STGIN: Spatial-Temporal Graph Interaction Network for Large-scale POI Recommendation

Shaohua Liu* Meituan Shanghai, China liushaohua07@meituan.com

Mingjian Chen Meituan Shanghai, China chenmingjian@meituan.com Yu Qi* Meituan Shanghai, China qiyu07@meituan.com

Teng Zhang Meituan Shanghai, China zhangteng09@meituan.com

> Jun Lei Meituan Shanghai, China leijun@meituan.com

ABSTRACT

In Location-Based Services, Point-Of-Interest(POI) recommendation plays a crucial role in both user experience and business opportunities. Graph neural networks have been proven effective in providing personalized POI recommendation services. However, there are still two critical challenges. First, existing graph models attempt to capture users' diversified interests through a unified graph, which limits their ability to express interests in various spatial-temporal contexts. Second, the efficiency limitations of graph construction and graph sampling in large-scale systems make it difficult to adapt quickly to new real-time interests. To tackle the above challenges, we propose a novel Spatial-Temporal Graph Interaction Network. Specifically, we construct subgraphs of spatial, temporal, spatial-temporal, and global views respectively to precisely characterize the user's interests in various contexts. In addition, we design an industry-friendly framework to track the user's latest interests. Extensive experiments on the real-world dataset show that our method outperforms state-of-the-art models. This work has been successfully deployed in a large e-commerce platform, delivering a 1.1% CTR and 6.3% RPM improvement.

CCS CONCEPTS

- Information systems \rightarrow Retrieval models and ranking.

KEYWORDS

POI recommendation, spatial-temporal, graph neural network

CIKM '23, October 21-25, 2023, Birmingham, United Kingdom

© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 979-8-4007-0124-5/23/10...\$15.00 https://doi.org/10.1145/xxxxxxxxxxx

ACM Reference Format:

Shaohua Liu, Yu Qi, Gen Li, Mingjian Chen, Teng Zhang, Jia Cheng, and Jun Lei. 2023. STGIN: Spatial-Temporal Graph Interaction Network for Largescale POI Recommendation. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM '23), October 21–25, 2023, Birmingham, United Kingdom.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/xxxxxxxxxxx

Gen Li*

Meituan

Shanghai, China

ligen08@meituan.com

Jia Cheng

Meituan

Shanghai, China

jia.cheng.sh@meituan.com

1 INTRODUCTION

In recent years, Location-Based Service providers including Facebook, Foursquare, and UberEats have become more and more popular. As one of the key services of LBS providers, the POI recommendation utilizes past user behaviors and contextual POI information to make a personalized recommendation. Unlike the traditional recommendation system, the quality of the POI recommendation is intrinsically linked to three dimensions of data: personal, spatial, and temporal, as well as their mutual interactions[11]. For example, a user may order coffee in the morning on business days while looking for a gym on Saturday afternoons.

Many approaches based on sequential user behavior data have been proposed to recommend POI candidates. LSTPM[17] explores the temporal and spatial correlations from the long-term behavior sequence and captures geographical influence from short-term sequences. STPIL[4] constructs various sequences to acquire the spatial-temporal periodic interests of different granularities, then applies integration for multiple interests. CatDM[21] divides a user's check-in history into several time windows and applies a personalized attention mechanism for each time window. However, these works only take into account the user's own behaviors which could cause data sparsity issues in certain spatial-temporal contexts.

Inspired by the idea of collaborative filtering[19] that similar users tend to make similar choices, Graph Neural Network[7, 13] has proven to be an effective way to mitigate data sparsity. GE[20] and JLGE[2] jointly learn the embeddings of multiple bipartite graphs into the same latent space. STGCN[8] considers both userregion periodic pattern and user-POI periodic pattern and fuses

^{*}These authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

all the context information into a unified graph. STPUDGAT[14] leverages spatial, temporal, and preference factors from both local and global views to learn POI-POI relationships. While these methods take advantage of graph structure to improve representation learning, they fail to explicitly model the user's changing interests across different spatial-temporal contexts. Moreover, due to the heavy cost of graph construction and sampling in large-scale industrial systems, these methods have great difficulty in integrating graph representation with the user's real-time behaviors, which can lead to performance degradation in online services.

Motivated by the above analysis, we propose a novel Spatial-Temporal Graph Interaction Network for large-scale POI recommendation. Specifically, a subgraph representation framework has been proposed to learn the user's interests from the spatial view, temporal view, spatial-temporal view, and global view respectively. In addition, to capture the user's latest interests in a real-time manner, we devise a flexible mechanism to combine the interests of multiple views with real-time behaviors in an industrial-friendly way. Our main contributions can be summarized as follows:

- We propose a spatial-temporal graph interaction method to capture the user's diverse interests under different spatial-temporal contexts. To the best of our knowledge, it is the first work to use multi-view spatial-temporal subgraphs for POI recommendation.
- We propose an industrial-friendly framework that combines the spatial-temporal graph learning and the user's real-time behaviors in an end-to-end manner and is able to track the user's latest preferences.
- We have successfully deployed this work in a large locationbased e-commerce platform and achieved encouraging results in both online and offline experiments.

2 PROBLEM DEFINITION

Generally speaking, a recommender system consists of two stages: matching and ranking[3]. In this paper, we implement our method in the matching stage by learning the embedding representations of queries and POIs respectively. With very few modifications, our model is also available in the ranking stage.

Let \mathcal{U} denote the set of users, \mathcal{P} denote the set of candidate POIs, \mathcal{S} denote the set of locations, and \mathcal{T} denote the set of time slots that are divided by a specific pattern(like an hour, day, etc). $p^u = \{p_1^u, p_2^u, ..., p_L^u\}$ denotes the behavior sequence of the user u. It is a list of POIs that are ordered by the corresponding behavior timestamps. Here L denotes the length of the user u's behavior sequence. In this work, we consider clicking as the behavior type.

Since users are consistently interacting with new POIs, we divide the behavior sequence p^u into two parts, one is the real-time sequence (p_r^u) which grows with new click behaviors, the other behaviors are considered as the history sequence (p_h^u) .

For a query q = (u, s, t) requested by a user u under current location s and time t, the goal of the POI recommendation is to select the top K POIs (\mathcal{P}') that the user would be interested in. It can be formulated as

$$\underset{\mathcal{P}' \subset \mathcal{P}, |\mathcal{P}'|=K}{\arg \max} \sum_{\boldsymbol{p} \in \mathcal{P}'} sim(\boldsymbol{e_q}, \boldsymbol{e_p}), \tag{1}$$

where e_q is the vector representation of the query q, and e_p is the vector representation of the POI p. The function sim() calculates the similarity of two vectors.

3 METHODOLOGY

In this section, we elaborate on the proposed Spatial-Temporal Graph Interaction Network. Model structure is shown in Figure 1(a).

3.1 Spatial-Temporal Graph Learning

Users tend to show different preferences in different spatial-temporal contexts. Since a unified graph mixes information from various contexts, it is difficult to accurately express the user's interest in a specific spatial-temporal context. Therefore, we use multi-view sub-graphs for context-specific interest learning.

Graph Construction. Based on the user's history sequence p_h^u , we construct subgraphs from the four types of views: global view, spatial view, temporal view, and spatial-temporal view. Each subgraph has two types of nodes, representing users and POIs in our task, connected by two types of heterogeneous edges. *Clicking edges* represent users' explicit interests. As shown in Figure 1(b), we extract POIs from behavior sequences that occurred in the corresponding spatial-temporal contexts and build an edge between the user and each POI. Note that, clicking edges in the global view are constructed with POIs in all contexts. *Co-clicking edges* represent the implicit relation between POIs. An edge is constructed between two POIs that have been clicked by the same user within a session[6]. It should be emphasized that co-clicking edges are shared among all subgraphs.

Graph Sampling & Aggregation. To precisely capture the user's interest in each view, we devise a spatial-temporal oriented mechanism for node aggregation. Specifically, for each user $u \in \mathcal{U}$, we adopt meta-path[5] "user \xrightarrow{click} POI $\xrightarrow{co-click}$ POI" to generate neighbors. Such a sampling strategy not only preserves the uniqueness of each view but also enhances the representation through shared co-clicking edges. Taking the temporal view $t_i \in \mathcal{T}$ as an example, neighbor set $B_{t_i}^u$ contains both POIs clicked by the user u at time t_i and POIs co-clicked by others within a short period. Then we apply a two-layer GAT[18] mechanism to get the user's interest in this temporal view as follows

$$\boldsymbol{u}_{t_i} = GAT_AGG(\{p | p \in B_{t_i}^u\}).$$
(2)

Similarly, we can get the user's interest from the spatial view u_{s_j} , the spatial-temporal view $u_{s_jt_i}$, and the global view u_g . The neighbors for each POI are generated from the meta-path "POI $\xrightarrow{co-click}$ POI", note that we skip the user node when generating POI's neighbors, since a user may have different types of behaviors in the sequence, which may introduce noise for learning the POI. Again, we use a two-layer GAT mechanism to get the representation of the POI as e_p , which is shared across all views. **Multi-view Interests Interaction**. Different contexts play different roles in determining each user's interests. Therefore, we design a user-dependent attention module to learn the relations among different views. For all spatial views, since POIs located far away are hardly reachable, we fetch the user's spatial views, we apply

STGIN: Spatial-Temporal Graph Interaction Network for Large-scale POI Recommendation

CIKM '23, October 21-25, 2023, Birmingham, United Kingdom



(a) Model Architecture

(b) Multi-View Subgraphs

Figure 1: Overall network structure and construction of multi-view subgraphs.

an attention aggregation process to learn their mutual influence guided by the user's global view. The process is defined as

$$\boldsymbol{u}_t = \sum_{i=1}^T \alpha_{t_i} \cdot \boldsymbol{u}_{t_i}.$$
 (3)

Note that α_{t_i} is the attention weight and is defined as

$$\alpha_{t_i} = \frac{exp(f(\mathbf{W}_t^T[\mathbf{e}_{|t_i - t_q|} || \mathbf{u}_{t_i} || \mathbf{u}_g]))}{\sum_{k=1}^T exp(f(\mathbf{W}_t^T[\mathbf{e}_{|t_k - t_q|} || \mathbf{u}_{t_k} || \mathbf{u}_g]))},$$
(4)

where f(.) is the activation function, || indicates concatenation, $e_{|t_i - t_q|}$ denotes the embedding of the absolute difference between query time t_q and the time of the corresponding temporal view t_i , T denotes the number of time slots, W_t^T is the weighting matrix.

The same process can be applied for spatial-temporal views and get u_{st} as the desired output.

Finally, we get the user's multi-view interest through all views

$$\boldsymbol{u}_{\boldsymbol{h}} = f(\boldsymbol{W}_{\boldsymbol{h}}^{\boldsymbol{I}} \left[\boldsymbol{u}_{\boldsymbol{g}} || \boldsymbol{u}_{\boldsymbol{s}} || \boldsymbol{u}_{\boldsymbol{t}} || \boldsymbol{u}_{\boldsymbol{s}\boldsymbol{t}} \right]).$$
(5)

3.2 Real-time Interests Perception

Users are constantly interacting with new POIs, and these real-time behaviors are essential for predicting users' subsequent interests. Therefore, we need to develop a flexible real-time interest tracking mechanism compatible with spatial-temporal graph learning.

Real-time Behaviors Modeling. Users' real-time behaviors (p_r^u) indicate their latest needs, but historical habits and current context should not be overlooked. First, historical habits such as brands and lifestyles often play an important role in current decisions. Second, given the context of the current time and location, users' next visit is likely to be related to a subset of real-time behaviors. Consequently, we model the user's real-time interest evolution based on the above factors

$$\boldsymbol{u}_{\boldsymbol{r}} = \sum_{i=1}^{M} \alpha_{r_i} \cdot \boldsymbol{e}_{\boldsymbol{p}_{r_i}}.$$
 (6)

 e_{pr_i} denotes the representation of a recently clicked POI. *M* is the length of real-time behaviors. α_{r_i} is the attention weight, which is formalized as

$$\alpha_{r_{i}} = \frac{exp(f(\mathbf{W}_{r}^{T}[\boldsymbol{e}_{|t_{r_{i}}-t_{q}|}||\boldsymbol{e}_{|s_{r_{i}}-s_{q}|}||\boldsymbol{e}_{p_{r_{i}}}||\boldsymbol{u}_{h}]))}{\sum_{k=1}^{M} exp(f(\mathbf{W}_{r}^{T}[\boldsymbol{e}_{|t_{r_{k}}-t_{q}|}||\boldsymbol{e}_{|s_{r_{k}}-s_{q}|}||\boldsymbol{e}_{p_{r_{k}}}||\boldsymbol{u}_{h}]))}, \quad (7)$$

where || indicates concatenation, $e_{|s_{r_i}-s_q|}$ denotes the embedding of the distance between query location s_q and the location of clicked POI s_{r_i} , W_r^T is the weighting matrix.

Adaptive Fusion Mechanism. we conduct a gating-based fusion to generate the final query representation e_q

$$\boldsymbol{e}_{\boldsymbol{q}} = \alpha_1 \cdot \boldsymbol{u}_{\boldsymbol{h}} + \alpha_2 \cdot \boldsymbol{u}_{\boldsymbol{r}} + \alpha_3 \cdot [\boldsymbol{e}_{\boldsymbol{t}_{\boldsymbol{q}}} || \boldsymbol{e}_{\boldsymbol{s}_{\boldsymbol{q}}}]. \tag{8}$$

 $\alpha_1, \alpha_2, \alpha_3$ are generated by the following softmax function

$$[\alpha_1, \alpha_2, \alpha_3] = Softmax([\boldsymbol{w}_h^T \boldsymbol{u}_h, \boldsymbol{w}_r^T \boldsymbol{u}_r, \boldsymbol{w}_c^T[\boldsymbol{e}_{t_q} || \boldsymbol{e}_{s_q}]]), \quad (9)$$

where w_h^T , w_r^T , and w_c^T are weighting vectors, e_{t_q} and e_{s_q} are the embeddings of query time t_q and query location s_q .

3.3 Optimization

After obtaining the representations e_q and e_p , we adopt the pairwise training method[1] and minimize the hinge loss as follows

$$Loss = \sum_{i=1}^{N} \sum_{j=1}^{K} max(0, margin - sim(e_{q}^{i}, e_{p_{i}}^{+}) + sim(e_{q}^{i}, e_{p_{j}}^{-})), (10)$$

where *N* is the length of the training data and *K* is the number of negative samples for each positive sample.

4 IMPLEMENTATION

Training. Graph building and learning are the most time-consuming parts in STGIN. Though Cartesian combinations between spatial and temporal information may be numerous, the actual volume of graph data is acceptable since most users only visited a small number of locations. Thus the graph can be trained simultaneously with other parts. In practice, we divide timestamps into four slots (morning, noon, dinnertime, and night) and use geohash¹ of length 5 to divide a city into multiple locations. For our production dataset with over 300 million users, 20 million POIs, and 600 million clicking behaviors, it takes about four hours to train on 50 virtual machines, each providing 6 CPU cores and 40GB RAM.

Serving. Benefiting from the real-time interests perception module, the graph structure could be updated at a given frequency without real-time requirements. The learned embeddings are stored in a key-value table, thus we can skip the real-time inference of graphs,

¹https://en.wikipedia.org/wiki/Geohash

CIKM '23, October 21-25, 2023, Birmingham, United Kingdom

Table 1: Statistics of dataset

Data	Users	POIs	Records	BehaviorLen	Locations
Train	637563	108551	2023304	149.6	320
Test	217561	60982	532194	149.9	305

Table 2: Comparisons of different methods

Methods	HitRate@200	Recall@200
CatDM	21.37%	15.42%
STPIL	23.80%	17.45%
JLGE	19.96%	14.45%
STGCN	18.52%	13.14%
STPUDGAT	22.21%	16.04%
STGIN	25.87%	19.12%
STGIN-RT	19.71%	14.10%
STGIN-Temporal	22.62%	16.35%
STGIN-Spatial	25.12%	18.57%
STGIN-Interaction	22.56%	16.26%

which may become a bottleneck in online service. It's worth noting that having a new clicking POI will change a user's real-time sequence as well as the representation vector e_q . On average, it takes just 5.6 milliseconds to make an online recommendation.

5 EXPERIMENTS

We design experiments to answer the following research questions: (RQ1): How does STGIN perform compared to other state-of-the-art models? (RQ2): What are the effects of different components? (RQ3): How effective are the spatial-temporal views interacting with each other? (RQ4): How does STGIN perform in online A/B tests?

5.1 Experimental Setup

Dataset and Settings. We conduct offline experiments on industrial production data of a large e-commerce platform. It spans one month(2022-08-09 to 2022-09-08) and is sampled in a few cities. The statistics of dataset are shown in Table 1. We take behaviors over the last 24 hours as the real-time sequence. All experiments are trained on Tensorflow with graph learning engine Euler². We adopt Adam[12] as the optimizer and the dimension of each trainable feature embedding is 16. The batch size and learning rate are set to 1024 and 0.001. We apply LeakyReLU[16] as the activation function. The number of negative samples is 6.

Competitors. We implement several competitive baselines for evaluation. JLGE[2] expresses user, POI, location, and time in the same latent space through multiple bipartite graphs. STGCN[8] fuses all spatial-temporal information into a unified graph and applies a time-based neighborhood sampling algorithm. STPUDGAT[14] uses both spatial and temporal factors to model the POI-POI relationship. Besides, we also implement two popular sequential models. CatDM[21] and STPIL[4] both explore user interests from spatial-temporal behavior subsequences and integrate them through attention mechanisms. To further understand the contribution of each component, we design four variants of our model. STGIN-RT drops real-time behavior sequence. STGIN-Temporal drops subgraphs of

Table 3: Result for Online A/B Test

Metric	CTR	RPM
Relative Improvement	+1.1%	+6.3%

the temporal view. STGIN-Spatial drops subgraphs of the spatial view. STGIN-Interaction merely concatenates graph representation and real-time sequence without their mutual interaction. **Metrics**. We apply widely used **HitRate@K** and **Recall@K**[9, 10, 15] to evaluate the performance of all methods.

5.2 Offline Evaluation

Exp-RQ1. As shown in Table 2, STGIN outperforms other competitors in all metrics. Due to the heavy cost of graph construction and sampling, STGCN faces a major challenge in addressing users' realtime behaviors. JLGE and STPUDGAT try to solve this challenge with the POI-POI graph but pay little attention to the interest interaction. In addition, a unified graph cannot express fine-grained interests in different spatial-temporal contexts. CatDM and STPIL are good at dealing with users' real-time behaviors, but they cannot address the issues of data sparsity when users have limited behaviors.

Exp-RQ2. From table 2, all the components significantly benefit our task. Precisely, STGIN-RT demonstrates real-time behaviors are essential for POI recommendation. STGIN-Temporal and STGIN-Spatial show that it is useful to capture the interests of users from spatial-temporal views. STGIN-Interaction shows that the potential of real-time behaviors can be enhanced by interaction. Exp-RQ3. To validate the effects of interaction among spatialtemporal views, we design temporal-related experiments from coarse to fine-grained. STGIN_Only_Temporal only retrieves a temporal subgraph that matches the current time. STGIN_Sum_Temporal retrieves temporal subgraphs of all time slots and applies sum pooling operation. In Figure 2, STGIN-Temporal gets the lowest score because it omits temporal information. STGIN_Only_Temporal performs slightly better than STGIN_Sum_Temporal, for the latter method does not differentiate the user's attention on different temporal views. Consequently, STGIN achieves the best performance.



Figure 2: Comparisons of temporal-related subgraphs.

5.3 Online A/B Test (RQ4)

Based on the optimizations described in the section 4, we have successfully deployed the STGIN model in our production environment to handle the real traffic of a large e-commerce APP. Online evaluation metrics are CTR and RPM (Revenue Per Mille). As noted in Table 3, both metrics increased compared to the base model, demonstrating the effectiveness of STGIN in practical environments.

6 CONCLUSION

In this paper, we propose a novel approach for POI recommendation called STGIN. It can characterize a user's diverse interests from spatial-temporal subgraphs and capture a user's latest interests through real-time behaviors modeling. Both online and offline experiments demonstrate the effectiveness of our method.

²https://github.com/alibaba/euler

STGIN: Spatial-Temporal Graph Interaction Network for Large-scale POI Recommendation

CIKM '23, October 21-25, 2023, Birmingham, United Kingdom

REFERENCES

- Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. 2005. Learning to rank using gradient descent. In Proceedings of the 22nd international conference on Machine learning. 89–96.
- [2] Ĝiannis Christoforidis, Pavlos Kefalas, Apostolos Papadopoulos, and Yannis Manolopoulos. 2018. Recommendation of points-of-interest using graph embeddings. In 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA). IEEE, 31–40.
- [3] Paul Covington, Jay Adams, and Emre Sargin. 2016. Deep neural networks for youtube recommendations. In Proceedings of the 10th ACM conference on recommender systems. 191–198.
- [4] Qiang Cui, Chenrui Zhang, Yafeng Zhang, Jinpeng Wang, and Mingchen Cai. 2021. ST-PIL: Spatial-Temporal Periodic Interest Learning for Next Point-of-Interest Recommendation. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2960–2964.
- [5] Yuxiao Dong, Nitesh V Chawla, and Ananthram Swami. 2017. metapath2vec: Scalable representation learning for heterogeneous networks. In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. 135–144.
- [6] Yufei Feng, Fuyu Lv, Weichen Shen, Menghan Wang, Fei Sun, Yu Zhu, and Keping Yang. 2019. Deep session interest network for click-through rate prediction. In Proceedings of the 28th International Joint Conference on Artificial Intelligence. 2301–2307.
- [7] Will Hamilton, Zhitao Ying, and Jure Leskovec. 2017. Inductive representation learning on large graphs. Advances in neural information processing systems 30 (2017).
- [8] Haoyu Han, Mengdi Zhang, Min Hou, Fuzheng Zhang, Zhongyuan Wang, Enhong Chen, Hongwei Wang, Jianhui Ma, and Qi Liu. 2020. STGCN: A Spatial-Temporal Aware Graph Learning Method for POI Recommendation. In 2020 IEEE International Conference on Data Mining (ICDM). IEEE, 1052–1057.
- [9] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2015. Session-based recommendations with recurrent neural networks. arXiv preprint arXiv:1511.06939 (2015).
- [10] Jui-Ting Huang, Ashish Sharma, Shuying Sun, Li Xia, David Zhang, Philip Pronin, Janani Padmanabhan, Giuseppe Ottaviano, and Linjun Yang. 2020. Embeddingbased retrieval in facebook search. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2553–2561.

- [11] Md Ashraful Islam, Mir Mahathir Mohammad, Sarkar Snigdha Sarathi Das, and Mohammed Eunus Ali. 2022. A survey on deep learning based Point-of-Interest (POI) recommendations. *Neurocomputing* 472 (2022), 306–325.
- [12] Thomas N Kipf and Max Welling. 2015. Semi-Supervised Classification with Graph Convolutional Networks. In International Conference on Learning Representations.
- [13] Thomas N Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In International Conference on Learning Representations.
- [14] Nicholas Lim, Bryan Hooi, See-Kiong Ng, Xueou Wang, Yong Liang Goh, Renrong Weng, and Jagannadan Varadarajan. 2020. STP-UDGAT: Spatial-temporalpreference user dimensional graph attention network for next POI recommendation. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management. 845–854.
- [15] Fuyu Lv, Taiwei Jin, Changlong Yu, Fei Sun, Quan Lin, Keping Yang, and Wilfred Ng. 2019. SDM: Sequential deep matching model for online large-scale recommender system. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2635–2643.
- [16] Vinod Nair and Geoffrey E Hinton. 2010. Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10). 807–814.
- [17] Ke Sun, Tieyun Qian, Tong Chen, Yile Liang, Quoc Viet Hung Nguyen, and Hongzhi Yin. 2020. Where to go next: Modeling long-and short-term user preferences for point-of-interest recommendation. In Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34. 214–221.
- [18] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2018. Graph attention networks. In Proceedings of International Conference on Learning Representations.
- [19] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. 2019. Neural graph collaborative filtering. In Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval. 165–174.
- [20] Min Xie, Hongzhi Yin, Hao Wang, Fanjiang Xu, Weitong Chen, and Sen Wang. 2016. Learning graph-based poi embedding for location-based recommendation. In Proceedings of the 25th ACM international on conference on information and knowledge management. 15-24.
- [21] Fuqiang Yu, Lizhen Cui, Wei Guo, Xudong Lu, Qingzhong Li, and Hua Lu. 2020. A category-aware deep model for successive POI recommendation on sparse check-in data. In *Proceedings of the web conference 2020*. 1264–1274.