




Hybrid graph convolutional networks with multi-head attention for location recommendation

Ting Zhong¹ · Shengming Zhang¹ · Fan Zhou¹  · Kunpeng Zhang² · Goce Trajcevski³ · Jin Wu¹

Received: 16 October 2019 / Revised: 13 May 2020 / Accepted: 20 May 2020 /

Published online: 23 June 2020

© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Recommending yet-unvisited points of interest (POIs) which may be of interest to users is one of the fundamental applications in location-based social networks. It mainly relies on the understanding of users, POIs, and their interactions. Previous studies either develop matrix factorization-based approaches or utilize deep learning frameworks to learn better representation of users and POIs in order to estimate users' latent preference. However, most of existing methods still confront the challenges like in traditional recommender systems, such as data sparsity and cold-start. In particular, they have difficulties in fully utilizing rich semantic information, such as social influence, geographical constraints and interactions between users and POIs. To fill this research gap, we propose a new recommendation framework – Hybrid Graph convolutional networks with Multi-head Attention for POI recommendation (HGMAP). HGMAP constructs a spatial graph based on the geographical distance between pairs of POIs and leverages Graph Convolutional Networks (GCNs) to express the high-order connectivity among POIs, which not only incorporates the spatial constraints but also provides an effective way to alleviate the sparse check-in problem. In addition, HGMAP exploits the user social relationship with another GCN and differentiates user preference over different aspects of POIs with a multi-head attention mechanism. We conducted extensive experiments on three public datasets and the results demonstrate that HGMAP significantly improves the recommendation performance over several state-of-the-art models, for example, up to approximately 4.8% and 7% for *Precision@10* and *Recall@10*, respectively.

Keywords Location recommendation · Graph convolutional networks · Multi-head attention · Collaborative filtering · Social influence

1 Introduction

With the rapid development of mobile internet technology and the widespread use of GPS-enabled devices, Location-Based Social Networks (LBSNs) such as Foursquare and Yelp,

✉ Fan Zhou
fan.zhou@uestc.edu.cn

become ubiquitous and gain great popularity for users to leave their footprints and share their experience. Accordingly, this results in large amounts of user–location interaction data that contain various kinds of Point-of-Interests (POIs), such as restaurant, museums, shopping malls, parks, and many others. These user-generated content are usually associated with geo-tags. Analyzing such rich data can be beneficial for many downstream applications, for example, building personalized POI recommender systems. The POI recommendation has spurred a significant research interest in both industry and academic communities [1, 13, 31, 46, 54], as it can provide various value-added services, for example, recommending tourists' vacation rentals (e.g., Airbnb), advertising scenic areas (e.g., TripAdvisor), and promoting experiences (e.g., Mafengwo).

One of the objectives in POI recommendation is to discover the yet-unvisited places of potential interest to users. Unlike typical recommendation tasks (e.g., movie, music and e-commerce item recommendation), POI recommendation exhibits several special characteristics, such as strong spatial-temporal dependence among POIs and the geographical constraints on users. For example, the recommendation of restaurants to users should take into account the geographical location of both users and restaurants. Prior studies have shown that there is a spatial clustering phenomenon in user check-ins, i.e., people prefer to visit POIs close to their home locations. Individual visited POIs tend to cluster together. In addition, the social relationship and visiting time also play important roles in personalized POI recommendation. People prefer to visit places where their friends visited / recommended. People are more likely to visit recreation parks on weekends while go to tech / financial firms often in weekdays. How to incorporate these features into POI recommender systems to better understand the relations among users and POIs has become a trending research topic. Of course, POI recommender systems also confront several challenges commonly rooted in traditional recommender systems, such as *data sparsity*, where an user usually visits a very small number of locations among millions of POIs in a LBSN. As an example, the density of the user-POI check-in count matrix is about 0.1% [31]; *cold start* where some users have no visiting history or some (new) POIs have never been visited by any users.

Researchers have proposed various methods to improve the POI recommendation performance by mainly focusing on exploiting different implicit context features embedded in user check-ins. For example, Collaborative Filtering (CF)-like techniques such as Matrix Factorization (MF) [16, 19] are used to predict user rating on POIs through explicit/implicit feedback while taking into consideration various constraints such as social influence [53, 61], temporal features [56, 57], sequential dependence [33, 63], and geographical constraints [15, 22, 26, 28]. All of these methods follow a similar procedure where they first extract latent features underlying interactions between users and POIs, and then predict users' preference based on the inner product of latent factors. However, they may not fully discover the complicated user-POI interaction from the data, since the inner product combines latent features linearly and limits the expressiveness of the methods [31, 34].

Recent advances in deep learning have inspired efforts on applying various neural networks for discovering non-linear and non-trivial relationships between users and POIs. For example, word2vec [38] has been used for transforming users and POIs to vectorized representations [52], while recurrent neural networks (RNN) are used for learning sequential behavior of user check-ins [10, 37, 66]. To distinguish users' preference over different POIs, an attention-based model has been proposed in [34] where the denoising autoencoder is employed to capture the geographical influence. Another recent work [71] borrows the idea of adversarial learning from deep generative adversarial neural networks [12] and attempts to improve the POI recommendation performance by exploiting the social influence and

geographical information in a reinforcement learning manner. Although existing efforts have shown promising performance improvement and are able to handle non-linear interactions between users and POIs, most of them just integrate the auxiliary information (e.g., POI context, social influence and spatial-temporal characteristics) by transforming from pre-existing features with historical data, and thus fail to encode the high-order structure information and capture users' potential long-distance interest. All the while, the data sparsity and cold-start problems are still major challenges for existing solutions.

In this work, we propose Hybrid Graph convolutional networks with Multi-head Attention for POI recommendation (called HGMAP), a general and flexible framework that captures user-POI interactions effectively by mining the social influence and geographical attributes with graph-based neural networks. Inspired by the success of graph neural networks (GNNs) [6, 18, 51], we use two independent Graph Convolutional Networks (GCNs) to explicitly incorporate the spatial and social influence aspects of the auxiliary information into our POI recommender system. However, unlike the existing GNN-based recommendation models [9, 39, 43, 47–49, 55] that directly employ convolutional layers to exploit interactions between users and items, we alternatively use two independent GCNs to learn the geographical relationship and social influence, respectively. Specifically, we build a POI graph based on the pairwise distance of corresponding POIs with Radius Basis Function (RBF), and learn the geographical relationship and the implicit relations among POIs using a graph neural network. We also model the social networks of users and aggregate feature information of connected users from local neighborhood using another GCN. By recursively propagating the embedding of geographical and social information, HGMAP can conceptually capture the high-order connectivity in an efficient, explicit, and end-to-end manner. In addition, we leverage a multi-head attentive encoder to capture non-linear user-POI interactions while learning the importance of each POI during information retrieval for personalized recommendation. The proposed model has the ability to learn good user and POI representation, and recommend users in an efficient manner due to its inductive learning capability.

Overall, the main contributions of this paper are four-fold:

- First, we present a novel POI recommender system using hybrid graph neural networks to learn both users' and POIs' latent representations, which effectively encode: (1) the social influence and geographical constraints – the most important features in POI recommendation [31]; and (2) the underlying relationship between users and POIs. Moreover, the cold-start problem at both user and POI sides can be largely alleviated by aggregating features from two heterogeneous graphs.
- Second, we provide a new perspective of incorporating geographical locations into POI recommendation by constructing a POI adjacency graph and learning complex POI relations beyond the Euclidean distance via layered graph convolutions. By doing so, our model can sample neighboring POIs from the graph to augment data for each user while, to an extent, overcoming the data sparsity issue.
- Third, we introduce a multi-head attention encoder to adaptively compute a preference score for each check-in and obtain user latent preference representation over unvisited POIs. User preference representation and user social representation are used to model the user influence on POI recommendation. In addition, we leverage this preference score combined with POIs' location representation to learn the influence of checked-in POIs on unvisited POIs. This enables our model to capture non-linear user-POI interactions and nuances between different POIs while bounding the user preference with geographical regularization.

- Last, we conducted extensive experiments on several large-scale benchmark datasets, i.e., Gowalla, Foursquare and Yelp, demonstrating that HGMAP can significantly improve recommendation performance as compared to state-of-the-art POI recommendation baselines.

The remainder of this paper is organized as follows. We review the relevant works in Section 2. The details of our HGMAP model are presented in Section 3. Experimental evaluations demonstrating the superiority of our model are discussed in Section 4, followed by conclusion and direction for future work in Section 5.

2 Related work

In this section, we review the relevant studies in POI recommendation, the graph learning-based recommender systems, as well as the attention-based recommendation models, and position our work in that context.

2.1 Personalized POI recommendation

POI recommendation (a.k.a. location recommendation or venue recommendation) helps users to discover new POIs of their interest, which can be beneficial to both users and businesses [31]. Collaborative Filtering (CF)-like techniques such as Matrix Factorization (MF) [19], Bayesian probabilistic matrix factorization [42], and Bayesian Personalized Ranking (BPR) [41] are widely used in modern recommendation systems. Previous works on POI recommendation have shown that the contextual information associated with users (e.g., visiting time and social connections) and POIs (e.g., geographical locations) play important roles in enhancing the effectiveness of POI recommendation [22, 24, 26, 28, 53, 56–58, 61, 72]. These methods assume that users who have the same check-ins share similar preferences, so they are inclined to visit similar locations in the future, and therefore leveraging these learned latent features of users and POIs to predict user preference to unvisited locations may improve performance. In addition, some studies [24, 26] have shown that check-ins can be treated as implicit feedback, which can be incorporated into MF-based models to improve POI recommendation accuracy, while other research works [30] leverage Probabilistic Factor Models (PFM) [14] to consider auxiliary factors such as geographical influence. A comprehensive survey [31] compared representative CF-based POI recommendation models and summarized that (i) geographical information and social influence are the two most effective factors for capturing user preference; (ii) GeoMF [26] and RankGeoFM [22] exhibit superior performance on POI recommendation over other CF-based methods.

However, the performance of the CF-based recommendation methods often drops significantly when user-POI interactions are extremely sparse. Meanwhile, they cannot be directly used for recommending new POIs that have not been visited by any users or making recommendations to new users who have no visiting records, which is a well-known cold-start problem. More importantly, these latent factor models are inherently linear, which limits their modeling capacity to capture non-linear user-POI interactions. To overcome these issues, a growing body of recent works have applied deep neural networks to the collaborative filtering setting for POI recommendation [13, 32, 34, 37, 40, 52, 56, 64, 71]. For example, recent efforts [37, 52, 56, 64] use POI embedding and recurrent neural networks to capture the check-in context and user sequential visiting behavior. A translation-based POI

recommendation framework is proposed [40] to model the relations among users, POIs, and spatial-temporal context, where knowledge graph embedding techniques are used to encode users and POIs in a latent space. Similarly, a denoising autoencoder has been adopted [34] to capture spatial-temporal context and interactions among users and POIs. Adversarial learning [12] has also been employed to learn underlying user preference distribution in [32, 71], unifying the reinforcement learning and matrix factorization methods into an adversarial learning framework for POI recommendation.

Despite their effectiveness and some inspiring results, existing methods are not able to yield optimal recommendation performance, in part due to the data sparsity and cold-start issues in POI recommendation. In addition, the aforementioned methods mainly focus on exploiting deep learning techniques to enhance the interaction function, so as to capture the nonlinear relations between users and POIs, which, however, neither explicitly captures the transitivity property of both users and POIs, nor guarantees the closeness of similar users and POIs in the embedded space.

2.2 GNNs in recommender systems

Graph Neural Networks (GNNs) [6, 18, 45, 51] which aggregate node features from the locally connected neighbors of nodes using deep neural networks, have attracted a considerable attention in recent years due to their effectiveness and remarkable success in various tasks such as graph classification, semi-supervised node classification, traffic forecasting [74], meta-graph learning [67], information cascade prediction [4], network alignment [68] and image segmentation [35]. The main idea of GNNs is to generate graph convolutional layers based on graph spectral theory, and adaptively transform node feature vectors with different neighborhood aggregation and graph-level pooling schemes. Most recently, several works leverage GNN architectures for building recommender systems [9, 39, 43, 47–49, 55]. GC-MC [43] first applies the Graph Convolution Networks (GCNs) [18] on the user-item interaction graph. PinSage is an industrial solution that combines random walks and GCNs to generate nodes' embeddings for a bipartite graph in Pinterest. NGCF [50] exploits the user-item graph by expressively modeling high-order connectivity in user-item interactions with GCNs, which can inject collaborative signals into the procedure of propagating embedding on the graph. Another category of works [47–49] try to apply GNNs on knowledge graphs in order to provide additional guidance for recommendations, which relies heavily on external knowledge graphs and manual design of meta-paths/meta-graphs, and thus are hard to implement in practice.

Existing GNN-based recommendation models mainly focus on exploiting the CF signals from user-item interaction graphs. Although the CF effect between users and items can be efficiently captured, they cannot be directly applied to POI recommendation due to the extremely sparse user-POI interactions [8, 31]. In this spirit, our approach is different from existing works in that we sidestep the graph convolutions on a sparse user-POI interaction graph but alternatively capture the implicit user social relationship and POI connections, which not only provides useful information for cold-start users/POIs, but also alleviates the sparsity problem with the constructed geographically adjacent POI graphs.

2.3 Attention mechanism for recommendation

Recently, attention mechanism allows us to learn the importance of specific positions of the input. It, combined with various neural network architectures, has been proven to be effective in many tasks, such as machine translation [2], human mobility learning [11, 72],

image retrieval [62], object detection [59], as well as recommender systems [3, 34, 37, 49, 65, 69, 73]. Earlier works [3, 37] utilize vanilla attention vectors to dynamically model the influence of items and learn the interactions between users and items. However, these recommendation models rely on standard attention mechanism and can only capture single aspect of the item importance and linear interactions among items. Nevertheless, users' preference is too complex to be captured by a single importance vector, while high-order item feature interactions are essential for improving recommendation performance [30].

Multi-head self-attention mechanism [44] is a natural language processing (NLP) model fully relying on self-attention module to learn structures of sentences and complex word representations. It has achieved state-of-the-art performance on a wide range of NLP tasks (e.g., translation, word-embedding, etc.) and inspired a variety of excellent models such as BERT [7] and ALBERT [20]. In this work, we utilize multi-head self-attention to learn users' multiple-aspects preference over POIs. By projecting POI embedding into multiple subspaces, different interactions between different subspaces can be retrieved to reflect users' various aspects of preference over POIs. In addition, it helps us better differentiate users that have similar preference while making more personalized recommendations.

3 HGMAP recommendation framework: model and methodology

We now proceed with details of our model HGMAP for POI recommendation. First we define some basic terminology used throughout this paper and formally introduce the POI recommendation problem. Subsequently, we discuss the basic aspects of HGMAP, which consists of four components: two graph convolutional networks, a multi-head attentive encoder and a prediction module (cf. Figure 1). Specifically, we utilize GCN_{location} to learn POIs' location representation from a POI location network constructed based on POI geographic coordinates. In order to get users' social representation and incorporate social influence and check-in similarity information, GCN_{social} is employed on user social network and user-POI interactions. Then we implement a multi-head attentive encoder to learn users' preference representation and a preference score for every check-in from user-POI interactions. Users' social representation and users' preference representation are combined to learn SIP (Social Influence on user Preference) ratings on POIs. Meanwhile, we utilize POIs' location representation with the preference score to learn GIP (Geographical Influence on user Preference) ratings. Finally, we make predictions based on the learned SIP ratings and GIP ratings. In this section, we also present three loss functions regarding how to optimize the proposed model HGMAP.

3.1 Preliminaries

Definition 1 POI recommendation: Let $\mathcal{U} = \{u_1, \dots, u_m\}$ denote a set of users, $\mathcal{P} = \{p_1, \dots, p_n\}$ be a set of POIs and $\mathcal{D} = \{d_1, \dots, d_n\}$ be a set of corresponding geographical coordinates (latitude and longitude) of POIs. Let $\mathbf{c}_i = [c_1, \dots, c_n]$ be the POIs that user u_i checked in. Given historical check-in information for all m users $\mathcal{C} = \{\mathbf{c}_1, \dots, \mathbf{c}_m\}$, POI recommendation is to recommend a list of POIs for each user that the user might be interested in but never visited.

POI recommendation is commonly studied using an user-POI check-in frequency matrix $\mathbf{G} \in \mathbb{R}^{m \times n}$ constructed from interaction between m users and n locations. Each element $g_{i,j} \in \mathbf{G}$ represents the number of times that user u_i has been to location p_j . In this work,

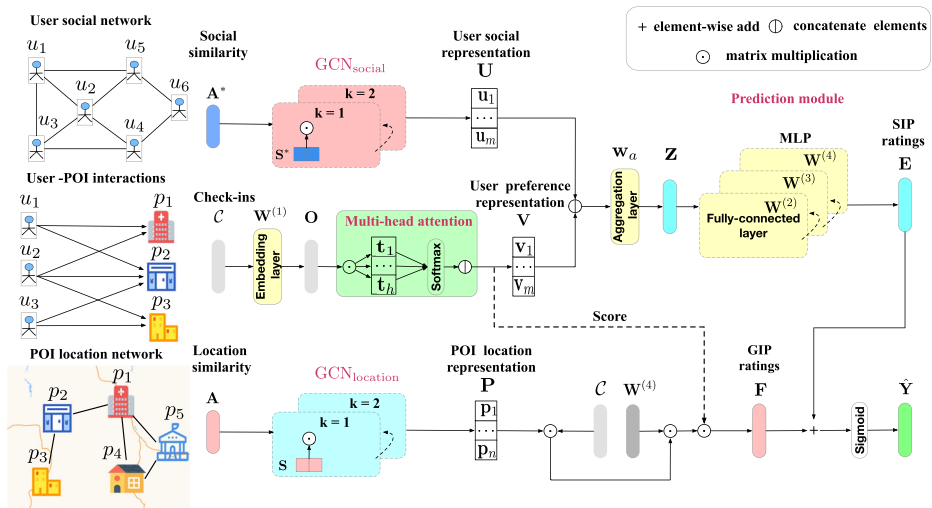


Figure 1 Overview of our proposed HGMAP model

we make the user-POI check-in binary matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$, where each element $b_{i,j} \in \{0, 1\}$ represents whether user u_i has been to location p_j . All notations used throughout the paper are listed in Table 1.

3.2 Learning POI and User Representation via Hybrid GNNs

We now describe how to leverage a variant of GNNs (i.e., GCNs) to learn POI and user representation. We have a connectivity network each for users and POIs used for capturing

Table 1 List of notations

Notation	Description
$\mathbf{Y}/\hat{\mathbf{Y}}$	The ground-truth value and predicted value
\mathbf{E}/\mathbf{F}	Users' rating matrix and locations' rating matrix
$\mathbf{W}^{(*)}$	The weight matrix
\mathbf{B}	User-POI check-in binary matrix
\mathbf{G}	Users' check-in frequency matrix
\mathbf{Q}	The confidence matrix
\mathbf{V}	Users' preference representation
\mathbf{U}	Users' social representation
\mathbf{P}	POIs' geographical location representation
$m/n/h$	The number of users / POIs / attention heads
\mathcal{C}/\mathcal{D}	Users' check-in set / POIs' geographical coordinates set
\mathcal{U}/\mathcal{P}	User / POI set
K	The number of layers in GCNs
D	The latent dimension of hidden layer
L	The latent dimension of POIs and users
R	The latent dimension of preference vectors

similarities among users and POIs, respectively. Figure 2 is toy example to show the connectivity for a POI p_1 and a user u_1 .

3.2.1 Modeling POI location representation

To learn POI representation, we leverage $GCN_{location}$ to capture local and global structural information in a network, especially the geographic relations among POIs (e.g., distant or close POIs). Thus, we first construct a POI geographic location network $\mathcal{G} = (\mathcal{P}, \mathbf{A})$, where $\mathcal{P} = \{p_1, \dots, p_n\}$ represents a set of POIs and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a sparse adjacency matrix and $a_{i,j}$ denotes the location similarity for a pair of POIs p_i and p_j . In this study, we choose a Gaussian Radial Basis Function (RBF) kernel to measure the location similarity $a_{i,j} \in [0, 1]$ for POI p_i and POI p_j , as follows:

$$a_{i,j} = \exp(-\eta \|d_i - d_j\|^2), \tag{1}$$

where d_i and d_j are the geographic coordinates of two POIs p_i and p_j , and $\eta > 0$ is a hyperparameter to control the level of geographical relevance between two given POIs. A larger value of $a_{i,j}$ indicates two POIs' geo-locations are closer. For the purpose of simplicity, we set $a_{i,j} = 0$ if it is less than a threshold value λ (i.e., $\lambda = 0.125$).

For $GCN_{location}$ with K layers, we take the location similarity matrix \mathbf{A} as an input to the first layer:

$$\mathbf{H}^{(0)} = \mathbf{A}, \tag{2}$$

The multi-layer $GCN_{location}$ follow the layer-wise propagation rule. Let \mathbf{S} denote the normalized adjacency matrix:

$$\mathbf{S} = \tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}}, \tag{3}$$

where $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$ is the adjacency matrix of the graph \mathcal{G} with added self-connections, \mathbf{I} is the identity matrix and $\tilde{\mathbf{D}}$ is the degree matrix of $\tilde{\mathbf{A}}$. The representation update of all location nodes becomes a simple sparse matrix multiplication:

$$\overline{\mathbf{H}}^{(k)} \leftarrow \mathbf{S} \mathbf{H}^{(k-1)}, \tag{4}$$

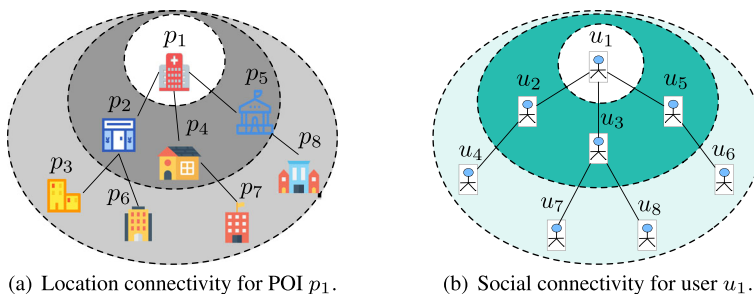


Figure 2 A toy example showing the connectivity of POI p_1 and user u_1

We adopt *ReLU*, which is a non-linear activation function, to optimize each layer. The updating rule of the k -th layer is as follows:

$$\mathbf{H}^{(k)} \leftarrow \text{ReLU} \left(\overline{\mathbf{H}}^{(k)} \Theta^{(k)} \right) \tag{5}$$

Following [51], our GCN_{location} module is a 2-layer simple graph convolution (SGC), which is a variant of GCNs and can compute more efficiently with significantly fewer parameters than traditional GCNs. The K^{th} layer output $\mathbf{H}^{(K)}$ is considered as the final location representation $\mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_n]^T \in \mathbb{R}^{n \times L}$ and L is the latent dimension of POI location representation.

3.2.2 Modeling user social representation

To learn user representation, we take a user social network and social similarity as input of GCN_{social} with multiple layers. The user social network is defined as $\mathcal{G}^* = (\mathcal{U}, \mathbf{A}^*)$, where \mathcal{U} represents a set of users $\{u_1, \dots, u_m\}$, $\mathbf{A}^* \in \mathbb{R}^{m \times m}$ is a sparse adjacency matrix, and $a_{i,j}^*$ is the edge weight between users u_i and u_j , representing how close two users are.

Based on the idea of collaborative filtering, user preference can be discovered by aggregating the behavior from similar users [53], which cannot be fully achieved by traditional GCNs where the edge weight is binary $a_{i,j}^* \in \{0, 1\}$. In order to understand the relationship among users, we make weights continuous $a_{i,j}^* \in [0, 1]$ and compute them by incorporating two semantic information: check-in and friendship. Let \mathcal{F} and \mathcal{R} denote users' friend set and users' check-in set in a LBSN. Then the edge weight $a_{i,j}^*$ between user u_i and user u_j is calculated as follows if they are not friends:

$$a_{i,j}^* = \frac{|\mathcal{R}_i \cap \mathcal{R}_j|}{|\mathcal{R}_i \cup \mathcal{R}_j|} \tag{6}$$

If user u_i and u_j are friends, $a_{i,j}^*$ is:

$$a_{i,j}^* = \beta \cdot \frac{|\mathcal{F}_i \cap \mathcal{F}_j|}{|\mathcal{F}_i \cup \mathcal{F}_j|} + (1 - \beta) \cdot \frac{|\mathcal{R}_i \cap \mathcal{R}_j|}{|\mathcal{R}_i \cup \mathcal{R}_j|} \tag{7}$$

where $\beta > 0$ is a tunable hyper-parameter with a range of $[0, 1]$ that is used to balance the relative weight of friend circle similarity and user visiting similarity. We denote the output of GCN_{social} $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]^T \in \mathbb{R}^{m \times L}$ as social representation for all m users. L is the dimension of the representation. Note that there is no social relationship for the Foursquare data where we only incorporate the check-in information.

3.3 User preference learning with multi-head attention

We now have obtained user representation and POI representation. Since our goal is to efficiently and comprehensively learn user preference over different POIs, it is requisite to measure the relevance between users and POIs while capturing the joint effect on user-POI interactions. Recently, attention mechanism has been widely used for recommender systems [3, 34, 37, 49]. For example, SAE-NAD [34] exploited the self-attentive autoencoders to learn complex user preference for POIs. However, the standard attention mechanism usually assigns a single importance value to a POI, which makes the model focus on only one (latent) aspect of POIs. This is not sufficient to reflect the sophisticated human sentiment

on POIs [30]. Particularly, some important (latent) aspects of POIs that might directly or indirectly influence user preference are missed.

The above-mentioned evidence inspires us to learn various aspects of user preference, which, through assigning multiple scores to each POI that user has visited, allow us to model the dependencies and importance of long-short term POI interactions. Towards this goal, we adopt multi-head self-attention [44] to effectively capture high-order interactions between POIs and retrieve the multi-aspect preference of users over POIs.

Technically, we first obtain a POI embedding matrix denoted by $\mathbf{W}^{(1)} \in \mathbb{R}^{L \times n}$. It is also the weight matrix of the embedding layer. Then, we utilize the multi-head attention mechanism with h attention heads to learn the preference over visited POIs for each user. The h attention matrices are:

$$\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_h]^\top \tag{8}$$

where $\mathbf{T} \in \mathbb{R}^{R \times L}$ and $\mathbf{t}_1 \in \mathbb{R}^{[R/h] \times L}$ represents the 1st attention head that learns user preference on POIs on some dimensions, i.e., traffic, food and scenery. $\mathbf{t}_2 \in \mathbb{R}^{[R/h] \times L}$ learns preference on different dimensions, and so on. R is the latent dimension of preferences.

$\mathbf{c}_i = [c_1, \dots, c_n]$ is a binary vector representing the set of check-in POIs for user u_i , where $c_j (1 \leq j \leq n)$ is 1 if user u_i has visited POI p_j and 0 otherwise. We utilize \mathbf{c}_i and $\mathbf{W}^{(1)}$ to get check-in POI representation of user u_i .

$$\hat{\mathbf{o}}_i = [c_1 \mathbf{W}_{(*,1)}^{(1)}, \dots, c_n \mathbf{W}_{(*,n)}^{(1)}] \tag{9}$$

where $\mathbf{W}_{(*,n)}^{(1)}$ is the n^{th} column of $\mathbf{W}^{(1)}$ and is the representation of the n^{th} POI. Note that $\hat{\mathbf{o}}_i \in \mathbb{R}^{L \times n}$ might have some zero columns. We delete them and get $\mathbf{o}_i \in \mathbb{R}^{L \times n_i}$, where n_i is the number of check-in POIs of user u_i and is the same as the number of non-zero columns in $\hat{\mathbf{o}}_i$. The set of check-in POI representation is denoted as $\mathbf{O} = \{\mathbf{o}_1, \dots, \mathbf{o}_m\}$.

Then, we learn the user u_i preference using h attention heads \mathbf{T} and check-in POI representation \mathbf{O} :

$$\begin{cases} \mathbf{s}_r = \text{softmax}(\tanh(\mathbf{t}_r \cdot \mathbf{o}_i)), r = 1, \dots, h \\ \mathbf{Score} = [\mathbf{s}_1, \dots, \mathbf{s}_h]^\top \end{cases} \tag{10}$$

where $\mathbf{Score} \in \mathbb{R}^{R \times n_i}$ is the user preference score matrix. Lastly, the preference of user u_i over check-in POIs can be computed as follows:

$$\mathbf{v}_i = \mathbf{Score} \cdot \mathbf{o}_i^\top \tag{11}$$

where $\mathbf{v}_i \in \mathbb{R}^{R \times L}$ denotes a preference representation of user u_i . We use $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_m]^\top \in \mathbb{R}^{m \times R \times L}$ to denote the preference representation of all m users.

3.4 Prediction module

POI recommendation in LBSNs is different from other recommendation tasks [34] in that there exist physical distances between users and POIs, and such a unique property spurs a well-known geographical clustering phenomenon – users usually appear in several specific areas and prefer to visit unvisited POIs that are around their checked-in POIs. Incorporating such a property is likely to improve the POI recommendation performance [23, 26, 34].

According to this clustering phenomenon, we speculate that check-in POIs of each user may affect other unvisited POIs with respect to geographic locations. Different from most of the previous studies that mainly exploit geographical influence from the perspective of POIs, our model combines user preference and geographical influence from both users and POIs. Specifically, we construct two ratings: (1) GIP (Geographical Influence on user Preference) rating with location influence and preference influence included; (2) SIP (Social Influence on user Preference) rating with social relationship and user preference considered.

- (1) From the perspective of POI geographic location, we first get check-in POI representation of user u_i from $\mathbf{c}_i = [c_1, \dots, c_n]$ and POI location representation \mathbf{P} .

$$\hat{\mathbf{j}}_i = [c_1 \mathbf{P}_{(*,1)}, \dots, c_n \mathbf{P}_{(*,n)}] \tag{12}$$

where $\mathbf{P}_{(*,n)}$ is the n^{th} column of \mathbf{P} and represents the n^{th} POI location representation. Note that $\hat{\mathbf{j}}_i \in \mathbb{R}^{L \times n}$ might have some zero columns. We also delete them in $\hat{\mathbf{j}}_i$ and get $\mathbf{j}_i \in \mathbb{R}^{L \times n_i}$. n_i , the number of check-in POIs of user u_i , is the same as the number of non-zero columns in $\hat{\mathbf{j}}_i$. We leverage check-in POI representation \mathbf{j}_i to compute the influence of check-in POI on unvisited POIs and incorporate the influence of user preference into the geographical influence as follows:

$$\mathbf{f}_i = \text{sum}(\mathbf{Score} \cdot (\mathbf{j}_i^\top \cdot \mathbf{W}^{(4)})) \tag{13}$$

where sum is an addition function that adds elements by row. $\mathbf{Score} \in \mathbb{R}^{R \times n_i}$ denotes the user preference matrix, $\mathbf{j}_i^\top \in \mathbb{R}^{n_i \times L}$ and $\mathbf{W}^{(4)} \in \mathbb{R}^{L \times n}$ is the parameter matrix in the MLP. Each $\mathbf{f}_i \in \mathbb{R}^{1 \times n}$ denotes GIP rating vector of user u_i and $\mathbf{F} = [\mathbf{f}_1, \dots, \mathbf{f}_m]^\top \in \mathbb{R}^{m \times n}$ denotes the all users' GIP ratings.

- (2) From the perspective of the user, we leverage user preference representation $\mathbf{v}_i \in \mathbb{R}^{R \times L}$ of user-POI interactions, combined with user social representation $\mathbf{u}_i \in \mathbb{R}^{1 \times L}$, to compute a rating vector of users on POIs as follows:

$$\begin{cases} \mathbf{z}_i = \mathbf{w}_a \cdot \text{Concat}(\mathbf{v}_i, \mathbf{u}_i) \\ \mathbf{e}_i = \text{MLP}(\mathbf{z}_i) \end{cases} \tag{14}$$

where $\mathbf{w}_a \in \mathbb{R}^{(R+1)}$ is the parameter vector of the aggregation layer. We use \mathbf{z}_i as the input of MLP to get a SIP rating vector $\mathbf{e}_i \in \mathbb{R}^{1 \times n}$ of user u_i for all POIs. $\mathbf{W}^{(2)} \in \mathbb{R}^{L \times D}$, $\mathbf{W}^{(3)} \in \mathbb{R}^{D \times L}$ and $\mathbf{W}^{(4)} \in \mathbb{R}^{L \times n}$ are the parameter matrices of the MLP. D is the latent dimension of hidden layer. $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_m]^\top \in \mathbb{R}^{m \times L}$ and $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_m]^\top \in \mathbb{R}^{m \times n}$ denotes all users' SIP ratings for POIs.

- (3) Finally, we combine the GIP rating \mathbf{f}_i and the SIP rating \mathbf{e}_i to get a final rating $\hat{\mathbf{y}}_i$, which is used to recommend a list of POIs for user u_i .

$$\hat{\mathbf{y}}_i = \text{sigmoid}(\mathbf{f}_i + \mathbf{e}_i) \tag{15}$$

where \mathbf{e}_i captures user u_i 's preference from user-POI interactions and social influence, and \mathbf{f}_i models the influence of geographic location and preference influence. sigmoid is a activation function and $\hat{\mathbf{Y}} = [\hat{\mathbf{y}}_1, \dots, \hat{\mathbf{y}}_m]^\top \in \mathbb{R}^{m \times n}$ denotes the predicted ratings for all m users.

Algorithm 1 Training process of GCN_{location} and GCN_{social}.

Input: POIs’ geographical coordinates set \mathcal{L} ; Users’ friend set \mathcal{F} ; Users’ check-in set \mathcal{R} .
Output: POIs’ location representation \mathbf{P} and users’ social representation \mathbf{U} .

```

1 while  $\mathcal{L}_{GCN\text{-}location}$  not converged do
2   foreach  $p_i \in \mathcal{P}$  do
3     Calculate Location similarity matrix  $\mathbf{A}$  via (1);
4     Calculate POI  $p_i$  location representation  $\mathbf{p}_i$  via (2), (3), (4) and (5);
5     Optimize the cross-entropy function (16) and update parameters by gradient
      descent.
6   end
7 end
8 while  $\mathcal{L}_{GCN\text{-}social}$  not converged do
9   foreach  $u_i \in \mathcal{U}$  do
10    Calculate similarity matrix  $\mathbf{A}^*$  via (6) or (7);
11    Calculate user  $u_i$  social representation  $\mathbf{u}_i$  from GCNsocial;
12    Optimize the cross-entropy function (17) and update parameters by gradient
      descent.
13  end
14 end

```

3.5 Optimization

We now turn towards optimizing three components: two GCNs and HGMAP. To do so, we first need to define the objective loss function of each and the overall. The training processes are summarized in algorithm 1 and algorithm 2.

- (1) For the GCNs learning POI location representation (denoted by GCN_{location}), we utilize the Cross Entropy loss to capture both POIs’ location similarity and POIs’ geographic location representation.

$$\begin{cases} \mathbf{X}_1 = \mathbf{P} \cdot \mathbf{P}^\top \\ \mathcal{L}_{GCN_{location}} = \sum_{i=1}^n - [\mathbf{A} \log(\mathbf{X}_1) + (1 - \mathbf{A}) \log(1 - \mathbf{X}_1)] \end{cases} \tag{16}$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the location similarity matrix and $\mathbf{P} \in \mathbb{R}^{n \times L}$ is the location representation.

- (2) For the GCNs learning user social representation (denoted by GCN_{social}), we incorporate users’ social similarity and users’ social representation into the loss function.

$$\begin{cases} \mathbf{X}_2 = \mathbf{U} \cdot \mathbf{U}^\top \\ \mathcal{L}_{GCN_{social}} = \sum_{i=1}^m - [\mathbf{A}^* \log(\mathbf{X}_2) + (1 - \mathbf{A}^*) \log(1 - \mathbf{X}_2)] \end{cases} \tag{17}$$

where $\mathbf{A}^* \in \mathbb{R}^{m \times m}$ is the social similarity matrix and $\mathbf{U} \in \mathbb{R}^{m \times L}$ is the social representation. During the GCNs training process, we take 5,000 POI locations (or 5,000 users) in each batch to calculate their corresponding representation.

- (3) Following prior work [34], the Mean Square Error (MSE) loss is commonly used to optimize MLP. In this study, we leverage a general weighting scheme [16] to distinguish visited and unvisited POIs, where we provide a confidence level for each POI [34] to tackle the One Class Collaborative Filtering (OCCF) problem. $\mathbf{Q} \in \mathbb{R}^{m \times n}$ denotes the confidence matrix and is computed using the observed check-in frequency

matrix $\mathbf{G} \in \mathbb{R}^{m \times n}$. This can calculate loss values more accurately and optimize our model better.

$$q_{i,j} = \begin{cases} \log(1 + g_{i,j}/\xi) & \text{if } g_{i,j} > 0 \\ 1 & \text{otherwise} \end{cases} \tag{18}$$

where ξ is a hyper-parameter. The objective function $\mathcal{L}_{\text{HGMAP}}$ for optimizing MLP is to measure the discrepancy between predicted value $\hat{\mathbf{Y}}$ and ground-truth value \mathbf{Y} .

$$\begin{aligned} \mathcal{L}_{\text{HGMAP}} &= \sum_{i=1}^m \sum_{j=1}^n \|q_{i,j} (\mathbf{y}_{i,j} - \hat{\mathbf{y}}_{i,j})\|_2^2 + \gamma (\|\mathbf{W}^{(*)}\|_F^2 + \|\mathbf{w}_a\|_2^2) \\ &= \|\mathbf{Q} \otimes (\mathbf{Y} - \hat{\mathbf{Y}})\|_F^2 + \gamma (\|\mathbf{W}^{(*)}\|_F^2 + \|\mathbf{w}_a\|_2^2) \end{aligned} \tag{19}$$

where \otimes is the element-wise multiplication and $\|\cdot\|_F$ is the Frobenius Norm. γ is the regularization parameter and $\mathbf{W}^{(*)}$ includes $\mathbf{W}^{(1)}$, $\mathbf{W}^{(2)}$, $\mathbf{W}^{(3)}$ and $\mathbf{W}^{(4)}$. $\mathbf{W}^{(1)}$ is the parameter matrix of the embedding layer and \mathbf{w}_a is the learned parameter vector in the aggregation layer. $\mathbf{W}^{(2)}$, $\mathbf{W}^{(3)}$ and $\mathbf{W}^{(4)}$ are the parameter matrices of the MLP. We leverage Adam [17] to automatically adjust the learning rate during learning.

Algorithm 2 Training process of HGMAP.

```

Input: User' historical check-ins set  $\mathcal{C}$ ; User set  $\mathcal{U}$ ; POI set  $\mathcal{P}$ .
Output: final rating matrix  $\hat{\mathbf{Y}}$ .
1 Initialize parameter matrices  $\mathbf{W}^{(1)}$ ,  $\mathbf{W}^{(2)}$ ,  $\mathbf{W}^{(3)}$ ,  $\mathbf{W}^{(4)}$  and  $\mathbf{w}_a$  randomly.
2 while  $\mathcal{L}_{\text{HGMAP}}$  not converged do
3   foreach  $u_i \in \mathcal{U}$  do
4     Calculate the check-in POIs' representation  $\hat{\mathbf{o}}_i$  via (9);
5     foreach  $t_r \in \mathbf{T}$  do
6       Employ multi-head attention mechanism to learn preference score matrix  $\mathbf{s}_r$ 
7       of user  $u_i$  via (10);
8     end
9     Calculate user  $u_i$  preference representation  $\mathbf{v}_i$  via (11);
10    Calculate GIP rating  $\mathbf{f}_i$  and SIP rating  $\mathbf{e}_i$  via (13) and (14);
11    Calculate final rating for POIs  $\hat{\mathbf{y}}_i$  via (15);
12    Calculate the confidence matrix  $\mathbf{Q}$  via (18);
13    Optimize the Mean Square Error(MSE) function (19) and update parameters by
14    gradient descent.
15  end
16 end

```

4 Experiments

In this section, we report observations from experiments conducted on three real-world datasets to quantitatively address the following questions:

- **Q1.** How does HGMAP perform compared with the state-of-the-art POI recommendation models?
- **Q2.** How do the hybrid GCNs in HGMAP affect the recommendation performance?
- **Q3.** How do the key hyper-parameters affect HGMAP's performance?

Table 2 Descriptive statistics of three datasets

Dataset	#Users	#POIs	#Check-ins	Density
Yelp	30,887	18,995	860,888	0.1399%
Gowalla	43,074	46,234	1,720,082	0.0500%
Foursquare	24,941	28,593	1,196,248	0.1006%

- **Q4.** Can HGMAP provide reasonable interpretability regarding user preference towards POIs?

4.1 Dataset and Evaluation Metric

To evaluate the effectiveness of HGMAP, we conducted experiments on three benchmark LBSN datasets, including:

- **Yelp dataset.** It is obtained from the Yelp challenge.¹ This dataset does not provide the exact check-in times but coarse check-in dates.
- **Gowalla dataset.** It is a widely used for POI recommendation and was collected between February 2009 and October 2010.
- **Foursquare dataset.** It is collected between April 2012 to September 2013 within the mainland of United States. Note that this data does not have social information, thus we do not model the social influence for this data.

Following the settings in [31, 34], we filter out those users with fewer than 20 check-in POIs and those POIs with fewer than 20 visitors for the Gowalla dataset. For Foursquare and Yelp datasets, we discard those users with fewer than 10 check-in POIs and those POIs with fewer than 10 visitors. We also partition each dataset into training set and test set. For example, for each user, we randomly select the 80% check-ins into the training, and treat the remaining as the testing. The descriptive statistics of three datasets after pre-processing are described in Table 2, from which we can see that they are all extremely sparse, i.e., the frequency of most POIs being visited is about 0.1%.

Similar to previous works [31, 34], we use three standard metrics, i.e., precision ($P@k$), recall ($R@k$) and mean average precision ($M@k$), to evaluate models. $P@k$ is the percentage of locations that are visited by user in the $top-k$ recommended locations. $R@k$ indicates the ratio of recovered POIs to visited locations and $M@k$ considers the rank of recommendations by assigning higher score to hits at higher positions.

4.2 Baselines

We conduct extensive comparisons to the following 12 state-of-the-art POI recommendation models:

- **MGMMF** [5] is a multi-center Gaussian model fused with matrix factorization, taking into account social influence and incorporating multi-center geographical influence into the fused framework. The main idea is based on the observation that a user tends to check-in around several geographical centers.

¹<https://www.yelp.com/dataset/challenge>

- **BPRMF** [41] is a Bayesian personalized ranking with matrix factorization method. It adopts a generic optimization criterion and models the implicit feedback to recommend *top-N* items. Note that BPRMF only focuses on user preference modeling, without utilizing any context information.
- **WRMF** [16] is a weighted regularized matrix factorization model. It couples the estimate of user preference to items with a confidence level based on matrix factorization while minimizing the square mean error. It assigns both observed and unobserved check-ins with different confidence values.
- **IRenMF** [29] is based on weighted matrix factorization and incorporates the geographical characteristics of neighboring POIs in both individual level (i.e., user has similar preference on neighboring POIs) and region level (i.e., POIs that are geographically close may share similar user preference) into the model.
- **GeoMF** [26] is a state-of-the-art MF-based POI recommendation model based on weighted matrix factorization. It considers check-ins as an implicit feedback and incorporates geographical influence by fitting nonzero check-ins with large weights and zero check-ins with smaller weights.
- **RankGeoFM** [22] is a ranking based geographical factorization method that incorporates the geographical influence of neighboring POIs to learn user preference rankings for POIs. It uses another latent matrix to represent user geographical preference, in addition to user preference matrix.
- **PACE** [52] is a deep neural architecture based on user preference and context embedding with representation methods [38]. It is a general semi-supervised learning framework that jointly models social influence and user trajectory behavior to predict both user preference over POIs and various context associated with users and POIs.
- **SAE-NAD** [34] is an attention-based POI recommendation model consisting of a self-attentive encoder and a neighbor-aware decoder. It uses a self-attentive encoder to differentiate the user preference, and adopts the neighbor-aware decoder to model the geographical influence of POIs.
- **STGN** [64] is a Spatio-Temporal Gated Network towards enhancing long-short term memory of the sequential visiting behavior learning. It uses coupled gates, i.e., time gate and distance gate, to capture the spatial-temporal relationship among successive check-ins.
- **APOIR** [71] is the first adversarial learning-based POI recommendation model. It consists of two parts, a recommender and a discriminator, which are jointly trained for learning user preference by playing a minimax game considering geographical influence and social relation as rewards in a reinforcement learning manner.
- **Geo-ALM** [32] is a geographical information based adversarial learning model which is very similar to APOIR, except that Geo-ALM directly fuses geographical features (both POI features and region features) and uses generative adversarial networks [12] without explicitly considering the social influence.
- **NGCF** [50], Neural Graph Collaborative Filtering, is the most recent item-based recommendation model built upon graph convolutional networks. NGCF only focuses on convolutional operations on user-item interactions while HGMAP learns additional information from both the user side and the POI side.

4.3 Parameter setting

We implement our HGMAP with Pytorch on a machine with NVIDIA GeForce GTX 1080Ti. In our experiments, the latent dimension L of both users' social representation and

POIs' location representation is set to 200. For two GCNs, the minimum value λ regarding user similarity and location similarity are both set to 0.125 unless otherwise specified. The geographical relevance level η is set to 60 in $\text{GCN}_{\text{location}}$ and the parameter β used for balances the importance of friend circle similarity and user visiting similarity is set to 0.3. The latent dimension R of the user preference vector and the number of attention heads are set to 36 and 6. The batch size of HGMAP is set to 256. The learning rate and regularization parameter γ are set to 0.001 and 0.001 respectively. We set the architecture of two-layer $\text{GCN}_{\text{social}}$ as $[m, 3000, 200]$. $\text{GCN}_{\text{location}}$ with two-layer has architecture as $[n, 3000, 200]$. m and n are the number of users and POIs in the input layer, respectively. For three datasets, we use an embedding layer and a 3-layer MLP as $[200, n]$ and $[200, 50, 200, n]$.

4.4 Performance comparison (Q1)

Tables 3, 4 and 5 illustrate the performance of HGMAP in comparison to the existing state-of-the-art POI recommendation models for top- K POI recommendation on Gowalla, Foursquare and Yelp, respectively. A pair t - $test$ is performed and the results are statistically significant ($p < 0.005$). By scrutinizing the results, we can make the following observations:

- (O1): General MF-based models, such as WRMF and BPRMF, achieve poor performance on three datasets, because they ignore the context information, e.g., social influence and geographical constraints. Meanwhile, simply incorporating geographical clustering phenomena of check-ins (e.g., MGMMF) does not perform well, since it fails to overlook the fine-grained POI-level context. In contrast, geographical MF-based implicit ranking methods, such as IRenMF, GeoMF and RankGeoFM, perform relatively well, which indicates that modeling user check-ins as implicit feedback is more appropriate in POI recommendation and that geographical influence is the most important factor for POI recommendation.x

Table 3 Performance comparison between HGMAP and baselines on the Gowalla dataset

Method	Gowalla								
	P@5	P@10	P@20	R@5	R@10	R@20	M@5	M@10	M@20
MGMMF	0.0437	0.0336	0.0261	0.0653	0.0991	0.1480	0.0459	0.0480	0.0517
BPRMF	0.0663	0.0495	0.0402	0.0895	0.1272	0.1740	0.0611	0.0621	0.0668
WRMF	0.0635	0.0483	0.0360	0.0753	0.1111	0.1617	0.0567	0.0555	0.0583
IRenMF	0.0817	0.0621	0.0489	0.1006	0.1547	0.2264	0.0767	0.0752	0.0776
GeoMF	0.0902	0.0675	0.0496	0.1042	0.1576	0.2286	0.0789	0.0796	0.0821
RankGeoFM	0.0898	0.0687	0.0505	0.1079	0.1613	0.2312	0.0815	0.0812	0.0856
PACE	0.0918	0.0697	0.0515	0.1129	0.1713	0.2412	0.0835	0.0832	0.0876
SAE-NAD	0.0976	0.0748	0.0549	0.1227	0.1822	0.2593	0.0924	0.0930	0.0986
STGN	0.0920	0.0701	0.0517	0.1131	0.1715	0.2415	0.0837	0.0834	0.0878
APOIR	0.0983	0.0763	0.0565	0.1243	0.1844	0.2605	0.0939	0.0948	0.0995
Geo-ALM	0.0922	0.0703	0.0519	0.1134	0.1718	0.2419	0.0839	0.0837	0.0881
NGCF	0.0978	0.0750	0.0551	0.1229	0.1825	0.2596	0.0927	0.0933	0.0990
HGMAP	0.1037	0.0798	0.0583	0.1316	0.1976	0.2814	0.0981	0.0993	0.1052

Table 4 Performance comparison between HGMAP and baselines on the foursquare dataset

Method	Fourquare								
	P@5	P@10	P@20	R@5	R@10	R@20	M@5	M@10	M@20
MGMMF	0.0415	0.0317	0.0240	0.0408	0.0616	0.0934	0.0300	0.0287	0.0310
BPRMF	0.0734	0.0552	0.0410	0.0716	0.1062	0.1550	0.0556	0.0529	0.0569
WRMF	0.0749	0.0560	0.0411	0.0727	0.1083	0.1550	0.0577	0.0546	0.0569
IRenMF	0.0814	0.0594	0.0422	0.0761	0.1101	0.1545	0.0605	0.0560	0.0619
GeoMF	0.0818	0.0632	0.0466	0.0776	0.1164	0.1697	0.0603	0.0573	0.0592
RankGeoFM	0.0829	0.0644	0.0478	0.0781	0.1190	0.1747	0.0594	0.0571	0.0619
PACE	0.0825	0.0654	0.0488	0.0791	0.1204	0.1807	0.0614	0.0581	0.0629
SAE-NAD	0.0912	0.0698	0.0518	0.0873	0.1317	0.1927	0.0666	0.0641	0.0698
STGN	0.0828	0.0658	0.0490	0.0794	0.1207	0.1809	0.0617	0.0584	0.0632
APOIR	0.0921	0.0710	0.0527	0.0881	0.1324	0.1937	0.0672	0.0650	0.0705
Geo-ALM	0.0832	0.0663	0.0493	0.0797	0.1210	0.1813	0.0621	0.0587	0.0635
NGCF	0.0913	0.0700	0.0521	0.0875	0.1320	0.1923	0.0669	0.0643	0.0701
HGMAP	0.0943	0.0729	0.0537	0.0899	0.1374	0.1996	0.0675	0.0654	0.0711

(O2): Compared to MF-based models, neural networks-based methods, including HGMAP, exhibit better performance. This demonstrates the importance of non-linear feature interactions between users and POI embeddings. In other words, the inner product in MF-based methods is insufficient to capture the complex interactions between users and POIs.

(O3): Among the deep recommendation models, PACE does exhibit the performance as expected, because it only learns the *shallow* embedding of users and POIs, while the collaborative filtering signals are not fully exploited. Similarly, STGN, mainly

Table 5 Performance comparison between HGMAP and baselines on the Yelp dataset

Method	Yelp								
	P@5	P@10	P@20	R@5	R@10	R@20	M@5	M@10	M@20
MGMMF	0.0210	0.0178	0.0146	0.0225	0.0381	0.0621	0.0142	0.0147	0.0163
BPRMF	0.0417	0.0351	0.0287	0.0408	0.0684	0.1112	0.0283	0.0286	0.0311
WRMF	0.0431	0.0374	0.0309	0.0445	0.0719	0.1189	0.0310	0.0315	0.0334
IRenMF	0.0487	0.0397	0.0319	0.0481	0.0771	0.1230	0.0352	0.0349	0.0376
GeoMF	0.0498	0.0410	0.0324	0.0478	0.0783	0.1233	0.0353	0.0350	0.0375
RankGeoFM	0.0503	0.0418	0.0330	0.0482	0.0791	0.1237	0.0355	0.0352	0.0376
PACE	0.0501	0.0411	0.0328	0.0481	0.0790	0.1207	0.0350	0.0350	0.0370
SAE-NAD	0.0517	0.0424	0.0334	0.0514	0.0829	0.1287	0.0373	0.0373	0.0401
STGN	0.0502	0.0413	0.0330	0.0484	0.0796	0.1215	0.0352	0.0351	0.0372
APOIR	0.0532	0.0436	0.0348	0.0526	0.0843	0.1312	0.0382	0.0391	0.0421
Geo-ALM	0.0506	0.0416	0.0332	0.0492	0.0801	0.1222	0.0354	0.0355	0.0375
NGCF	0.0515	0.0431	0.0330	0.0518	0.0831	0.1293	0.0376	0.0377	0.0403
HGMAP	0.0564	0.0467	0.0368	0.0568	0.0927	0.1434	0.0406	0.0410	0.0441

focusing on sequential check-in behavior of users, does not show expected performance. The possible reason is that STGN fails to explicitly explore the important interactions between users and POIs, as well as other user and POI contexts, e.g., social influence and POI-level neighboring information.

- (O4): Furthermore, SAE-NAD shows good performance on POI recommendation, mainly because it captures the non-linear interactions between users and POIs with deep autoencoder and attention mechanism. However, it ignores the social influence, as well as the high-order connectivity among POIs. In addition, two adversarial POI recommendation models, APOIR and Geo-ALM, generally achieve better performance than SAE-NAD, due to their high-quality negative sampling and capability of general user preference learning. The slight improvement of APOIR over Geo-ALM indicates the effectiveness of social influence modeling in APOIR.
- (O5): Our HGMAP consistently yields the best performance across all datasets. For example, HGMAP improves over the second best baseline w.r.t. $R@10$ by 7.2%, 3.8% and 10% on Gowalla, Foursquare and Yelp datasets, respectively. Compared to APOIR and SAE-NAD – two representative non-linear interaction learning methods – HGMAP explicitly models the POI adjacent graph by propagating the connectivity over the graph. Note that although SAE-NAD considers the POI distance, it neither learns high-order connectivity among POIs, nor does it incorporate the social influence. This result also demonstrates the effectiveness of our graph convolutions on both social graph and POI graph.
- (O6): Lastly, we note that NGCF does not perform well on POI recommendation, although it adopts the graph convolution for non-linear user-POI interaction learning. The performance gain of HGMAP over NGCF demonstrates the effectiveness of social influence learning in HGMAP. Moreover, our method does not learn the collaborative interactions via graph neural networks – which is the case of NGCF, but instead applies graph learning on social relationship and POI neighboring connection. This result also provides another perspective of incorporating graph neural networks into recommender systems. Due to the extremely sparse check-ins, the collaborative signal, arguably, cannot be effectively captured only by graph neural networks.

4.5 Ablation study (Q2)

To investigate the impact of social influence and geographical constraints, we conducted an ablation study by comparing to three variants of HGMAP. In particular: the first variant **HGMAP-I** is formed by disabling the graph convolutional networks modeling social influence – note that there are no social relationship in the Foursquare data; the second variant **HGMAP-II** replaces the POI adjacent graph neural networks with a simple distance matrix, as used in [34]; the third variant **HGMAP-III** replaces the multi-head attention module in HGMAP with another GCN, which propagates the user interest over POIs in the user-POI interaction graph, similar to the GCN used in NGCF [50]. We summarize the experimental results in Table 6, from which we have the following findings:

- (F1): The discrepancy between HGMAP and HGMAP-I implies the effectiveness of social influence, which makes sense since social relationship plays an important role in (POI) recommendation [31, 60], especially for those cold-start users who have less and even no check-in records. This result also explains that why those deep recommendation methods, such as PACE, STGN, Geo-ALM and NGCF, do not perform

Table 6 Ablation study of HGMAP

Method	Yelp		
	P@10	R@10	M@10
HGMAP-I	0.0421	0.0883	0.0382
HGMAP-II	0.0432	0.0895	0.0390
HGMAP-III	0.0450	0.0912	0.0401
HGMAP	0.0467	0.0927	0.0410
Method	Gowalla		
HGMAP-I	0.0756	0.1913	0.0948
HGMAP-II	0.0770	0.1937	0.0961
HGMAP-III	0.0782	0.1953	0.0975
HGMAP	0.0798	0.1976	0.0993
Method	Fourquare		
HGMAP-II	0.0690	0.1327	0.0618
HGMAP-III	0.0705	0.1345	0.0637
HGMAP	0.0723	0.1365	0.0655

Note that there is no social relationship in the Foursquare data

well. Note that there exist many social graph learning models such as DeepWalk and node2vec that explore the local connectivity among nodes. However, these methods mainly focus on preserving the local structure, therefore ignoring high-order connectivity among nodes.

- (F2): Compared to HGMAP-II, HGMAP yields remarkable improvements, which demonstrates the effectiveness of the proposed POI graph neural networks in HGMAP. It is commonly acknowledged that geographical influence is one of the most important factors in POI recommendation [22, 26, 31]. However, existing methods vary significantly from each other on how to incorporate this constraint. While earlier efforts have incorporated the geographical information into MF which are limited by the non-linear interactions of inner product, the recent deep learning-based methods either simply compute the POI distance [34] or model it as a reward function [71] – both of which are not sufficient to capture the implicit connections and possible patterns among POIs. In contrast, HGMAP explicitly learns the relationship from the POI graph, which not only captures meaningful but non-existing check-in behavior of users, but also provides a way of augmenting the sampling data by propagating the information on the POI graph. In this vein, it can be considered as a LBSN data augmentation to alleviate the sparse check-in problem [70].
- (F3): Moreover, HGMAP-III does not show comparable performance even with another graph convolution on user-POI interactions. This result proves our conjecture that the sparse check-in problem in LBSN dataset renders the graph collaborative filtering method inapplicable for capturing user-POI interactions. The reason behind this phenomena can be understood intuitively. That is, aggregating the embeddings of the interactions between users and POIs would be largely hindered for users with few check-ins or POIs with few visitors. Therefore, the collaborative signals would be easily “blocked” for cold-start users and/or POIs when embedding propagation, which could be further aggravated by stacking multiple layers of graph convolutions for sparse check-in data.

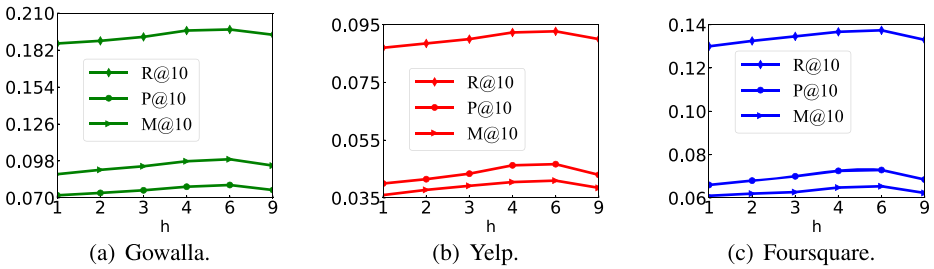


Figure 3 Impact of number of heads h

4.6 Sensitivity of parameters (Q3)

Now we investigate several important parameters of HGMAP, i.e., the number of attention heads h and the parameter λ which is the threshold value of user and POI similarity in GCNs.

Effect of h : HGMAP adapts a multi-head self-attention mechanism to capture the multi-aspects of user-POI interactions. Figure 3a, b and c plot the influence of the number of heads, where we can observe that 4 or 6 heads are enough for our model to achieve good performance.

Effect of λ : Parameter λ specifies the lower bound value of identifying similar users and POIs, below which the similarity between two users (or POIs) is to 0, i.e., the lower the value, the more non-zero similarity scores, and therefore more computation required in the model. Figure 4a, b and c show the effect of λ , which indicates that HGMAP attains the best performance when $\lambda = 0.125$. Note that it is better to distinguish this hyper-parameter for users and POIs. However, we found that the difference is very nuance in our experiments.

Effect of η : Parameter η is used to control the geographical relevance level between POIs in $GCN_{location}$, which can be used to jointly capture both POIs' location similarity and POIs' geographic location representation. Figure 5a, b and c show the impact of η on three datasets, which indicates that HGMAP attains the best performance when η is within the range of [60-80].

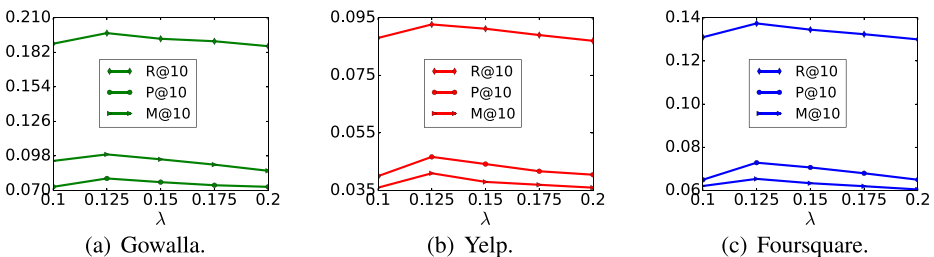


Figure 4 Impact of similarity lower bound value λ

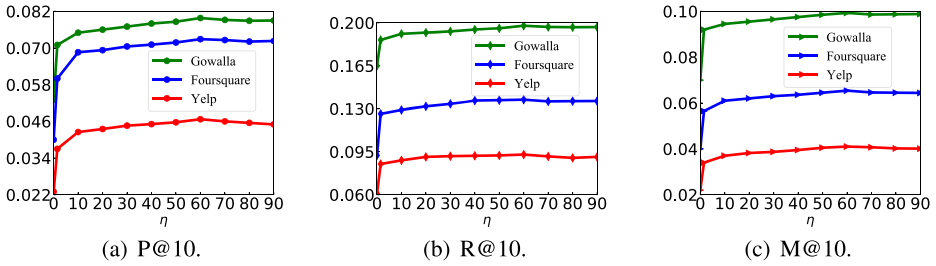


Figure 5 Impact of geographical relevance level control value η

Effect of β : Parameter β balances the relative weight of friend circle similarity and user visiting similarity in GCN_{social} . Figure 6a and b reveal the influence of β on model performance. Clearly, HGMAP achieves the best performance when $\beta=0.3$. This demonstrates that visiting similarity of users has a higher influence score than friend circle similarity for modeling user presentation. Note that foursquare dataset has no user social information.

Convergence : Another merit of HGMAP is the high computational efficiency. HGMAP consists of three main components, i.e., two GCNs for social influence and geographical influence learning, and one multi-head attention encoder for user-POI interaction learning. For the two GCNs, they only have 2-layer convolutions without non-linear transformation in the first layer – which yields improvements in computational efficiency. In addition, HGMAP also consists of 3-layer MLPs, which also has a fast converge rate. Figure 7 illustrates the training of HGMAP, which indicates that our model can fast converge to optimal performance. For example, it achieves the best performance on Precision and MAP with around 40 epochs.

4.7 Interpretability (Q4)

To better understand HGMAP, we visualize the user and POI embeddings learned from HGMAP using t -SNE [36]. Figure 8 plots the 2D visualization of the representation derived from the training of Yelp, Foursquare, and Gowalla. Obviously, the closeness of users and POIs are well reflected in the learned representation space, and users (POIs) of the same type are usually mapped to close positions in two-dimensional space. Each point denotes a user in Fig. 8a, c and e; and a POI in Fig. 8b, d and f, respectively. Figure 8a, c and e show that the embeddings of users are well clustered, meaning that our model can distinguish

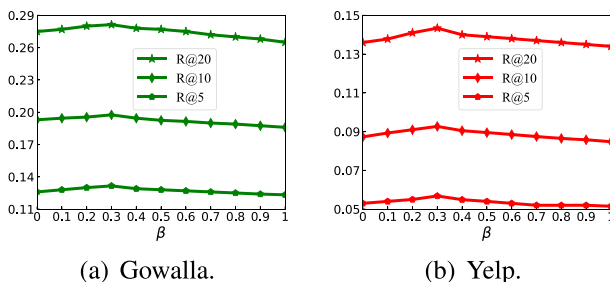


Figure 6 Impact of relative weight value β between user social similarity and visiting similarity

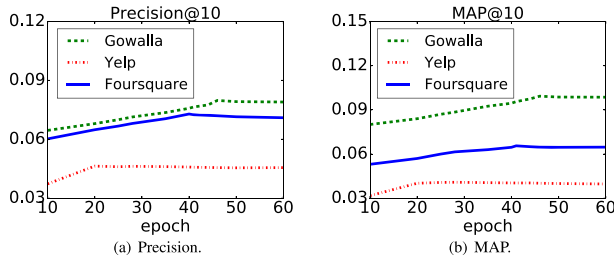


Figure 7 Convergence of HGMAP

users. Additionally, each color represents a type of users who have a similar circle of friends and visiting record. In other words, users do exhibit certain discernible patterns in their POI check-ins which our HGMAP aims to capture. Similarly, we observe that the proximity of POI embeddings corresponds well with the similarity of user check-ins. In the same fashion, each color denotes a type of POIs that have a similar geographical position in Fig 8b, d and f. It means that a given POI presented to a user was relevant enough for that user to check-in this POI, so that HGMAP can retrieve it later, i.e., it is beneficial for the accurate recommendation of HGMAP.

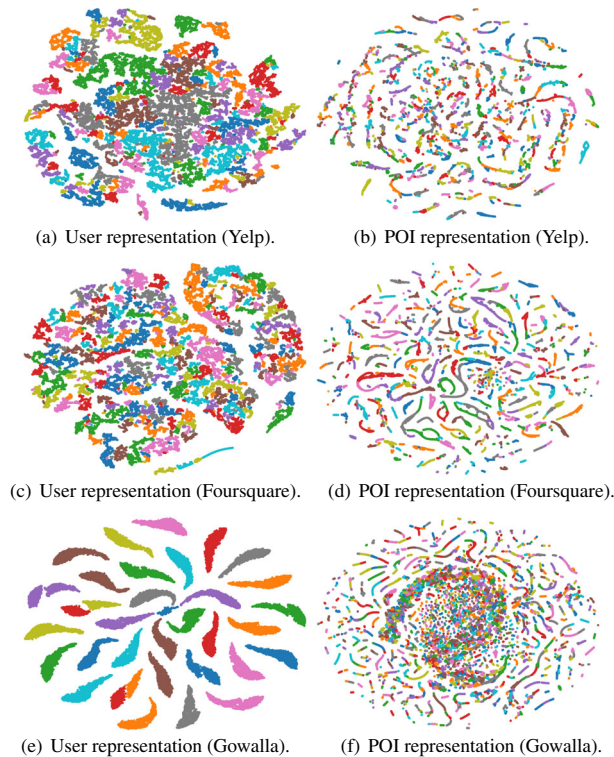


Figure 8 Visualization of the learned user and POI representation on the Yelp, Foursquare and Gowalla datasets

5 Conclusion

In this study, we present a novel hybrid graph-based model HGMAP for POI recommendation, which consists of two graph neural networks and one multi-head attention encoder. Instead of only modeling user-item (POI) interactions as previous works do, we exploit the graph neural networks for capturing auxiliary information including social influence and geographical constraints. A POI adjacent graph is constructed to capture the implicit user mobility patterns by propagating the check-in embeddings on the POI graph. The experimental results based on three real-world datasets demonstrate that the proposed model outperforms the state-of-the-art baselines, and the latent space learned from both user and POI embedding propagation can well reflect discernible clustering patterns. This, in turn, indicates a promising direction that training and optimizing recommendation tasks with graph-based auxiliary information learning, especially for sparse data and cold-start users (items).

One of our immediate future works is to incorporate other auxiliary information for better POI recommendation, such as temporal features, POI categories and sequential check-in behavior. An important question that we plan to tackle is the shallow issue of graph neural networks due to the vanishing gradient problem in stacking multiple layers. We also plan to investigate methods against the sparse user-POI interactions by leveraging deep generative models [21, 25, 27] to discover underlying non-linear user-POI interactions while improving the recommendation performance.

Acknowledgments This work was supported by National Natural Science Foundation of China (Grant No.61602097 and No.61472064), NSF grant CNS 1646107.

References


1. Altaf, B., Yu, L., Zhang, X.: Spatio-Temporal Attention Based Recurrent Neural Network for Next Location Prediction. In: IEEE International Conference on Big Data, Big Data 2018, Seattle, pp. 937–942 (2018)
2. Bahdanau, D., Cho, K., Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate. In: ICLR (2015)
3. Chen, J., Zhang, H., He, X., Nie, L., Liu, W., Chua, T.S.: Attentive collaborative filtering: Multimedia recommendation with item-and component-level attention. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 335–344 (2017)
4. Chen, X., Zhou, F., Zhang, K., Trajcevski, G., Zhong, T., Zhang, F.: Information Diffusion Prediction via Recurrent Cascades Convolution. In: 2019 IEEE 35Th International Conference on Data Engineering (ICDE), pp. 770–781. IEEE (2019)
5. Cheng, C., Yang, H., King, I., Lyu, M.R.: Fused matrix factorization with geographical and social influence in location-based social networks. In: Proceedings of the AAAI International Conference on Artificial Intelligence (2012)
6. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. In: Advances in Neural Information Processing Systems (NIPS), pp. 3844–3852 (2016)
7. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 4171–4186 (2019)
8. Eom, C.S., Lee, C.C., Lee, W., Leung, C.K.: Effective privacy preserving data publishing by vectorization. *Information Sciences* (2019)
9. Fan, S., Zhu, J., Han, X., Shi, C., Hu, L., Ma, B., Li, Y.: Metapath-guided heterogeneous graph neural network for intent recommendation. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 2478–2486. ACM (2019)

10. Gao, Q., Zhou, F., Zhang, K., Trajcevski, G., Luo, X., Zhang, F.: Identifying human mobility via trajectory embeddings. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), pp. 1689–1695 (2017)
11. Gao, Q., Zhou, F., Trajcevski, G., Zhang, K., Ting, Z., Zhang, F.: Predicting human mobility via variational attention. In: Proceedings of the International Conference on World Wide Web (WWW), pp. 2750–2756. ACM (2019)
12. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Nets. In: Advances in Neural Information Processing Systems (NIPS), pp. 2672–2680 (2014)
13. Hang, M., Pytlarz, I., Neville, J.: Exploring student check-in behavior for improved point-of-interest prediction. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 321–330. ACM (2018)
14. Hao, M., Chao, L., King, I., Lyu, M.R.: Probabilistic factor models for web site recommendation. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 265–274. ACM (2011)
15. Hosseini, S., Yin, H., Zhou, X., Sadiq, S., Kangavari, M.R., Cheung, N.M.: Leveraging multi-aspect time-related influence in location recommendation. *World Wide Web* **22**(3), 1001–1028 (2019)
16. Hu, Y., Koren, Y., Volinsky, C.: Collaborative filtering for implicit feedback datasets. In: Proceedings of the International Conference on Data Mining (ICDM), pp. 263–272. IEEE (2008)
17. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv:1412.6980 (2014)
18. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: Proceedings of the International Conference on Learning Representations (ICLR) (2017)
19. Koren, Y., Bell, R., Volinsky, C.: Matrix factorization techniques for recommender systems. *J. Comput.* **48**(8), 30–37 (2009)
20. Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P., Soricut, R.: Albert: a Lite Bert for Self-Supervised Learning of Language Representations. In: International Conference on Learning Representations (2019)
21. Lee, W., Song, K., Moon, I.C.: Augmented variational autoencoders for collaborative filtering with auxiliary information. In: Proceedings of the International Conference on Information and Knowledge Management (CIKM), pp. 1139–1148. ACM (2017)
22. Li, X., Cong, G., Li, X.L., Pham, T.A.N., Krishnaswamy, S.: Rank-geofm: a ranking based geographical factorization method for point of interest recommendation. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 433–442 (2015)
23. Li, X., Cong, G., Li, X.L., Pham, T.A.N., Krishnaswamy, S.: Rank-geofm: a ranking based geographical factorization method for point of interest recommendation. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 433–442. ACM (2015)
24. Li, H., Ge, Y., Hong, R., Zhu, H.: Point-of-interest recommendations: Learning potential check-ins from friends. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 975–984. ACM (2016)
25. Li, X., She, J.: Collaborative variational autoencoder for recommender systems. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 305–314. ACM (2017)
26. Lian, D., Zhao, C., Xie, X., Sun, G., Chen, E., Rui, Y.: Geomf: joint geographical modeling and matrix factorization for point-of-interest recommendation. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 831–840. ACM (2014)
27. Liang, D., Krishnan, R.G., Hoffman, M.D., Jebara, T.: Variational autoencoders for collaborative filtering. In: Proceedings of the International Conference on World Wide Web (WWW), pp. 689–698. ACM (2018)
28. Liu, B., Fu, Y., Yao, Z., Xiong, H.: Learning geographical preferences for point-of-interest recommendation. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 1043–1051. ACM (2013)
29. Liu, Y., Wei, W., Sun, A., Miao, C.: Exploiting geographical neighborhood characteristics for location recommendation. In: Proceedings of the International Conference on Information and Knowledge Management (CIKM), pp. 739–748. ACM (2014)
30. Liu, B., Xiong, H., Papadimitriou, S., Fu, Y., Yao, Z.: A general geographical probabilistic factor model for point of interest recommendation. *IEEE Trans. Knowl. Data Eng. (TKDE)* **27**(5), 1167–1179 (2015)
31. Liu, Y., Pham, T.A.N., Cong, G., Yuan, Q.: An experimental evaluation of point-of-interest recommendation in location-based social networks. *Proc. VLDB Endowment* **10**(10), 1010–1021 (2017)

32. Liu, W., Wang, Z.J., Yao, B., Yin, J.: Geo-alm: Poi recommendation by fusing geographical information and adversarial learning mechanism. In: Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), pp. 1807–1813 (2019)
33. Lu, Y.S., Shih, W.Y., Gau, H.Y., Chung, K.C., Huang, J.L.: On successive point-of-interest recommendation. *World Wide Web* **22**(3), 1151–1173 (2019)
34. Ma, C., Zhang, Y., Wang, Q., Liu, X.: Point-of-interest recommendation: Exploiting self-attentive autoencoders with neighbor-aware influence. In: Proceedings of the International Conference on Information and Knowledge Management (CIKM), pp. 697–706. ACM (2018)
35. Ma, F., Gao, F., Sun, J., Zhou, H., Hussain, A.: Attention graph convolution network for image segmentation in big SAR imagery data. *Remote. Sens.* **11**(21), 2586 (2019)
36. Maaten, L.v.d., Hinton, G.: Visualizing data using t-sne. *J. Mach. Learn. Res. (JMLR)* **9**, 2579–2605 (2008)
37. Manotumruksa, J., Macdonald, C., Ounis, I.: A contextual attention recurrent architecture for context-aware venue recommendation. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 555–564. ACM (2018)
38. Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed Representations of Words and Phrases and Their Compositionality. In: Advances in Neural Information Processing Systems (NIPS), pp. 3111–3119 (2013)
39. Monti, F., Bronstein, M., Bresson, X.: Geometric Matrix Completion with Recurrent Multi-Graph Neural Networks. In: Advances in Neural Information Processing Systems (NIPS), pp. 3697–3707 (2017)
40. Qian, T., Liu, B., Nguyen, Q.V.H., Yin, H.: Spatiotemporal representation learning for translation-based poi recommendation. *ACM Trans. Inf. Syst. (TOIS)* **37**(2), 18:1–18:24 (2019)
41. Rendle, S., Freudenthaler, C., Gantner, Z., Schmidt-Thieme, L.: Bpr: Bayesian personalized ranking from implicit feedback. In: Proceedings of International conference on uncertainty in artificial intelligence (UAI), pp. 452–461. AUAI Press (2009)
42. Salakhutdinov, R., Mnih, A.: Bayesian Probabilistic Matrix Factorization Using Markov Chain Monte Carlo. In: International Conference on Machine Learning (ICML), pp. 880–887 (2008)
43. van den Berg, R., Kipf, T.N., Welling, M.: Graph convolutional matrix completion. arXiv:1706.02263v2 (2017)
44. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is All You Need. In: Advances in Neural Information Processing Systems (NIPS), pp. 5998–6008 (2017)
45. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., Bengio, Y.: Graph attention networks. In: Proceedings of the International Conference on Learning Representations (ICLR) (2018)
46. Wang, W., Yin, H., Du, X., Nguyen, Q.V.H., Zhou, X.: TPM: a temporal personalized model for spatial item recommendation. *ACM Trans. Intell. Syst. Technol. (TIST)* **9**(6), 61:1–61:25 (2018)
47. Wang, H., Zhang, F., Zhang, M., Leskovec, J., Zhao, M., Li, W.: Wang, Z.: Knowledge-aware Graph Neural Networks with Label Smoothness Regularization for Recommender Systems. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 968–977. ACM (2019)
48. Wang, H., Zhao, M., Xie, X., Li, W., Guo, M.: Knowledge graph convolutional networks for recommender systems. In: Proceedings of the International Conference on World Wide Web (WWW), pp. 3307–3313. ACM (2019)
49. Wang, X., He, X., Cao, Y., Liu, M., Chua, T.: KGAT: knowledge graph attention network for recommendation. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 950–958. ACM (2019)
50. Wang, X., He, X., Wang, M., Feng, F., Chua, T.: Neural graph collaborative filtering. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 165–174. ACM (2019)
51. Wu, F., Zhang, T., Souza, Jr., A.H.d., Fifty, C., Yu, T., Weinberger, K.Q.: Simplifying Graph Convolutional Networks. In: International Conference on Machine Learning (ICML), pp. 6861–6871 (2019)
52. Yang, C., Bai, L., Zhang, C., Yuan, Q., Han, J.: Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation. In: Proceedings of the International Conference on Knowledge Discovery & Data Mining (SIGKDD), pp. 1245–1254. ACM (2017)
53. Ye, M., Yin, P., Lee, W.C., Lee, D.L.: Exploiting geographical influence for collaborative point-of-interest recommendation. In: Proceedings of the International conference on Research and development in information retrieval (SIGIR), pp. 325–334. ACM (2011)

54. Yin, H., Zhou, X., Cui, B., Wang, H., Zheng, K., Hung, N.Q.V.: Adapting to user interest drift for poi recommendation. *IEEE Trans. Knowl. Data Eng. (TKDE)* **28**(10), 2566–2581 (2016)
55. Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W.L., Leskovec, J.: Graph convolutional neural networks for web-scale recommender systems. In: *Proceedings of the SIGKDD International Conference on Knowledge Discovery & Data Mining (SIGKDD)*, pp. 974–983. ACM (2018)
56. Ying, H., Wu, J., Xu, G., Liu, Y., Liang, T., Zhang, X., Xiong, H.: Time-aware metric embedding with asymmetric projection for successive poi recommendation. *World Wide Web* **22**(5), 2209–2224 (2019)
57. Yuan, Q., Cong, G., Ma, Z., Sun, A., Thalmann, N.M.: Time-aware point-of-interest recommendation. In: *Proceedings of the International conference on Research and development in information retrieval (SIGIR)*, pp. 363–372. ACM (2013)
58. Zhang, J.D., Chow, C.Y.: Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In: *Proceedings of the International conference on Research and development in information retrieval (SIGIR)*, pp. 443–452. ACM (2015)
59. Zhang, C., Kim, J.: Object Detection with Location-Aware Deformable Convolution and Backward Attention Filtering. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 9452–9461 (2019)
60. Zhang, S., Yao, L., Sun, A., Tay, Y.: Deep learning based recommender system: A survey and new perspectives. *ACM Comput. Surv.* **52**(1), 5:1–5:38 (2019)
61. Zhang, Z., Liu, Y., Zhang, Z., Shen, B.: Fused matrix factorization with multi-tag, social and geographical influences for poi recommendation. *World Wide Web* **22**(3), 1135–1150 (2019)
62. Zhang, Y., Feng, Y., Shang, J., Zhou, M., Qiang, B.: Attention-aware joint location constraint hashing for multi-label image retrieval. *IEEE Access* **8**, 3294–3307 (2020)
63. Zhao, S., Zhao, T., King, I., Lyu, M.R.: Geo-teaser: Geo-temporal sequential embedding rank for point-of-interest recommendation. In: *Proceedings of the International Conference on World Wide Web (WWW)*, pp. 153–162. ACM (2017)
64. Zhao, P., Zhu, H., Liu, Y., Xu, J., Li, Z., Zhuang, F., Sheng, V.S., Zhou, X.: Where to go next: a spatio-temporal gated network for next poi recommendation. In: *Proceedings of the AAAI International Conference on Artificial Intelligence*, pp. 5877–5884 (2019)
65. Zhong, T., Wen, Z., Zhou, F., Trajcevski, G., Zhang, K.: Session-based recommendation via flow-based deep generative networks and bayesian inference. *Neurocomputing* (2020)
66. Zhou, F., Gao, Q., Zhang, K., Trajcevski, G., Ting, Z., Zhang, F.: Trajectory-user linking via variational autoencoder. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 3212–3218 (2018)
67. Zhou, F., Cao, C., Zhang, K., Trajcevski, G., Zhong, T., Geng, J.: Meta-gnn: on few-shot node classification in graph meta-learning. In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, pp. 2357–2360 (2019)
68. Zhou, F., Wen, Z., Trajcevski, G., Zhang, K., Zhong, T., Liu, F.: Disentangled Network Alignment with Matching Explainability. In: *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pp. 1360–1368. IEEE (2019)
69. Zhou, F., Wen, Z., Zhang, K., Trajcevski, G., Zhong, T.: Variational session-based recommendation using normalizing flows. In: *Proceedings of the International Conference on World Wide Web (WWW)*, pp. 3476–3475. ACM (2019)
70. Zhou, F., Yin, R., Trajcevski, G., Zhang, K., Wu, J., Khokhar, A.: Improving human mobility identification with trajectory augmentation. *GeoInformatica* (2019)
71. Zhou, F., Yin, R., Zhang, K., Trajcevski, G., Zhong, T., Wu, J.: Adversarial point-of-interest recommendation. In: *Proceedings of the International Conference on World Wide Web (WWW)*, pp. 3462–34618. ACM (2019)
72. Zhou, F., Yue, X., Trajcevski, G., Zhong, T., Zhang, K.: Context-aware variational trajectory encoding and human mobility inference. In: *Proceedings of the International Conference on World Wide Web (WWW)*, pp. 3469–3475. ACM (2019)
73. Zhou, F., Mo, Y., Trajcevski, G., Zhang, K., Wu, J., Zhong, T.: Recommendation via collaborative autoregressive flows. *Neural Networks* (2020)
74. Zhou, F., Yang, Q., Zhang, K., Trajcevski, G., Zhong, T., Khokhar, A.: Reinforced spatio-temporal attentive graph neural networks for traffic forecasting. *IEEE Internet of Things Journal* (2020)

Affiliations

Ting Zhong¹ · Shengming Zhang¹ · Fan Zhou¹  · Kunpeng Zhang² ·
Goce Trajcevski³ · Jin Wu¹

Ting Zhong
zhongting@uestc.edu.cn

Shengming Zhang
shmizhang@gmail.com

Kunpeng Zhang
kpzhang@umd.edu

Goce Trajcevski
gocet25@iastate.edu

Jin Wu
wj@uestc.edu.cn

- ¹ School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, 610054, China
- ² Department of Decision, Operations & Information Technologies, University of Maryland, College Park, MD 20742, USA
- ³ Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa 50011 USA