentions, which vary over time. For example, users usually search for restaurants to eat after work or look for hotels for accommodation after leaving airports. (2) static intentions, which remain stable and reflect the overall distributions of the user needs. A case in point is that young people tend to seek entertainment in bars, while elder people prefer to spend time in parks. Both types of user intentions are crucial for accurate recommendation. In practice, we derive user intentions from the category information of POIs, as it conveys the service types offered by POIs, e.g., restaurants provide the food services and bars provide the entertainment services. In IGE, dynamic intentions are captured by modeling category transition patterns in the sequential category features, while static intentions are extracted from users' distributional category features.

To be specific, we once again employ a L-layer transformer to encode the sequential category feature, i.e., the category sequence S_c^u, for dynamic intention modeling. By convention, we first generate an input sequence

representation $[\mathrm{c}_1; \mathrm{c}_2, ..., \mathrm{c}_k]$ for S_c^u through the initial category embedding matrix $\mathrm{C} \in \mathbb{R}^{|\mathcal{C}|\times d}$. Then we use the L-layer transformer to process the sequence representation and obtain an output vector $\mathrm{h}_{c,k}^{(L)}$, which encodes the dynamic user intention. For static intention modeling, we straightforwardly aggregate category embeddings by the distributional category feature $\mathrm{f}_u \in \mathbb{R}^{|\mathcal{C}|}$, which is formulated as:

$$\mathrm{h}_u = \mathrm{f}_u^\top \mathrm{C}, \tag{12}$$

where h_u encodes the static intention and the i-th index in f_u denotes the normalized frequency of user u visiting category c_i.

After obtaining $\mathrm{h}_{c,k}^{(L)}$ and h_u, we predict the next new POI on a global scale. The predicted score $\mathrm{o}_{g,k+1} \in \mathbb{R}^{|\mathcal{V}|}$ and the probability distribution $\hat{\mathrm{y}}_{g,k+1}$ for global exploration are:

$$\mathrm{o}_{g,k+1} = (\frac{1}{2}\mathrm{V}_g + \frac{1}{2}\mathrm{V}_c)(\frac{1}{2}\mathrm{h}_{c,k}^{(L)} + \frac{1}{2}\mathrm{h}_u), \tag{13}$$

$$\hat{\mathrm{y}}_{g,k+1} = \mathrm{Softmax}(\mathrm{o}_{g,k+1}), \tag{14}$$

where $\mathrm{V}_g \in \mathbb{R}^{|\mathcal{V}|\times d}$ is another POI embedding matrix for the global exploration module IGE, which is distinguished from V_l in the local exploration module ZLE. $\mathrm{V}_c \in \mathbb{R}^{|\mathcal{V}|\times d}$ is a category embedding matrix where the i-th row corresponds to the category embedding of POI v_i. Apparently, no geographical information is involved in the IGE module, making it possible to recommend distant but suitable POIs. To optimize the IGE module, we use the Cross-Entropy loss function over $\hat{\mathrm{y}}_{g,k+1}$ and get the objective $\mathcal{L}_g$.

4.3 Fusion

The ZLE and IGE modules provide different recommendation outputs from different perspectives. To generate the final result $\hat{\mathrm{y}}_{k+1}$, we propose two fusion strategies: Mean Pooling (MP) and Weighted Pooling (WP).

The MP strategy optimizes $\mathcal{L}_l$ and $\mathcal{L}_g$ individually at the training stage, and fuse the probability distributions $\hat{\mathrm{y}}_{l,k+1}$ and $\hat{\mathrm{y}}_{g,k+1}$ by a mean pooling operation at the test stage:

$$\hat{\mathrm{y}}_{k+1} = \hat{\mathrm{y}}_{l,k+1} + \hat{\mathrm{y}}_{g,k+1}. \tag{15}$$

However, the mean pooling operation could output a noisy result, as users might already have determined whether to explore nearby or distant POIs. For example, when a user decides to travel the next day, we should provide the IGE output to help the user choose a distant destination. Once the user has reached the destination, we should present the ZLE output to assist in exploring the local area. To achieve this, we propose the Weighted Pooling (WP) strategy, which first predicts the user's next exploration type and then outputs a weighted sum of the ZLE and IGE outputs.

More specifically, the WP strategy trains a decision prediction network $\Phi(\cdot)$, which could be a simple MLP together with a Softmax layer. $\Phi(\cdot)$ takes as input the current context of the target user, i.e., time, POI ID, POI category, and user ID, and outputs the probability score for each type of exploration. We can formulate this strategy as:

$$[\hat{m}_{l,k}, \hat{m}_{g,k}] = \Phi(\mathrm{t}_k, \mathrm{v}_{g,k}, \mathrm{c}_k, \mathrm{u}), \tag{16}$$

where $\hat{m}_{l,k}$ and $\hat{m}_{g,k}$ are predicted probability scores for local and global exploration, respectively. $\mathrm{t}_k \in \mathrm{T}$ denotes the hour embedding for the current timestamp t_k. Here we divide a day into 24 slots with one hour per slot and introduce a time embedding matrix $\mathrm{T} \in \mathbb{R}^{24\times d}$ for these 24 slots. Then we project t_k into one of the slots and create the time embedding t_k by looking up the time embedding matrix T. $\mathrm{v}_{g,k} \in \mathrm{V}_g$ denotes the POI embedding for v_k. We use V_g because the user's next decision is mainly dependent on his/her visit intention, which has been implicitly encoded in the IGE module. c_k is the category embedding for c_k and $\mathrm{u} \in \mathrm{U}$ represents the user embedding for user u. $\mathrm{U} \in \mathbb{R}^{|\mathcal{U}|\times d}$ is the initial user embedding matrix.

To obtain an accurate decision predictor $\Phi(\cdot)$, one obstacle that we face is the lack of user decision labels. Hence, we propose to manually annotate the decision labels. In real life, users often plan their activities for the next day at the end of the current day. Furthermore, users often begin with global exploration in the early stage of the day and then proceed to conduct local exploration after traveling to the target area. To verify this point, we analyze the relationship between users' movement distances and movement sequences during a day. Figure 4 shows the results of four cities across Foursquare and Gowalla platforms. From the results, we observe that users from both platforms are inclined to make long-distance visits at the beginning of their daily routes. Meanwhile, users' movement distances exhibit a general declining trend as the routes continue. This demonstrates the existence of the "first global exploration and then local exploration" pattern in users' daily routes.

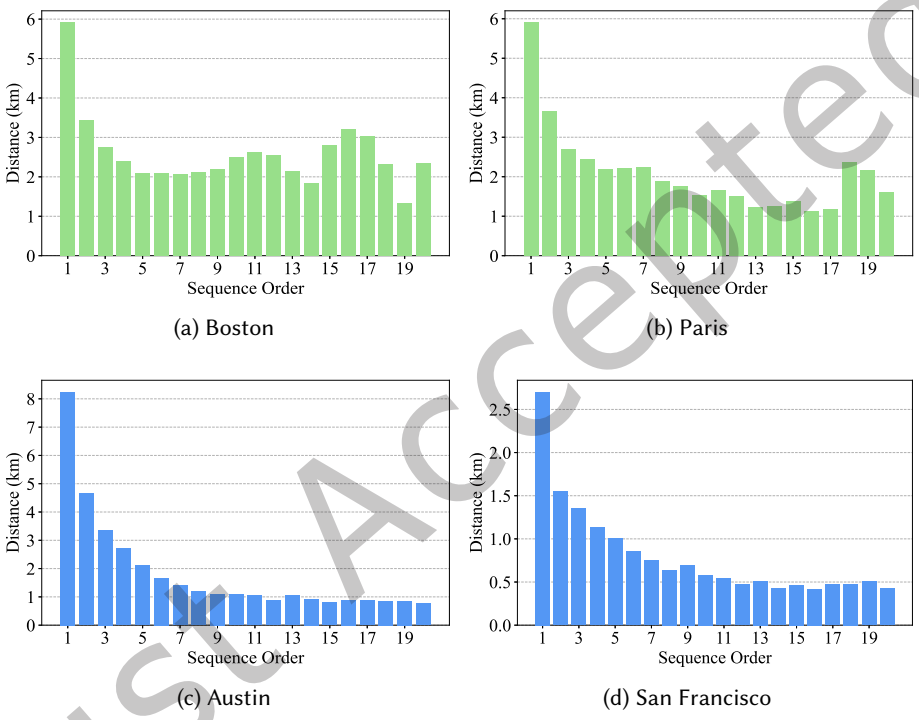


Fig. 4. The relationship between users' movement distances and the movement sequences during a day. In each sub-figure, a smaller order number indicates an earlier movement in the user's daily route sequence. For a fair comparison, we consider the first 20 visits in the users' daily routes across all platforms. Boston and Paris are from the Foursquare platform, Austin and San Francisco are from the Gowalla platform.

Consequently, we raise an annotation rule that users need global exploration recommendation after they finish a day's trajectory, and need local exploration recommendation during the trajectory. The labeling process can be defined as:

$$m_k = \begin{cases} 0 & day(t_{k+1}) = day(t_k) \\ 1 & day(t_{k+1}) \neq day(t_k) \end{cases}, \tag{17}$$

where m_k is the label for training $\Phi(\cdot)$. The function $day(\cdot)$ maps the input timestamp to a specific date.

With the annotated label m_k and the predicted score $[\hat{m}_{l,k}, \hat{m}_{g,k}]$, we turn m_k into the one-hot label and employ the Cross-Entropy loss to calculate the objective $\mathcal{L}_\Phi$ for training $\Phi(\cdot)$.

Finally, we fuse the outputs of the ZLE and IGE modules for the final scores $\hat{y}_{k+1}$ by the weights $\hat{m}_{l,k}$ and $\hat{m}_{g,k}$:

$$o_{k+1} = \hat{m}_{l,k} o_{l,k+1} + \hat{m}_{g,k} o_{g,k+1}, \tag{18}$$

$$\hat{y}_{k+1} = \text{Softmax}(o_{k+1}). \tag{19}$$

With $\hat{y}_{k+1}$, we apply the Cross-Entropy loss to get a joint loss $\mathcal{L}_{fus}$ for adaptively optimizing the ZLE and IGE modules.

4.4 Training

The training of LGE depends on the fusion strategy we use. **LGE (MP)** adopts the Mean Pooling strategy. In LGE (MP), the ZLE and IGE modules are trained separately and the overall loss is defined as:

$$\mathcal{L}_{MP} = \mathcal{L}_l + \mathcal{L}_z + \lambda \mathcal{L}_{con} + \beta \mathcal{L}_{reg} + \mathcal{L}_g, \tag{20}$$

where $\mathcal{L}_l$, $\mathcal{L}_z$, $\mathcal{L}_{con}$, and $\mathcal{L}_{reg}$ are for training the ZLE module. $\mathcal{L}_g$ is for training the IGE module. λ and β are hyper-parameters to control the impacts of contrastive loss and regularization loss in the ZLE module.

LGE (WP) utilizes the WP strategy, where the ZLE and IGE modules are jointly optimized by minimizing $\mathcal{L}_{fus}$ and the decision prediction network also needs to be trained by optimizing $\mathcal{L}_\Phi$. Finally, the overall loss is formulated as:

$$\mathcal{L}_{WP} = \mathcal{L}_{fus} + \mathcal{L}_z + \lambda \mathcal{L}_{con} + \beta \mathcal{L}_{reg} + \mathcal{L}_\Phi. \tag{21}$$

5 Experiment

5.1 Experimental Setup

5.1.1 Datasets. We evaluate LGE on two real-world datasets collected by the Foursquare [3] and Gowalla [4] platforms, both of which contain check-ins from numerous cities. The Foursquare dataset was collected from April 2012 to September 2013. The Gowalla dataset consists of check-ins from February 2009 to October 2010. Each check-in record is formulated as a tuple of user ID, POI ID, and timestamp. Each POI is associated with a category and a pair of GPS coordinates. To release the computational burden, we select two representative city datasets from each large dataset for evaluation, including **Foursquare-Boston (FB)**, **Foursquare-Paris (FP)**, **Gowalla-Austin (GA)**, and **Gowalla-San Francisco (GS)**. The statistics of these city datasets are shown in Table 1.

For dataset partition, we follow the traditional way to keep consistent with existing studies. Specifically, we divide each user's check-in sequence into train/validation/test sets in chronological order. We select the first 70% check-ins as the training set, the middle 10% as the validation set, and the remaining 20% as the test set. Table 1 further presents proportions of new POI visits in train, validation, and test sets.

5.1.2 Baselines. We compare LGE with the following 10 baselines which are categorized into six types:

Traditional NPR method:

- **FPMC** [28]: A classic sequential model that combines matrix factorization and Markov Chain to capture low-order transition patterns.

RNN based NPR method:

- **STLSTM** [11]: This is a classic sequential POI recommendation model based on LSTM, which refines the gate mechanism in LSTM with temporal and spatial information.

[3] https://sites.google.com/site/yangdingqi/home
[4] http://snap.stanford.edu/data/loc-gowalla.html

Table 1. Dataset Statistics. "#NP" denotes the record number of new POI visits. "#ND" denotes the data density w.r.t new POI visits. "#PTrain", "#PValid", and "#PTest" denote the proportions of new POI visits in train, validation, and test sets, respectively.

Dataset	#User	#POI	#Record	#NP	#ND	#PTrain	#PValid	#PTest
Foursquare-Boston (FB)	4,595	14,445	104,181	55,315	0.0008	0.5645	0.4031	0.4406
Foursquare-Paris (FP)	6,904	19,838	111,277	63,519	0.0005	0.6054	0.4267	0.4741
Gowalla-Austin (GA)	7,841	18,252	317,199	203,860	0.0014	0.6779	0.5489	0.5410
Gowalla-San Francisco (GS)	6,105	10,631	144,505	97,029	0.0015	0.7055	0.5550	0.4406

Attention based NPR methods:

- **TiSASRec** [13]: A classic attention-based model incorporates temporal interval information into the calculation of attention weights.
- **STiSAN** [34]: A recent attention-based model adapts the positional encoder to be time-aware and uses scaled spatio-temporal intervals to adjust the attention weights.

Graph based NPR methods:

- **GETNext** [34]: A recent graph-based model aggregates all users' trajectories into a global flow map, which enhances the learning of collaborative signals from all users.
- **Graph-FlashBack** [27]: A representative graph-based model learns a weighted POI transition graph from a customized spatial-temporal knowledge graph, which can be utilized to enrich POI representations.
- **ARGAN** [40]: A recent graph-based model employs the graph structure learning method to adaptively learn an optimized POI-POI graph structure, which suits the task of the next POI recommendation.
- **EEDN** [36]: A representative hypergraph-based model utilizes item and user hypergraphs to enhance user and POI representations, which can mitigate the cold-start problem.

Graph based N^2PR method:

- **JTLL** [20]: A state-of-the-art method for N^2PR, which extends a graph-based NPR model HMT-GRN [19] by adding a triplet loss to learn relations between users and unvisited POIs, which boosts HMT-GRN's ability to recommend new POIs.

LLM based NPR method:

- **GNPR-SID** [20]: A state-of-the-art LLM based method for NPR. GNPR-SID first generates semantic IDs for POIs based on their semantic and collaborative features, and then constructs a sequential POI prediction prompt to fine-tune LLMs for recommendation. GNPR-SID chooses Llama3-8B as its backbone.

5.1.3 Evaluation Metric. We use two widely adopted evaluation metrics: Hit Rate at a cutoff top k (Recall@k) and Normalized Discounted Cumulative Gain at a cutoff top k (NDCG@k). HR@k measures the proportion of the ground-truth POI being included in the top-k recommendation list, while NDCG@k measures the quality of the result recommendation list and assigns a higher weight to a higher rank position. Considering a test set with n samples, the Recall@k and NDCG@k are formulated as:

$$HR@k = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(rank_i < k),$$

$$NDCG@k = \frac{1}{n}\sum_{i=1}^{n} \frac{\mathbb{1}(rank_i < k)}{log_2(rank_i + 1)},$$

where n denotes the number of samples, $rank_i$ is the rank of the target POI in the recommendation list for sample i, and $\mathbb{1}$ is an indicator function that only returns 1 when the condition is met. In our reports, k is from {5, 10, 20, 50}

5.1.4 Implementation Details. The key hyper-parameters and training settings of LGE are listed below. We set the dimensions of all the embedding matrices to 256. Each transformer has 4 layers and 4 heads. λ is set to 20 for Foursquare-Boston, Foursquare-Paris, and Gowalla-San Francisco, and 5 for Gowalla-Austin. β is set to 10 for Foursquare-Boston and Foursquare-Paris, 1 for Gowalla-Austin, and 5 for Gowalla-San Francisco. The initial cluster size n for K-Means is set to 1000 for all city datasets. We use Adam optimizer with the learning rate of $1e^{-4}$ and the batch size of 256 to train our model. For all baselines, we perform a grid search to find the optimized setting of common parameters, including the learning rate from $\{1e^{-1}, 1e^{-2}, 1e^{-3}, 1e^{-4}\}$, the batch size from {64, 128, 256}, and the embedding size from {64, 128, 256}. To fully exploit the potential to recommend new POIs, all models are trained only on visits to new POIs. We implement all models with Pytorch and conduct experiments on Nvidia GeForce RTX 4090 (24GB memory). We repeat each experiment five times with different random seeds and report the average score.

5.2 Performance Comparison

Table 2 shows the experimental results of all models on four city datasets. From these results, we have the following observations.

Table 2. N^{2}PR Performance comparison with baselines on four city datasets. For each row, the best results are highlighted in bold, and the second best scores are underlined. LGE (MP) uses the Mean Pooling strategy, while LGE (WP) uses the Weighted Pooling strategy. * denotes the statistically significant improvement compared to the strongest baseline ($p < 0.01$).

Data	Metric	FPMC	STLSTM	TiSASRec	STiSAN	GETNext	Graph FlashBack	ARGAN	EEDN	JTLL	GNPR SID	LGE (MP)	LGE (WP)	Imp.
FB	HR@5	0.0439	0.0430	0.0435	0.0366	0.0293	0.0448	0.0401	0.0458	0.0442	0.0408	**0.0597***	0.0569	30.3%
	NDCG@5	0.0311	0.0314	0.0324	0.0284	0.0204	0.0329	0.0270	0.0333	0.0315	0.0318	**0.0412***	0.0387	23.7%
	HR@10	0.0590	0.0583	0.0600	0.0483	0.0451	0.0587	0.0571	0.0597	0.0574	0.0492	**0.0851***	0.0815	42.5%
	NDCG@10	0.0359	0.0363	0.0376	0.0321	0.0255	0.0373	0.0324	0.0337	0.0357	0.0346	**0.0493***	0.0466	31.1%
FP	HR@5	0.0353	0.0485	0.0419	0.0x411	0.0355	0.0453	0.0429	0.0507	0.0345	0.0464	**0.0568***	0.0534	12.0%
	NDCG@5	0.0220	0.0296	0.0259	0.0275	0.0222	0.0281	0.0277	0.0332	0.0238	0.0317	**0.0343***	0.0331	3.31%
	HR@10	0.0612	0.0668	0.0648	0.0601	0.0596	0.0674	0.0653	0.0683	0.0494	0.0634	**0.0886***	0.0835	29.7%
	NDCG@10	0.0303	0.0357	0.0332	0.0336	0.0300	0.0352	0.0350	0.0389	0.0286	0.0372	**0.0445***	0.0427	14.4%
GA	HR@5	0.0233	0.0510	0.0620	0.0437	0.0601	0.0333	0.0700	0.0272	0.0653	0.0394	0.0817	**0.0827***	18.1%
	NDCG@5	0.0170	0.0340	0.0320	0.0272	0.0416	0.0202	0.0385	0.0204	0.0431	0.0272	0.0479	**0.0490***	13.7%
	HR@10	0.0299	0.0751	0.1007	0.0710	0.0801	0.0577	0.1086	0.0383	0.0944	0.0529	0.1295	**0.1313***	20.9%
	NDCG@10	0.0191	0.0418	0.0445	0.0360	0.0481	0.0280	0.0509	0.0240	0.0524	0.0315	0.0633	**0.0647***	23.5%
GS	HR@5	0.0417	0.0430	0.0684	0.0392	0.0603	0.0398	0.0725	0.0262	0.0698	0.0168	0.0888	**0.0914***	26.1%
	NDCG@5	0.0313	0.0305	0.0359	0.0222	0.0441	0.0250	0.0425	0.0169	0.0502	0.0115	0.0544	**0.0561***	11.8%
	HR@10	0.0572	0.0638	0.1041	0.0702	0.0794	0.0673	0.1096	0.0383	0.1022	0.0279	0.1392	**0.1427***	30.2%
	NDCG@10	0.0363	0.0372	0.0474	0.0321	0.0502	0.0338	0.0544	0.0208	0.0606	0.0151	0.0703	**0.0726***	19.8%

(1) Our proposed models, including LGE (MP) and LGE (WP), outperform all baselines for all evaluation metrics by a significant margin. Compared with the best baseline results, our models achieve an average improvement of 21.9%. In detail, the improvement ratios are 21.6%, 13.1%, 30.8%, and 22.2% in terms of HR@5, NDCG@5, HR@10, and NDCG@10, respectively. Overall, the results have clearly demonstrated the superiority of our model in recommending next new POIs.

(2) Graph-based models are more promising for the N^{2}PR task than RNN- and attention-based models. In particular, we observe that EEDN is the best baseline on FB and FP datasets, while ARGAN and JTLL are two strongest

baselines on the GA and GS datasets. This phenomenon is mainly due to the fact that their constructed graphs establish connections between users' visited POIs and unvisited POIs. Consequently, modeling these graphs increases the exposure probability of new POIs. However, these graph-based models are solely supervised by the labels of users' next visits, which are often sparse and weak, causing the insufficiency in finding new POIs. In contrast, our proposed LGE introduces informative supervisory signals, e.g., zone-aware contrastive learning and zone prediction regularizer, which enhance LGE's ability to perceive new POIs.

(3) Compared with the RNN-based model STLSTM, the advantages of attention-based models TiSASRec and STiSAN are not obvious. One main reason is that the attention mechanisms in TiSASRec and STiSAN are only utilized for capturing visited POI dependencies within users' historical POI sequences. Such conventional operation has the same drawback as the RNN-based model that the dependencies in unvisited POIs are totally ignored. Hence TiSASRec and STiSAN cannot consistently outperform the STLSTM.

(4) Comparing LGE (MP) and LGE (WP), we find that the Weighted Pooling strategy produces a positive impact on Gowalla datasets (i.e., GA and GS) but fails to work on Foursquare datasets (i.e., FB and FP), suggesting distinct data characteristics. In the Gowalla platform, users are encouraged to share traveling experiences, which often follow a pattern of first global exploration and then local exploration. Related evidence can be drawn from Figure 4, which shows the relationship between user's movement distances and movement sequences. In Figure 4, the histograms of Gowalla are noticeably more left-skewed than those of Foursquare. Meanwhile, the histograms of Gowalla have a lower average visit distance than those of Foursquare (2.59km for FB, 2.10km for FP, 1.82km for GA, 0.80km for GS). These observations indicate that the pattern of first global exploration and then local exploration is more pronounced for Gowalla. Hence predicting next exploration type is much easier on Gowalla datasets for the decision prediction network $\Phi(\cdot)$, leading to a better performance of LGE (WP). In contrast, users' exploration decisions on the Foursquare platform are more casual, resulting in an inferior performance. To show more evidence, we will further analyze the decision accuracy of $\Phi(\cdot)$ in the following in-depth analysis.

(5) EEDN and GraphFlashBack are competitive on Foursquare datasets (i.e., FB and FP) but produce unsatisfactory results on Gowalla datasets (i.e., GA and GS). In the Gowalla platform, users' next movements are closely correlated with the recent visits during their travels. However, EEDN lacks the capability of encoding historical sequences and GraphFlashBack merely relies on a simple RNN for sequence modeling. Hence, they are unable to fully learn the latent clues from the recent visited POIs, ultimately leading to poor outcomes.

(6) LLM based method GNPR-SID still underperforms LGE in terms of next new POI recommendation across all datasets, verifying the consistent superiority of our proposed method LGE. This finding further supports our claim that users' exploration behaviors should be handled with specific consideration.

5.3 Ablation Study

We conduct an ablation study to verify the contribution of each proposed component for N^2PR. To compare with the full version, we implement six degraded variants of our model, including:

- **w/o ZLE** removes the zone-aware local exploration module and only recommends POIs on a global scale.
- **w/o ZLE**$_{cl}$ removes the contrastive loss for learning zone-aware POI representations in the ZLE module.
- **w/o ZLE**$_{rl}$ removes the regularization loss for supervising the POI prediction process in the ZLE module.
- **w/o IGE** removes the intention-aware global exploration module and only recommends POIs in the local area.
- **w/o IGE**$_{st}$ removes the modeling of static user intentions in the IGE module and only models dynamic user intentions.
- **w/o IGE**$_{dy}$ removes the modeling of dynamic user intentions in the IGE module and only models static user intentions.

The HR@5 and NDCG@5 results on four datasets are shown in Table 3. From the results, we have the following findings.

Table 3. Ablation study on key components of LGE. "Full" denotes the full model that produces the best results.

Model	FB		FP		GA		GS	
	HR@5	NDCG@5	HR@5	NDCG@5	HR@5	NDCG@5	HR@5	NDCG@5
Full	**0.0597**	**0.0412**	**0.0568**	**0.0343**	**0.0827**	**0.0490**	**0.0914**	**0.0561**
w/o ZLE	0.0438	0.0325	0.0455	0.0288	0.0421	0.0263	0.0403	0.0272
w/o ZLE$_{cl}$	0.0512	0.0346	0.0504	0.0305	0.0820	0.0478	0.0874	0.0518
w/o ZLE$_{rl}$	0.0562	0.0392	0.0553	0.0340	0.0797	0.0477	0.0886	0.0546
w/o IGE	0.0587	0.0392	0.0532	0.0325	**0.0827**	0.0485	0.0894	0.0533
w/o IGE$_{st}$	0.0583	0.0400	0.0548	0.0333	0.0820	0.0487	0.0898	0.0547
w/o IGE$_{dy}$	0.0585	0.0400	0.0541	0.0330	**0.0827**	0.0488	0.0897	0.0549

(1) After removing the ZLE module, the w/o ZLE model suffers from an obvious decrease in recommendation performance on four datasets. This shows that encouraging users to explore nearby POIs is apparently beneficial for accurate new POI recommendations.
(2) Removing the contrastive loss or the regularization loss will cause a performance decrease, confirming their positive impacts on the recommendation. This also implies that zone information indeed helps predict users' future movements in the local area. Furthermore, the contrastive loss is more vital than the regularization loss, as w/o ZLE$_{cl}$ performs worse than w/o ZLE$_{rl}$ in most cases. This could be attributed to the inaccuracies in the zone predictions, which mislead the POI predictions.
(3) The w/o IGE, w/o IGE$_{st}$, and w/o IGE$_{dy}$ models all experience a performance drop, showing that recommending distant POIs that meet the static and dynamic intentions of users is useful. However, compared to the ZLE module, the contribution of the IGE module is relatively modest. In some cases, removing IGE has little impact. Moreover, removing only one of IGE's components (e.g., w/o IGE$_{st}$ or w/o IGE$_{dy}$) results in a larger performance drop than removing the entire IGE. We attribute this abnormal phenomenon to the seesaw effect between the local exploration and the global exploration. That is to say, IGE can strengthen the model's capability to accurately recommend global POIs, but may simultaneously impair its ability to effectively recommend local POIs. This leads to unexpected ablation results for the IGE module.

To verify the effectiveness of the IGE module and to support the seesaw effect in the third finding, we further evaluate it on local and global new POIs. Specifically, we divide the test set into three groups based on the distance between the target POI and the last visited POI. Each group corresponds to a unique distance range, with higher group numbers indicating larger distances and stronger global exploration effects. Considering that the abnormal fluctuation mainly occurs on the HR@5 metric, we show the HR@5 results in Figure 5 below. From the results, we have the following findings.

(1) Our proposed IGE module and its each component bring benefits to the complete model LGE in terms of recommending global new POI. On group G2 which has the largest global exploration effects, w/o IGE, w/o IGE$_{st}$, and w/o IGE$_{dy}$ all underperform LGE. At the same time, removing only one component from IGE (e.g., w/o IGE$_{st}$ or w/o IGE$_{dy}$) results in a smaller performance drop on group G2 than removing the entire IGE, verifying the effectiveness of the IGE module. Furthermore, w/o IGE$_{dy}$ suffers from a more severe performance decline than w/o IGE$_{st}$. This phenomenon shows that the dynamic intention plays a more essential role than the static intention.
(2) The existence of the local-global exploration seesaw effect is observed. Compared with the complete model LGE, w/o IGE experiences not only a clear performance drop on group G2, but also a large performance increase on

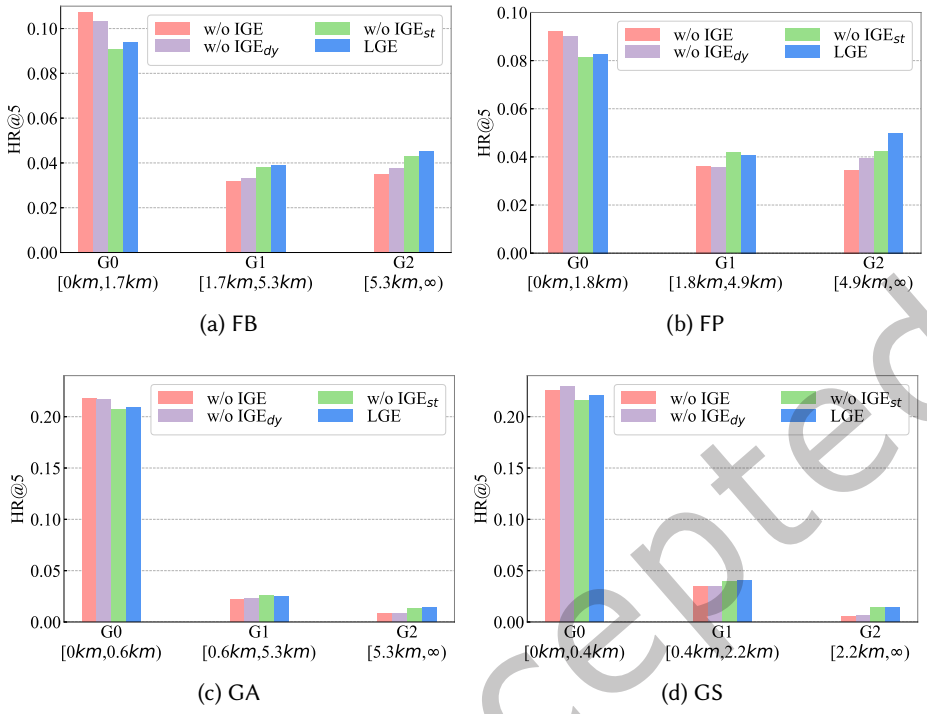


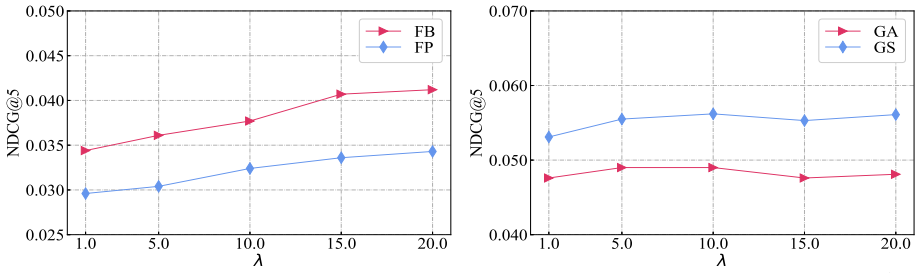
Fig. 5. The ablation study for the IGE module on different distance groups. A larger group number implies a stronger effect of the global exploration.

group G0. This explains why the removal of IGE has little impact on the overall performance in some datasets, e.g., Gowalla-Austin. Meanwhile, w/o IGE$_{st}$ and w/o IGE$_{dy}$ perform better on group G2 but worse on group G0 than w/o IGE. Therefore, it is reasonable that removing only one of IGE's components causes a larger performance drop than removing the entire IGE in some cases, e.g., Foursquare-Boston and Gowalla-Austin.

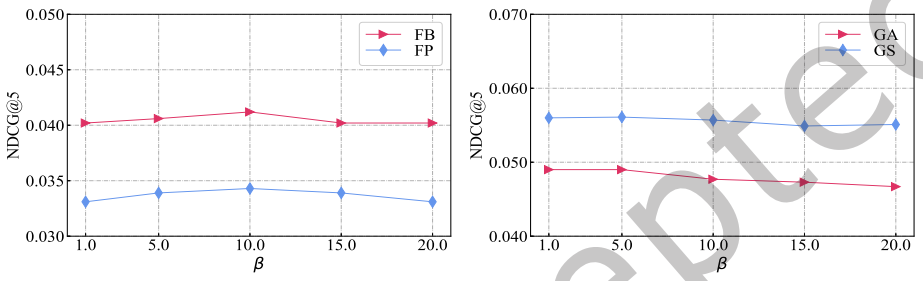
5.4 Hyperparameter Study

We examine the sensitivity of LGE to three pivotal hyperparameters: λ and β, which modulate the contrastive loss $\mathcal{L}_{con}$ and regularization loss $\mathcal{L}_{reg}$ respectively, and the initial K-Means cluster size n. Specifically, we vary λ and β within the set $\{1, 5, 10, 15, 20\}$ and range n from 200 to 2000 with a step size of 200. The corresponding NDCG@5 performance is shown in Figure 6.

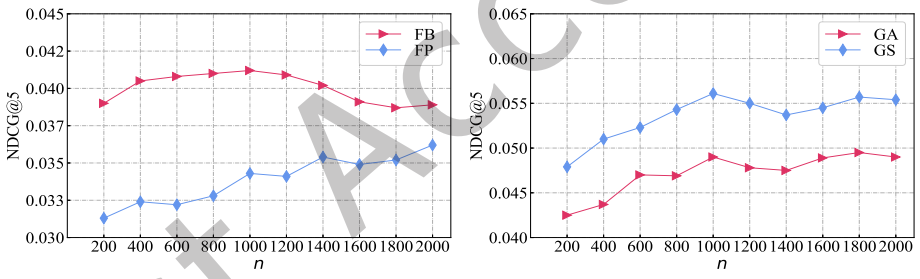
5.4.1 Impact of λ for Contrastive Loss. From Figure 6a, we find that increasing λ from 1 to 20 boosts the performance on Foursquare datasets (i.e., FB and FP). This verify the importance of learning the zone-aware POI representations for the local exploration. Meanwhile, on Gowalla datasets (i.e., GA and GS), the performance curves first experience a slow improvement, then reach a plateau, or show a downward trend. This is because the Gowalla data is denser than the Foursquare data, as shown in Table 1. Learning user behavior to new POIs from dense data is much easier, hence a relatively weaker contrastive signal, i.e., lower λ, could be sufficient.



(a) λ for contrastive loss



(b) β for regularization loss



(c) n for K-Means

Fig. 6. Influence of key hyperparameters.

5.4.2 Impact of β for Regularization Loss. From Figure 6b, it is observed that LGE achieves its best performance when the β is set to a moderate value (e.g., 10 for FB and FP, 1 for GA, and 5 for GS). Similar to λ, the optimal value of β on the denser Gowalla dataset is smaller than that on Foursquare. Furthermore, when comparing Figure 6b and Figure 6a, LGE is less sensitive to β than λ. This observation is also consistent with our conclusion in Section 5.3 that the regularization loss is less vital than the contrastive loss.

5.4.3 Impact of n for K-Means. From Figure 6c, it can be observed that the performance curve across the four datasets exhibits a consistent trend. Specifically, the performance improves as n increases in the early stage, but plateaus or declines when n reaches approximately 1000. Therefore, selecting an appropriate number of clusters

is crucial for LGE. Based on the empirical results and the total number of POIs in each dataset, we recommend maintaining approximately 10 to 20 POIs per cluster.

5.5 In-depth Analysis

We conduct an in-depth analysis on LGE's recommendation ability.

5.5.1 Repeat POI Recommendation. LGE is specially designed for the task of new POI recommendation. However, real-world user behavior often involves a mix of exploring new POIs and revisiting repeat ones. To this end, we investigate LGE's potential for repeat POI recommendation.

First, we investigate whether LGE is suitable for recommending repeat POIs. We train LGE on repeat POIs, namely LGE (Rep), and compare its performance on next repeat POI recommendation with the best baselines, i.e., EEDN for Foursquare and ARGAN for Gowalla. From the results presented in Table 4 below, we surprisingly observe that, in most cases, LGE (Rep) outperforms the best baseline methods, suggesting that LGE has the potential to conduct repeat POI recommendation.

Table 4. Performance comparison on next repeat POI recommendation. The baseline model and LGE (Rep) are trained on repeat POIs. Mention that LGE (Rep) adopts the MP strategy for Foursquare and the WP strategy for Gowalla.

Model	FB		FP		GA		GS	
	HR@5	NDCG@5	HR@5	NDCG@5	HR@5	NDCG@5	HR@5	NDCG@5
Best Baseline	0.4865	0.3522	**0.5198**	0.3770	0.3245	0.2497	0.4061	0.3094
LGE (Rep)	**0.4963**	**0.3912**	0.5092	**0.4021**	**0.3779**	**0.2874**	**0.4208**	**0.3326**

Table 5. Performance comparison on next new POI recommendation. The baseline models, LGE (MP), and LGE (WP) are trained on new POIs only, while LGE (Rep_New) is trained on both repeat and new POIs. Mention that LGE (Rep_New) adopts the MP strategy for Foursquare and the WP strategy for Gowalla.

Model	FB		FP		GA		GS	
	HR@5	NDCG@5	HR@5	NDCG@5	HR@5	NDCG@5	HR@5	NDCG@5
Best Baseline	0.0458	0.0333	0.0507	0.0332	0.0700	0.0431	0.0725	0.0502
LGE (MP)	**0.0597**	**0.0412**	**0.0568**	**0.0343**	0.0817	0.0479	0.0888	0.0544
LGE (WP)	0.0569	0.0387	0.0534	0.0331	**0.0827**	**0.0490**	**0.0914**	**0.0561**
LGE (Rep_New)	0.0467↓	0.0338↓	0.0506↓	0.0313↓	0.0666↓	0.0384↓	0.0795↓	0.0479↓

Second, we investigate whether integrating repeat POIs at the final prediction stage would enhance new POI recommendations. we train LGE on both repeat and new POIs, namely LGE (Rep_New), and compare its performance on next new POI recommendation with three methods, e.g., the best baseline, LGE (MP), and LGE (WP). These three methods are trained on new POIs only. Table 5 presents the results. We observe that LGE (Rep_New) experiences a significant performance drop compared to LGE (MP) and LGE (WP). This indicates the repeat POI signals are harmful to recommending new POIs. We attribute this phenomenon to the fact that the task of repeat POI recommendation is much easier than that of new POI recommendation. Simply mixing them during training might cause the model biased to the easier repeat POI recommendation. Nevertheless, LGE (Rep_New) is still competitive when compared to the best baseline trained on new POIs only, further demonstrating the effectiveness of our proposed method.

5.5.2 Local and Global New POI Recommendation. We compare LGE with baselines in recommending new local and global POIs. We follow the experimental setting in the ablation study, and evenly divide the test set into three groups, with higher group numbers indicating larger distances and stronger global exploration effects. We then separately test the performance of LGE, w/o ZLE, w/o IGE, and several strong baselines (i.e., ARGAN, EEDN, and JTLL) in each group. The results are shown in Figure 7.

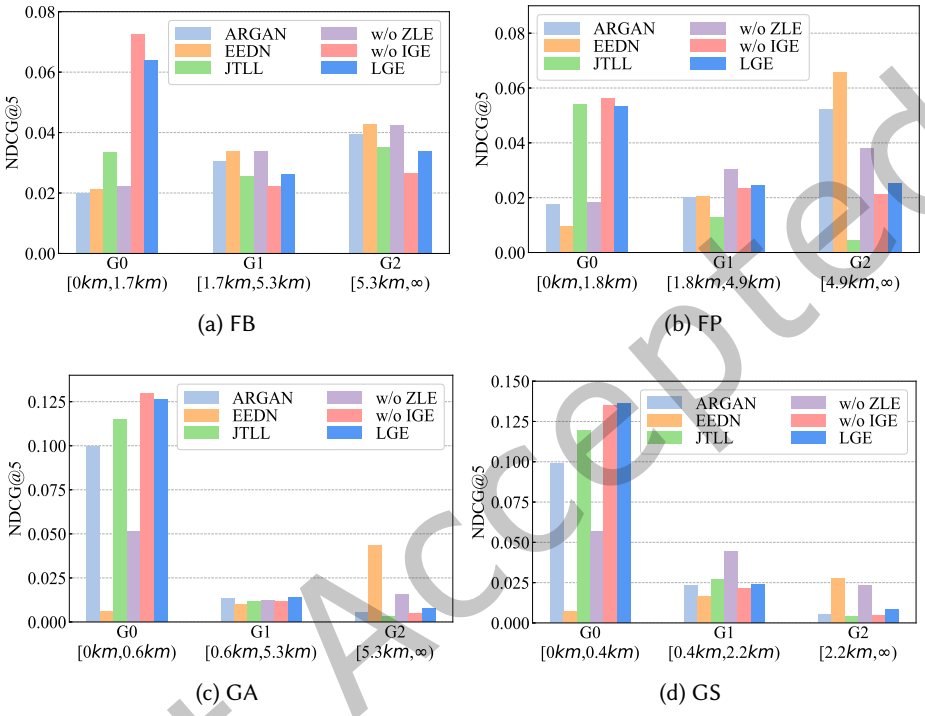


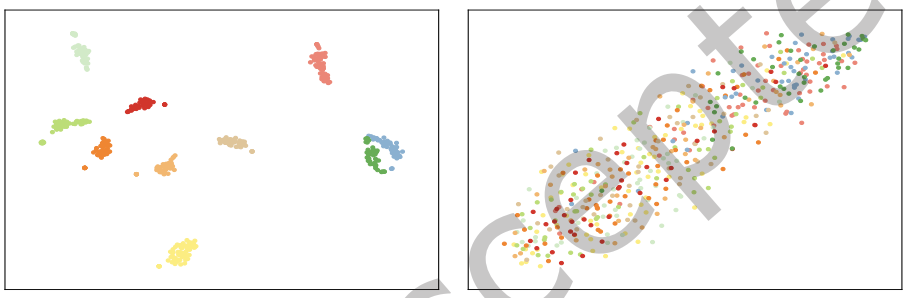
Fig. 7. Performance on different distance groups. A larger group number implies a stronger effect of the global exploration.

We observe that the w/o IGE model achieves the best or the second best performance on group G0, indicating the ability of the ZLE module to recommend new POIs in the local area. We also find that the performance of w/o ZLE is competitive on groups G1 and G2, demonstrating the potential of the IGE module to recommend distant new POIs. When comparing LGE with baselines, we can draw a conclusion that the overall superiority of LGE stems from a balanced consideration of recommending new POIs from both the local and global perspectives of exploration. However, it is also observed that LGE is inferior to EEDN on group G2 across all datasets, we argue this is understandable because EEDN has leveraged various types of features with different networks, e.g., collaborative information, sequential information, channel information, and rating information. More importantly, EEDN constructs user-POI interaction hypergraphs to capture global visit patterns. Convolution over these hypergraphs can easily integrate related but distant POIs. In contrast, the IGE module of LGE delivers a new idea to recommend global new POIs based on user dynamic and static intentions. It merely utilizes the category information without any complex network to model user intentions, but surprisingly achieves the most competitive performance against EEDN in long-distance POI recommendation.

5.5.3 Decision Accuracy of $\Phi(\cdot)$. As claimed in Section 5.2, the decision prediction network $\Phi(\cdot)$ is more suitable for Gowalla datasets, leading to a better performance of LGE (WP). To further verify the claim, we analyze the decision accuracy of $\Phi(\cdot)$ and adopt the widely used F1 score to evaluate the prediction performance. From the results in Table 6, we find that the F1 score on Gowalla datasets is much higher than that on Foursquare datasets. This phenomenon reflects the better compatibility of $\Phi(\cdot)$ with Gowalla datasets than Foursquare datasets, which aligns with our expectation. Hence, it is reasonable that LGE (WP) performs well on Gowalla datasets but less effective on Foursquare datasets.

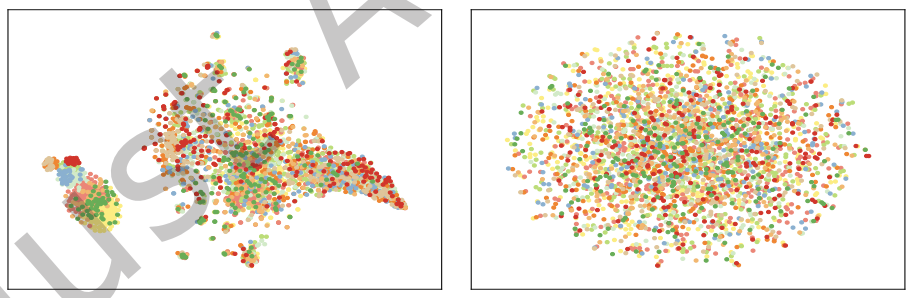
Table 6. The F1 score of the decision predictor $\Phi(\cdot)$ on four datasets.

Metric	FB	FP	GA	GS
F1	0.5488	0.6012	0.7579	0.7680



(a) POI embeddings from our ZLE

(b) POI embeddings from the baseline method JTLL



(c) POI embeddings from our IGE

(d) POI embeddings from the baseline method EEDN

Fig. 8. POI Embedding Visualization. Each color denotes a specific zone in (a) and (b), and a specific category in (c) and (d).

5.5.4 POI Embedding Visualization. We conduct two groups of visualization analysis with t-SNE [31] to show the advantages of LGE. The first group aims to analyze POI embeddings in the zone-aware local exploration module, i.e., V_l, while second group is to analyze POI embeddings in the intention-aware global exploration module, i.e., V_g. In both groups, we take the FB dataset as an example.

For the first group, we select 10 zones with the highest POI coverage for visualization. Then we compare the POI embeddings from our ZLE with those from the baseline method JTLL, since JTLL is competitive in recommending new POIs in the local area according to Figure 7. Figures 8a and 8b plot the results. It is clear that POIs located in the same zone are tightly grouped in LGE. In contrast, these POIs are totally mixed up in JTLL. This explains LGE's advantage over JTLL in recommending nearby POIs.

For the second group, we choose to display 10 categories that contain the most POIs. Then we compare the POI embeddings from our IGE with those from the baseline method EEDN, because EEDN has the ability to recommend remote new POIs according to Figure 7. Figures 8c and 8d plot the results. We observe a clustering phenomenon in LGE for POIs belonging to the same category (e.g., the left part of Figure 8c), which is absent in Figure 8d. This distinguishes LGE from EEDN and verifies its ability to recommend POIs that meet the needs of users.

6 conclusion

In this paper, we uncover the coexistence of local and global exploration patterns in users' visits to new POIs, and propose a new model for the next new POI recommendation. We design a zone-aware local exploration module to recommend new POIs in the local area by enriching POI representations and regularizing the POI prediction process with zone information. We further develop an intention-aware global exploration module to extract user static and dynamic intentions from category features and recommend new POIs that meet user intentions. We finally present a fusion module with an average pooling strategy and a weighted pooling strategy to combine the outputs of two exploration modules for the unified recommendation list. Empirical results on four real-world city datasets demonstrate the effectiveness of our proposed model via the superior performance over the state-of-the-art methods in recommending the next new POIs.

7 Acknowledgement

This work is supported in part by the National Natural Science Foundation of China under Grant 62276193 and 62576254. It is also supported by the Joint Lab. on Credit Sci. and Tech. of CSCI-Wuhan University.

References

[1] Wei Chen, Haoyu Huang, Zhiyu Zhang, Tianyi Wang, Youfang Lin, Liang Chang, and Huaiyu Wan. 2025. Next-POI Recommendation via Spatial-Temporal Knowledge Graph Contrastive Learning and Trajectory Prompt. *IEEE Transactions on Knowledge and Data Engineering* (2025), 3570–3582.

[2] Yile Chen, Yicheng Tao, Yue Jiang, Shuai Liu, Han Yu, and Gao Cong. 2025. Enhancing Large Language Models for Mobility Analytics with Semantic Location Tokenization. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*. 262–273.

[3] Chen Cheng, Haiqin Yang, Michael R Lyu, and Irwin King. 2013. Where you like to go next: Successive point-of-interest recommendation. In *Proceedings of the 23th International Joint Conference on Artificial Intelligence*.

[4] Jiawei Cheng, Jingyuan Wang, Yichuan Zhang, Jiahao Ji, Yuanshao Zhu, Zhibo Zhang, and Xiangyu Zhao. 2025. Poi-enhancer: An llm-based semantic enhancement framework for poi representation learning. In *Proceedings of the 39th AAAI conference on artificial intelligence*. 11509–11517.

[5] Qiang Cui, Chenrui Zhang, Yafeng Zhang, Jinpeng Wang, and Mingchen Cai. 2021. ST-PIL: Spatial-Temporal Periodic Interest Learning for Next Point-of-Interest Recommendation. In *Proceedings of the 30th ACM International conference on information and knowledge management*. 2960–2964.

[6] Jie Feng, Yong Li, Chao Zhang, Funing Sun, Fanchao Meng, Ang Guo, and Depeng Jin. 2018. Deepmove: Predicting human mobility with attentional recurrent networks. In *Proceedings of the 27th International Conference on World Wide Web*. 1459–1468.

[7] Shanshan Feng, Xutao Li, Yifeng Zeng, Gao Cong, and Yeow Meng Chee. 2015. Personalized ranking metric embedding for next new poi recommendation. In *Proceedings of the 24th International Conference on Artificial Intelligence*. 2069–2075.

[8] Shanshan Feng, Haoming Lyu, Fan Li, Zhu Sun, and Caishun Chen. 2024. Where to move next: Zero-shot generalization of llms for next poi recommendation. In *ieee conference on artificial intelligence*. IEEE, 1530–1535.

[9] Jing He, Xin Li, Lejian Liao, Dandan Song, and William Cheung. 2016. Inferring a personalized next point-of-interest recommendation model with latent behavior patterns. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. 137–143.
[10] Xiaowen Huang, Jitao Sang, Jian Yu, and Changsheng Xu. 2022. Learning to learn a cold-start sequential recommender. *ACM Transactions on Information Systems* 40, 2 (2022), 1–25.
[11] Dejiang Kong and Fei Wu. 2018. HST-LSTM: A hierarchical spatial-temporal long-short term memory network for location prediction.. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. 2341–2347.
[12] Yantong Lai, Yijun Su, Lingwei Wei, Tianqi He, Haitao Wang, Gaode Chen, Daren Zha, Qiang Liu, and Xingxing Wang. 2024. Disentangled contrastive hypergraph learning for next POI recommendation. In *Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval*. 1452–1462.
[13] Jiacheng Li, Yujie Wang, and Julian McAuley. 2020. Time interval aware self-attention for sequential recommendation. In *Proceedings of the 13th ACM International Conference on Web Search and Data Mining*. 322–330.
[14] Peibo Li, Maarten de Rijke, Hao Xue, Shuang Ao, Yang Song, and Flora D Salim. 2024. 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*. 1463–1472.
[15] Ranzhen Li, Yanyan Shen, and Yanmin Zhu. 2018. Next point-of-interest recommendation with temporal and multi-level context attention. In *Proceedings of the 2018 IEEE 18th International Conference on Data Mining*. 1110–1115.
[16] Xin Li, Dongcheng Han, Jing He, Lejian Liao, and Mingzhong Wang. 2019. Next and next new POI recommendation via latent behavior pattern inference. *ACM Transactions on Information Systems* 37, 4 (2019), 1–28.
[17] Yang Li, Tong Chen, Hongzhi Yin, and Zi Huang. 2021. Discovering Collaborative Signals for Next POI Recommendation with Iterative Seq2Graph Augmentation. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence*. 1491–1497.
[18] Defu Lian, Yongji Wu, Yong Ge, Xing Xie, and Enhong Chen. 2020. Geography-aware sequential location recommendation. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*. 2009–2019.
[19] Nicholas Lim, Bryan Hooi, See-Kiong Ng, Yong Liang Goh, Renrong Weng, and Rui Tan. 2022. Hierarchical multi-task graph recurrent network for next poi recommendation. In *Proceedings of the 45th international ACM SIGIR conference on Research and development in Information Retrieval*.
[20] Nicholas Lim, Bryan Hooi, See-Kiong Ng, Yong Liang Goh, Renrong Weng, and Rui Tan. 2023. Learning Hierarchical Spatial Tasks with Visiting Relations for Next POI Recommendation. *ACM Transactions on Recommender Systems* 1, 4 (2023), 1–26.
[21] Nicholas Lim, Bryan Hooi, See-Kiong Ng, Xueou Wang, Yong Liang Goh, Renrong Weng, and Jagannadan Varadarajan. 2020. 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*. 845–854.
[22] Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. 2016. Predicting the next location: A recurrent model with spatial and temporal contexts. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. 194–200.
[23] Yingtao Luo, Qiang Liu, and Zhaocheng Liu. 2021. STAN: Spatio-Temporal Attention Network for Next Location Recommendation. In *Proceedings of the 30th International Conference on World Wide Web*. 2177–2185.
[24] Xingyu Pan, Yushuo Chen, Changxin Tian, Zihan Lin, Jinpeng Wang, He Hu, and Wayne Xin Zhao. 2022. Multimodal meta-learning for cold-start sequential recommendation. In *Proceedings of the 31st ACM international conference on information and knowledge management*. 3421–3430.
[25] Anton Pembek, Artem Fatkulin, Anton Klenitskiy, and Alexey Vasilev. 2025. Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations. In *19th ACM Conference on Recommender Systems*. 626–631.
[26] Yifang Qin, Yifan Wang, Fang Sun, Wei Ju, Xuyang Hou, Zhe Wang, Jia Cheng, Jun Lei, and Ming Zhang. 2023. DisenPOI: Disentangling Sequential and Geographical Influence for Point-of-Interest Recommendation. In *Proceedings of the 16th ACM International Conference on Web Search and Data Mining*. 508–516.
[27] Xuan Rao, Lisi Chen, Yong Liu, Shuo Shang, Bin Yao, and Peng Han. 2022. Graph-flashback network for next location recommendation. In *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. 1463–1471.
[28] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web*. 811–820.
[29] Ke Sun, Chenliang Li, and Tieyun Qian. 2024. City matters! a dual-target cross-city sequential poi recommendation model. *ACM Transactions on Information Systems* 42, 6 (2024), 1–27.
[30] Ke Sun, Tieyun Qian, Tong Chen, Yile Liang, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2020. Where to go next: Modeling long-and short-term user preferences for point-of-interest recommendation. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence*. 214–221.
[31] Laurens Van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, 11 (2008).
[32] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Proceedings of the Advances in Neural Information Processing Systems*. 5998–6008.

[33] Dongsheng Wang, Yuxi Huang, Shen Gao, Yifan Wang, Chengrui Huang, and Shuo Shang. 2025. Generative Next POI Recommendation with Semantic ID. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*. 2904–2914.
[34] En Wang, Yiheng Jiang, Yuanbo Xu, Liang Wang, and Yongjian Yang. 2022. Spatial-Temporal Interval Aware Sequential POI Recommendation. In *Proceedings of the 2022 IEEE 38th International Conference on Data Engineering*. IEEE, 2086–2098.
[35] Jianling Wang, Kaize Ding, and James Caverlee. 2021. Sequential recommendation for cold-start users with meta transitional learning. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1783–1787.
[36] Xinfeng Wang, Fumiyo Fukumoto, Jin Cui, Yoshimi Suzuki, Jiyi Li, and Dongjin Yu. 2023. EEDN: Enhanced Encoder-Decoder Network with Local and Global Context Learning for POI Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 383–392.
[37] Xiaolin Wang, Guohao Sun, Xiu Fang, Jian Yang, and Shoujin Wang. 2022. Modeling Spatio-temporal Neighbourhood for Personalized Point-of-interest Recommendation. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*. 3530–3536.
[38] Zhiqiang Wang, Jiayi Pan, Xingwang Zhao, Jianqing Liang, Chenjiao Feng, and Kaixuan Yao. 2025. Counterfactual Task-augmented Meta-learning for Cold-start Sequential Recommendation. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence*. 12801–12809.
[39] Zhaobo Wang, Yanmin Zhu, Haobing Liu, and Chunyang Wang. 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*. 1154–1163.
[40] Zhaobo Wang, Yanmin Zhu, Chunyang Wang, Wenze Ma, Bo Li, and Jiadi Yu. 2023. Adaptive Graph Representation Learning for Next POI Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 393–402.
[41] Xiaolong Xu, Hongsheng Dong, Lianyong Qi, Xuyun Zhang, Haolong Xiang, Xiaoyu Xia, Yanwei Xu, and Wanchun Dou. 2024. Cmclrec: Cross-modal contrastive learning for user cold-start sequential recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1589–1598.
[42] Xiaodong Yan, Tengwei Song, Yifeng Jiao, Jianshan He, Jiaotuan Wang, Ruopeng Li, and Wei Chu. 2023. Spatio-Temporal Hypergraph Learning for Next POI Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 403–412.
[43] Song Yang, Jiamou Liu, and Kaiqi Zhao. 2022. 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*. 1144–1153.
[44] Jing Yuan, Yu Zheng, and Xing Xie. 2012. Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 186–194.
[45] Jia-Dong Zhang, Chi-Yin Chow, and Yanhua Li. 2014. Lore: Exploiting sequential influence for location recommendations. In *Proceedings of the 22nd ACM SIGSPATIAL international conference on advances in geographic information systems*. 103–112.
[46] Lu Zhang, Zhu Sun, Ziqing Wu, Jie Zhang, Y Ong, and Xinghua Qu. 2022. Next Point-of-Interest Recommendation with Inferring Multi-step Future Preferences. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence*. 3751–3757.
[47] Mingwei Zhang, Yang Yang, Rizwan Abbas, Ke Deng, Jianxin Li, and Bin Zhang. 2021. SNPR: A Serendipity-Oriented Next POI Recommendation Model. In *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*. 2568–2577.
[48] Pengpeng Zhao, Haifeng Zhu, Yanchi Liu, Jiajie Xu, Zhixu Li, Fuzhen Zhuang, Victor S Sheng, and Xiaofang Zhou. 2019. Where to Go Next: A Spatio-Temporal Gated Network for Next POI Recommendation. In *Proceedings of the 33rd AAAI Conference on Artificial Intelligence*. 5877–5884.
[49] Shenglin Zhao, Tong Zhao, Haiqin Yang, Michael Lyu, and Irwin King. 2016. STELLAR: Spatial-temporal latent ranking for successive point-of-interest recommendation. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*. 315–322.
[50] Yujia Zheng, Siyi Liu, Zekun Li, and Shu Wu. 2021. Cold-start sequential recommendation via meta learner. In *Proceedings of the 35th AAAI conference on artificial intelligence*. 4706–4713.
[51] Hongli Zhou, Zhihao Jia, Haiyang Zhu, and Zhizheng Zhang. 2024. Cllp: Contrastive learning framework based on latent preferences for next poi recommendation. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1473–1482.
[52] Hailun Zhou, Jiajie Xu, Qiaoming Zhu, and Chengfei Liu. 2025. Disentangled Graph Debiasing for Next POI Recommendation. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 1779–1788.