2023 IEEE International Conference on Data Mining (ICDM) | 979-8-3503-0788-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICDM58522.2023.00100

IKGN: Intention-aware Knowledge Graph Network for POI Recommendation

Xiaoyu Zhu*†, Chenyang Bu*†, Bingbing Dong*†, Shengwei Ji‡, Yi He§, Xindong Wu¶ ✉
* Key Laboratory of Knowledge Engineering with Big Data (the Ministry of Education of China), Hefei University of Technology
† School of Computer Science and Information Engineering, Hefei University of Technology, China
‡ School of Artificial Intelligence and Big Data, Hefei University, China
§ Department of Computer Science, Old Dominion University, USA
¶ Research Center for Knowledge Engineering, Zhejiang Lab, China
*†{2021111066, blingdong}@mail.hfut.edu.cn, *†chenyangbu@hfut.edu.cn, ‡jisw@hfuu.edu.cn
§yihe@cs.odu.edu, ¶ ✉ xwu@zhejianglab.com

***Abstract*—Point-of-Interest (POI) recommendation, pivotal for guiding users to their next interested locale, grapples with the persistent challenge of data sparsity. Whereas knowledge graphs (KGs) have emerged as a favored tool to mitigate the issue, existing KG-based methods tend to overlook two crucial elements: the intention steering users' location choices and the high-order topological structure within the KG. In this paper, we craft an Intention-aware Knowledge Graph (IKG) that harmonizes users' visit histories, movement trajectories, and location categories to model user intentions. Building upon IKG, our novel Intention-aware Knowledge Graph Network (IKGN) delves deeper into the POI recommendation by weighing and propagating node embeddings through an attention mechanism, capturing the unique locational intent of each user. A sequential model like GRU is then employed to ensure a comprehensive representation of users' short- and long-term location preferences. An empirical study on two real-world datasets validates the effectiveness of our proposed IKGN, with it markedly outshining seven benchmark rival models in both Recall and NDCG metrics. The code of IKGN is available at https://github.com/Jungle123456/IKGN.**

***Index Terms*—Knowledge Graph, POI Recommendation, User Intention, Higher-Order Connectivity**

I. Introduction

Point-of-Interest (POI) recommendation aspires to recommend unvisited locales to users, drawing inference from their historical check-in data [1]. At its core, a POI recommender system not only seeks to aid users in identifying relevant POIs such as shopping centers, eateries, and recreational areas, but also endeavors to forecast a sequence of POIs, thereby crafting trajectories reflective of users' evolving interests. Such an endeavor weaves spatial and temporal dimensions, where both the chronology and geographic proximity of POIs play pivotal roles. Driven by the technological progression in smart devices and mobile location-based services, users are now more inclined to share their POI experiences on platforms like Foursquare and Yelp. This interplay of innovative technology, user engagement, and spatiotemporal aspects enriches the research landscape of POI recommendation.

In the nascent phase of POI recommendation research, seminal studies leveraged traditional Collaborative Filtering and Matrix Factorization algorithms [2], [3], dealing with static POI matrices only. To respect and capture the temporal patterns of user interests, later studies employed sequential models such as Markov chains [4] and Recurrent Neural Networks (RNNs), laying the groundwork for a series of subsequent innovations [5]–[10]. These studies unfurled RNN variants adept at extracting long- and short-term user trajectory features with various memory mechanisms, further bolstering POI recommendation efficacy. Despite so, the persistent issue of *data sparsity* remains a prominent challenge.

To mitigate this challenge, knowledge graphs (KGs) [11], [12] with extensive semantic contents has been incorporated into POI recommender systems to provide auxiliary information [13]–[16]. To wit, [13] proposed a meta-path-based random walk algorithm to identify superior neighbors on KGs, exploiting the topological structure to better the recommendation accuracy. Another thrust of research focused on learning KG embeddings [14]–[16], where KGs were trained using models such as TransR [17] to derive embedding representations of users and locations, which were then fed into a Gated Recurrent Unit (GRU) network for sequential modeling. Nevertheless, while these methods demonstrated performance enhancements, we argue that current KG-based methods for POI recommendation suffer two notable limitations: 1) the lack of consideration for the underlying reasons why users frequently visit specific location categories and 2) an oversight of high-order connectivity inherent to the knowledge graphs.

In response to the highlighted issues, we introduce an Intention-aware Knowledge Graph (IKG) that captures location categories intertwined with spatiotemporal user information. Our IKG emerges from a rich tapestry of user visit histories, movement trajectories, and location category data, as illustrated in Fig. 1. Central to our discourse is the concept of *intention*, which we delineate as the underlying reason steering users' location preferences. It is intuitive that the nuanced intentions underpin varied location preferences; consider, for instance, while User A might predominantly visit parks during weekends, User B may have a preference for cafés after work hours on weekdays. This diversity in choices suggests that

✉ Corresponding Author: Dr. Xindong Wu (xwu@zhejianglab.com)

979-8-3503-0788-7/23/$31.00 ©2023 IEEE
DOI 10.1109/ICDM58522.2023.00100

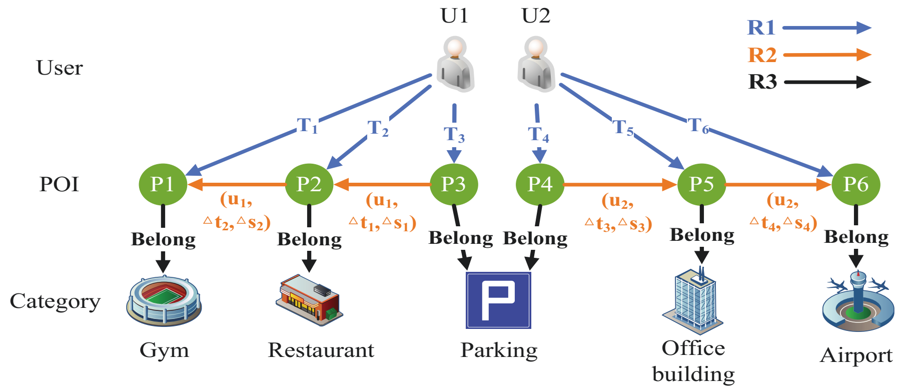

Fig. 1. An example of IKG. IKG includes three types of nodes and three types of relationships. Specifically, the three types of nodes are user nodes (U), POI nodes (P), and location category nodes (C), while the three types of relationships are user visit relationships (R_1), spatio-temporal cost relationships (R_2), and category correspondence relationships (R_3). The specific construction process of IKG will be detailed in Sec IV-A.

user intentions are intricately woven into the very fabric of the location categories they frequent. Upon this intuition, we propose to integrate IKG into POI recommendation through an end-to-end neural network model, coined *Intention-aware Knowledge Graph Network* (IKGN), which strives to capture the fine-grained user-location dynamics. Specifically, we leverage TransR [17] to initialize the embeddings of nodes and their relations. To capture the personalized intention of a user towards a certain location category, we tailor a knowledge-aware attention mechanism to further learn the node and relation embeddings via recursive propagation and aggregation of neighboring nodes, where the weight of each node neighbor is automatically learned during the learning process. Such learned embeddings encapsulate rich insights on users and locations, which are fed into a GRU network [14] for identifying both short- and long-term user intentions.

Specific contributions of this paper are as follows:

- This paper explores a novel POI recommendation problem by taking into account the user *intention*, striving to model the spatiotemporal dynamics between users and locations, aiming to discern the underlying reasons driving users to repeatedly opt for certain location categories.
- We propose an end-to-end IKGN model, devised around the Intention-aware Knowledge Graph (IKG). Our model adeptly captures high-order entity dependencies within the IKG, alongside user intentions. Using a knowledge graph-aware attention mechanism, the IKGN discerns the relative importance of neighboring node information and optimizes user and location representations, thereby enhancing the POI recommendation performance.
- Extensive experiments on two widely-used real-world POI benchmarks substantiate that our proposal markedly outperforms the state-of-the-art competitors.

II. Related Work

A. Knowledge Graph for Recommendation

Recommendation methods based on knowledge graphs (KG) can generally be divided into three types: embedding-based methods, path-based methods, and propagation-based methods [18]. The embedding-based methods involve using translation models such as TransE [19], TransH [20], and TransR [17] to learn vector representations of entities and relationships. The CKE method [21] utilizes TransE in KG to obtain entity embeddings, and then these embeddings are inputted into matrix factorization (FM) to enhance the item representation. On the other hand, the path-based methods involve mining various connection relationships between users and items based on the graph. For instance, the KPRN model [22] combines the semantics of entities and relationships to infer user preferences based on the order dependency in the paths, thus generating path representations. In recent years, propagation-based methods have achieved better results, which learn the representations of users and items combined with the above two methods. For example, KGAT [12] combines user-item interactions and knowledge graphs into a heterogeneous graph and then applies aggregation mechanisms to it. CKAN [18] uses two different strategies to propagate collaborative signals and knowledge-aware signals. KGIN [23] achieves better relationship modeling by identifying user intention and relational path-aware aggregation in two dimensions. MetaKG [24] proposes a new framework based on meta-learning to capture user preferences and entity knowledge to alleviate cold start recommendation problems.

B. POI Recommendation

As a frequently utilized approach for sequence prediction, initial research on recommending the next point of interest (POI) primarily relied on collaborative filtering and matrix factorization models [25]–[28]. Nonetheless, the dynamic changes of user preferences has impeded the practical implementation of such models. In recent years, Recurrent Neural Network (RNN) has been the foundation for most next POI recommendation models. STRNN [9] is an example of a model that captures local spatiotemporal context to improve the model's performance. Time-LSTM [8] adds a time gate to the LSTM structure to model the temporal and spatial

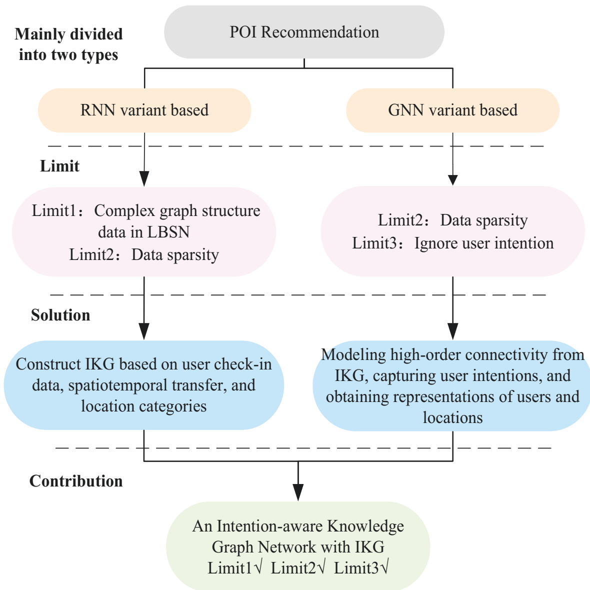

Fig. 2. Motivations and contributions of the proposed model.

intervals of user transitions. The DeepMove [5] model is a multimodal RNN model that combines sparse user features and attention modules to capture user mobility patterns. LSPTM [7] proposes a new model that includes a non-local network for modeling long-term preferences and a geographically extended RNN for learning short-term preferences. PLSPL [29] proposes a unified model for joint learning of users' preferences, which combines long-term and short-term preferences based on users' linear combination units to learn the personalized weights of different users in different parts. STAN [30] proposes a spatiotemporal attention network for location recommendation, which uses a dual-attention system to learn explicit spatiotemporal correlations within a user's trajectory. DRCF [31] is a deep recursive collaborative filtering framework with a paired ranking function. It utilizes multi-layer perceptron and recursive neural network architectures to capture user-place interactions in a CF manner from observed feedback sequences. ATCA-GRU [32] is an attention-based GRU model for the next POI category recommendation, which reduces the negative impact of sparse check-in data, captures long-distance dependencies between user check-ins, and selectively utilizes attention mechanisms to focus on relevant historical check-in trajectories in check-in sequences.

In the field of POI recommendation, a small amount of research has combined knowledge graphs to improve recommendation performance. SKGEM [33] is a new Scalable Knowledge Graph Embedding Model that aims to automatically learn low dimensional node representations to formularize and merge all heterogeneous factors into a context oriented graph, and to some extent solve the cold start problem for the next POI recommendation. ARNN [13] uses the external attributes of users and POIs to construct a knowledge graph and designs a random walk process on the KG based on meta-paths to effectively discover neighbors based on geographic, semantic, and user preference factors. The recent study, STKGRec [14], constructs a spatial-temporal knowledge graph (STKG) based on user movement trajectories and models long-term and short-term preferences based on STKG. However, these methods do not consider user intention and do not pay enough attention to the importance of high-order propagation in KG, so they fail to fully utilize the rich semantic information in KG as illustrated in Fig. 2. In this work, we propose a new method to construct the IKG based on users' spatiotemporal movement behavior and POI location categories. We use a knowledge-aware attention mechanism to obtain node neighbors' information and model high-order entity dependencies to capture user intention for location categories. Based on this, We update node representations of users and locations and model long-term and short-term preferences of users to recommend the next POI that they may be interested in.

III. Problem Formulation

In this section, we will introduce the notation definitions used in this paper, as well as the basic problem statement for the Next POI recommendation. Let U=$\{u_1, u_2, \ldots, u_{|U|}\}$ denote the set of users, P=$\{p_1, p_2, \ldots, p_{|P|}\}$ denote the set of POIs, C=$\{c_1, c_2, \ldots, c_{|C|}\}$ denote the set of location categories, where there is a many-to-many relationship between U and P, and a many-to-one relationship between P and C.

a) Definition 3.1: User Trajectory Sequence. For each user u_i in U, there is a corresponding trajectory sequence S^i=$\{s_1^i, s_2^i, \ldots, s_{n-1}^i, s_n^i\}$, where s_n^i is the current trajectory sequence and the others are historical trajectory sequences. For each s in S, it consists of a consecutive sequence of POI locations that are visited by the user u. i.e., s^i=$\{p_1, p_2, \ldots p_{|s^i|}\}$.

b) Definition 3.2: Intention-aware Knowledge Graph (IKG). IKG: $G = \{(h, r, t)|h, t \in \varepsilon, r \in R\}$, where ε and R are sets of entities and relations, respectively.

c) Definition 3.3: Problem Statement. The Next POI recommendation can be defined: for user u_i and a given trajectory sequence S_i, where the current trajectory s_n^i=$\{p_1, p_2, \ldots, p_{t-1}\}$($s_n^i \in S^i$) , p_{t-1} is the most recently visited POI by u_i and the goal is to recommend the top-N POIs to user u_i at the next time step t.

IV. Method

In this section, we provide a detailed description of how to construct IKG and our proposed model, IKGN. IKGN employs IKG in an end-to-end manner to achieve the next POI recommendation. Fig. 3 illustrates that the overall framework of IKGN consists of three primary components: 1) the knowledge graph embedding layer, which converts entities and relationships in IKG into embedding representations; 2) the intention-aware aggregation layer, which models higher-order entity dependencies, recursively propagates embeddings from nodes' neighbors, and uses a knowledge-aware attention mechanism to learn the weights of each neighbor during propagation to capture the user's intention towards location categories, ultimately optimizing and updating node representations; 3) the long-term and short-term preference modeling and prediction layer, which uses the latest node and relation representations to explore users' long-term and short-term preferences from their complete check-in sequences, ultimately merging vector representations of different user preferences and outputting prediction scores.

A. Knowledge Graph Construction Layer

In this paper, we present the IKG, which incorporates semantic information from location categories based on the user-POI interaction bipartite graph and considers users' visit histories and spatio-temporal transition behavior of users to extend the knowledge graph. IKG: $G = \{(h, r, t)|h, t \in \varepsilon, r \in R\}$. ε and R can be abstracted into three types of nodes and three types of relationships. Specifically, the three types of nodes are user nodes (U), POI nodes (P), and location category nodes (C), while the three types of relationships are user visit relationships (R_1), spatio-temporal cost relationships (R_2), and category correspondence relationships (R_3). The user visit relationship (R_1) and spatio-temporal transition relationship (R_2) are extracted from the user trajectory sequence. R_1 is established between U and P nodes and contains time attributes, indicating that a user visited a POI location at a certain time. R_2 appears between P_i and P_j nodes and contains user mapping as well as time and space change attributes, indicating the time and space cost spent by a user from P_i to P_j. Lastly, R_3 appears between P and C nodes and does not contain attributes, only indicating that P belongs to location category C, and the relationship between P and C is many-to-one.

Thus, the triples in IKG are classified into three types, specifically $P_1 = (U, T_i, P)$, $P_2 = (P_i, (U_i, \Delta S, \Delta T), P_j)$, and $P_3 = (P, Belong, C)$, where U, P, $C \in \varepsilon$, and T_i, $(U, \Delta S, \Delta T)$, $Belong \in R$. The user's movement trajectory in IKG is considered as a combination of T_1 and T_2, while the type of location that the user frequently visits can be observed from T_3. By integrating knowledge into the user's movement trajectory, the IKGN framework can identify the user's intention toward certain types of locations, which can aid in improving the results of the Next POI recommendation.

It is worth noting that this study follows the methodology presented in reference [7] which divides time into 48 time periods. The first 24 time periods correspond to the 24 hours of weekdays, while the last 24 time periods correspond to the 24 hours of weekends, without distinguishing holidays.

B. Knowledge Graph Embedding Layer

The knowledge graph embedding layer serves to parameterize entities and relationships into vector representations. In this study, we employ TransR [17] on IKG for this purpose. Specifically, given a triple (h, r, t) in IKG, the rationality score for the triple is computed using the following formula [17]:

$$g(h, r, t) = ||W_r e_h + e_r - W_r e_t||_2^2, \tag{1}$$

where the matrix $W_r \in R^{k*d}$ serves as the transformation matrix for relation r. A lower score for a triple (h, r, t) signifies a higher likelihood for the triple to be true and vice versa. During training, TransR considers the relative order between valid and corrupted triples, promoting their differentiation via pairwise ranking loss [17], formatted as follows.

$$\mathcal{L}_{IKG} = \sum_{(h,r,t,t') \in \Gamma} -ln\sigma \left(g(h, r, t') - g(h, r, t) \right), \tag{2}$$

Where $\Gamma = \{(h, r, t, t') \mid (h, r, t) \in G, \ (h, r, t') \notin G\}$ and (h, r, t') is a corrupted triplet constructed by replacing one entity in a valid triplet at random. Additionally, $\sigma(\cdot)$ represents the sigmoid function. This layer models entities and relations as their initial embedding representation. The reason for selecting TransR from the pool of crowdsourced embedding models lies in its incorporation of entity-relationship constraints. In TransR, every triplet consisting of a head entity, relationship, and tail entity necessitates a harmonized representation across both entity space and relationship space. By imposing such constraints, TransR facilitates the acquisition of precise embedded representations and enhances the predictive capability of relationships within the knowledge graph.

C. Intention-aware Aggregation Layer

The intention-aware aggregation layer aims to capture the correlation between POIs and users' intentions by mining the location category attributes on the IKG. This module recursively propagates embeddings from neighboring nodes, modeling the high-order connectivity of IKG. While performing the propagation, the concept of graph attention networks [34] is integrated to generate attention weights for cascade propagation. This integration helps differentiate the impact of various location categories on users. Furthermore, it empowers the capability to acquire precise embedding representations for both users and locations

As an example, user u_1 visited p_1, p_2, p_3, and p_4, where p_1 and p_2 belong to c_3, p_3 and p_4 belong to c_1 ($c_1, c_3 \in C$). The POI nodes take location categories as input to enrich their characteristics and simulate the intention of user u_1 by propagating information from c_1 and c_3 to u_1.

Regarding the entity h in IKG, it is necessary to define its neighborhood $Q_h = \{(h, r, t)|(h, r, t) \in G\}$. Here, h represents the head entity in the IKG, and Q_h denotes the neighborhood of h. The linear combination of h's neighbors can be defined through the following equation:

$$e_{Q_h} = \sum_{(h,r,t,) \in Q_h} \partial(h, r, t) e_t, \tag{3}$$

Where $\partial(h, r, t)$ is a controlling factor for information propagation, the amount of information transmitted from node t to h is regulated based on the relationship r. $\partial(h, r, t)$ can be expressed as follows [12]:

$$\partial(h, r, t) = (W_r e_t)^T tanh(W_r e_h + e_r). \tag{4}$$

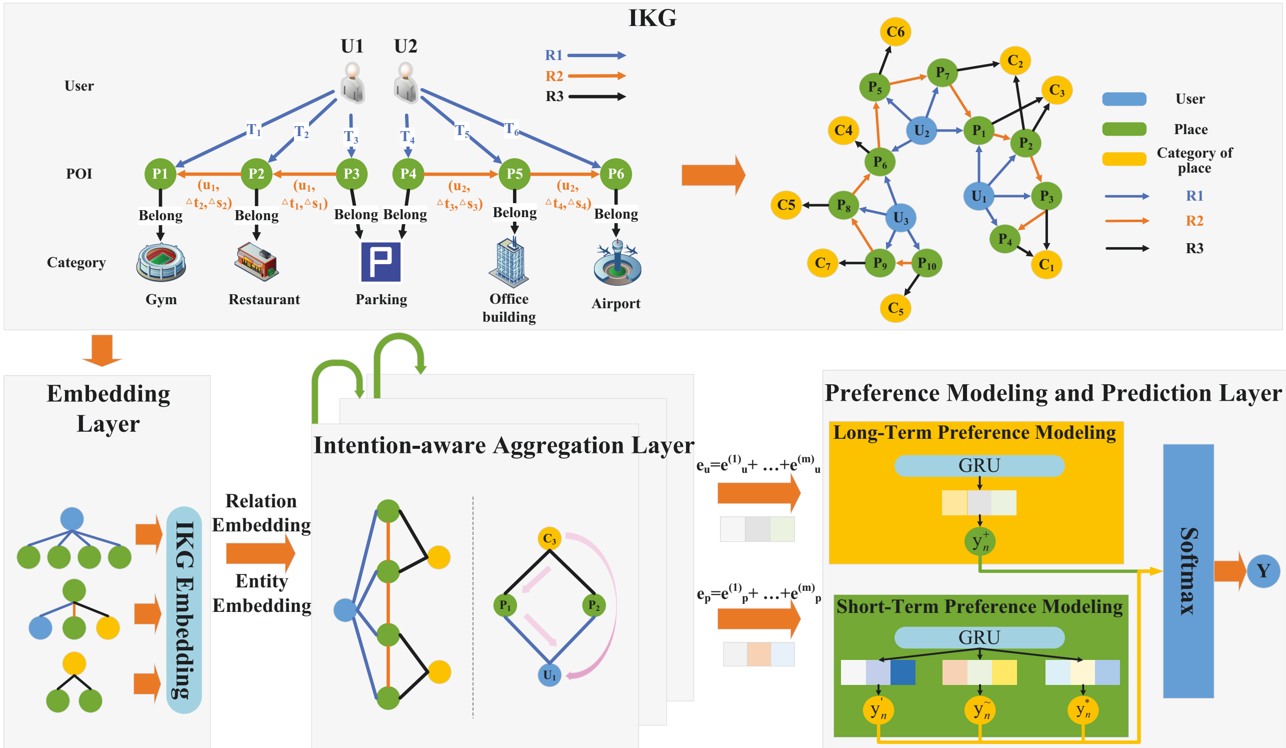

Fig. 3. The architecture of the IKGN model. The first part on the left represents an example of IKG, and the three parts on the right form the IKGN model, which consists of three components: knowledge graph embedding layer, intention-aware aggregation layer, and preference modeling and prediction layer.

Thereafter, the controlling factor is normalized using the softmax function, formulated as follows.

$$\partial(h,r,t)=\frac{exp(\partial(h,r,t))}{\sum_{(h,r',t',)\in Q_h} exp(\partial(h,r',t'))}. \tag{5}$$

This normalization process allows the final attention score to determine which neighboring nodes should receive more attention to capture user intention signals. Subsequently, we use e_{Q_h} and e_h to update the high-order representation of entity h. The embedding representation of entity h at layer m is recursively defined as follows:

$$e_h^{(m)}=f(e_h^{(m-1)},e_{Q_h}^{(m-1)}). \tag{6}$$

The definition of $e_{Q_h}^{(m-1)}$is based on Equation (3), while the aggregation function f [12] acts as an aggregator to combine the linear combination of h and its neighboring values. Specifically, this aggregator is defined as follows:

$$\begin{aligned} f =&LeakyReLU(W_1(e_h+e_{N_h})) \\ &+LeakyReLU(W_2(e_h\odot e_{N_h})), \end{aligned} \tag{7}$$

where $W_1, W_2 \in R^{d' * d}$ are trainable weight matrices and $\odot$ denotes element-wise multiplication.

As a result, the embedding propagation process captures propagation paths that are similar to $u_1 - p_1 - c_3$. Furthermore, $e_{u_1}^{(2)}$ explicitly encodes information from c_3, clearly, high-order embedding propagation injects user intention into the representation learning process. To facilitate propagation, we treat the relationships in IKG as bidirectional.

After m layers of propagation, we obtain multiple representations of user u and location p from e_h, which are denoted as $\{e_u^{(1)}, e_u^{(2)} \ldots, e_u^{(m)}\}$ and $\{e_p^{(1)}, e_p^{(2)}, \ldots, e_p^{(m)}\}$, respectively. The final representations of the user and location are defined as follows and are used as input for the preference modeling and prediction layer:

$$e_u = e_u^{(1)} + e_u^{(2)} + \ldots + e_u^{(m)}, \tag{8}$$

$$e_p = e_p^{(1)} + e_p^{(2)} + \ldots + e_p^{(m)}. \tag{9}$$

D. Preference Modeling and Prediction Layer

The backbone of the module is modified by the work of [14] with an innovative approach. We integrate intention-aware aggregation to acquire user and location vectors in order to capture user intention and collaboration signals. This method produces more accurate representations and inputs them into the long-term and short-term preference modeling module of [14], ultimately enhancing the final recommendation results.

After modeling both long-term and short-term preferences [14] in the layer, we derive the long-term preference representation y_n^+, the current sequence representation $y_n^\sim$, the representation of continuous spatial-temporal movement

y_n^*, and the representation of non-continuous spatial-temporal movement $y_n^{'}$. Using this information, we calculate the probability distribution for predicting POIs as follows [14]:

$$\hat{Y} = softmax(W_P[y_n^{+}||y_n^{\sim}||y_n^{*}||y_n^{'}]), \tag{10}$$

where $||$ represents a concatenation operation, while W_P is a trainable matrix for all POIs. Therefore, user intends to visit the POI with highest probability at time t.

To optimize the parameters of the recommendation model, we employ the logarithmic likelihood function as follows:

$$\mathcal{L}_{POI-Intention} = -\sum_{k=1}^{N} log\left(\hat{Y}_k\right). \tag{11}$$

Finally, we represent the target loss function using Equations (2) and (11):

$$\mathcal{L}_{loss} = \mathcal{L}_{IKG} + \mathcal{L}_{POI-Intention}. \tag{12}$$

Algorithm 1 provides a comprehensive overview of the detailed process employed in IKGN. During the training phase, we adopt an alternating optimization approach, optimizing $\mathcal{L}_{IKG}$ and $\mathcal{L}_{POI-Intention}$ iteratively, for optimizing both the embedding loss and the recommendation loss of IKGN.

Algorithm 1: IKGN

Input: Intention-aware Knowledge Graph G, Trajectory Sequence S, STKGRec Algorithm [14]
Output: prediction Result $\hat{Y}$

while *IKGN Not Converge* **do**
  for $triple \in G$ **do**
    Calculate triple score $\leftarrow$ use Equations (1) and (2);
  get e^h, e^r, e^t;
  for $h, t \in \varepsilon$ **do**
    $e_h^{(0)} \leftarrow e^h$, $e_t^{(0)} \leftarrow e^t$;
    for $m \leftarrow 1$ **to** M **do**
      $e_{Q_h} \leftarrow$ use Equations (3), (4), and (5);
      $e_h^{(m)}, e_t^{(m)} \leftarrow$ use Equations (6) and (7);
  $e_u \leftarrow$ use Equation (8);
  $e_p \leftarrow$ use Equation (9);
  y_n^{+}, $y_n^{\sim}$, y_n^{*}, $y_n^{'}$ $\leftarrow$ use STKGRec(e_u, e_p, S) [14];
  $\hat{Y} \leftarrow$ use Equation (10);
  return $\hat{Y}$;
STKGRec(e_u, e_p, S)
for $l \in S$ **do**
  $y_n^{+} \leftarrow$ Long-term Preference Modeling;
  $y_n^{\sim}$, y_n^{*}, $y_n^{'}$ $\leftarrow$ Short-term Preference Modeling;
return y_n^{+}, $y_n^{\sim}$, y_n^{*}, $y_n^{'}$;

V. Experiments

In this section, we conduct experiments on two real-world datasets of Foursquare check-in data to assess the efficacy of our proposed IKGN model. We aim to answer the following research questions:

- **RQ1**: How does IKGN perform compared to the state-of-the-art POI recommendation methods?
- **RQ2**: How do key components affect IKGN?
- **RQ3**: How does IKGN perform on sparser datasets?
- **RQ4**: How do different hyper-parameter settings (depth of the intention-aware aggregation layer and embedding size) affect IKGN?

A. Datasets and Preprocessing

We conduct an evaluation of our model on publicly available Foursquare check-in datasets[1] from two real-world cities: New York City (NYC) and Tokyo (TKY). These datasets have been widely utilized in related research papers. The records from these datasets are collected between April 2012 and February 2013 and include various data points, such as ID, POI ID, category ID, category name, latitude and longitude coordinates, and timestamps. To improve the quality of the dataset, we remove POIs that are visited less than 10 times in these two datasets and only include sessions with at least three check-ins. Users with less than five sessions are removed from our analysis. We then split the data into training, validation, and test sets, with the first 80% of each user's sessions used as the training set, the most recent 10% as the test set, and the remaining 10% as the validation set for hyperparameter tuning. In addition, we also extracted small-scale datasets from four countries on Foursquare check-in datasets, namely China (CN), Greece (GR), Panama (PA), and Paraguay (PY). These datasets are more sparse, and used to verify IKGN's ability to mitigate the impact of data sparsity on POI recommendations. We performed the same processing on these four datasets as described above. Table I provides a summary of the dataset's statistical information.

B. Methods for Comparison

To demonstrate the effectiveness of the model we proposed, we compare IKGN with the following baselines:

- STRNN [9] is presented as a recursive neural network that models spatio-temporal context based on time-specific and distance-specific transformation matrices. This model takes advantage of the inherent spatio-temporal dependencies in the data and a linear interpolation is applied for the training of transition matrices.
- DeepMove [5] is described as a model that uses a recursive neural network to learn user preferences from historical and current sequences. This model employs an attention mechanism to calculate the similarity between the current and historical states, allowing for more accurate preference modeling.

[1] https://sites.google.com/site/yangdingqi/home/foursquare-dataset

TABLE I
SUMMARY OF DATASET STATISTICS.

Dataset	User	POI	Category	Session
NYC	1020	14085	400	18459
TKY	2232	21139	385	50848
CN	354	5523	397	5440
GR	570	5642	387	8477
PA	369	2661	334	6086
PY	705	5118	386	9858

- STGN [35] is presented as a model that integrates time and distance intervals by adding gates, enabling it to capture the spatio-temporal context of the data.
- PLSPL [29] proposes a unified model for jointly learning users' long-term and short-term preferences. This model uses linear combination units to combine long-term and short-term preferences, enabling it to learn personalized weights for different users in different parts.
- LSPTM [7] is introduced as a model that integrates spatio-temporal context features into the RNN framework to model user mobility between Points of Interest (POIs).
- STAN [30] proposes a spatio-temporal attention network for location recommendation. This model uses a dual-attention system to learn explicit spatio-temporal correlations within trajectories.
- STKGRec [14] constructs a spatio-temporal knowledge graph (STKG) to capture users' transition patterns in trajectories. This model uses historical trajectory encoding module and knowledge graph path reasoning to capture users' long-term and short-term preferences.

C. Parameter Settings

The IKGN model is implemented in PyTorch[2], with an embedding size of 130. The Adam optimization algorithm is employed to optimize all model parameters, using a batch size of 128. For the TKY dataset, the depth of the intention-aware aggregation layer is set to 2, while the corresponding value for the NYC dataset is 3. To identify the most optimal learning rate, a grid search is conducted across a set of learning rates, specifically 0.00005, 0.0001, and 0.0005.

D. Analysis on Recommendation Effectiveness (RQ1)

The overall performance comparison of different models on the NYC and TKY datasets is shown in Table II and Fig. 4. We adopt two widely-used evaluation metrics [36], [37]: $R@K$ ($recall@K$) and $N@K$ ($ndcg@K$). Based on our experimental results, we have drawn the following conclusions:

- The IKGN model outperforms the compared methods on all metrics across two datasets. Taking the TKY dataset as an example, considering a class of works using recurrent neural networks, LSPTM [7] achieves optimal performance. However, compared with LSPTM, our model have achieved 29% higher R@k and 40% higher N@k. Similarly, the strongest baseline, STKGRec [14] integrates knowledge graphs and recurrent neural networks, but our model still outperforms it by 10.8% in terms of R@K, and 13.4% in terms of N@K.
- Our model and the strongest baseline, STKGRec [14], exhibit a significant improvement over other models. This may be attributed to the fact that both models incorporate knowledge graphs. Unlike the general user-item bipartite graph, knowledge graphs provide a richer source of semantic information, thereby mitigating issues related to data sparsity and improving experimental results.
- Our method outperforms STKGRec [14] possibly because we not only captured users' fine-grained intention for location categories but also learned high-order representations of users and items. This was demonstrated in subsequent ablation experiments, as shown in Sec V-E.

[2]https://pytorch.org

E. Analysis on Key Components (RQ2)

To confirm the impact of individual components within IKGN on enhancing performance, we execute two simplified versions of the model for ablation testing:

- IKGN-IA: This version has removed the location category nodes from the ontology of IKG while retaining the embedding layer, intention-aware aggregation layer, and preference modeling layer.
- IKGN-C: This version is a modified version of IKGN that retains the embedding representation layer and preference modeling layer but eliminates the intention-aware aggregation layer present in IKGN.

Table III presents the results of diverse versions of IKGN on TKY datasets. Our ablation testing enables us to discover the following observations:

- In the case of IKGN-C, some indicators have slightly lower performance than the strongest baseline. The reason may be that although this version has introduced location category information in the graph, it lacks an intention-aware aggregation layer and does not fully capture the user's intention.
- The performance of IKGN-IA is generally better than IKGN-C, but lower than IKGN. The reason may be that although the KG used in this version does not have category information, it still captures collaborative signals and obtains richer node representations of users and POIs through athen intention-aware aggregation layer.
- The IKGN model is a combination of the IKGN-C and IKGN-IA models, which demonstrates the best performance on two datasets. This indicates that enriching the knowledge graph with location category information and capturing user intention based on it have a positive impact on the next POI recommendation.

F. Analysis on Recommendation Effectiveness with Sparser Datasets (RQ3)

To confirm the effectiveness of IKGN in alleviating data sparsity issues, we conduct comparative experiments with the best baseline (STKGRec) [14] on four smaller and more sparse datasets, and the results are shown in Fig. 5.

TABLE II
THE PERFORMANCE COMPARISON ON NYC AND TKY. THE BEST RESULT IN EACH COLUMN IS INDICATED IN BOLD, WHILE THE SECOND BEST RESULT IS UNDERLINED.

Dataset	Metric	STRNN [9]	DeepMove [5]	STGN [35]	PLSPL [29]	LSPTM [7]	STAN [30]	STKGRec [14]	IKGN
NYC	R@1	0.0921	0.1189	0.0868	0.1171	0.155	0.1123	<u>0.1764</u>	**0.1823**
	R@5	0.1935	0.2511	0.1495	0.2916	0.3683	0.3318	<u>0.3927</u>	**0.3945**
	R@10	0.235	0.3062	0.1758	0.3558	0.4545	0.4726	<u>0.4773</u>	**0.4810**
	N@1	0.092	0.1186	0.0868	0.1171	0.1555	0.1123	<u>0.1764</u>	**0.1823**
	N@5	0.1459	0.189	0.1204	0.2028	0.2665	0.2274	<u>0.2901</u>	**0.2938**
	N@10	0.1593	0.206	0.1289	0.23	0.2946	0.2695	<u>0.3176</u>	**0.3219**
TKY	R@1	0.1229	0.147	0.1222	0.1278	0.1612	0.1123	<u>0.224</u>	**0.2635**
	R@5	0.2620	0.3066	0.2186	0.3105	0.3608	0.3418	<u>0.4142</u>	**0.4552**
	R@10	0.316	0.3691	0.2612	0.3808	0.4391	0.4512	<u>0.4834</u>	**0.5241**
	N@1	0.1229	0.147	0.1229	0.1278	0.1612	0.1123	<u>0.224</u>	**0.2635**
	N@5	0.1963	0.2314	0.1732	0.223	0.2660	0.2222	<u>0.325</u>	**0.3655**
	N@10	0.2137	0.2517	0.1869	0.2456	0.2914	0.2576	<u>0.3471</u>	**0.3879**

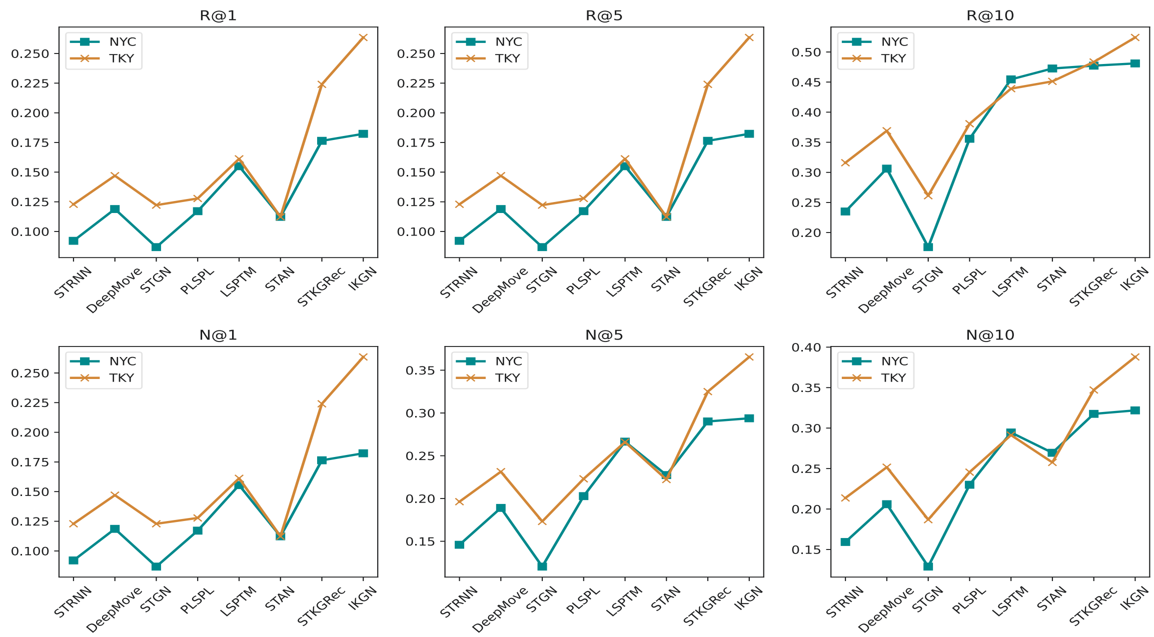

Fig. 4. The performance comparison on NYC and TKY with all metrics.

Compared to the best baseline [14], IKGN has demonstrated improvements across all metrics on the four datasets. This enhancement can be attributed to the richer semantic and contextual information provided by IKG. Furthermore, our method incorporates an intention-aware aggregation layer that effectively combines multiple layers of nodes within the IKG. This process enhances the representation of nodes and ultimately alleviates the issue of data sparsity.

TABLE III
THE PERFORMANCE OF DIFFERENT IKGN VARIANTS ON TKY.

Dataset	Metric	Best Baseline	IKGN	IKGN-IA	IKGN-C
TKY	R@1	0.224	**0.2635**	0.2441	0.2271
	R@5	0.4142	**0.4552**	0.4393	0.4142
	R@10	0.4834	**0.5241**	0.5125	0.4822
	N@1	0.224	**0.2635**	0.2441	0.2271
	N@5	0.325	**0.3655**	0.3481	0.3266
	N@10	0.3471	**0.3879**	0.3718	0.3486

G. Analysis on Hyper-parameter Setting (RQ4)

a) Effect of Intention-aware Aggregation Layer Depth: We vary the depth of the intention-aware Aggregation Layer to investigate the efficiency of multiple aggregation layers. In particular, the layer number is searched in the range of 1, 2, 3, 4; we use IKGN-1 to indicate the model with one layer and similar notations for others. We summarize the results in Table IV, and have the following observations:

TABLE IV
THE PERFORMANCE OF DIFFERENT AGGREGATION LAYERS ON TKY.

Dataset	Metric	IKGN-1	IKGN-2	IKGN-3	IKGN-4
TKY	R@1	0.2544	**0.2635**	0.2597	0.2563
	R@5	0.4451	**0.4552**	0.4526	0.4517
	R@10	0.5143	0.5241	**0.5246**	0.5214
	N@1	0.2544	**0.2635**	0.2597	0.2563
	N@5	0.3652	**0.3655**	0.3623	0.3610
	N@10	0.3787	**0.3879**	0.3857	0.3836

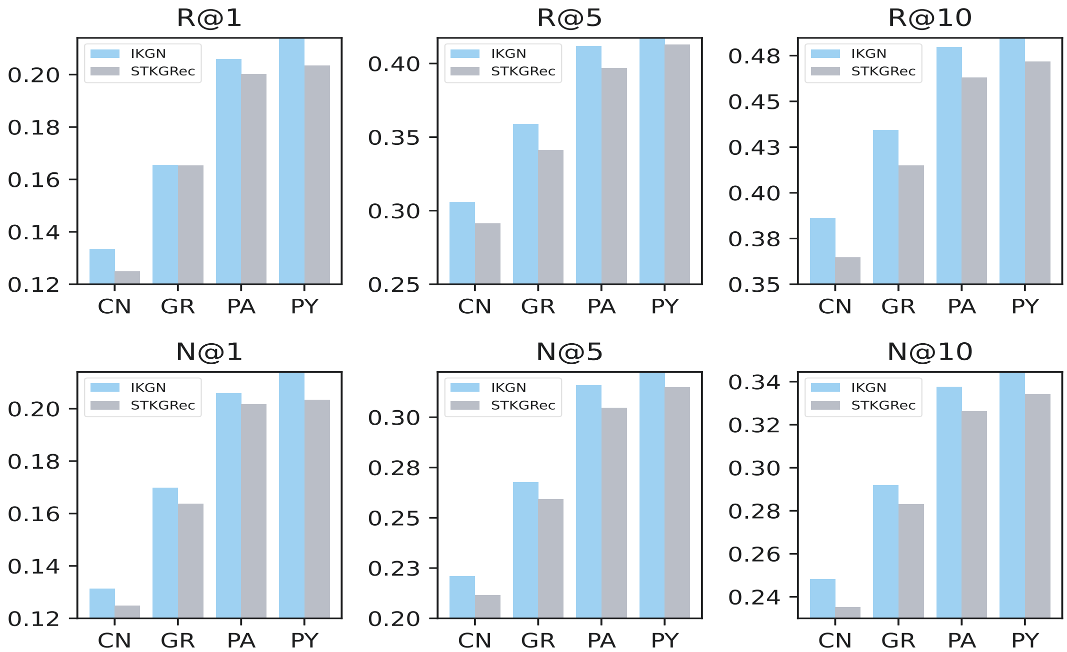

Fig. 5. The results compared to the best baseline on four datasets with all metrics.

- The performance of the model can be improved by increasing the depth of IKGN. Specifically, IKGN-2 and IKGN-3 demonstrate improvements over IKGN-1 in most metrics, which we attribute to their effective modeling of higher-order relationships that capture user intention.
- Upon comparing Table II and Table IV collectively, it is evident that IKGN-1 consistently outperforms other baseline methods. This further validates the effectiveness of knowledge propagation through better modeling of first-order relationships.
- All metrics of IKGN-4 are comparatively weaker than those of IKGN-3. We argue that although deeper propagation layers can integrate information from more faraway connections into node representations, it also introduces noise. Therefore, maintaining a reasonable depth of propagation layers can better capture user intention.

b) Effect of Embedding Size: We examine the impact of embedding dimensions on IKGN performance. Experiments are conducted on the NYC dataset with different embedding sizes, and the results are shown in Fig. 6.

Initially, increasing the embedding dimensions significantly improves performance, as larger vector dimensions can encode more entity and relationship information. When the dimensions reach a certain value, the model tends to stabilize. Similar trends are observed in the TKY dataset.

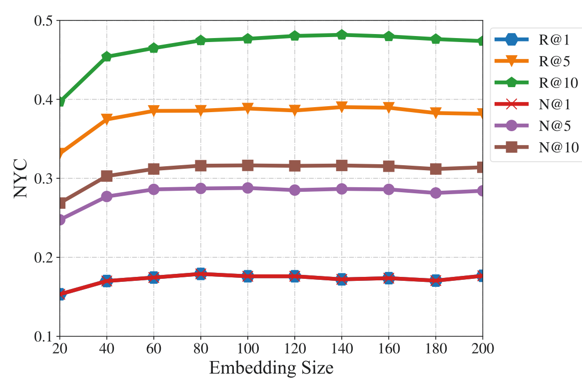

Fig. 6. The results of NYC dateset with different embedding size.

VI. Conclusion and Future Work

This paper introduced a novel IKGN model for the next POI recommendation, with its crux lying in an intention-aware knowledge graph enriched with location categories and spatiotemporal information. Specifically, IKGN comprises three main building blocks, namely the knowledge embedding layer, the intention-aware aggregation layer, and the preference modeling and prediction layer. This design elucidates user intentions within the knowledge graph, seamlessly extending it through high-order user and location representations. Extensive experimental results substantiate that our proposed IKGN outperforms seven state-of-the-art methods in terms of Recall and NDCG metrics on two real-world datasets.

In the future, we will explore two research directions. First, while this study centered on intention propagation, the innate subtlety of user intentions offers room for more explict extractions from frequently visited POI categories. Second, constructing a multimodal KG, which assimilates textual comments, images, and videos related to POIs, represents an exciting frontier for future POI recommendation research.

ACKNOWLEDGEMENT

This work is supported in part by the National Natural Science Foundation of China under Grants 62120106008 and 61806065, the Fundamental Research Funds for the Central Universities (No. JZ2022HGTB0239), the National Science Foundation under Grants IIS-2245946 and IIS-2236578, and the Open Fund of Infrared and Low Temperature Plasma Key Laboratory of Anhui Province, NUDT under grants IRKL2022KF06 and KYWX2023002.

REFERENCES

[1] M. Ye, P. Yin, W.C. Lee and D.L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," Proceedings of the International Conference on Research on Development in Information Retrieval, pp. 325-334, 2011.

[2] H. Huang and G. Gartner, "Using trajectories for collaborative filtering-based POI recommendation," International Journal of Data Mining, Modelling and Management, vol. 6, no. 4, pp. 333-346, 2014.

[3] Y. Koren, R. Bell and C. Volinsky, "Matrix factorization techniques for recommender systems," Computer, vol. 42, no. 8, pp. 30-37, 2009.

[4] C. Cheng, H. Yang, M.R. Lyu and I. King, "Where you like to go next: Successive point-of-interest recommendation," Proceedings of the 23th International Joint Conference on Artificial Intelligence, pp. 2605-2611, 2013.

[5] J. Feng, Y. Li and C. Zhang et al, "Deepmove: Predicting human mobility with attentional recurrent networks," Proceedings of the Web Conference, pp. 1459-1468, 2018.

[6] B. Hidasi, A. Karatzoglou, L. Baltrunas and D. Tikk, "Session-based recommendations with recurrent neural networks," arXiv preprint arXiv:1511.06939, 2015.

[7] K. Sun, T. Qian and T. Chen et al, "Where to go next: Modeling long- and short-term user preferences for point-of-interest recommendation," Proceedings of the 34th AAAI Conference on Artificial Intelligence, pp. 214-221, 2020.

[8] Y. Zhu, H. Li and Y, Liao et al, "What to Do Next: Modeling User Behaviors by Time-LSTM," Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp. 3602-3608, 2017.

[9] Q. Liu, S. Wu, L. Wang and T. Tan, "Predicting the next location: A recurrent model with spatial and temporal contexts," Proceedings of the 30th AAAI Conference on Artificial Intelligence, pp. 194-200, 2016.

[10] J. Manotumruksa, C. Macdonald and I. Ounis, "A contextual attention recurrent architecture for context-aware venue recommendation," ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 555-564, 2018.

[11] R. Shimizu, M. Matsutani and M. Goto, "An explainable recommendation framework based on an improved knowledge graph attention network with massive volumes of side information," Knowledge-Based Systems, Vol. 239, pp. 107970, 2022.

[12] X. Wang, X. He, Y. Cao, M. Liu and T.S. Chua, "Kgat: Knowledge graph attention network for recommendation," Proceedings of the 25th ACM Conference on Knowledge Discovery & Data Mining, pp. 950-958, 2019.

[13] Q. Guo, Z. Sun, J. Zhang and Y.L. Theng, "An attentional recurrent neural network for personalized next location recommendation," Proceedings of the 34th AAAI Conference on Artificial Intelligence, pp. 83-90, 2020.

[14] W. Chen, H. Wan and S. Guo et al, "Building and exploiting spatial–temporal knowledge graph for next POI recommendation," Knowledge-Based Systems, Vol. 258, pp. 109951, 2022.

[15] H. Wang, Q. Yu, Y. Liu, D. Jin and Y. Li, "Spatio-temporal urban knowledge graph enabled mobility prediction," ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 5, no. 4, pp. 1-24, 2021.

[16] T. Qian, B. Liu, Q.V.H. Nguyen and H. Yin, "Spatiotemporal representation learning for translation-based POI recommendation," ACM Transactions on Information Systems, Vol. 37, no. 2, pp. 1-24, 2019.

[17] Y. Lin, Z. Liu, M. Sun, Y. Liu and X. Zhu, "Learning entity and relation embeddings for knowledge graph completion," Proceedings of the 29th AAAI Conference on Artificial Intelligence, Vol. 29, no. 1, pp. 2181-2187, 2015.

[18] Z. Wang, G. Lin, H. Tan, Q. Chen and X. Liu, "CKAN: collaborative knowledge-aware attentive network for recommender systems," Proceedings of the International Conference on Research on Development in Information Retrieval, pp. 219-228, 2020.

[19] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," Advances in Neural Information Processing Systems, pp. 2787–2795, 2013.

[20] Z. Wang, J. Zhang, J. Feng and Z. Chen, "Knowledge graph embedding by translating on hyperplanes," AAAI Conference on Artificial Intelligence, pp. 1112–1119, 2014.

[21] F. Zhang, N.J. Yuan, D. Lian, X. Xie and W.Y. Ma, "Collaborative knowledge base embedding for recommender systems," ACM Knowledge Discovery and Data Mining, pp. 353-362, 2016.

[22] X. Wang, D. Wang and C. Xu et al, "Explainable reasoning over knowledge graphs for recommendation," AAAI conference on Artificial Intelligence, pp. 5329-5336, 2019.

[23] X. Wang, T. Huang and D. Wang et al, "Learning intents behind interactions with knowledge graph for recommendation," Proceedings of the Web Conference, pp. 878-887, 2021.

[24] Y. Du, X. Zhu, L. Chen, Z. Fang and Y. Gao, "Metakg: Meta-learning on knowledge graph for cold-start recommendation," IEEE Transactions on Knowledge and Data Engineering, pp. 2022.

[25] J.B. Griesner, T. Abdessalem and H. Naacke, "POI recommendation: Towards fused matrix factorization with geographical and temporal influences," ACM Conference on Recommender Systems, pp. 301-304, 2015.

[26] J.J. Levandoski, M. Sarwat, A. Eldawy and M.F. Mokbel, "Lars: A location-aware recommender system," International Conference on Data Engineering, pp. 450-461, 2012.

[27] D. Wu, X. Luo, M. Shang, Y. He, G. Wang, and X. Wu, "A data-characteristic-aware latent factor model for web services QoS prediction," IEEE Transactions on Knowledge and Data Engineering, pp. 2525-2538, 2020

[28] Y. He, S. Chen, T. Nguyen, B. A. Wade, and X. Wu, "Deep matrix tri-factorization: Mining vertex-wise interactions in multi-space attributed graphs," In Proceedings of the 2020 SIAM International Conference on Data Mining, pp. 334-342, 2020.

[29] Y. Wu, K. Li, G. Zhao and X. Qian, "Personalized long-and short-term preference learning for next POI recommendation," IEEE Transactions on Knowledge and Data Engineering, Vol. 34, no. 4, pp. 1944-1957, 2022.

[30] Y. Luo, Q. Liu and Z. Liu, "Stan: Spatio-temporal attention network for next location recommendation," Proceedings of the Web Conference, pp. 2177-2185, 2021.

[31] J. Manotumruksa, C. Macdonald and I. Ounis, "A deep recurrent collaborative filtering framework for venue recommendation," ACM on Conference on Information and Knowledge Management, pp. 1429-1438, 2017.

[32] Y. Liu, A. Pei and F. Wang, "An attention-based category-aware GRU model for the next POI recommendation," International Journal of Intelligent Systems, Vol. 36, no. 7, pp. 3174-3189, 2021.

[33] C. Ounoughi, A. Mouakher, M.I. Sherzad and S.B. Yahia, "A Scalable Knowledge Graph Embedding Model for Next Point-of-Interest Recommendation in Tallinn City," Research Challenges in Information Science: 15th International Conference, RCIS 2021, Limassol, Cyprus, May 11–14, 2021, Proceedings, pp. 435-451, 2021.

[34] P. Veličković, G. Cucurull and A. Casanova et al, "Graph attention networks," arXiv preprint arXiv:1710.10903, 2017.

[35] P. Zhao, A. Luo and Y. Liu et al, "Where to go next: A spatio-temporal gated network for next poi recommendation," IEEE Transactions on Knowledge and Data Engineering, Vol. 34, no. 5, pp. 2512-2524, 2020.

[36] R. Li, Y. Shen and Y. Zhu, "Next point-of-interest recommendation with temporal and multi-level context attention," 2018 IEEE International Conference on Data Mining (ICDM), pp. 1110-1115, 2018.

[37] W. Krichene and S. Rendle, "On sampled metrics for item recommendation," Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, pp. 1748-1757, 2020.