

RESEARCH

Open Access



BMAM: complete the missing POI in the incomplete trajectory via mask and bidirectional attention model

Jun Zeng^{*} , Yizhu Zhao, Yang Yu, Min Gao, Wei Zhou and Junhao Wen

^{*}Correspondence:
zengjun@cqu.edu.cn

School of Big Data and Software
Engineering, Chongqing
University, Chongqing, China

Abstract

Studies on the checked-in point-of-interests have become an important means to learn user's behavior. Nevertheless, users do not sign in to all visited locations. There are unobserved check-in locations in the generated POI trajectory. Such the trajectory is called an incomplete trajectory, and unobserved point is called missing point. However, incomplete trajectory has a negative impact on downstream tasks such as personalized recommendation system, criminal identification and next location prediction. It is a challenge to use the forward sequence and backward sequence information of the missing point to complete the missing POI. Therefore, we propose a bidirectional model based on mask and attention mechanism (BMAM) to solve the problem of missing POI completion in user's incomplete trajectory. The context information of trajectory checked in by user can be mined to connect the missing POI with the forward sequence and backward sequence information. Therefore, the model learns the order dependence between each location according to the user trajectory sequence and obtain the user's dynamic preference to identify the missing POI in the sequence. Besides, the attention mechanism is used to improve the user's representation feature, that is, the preference for POI categories. The experimental results demonstrate that our BMAM outperforms the state-of-the-art models for completion on missing POI of user's incomplete sequence.

Keywords: The missing POI, User behavior sequence, Attention mechanism

1 Introduction

In the era of information overload, Location-based social networks (LBSNs) such as GoWalla¹ and Foursquare² have grown rapidly and become increasingly popular. There is no doubt that human mobile behavior is easier to be digitized and shared with friends [1]. And with wireless sensor technology and global navigation satellite system used to navigate and help transfer data between high-speed mobile, communication between vehicles becomes easier and privacy is considered [2, 3]. What's more, studies on the checked in point-of-interests (POIs) has become an important means to learn user's behavior, which has attracted wide attention from both academia and industry [4, 5].

¹ <https://blog.gowalla.com/>.

² <https://foursquare.com/>.

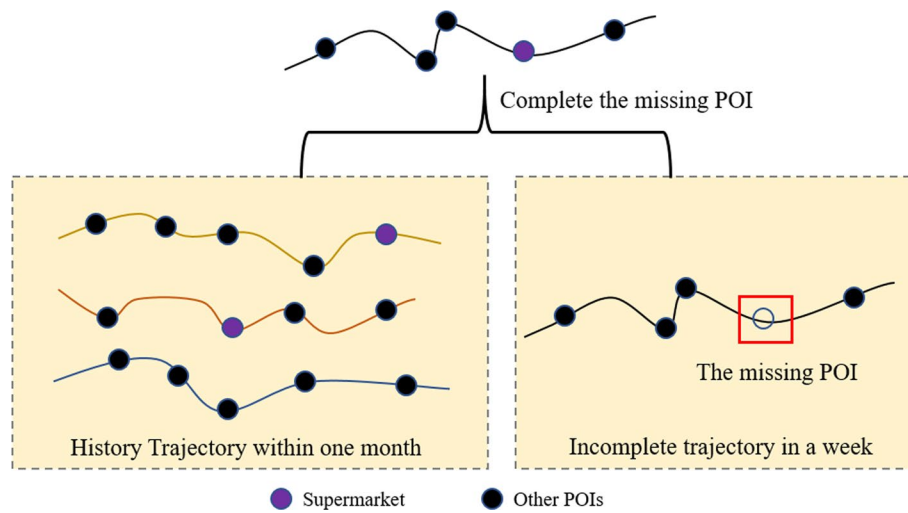


Fig. 1 Scene description

Widely adopted location-based services have accumulated large-scale human mobility data, which is widely used in various fields, such as personalized recommendation systems, location prediction and criminal identification [6–8]. Nevertheless, users do not access mobile services and contribute their data all the time [6]. In fact, the user provided check-in POIs is usually incomplete [9]. Such the trajectory is called an incomplete trajectory, in which the unobserved point is called missing POI. Because the user trajectory information recorded every day is limited and incomplete, it is difficult to complete downstream tasks such as POI recommendation, human mobility and crime analysis [10]. Therefore, it is of great importance to find the missing POI and complete incomplete check-in location trajectory.

When the epidemic breaks out, timely prevention and control is very important. Information technology can help countries control epidemic. However, incomplete information will increase the epidemic risk in some areas. As shown in Fig. 1, a confirmed patient is found, his historical trajectory within one week needs to be tracked and investigated to reduce the epidemic risk. When the time span of the confirmed patient checking in to two locations is large, and the continuity of the two locations does not conform to the daily habits of the confirmed patient, it means that there may be missing location between the two locations. We assume that the patient went to the supermarket 2 days before diagnosis, but did not check in location information on the location-based social network. The supermarket is the place that the patient habitually goes to, and there is check-in information within 1 month. Therefore, in the location-based social network, there is no supermarket in the patient's historical track in the last week. When tracking the patient's historical track through location-based social network information, the supermarket will be ignored. People who come into contact with the patient in the supermarket have potential risks. The epidemic risk in the area where the supermarket is located will increase. Therefore, it is very important to complete the missing location supermarket in the incomplete trajectory in epidemic prevention and control.

Unlike the next location prediction, industrial IoT API recommendation or POI recommendation [11–13], the missing POI problem in incomplete trajectory needs to be

combined with sequence correlation, that is, forward sequence and backward sequence. What's more, Due to the development of cloud computing, more and more applications in different fields are deployed to cloud computing and generate various strategy optimization methods [14, 15], but it is not applicable to the missing POI completion. POI recommendation is to analyze all historical check-in data of users, and use collaborative filtering [16], matrix decomposition [17] and other methods [1, 18–20] to mine the internal relationship between users, user and POI, POI and POI. These features are used to predict the user's next check-in location to recommend POI. However, to complete the missing POI, it is necessary to learn the sequence information of missing position forward sequence and backward sequence. To solve this problem, the model should not only mine the location relationship, but also use the sequence information to learn the bidirectional representation of incomplete trajectory. However, it is a challenge to discover and integrate user behavior sequence relationships to complete the missing POI. This is because the trajectory sequence is incomplete. It is very difficult to learn the information before and after the missing position in the user behavior sequence and establish the relationship between POIs.

Nevertheless, few studies focus on missing POI check-in identification, which is to identify where a user has visited at a specific time in the past. Their work only considers spatial–temporal information and user local preferences, simply splicing the left and right features of missing points, but ignoring the sequence relationship of user trajectory [10]. How to construct the order dependence in the user behavior sequence from the semantic level and obtain the user's dynamic preference to predict the missing POI is very important. The research on trajectory completion mainly focuses on the GPS trajectory completion of taxis [21, 22], generate the missing points of GPS trajectories over occupancy grid map. Xi et al. [10] utilize bidirectional global spatial and local temporal information of POI to capture the complex dependence relationships and user's dynamic preferences for the missing POI check-in identification. But due to the simple structure of neural network, the model cannot fully extract the feature representation of spatial–temporal dependence and user dynamic preference.

Keeping the above challenges in mind, we propose a bidirectional model based on mask and attention mechanism (BMAM), which is used to learn the cross representation of POI features in sequence and enhance user preference features to complete the missing POI. When modeling the sequence of user behavior, the bidirectional model is more appropriate than the unidirectional model, because all POIs in the bidirectional model can use the context information of the forward sequence and the backward sequence [23]. Due to the remarkable achievements of deep learning in the field of POI research, Recurrent Neural Network (RNN) [24] and other deep learning technologies gradually replace the simple method of Collaborative Filtering (CF) [16]. In particular, self-attention can mine important features from a large amount of information and capture the internal correlation of data, reducing the dependence on external information.

Therefore, we make use of the bidirectionality of Bert [25] to fully mine the relationship between the forward sequence and the backward sequence of the missing position in the trajectory sequence. Among them, the masked language model (masked LM) in Bert pre training task combines the context information between sentences through self-attention to predict the masked words. In particular, inspired

by the successful application of bert4rec [23] in item recommendation, we apply the masked LM to the problem of missing POI completion in incomplete trajectory sequences. The user behavior sequence is regarded as a paragraph, and the missing POI is regarded as the words to be predicted. Like filling in the blank, the missing part needs to be judged according to the semantic logic of the paragraph. Similarly, in the problem of missing POI completion, the model need to mine the order dependence between each location according to the user trajectory sequence and obtain the user's dynamic preference to identify the missing POI in the sequence. Therefore, the transformer [26] encoder is used to learn the relationship between the sequences around the masked POI and the dependence of long-distance information. The transformer encoder calculates the correlation between all locations in the trajectory sequence in parallel through the self-attention, so as to obtain the order dependence between locations. The order dependence combines the forward sequence and backward sequence information of missing point from the semantic level to find the missing POI.

Besides, in order to make full use of POI features and user preference features, POI category feature are used to mine more relationships between users and check-in POI. In the input, user check-in POI trajectory sequence and the corresponding category sequence are used to learn the cross feature information and the forward sequence and backward sequence information of the missing POI in the sequence. What's more, in order to enhance the user preference feature representation, the attention mechanism is used to mine the user's category preference for the check-in POI. In the checked in POI trajectory, the model can capture the categories of interest points concerned by users, so as to complete the missing POI in the incomplete trajectory according to the categories.

Overall, our contributions can be summarized as follows:

- (1) We present the BMAM framework, which can effectively solve the problem of missing POI completion in user's incomplete trajectory by learning the order dependence of POIs in trajectory sequence.
- (2) In order to effectively fuse the forward sequence and backward sequence information of the missing POI, we add the category information of POI to the trajectory sequence. In addition, the user's presentation feature is used to strengthen the completion of the missing POI.
- (3) Detailed experiments and deployment are conducted to prove the BMAM's effectiveness.

The rest of the paper is organized as follows. Section 2 introduces the research problem. Section 3 provides detailed methodology of our proposed model. Section 4 presents experiments and the results. Section 5 summarizes the related work, and Sect. 6 concludes this paper and outlines prospects for future study.

2 Methods

In this section, we first introduce the research problem, then we present the data pre-processing process.

2.1 Problem definition

Incomplete trajectory can cause difficulties for downstream tasks such as predicting user's next location and recommend proper points-of-interests. Users U do not check in to all visited locations L . There are unobserved check-in locations in the generated POI trajectory T . Specifically, the user set is $U = \{u_1, u_2, \dots, u_{|U|}\}$, the location set is $L = \{l_1, l_2, \dots, l_{|L|}\}$, and the category of location is $C = \{c_1, c_2, \dots, c_{|C|}\}$. All locations checked in by the user are arranged in chronological order to form a trajectory sequence. The model we proposed is mainly to mine the order dependence between each location according to the user trajectory sequence and obtain the user's dynamic preference to identify the missing POI in the sequence, so the time interval is not considered. The incomplete trajectory of user is $T_n = (l_1^u, l_2^u, \dots, l_{t-1}^u, l_{\text{missing}}^u, l_{t+1}^u, \dots, l_n^u)$. l_{missing}^u is the missing POI at time t in the trajectory, and n is the number of interest points checked in by user. The category sequence corresponding to the user check-in location is $C_n = (c_1^u, c_2^u, \dots, c_{t-1}^u, c_{\text{missing}}^u, c_{t+1}^u, \dots, c_n^u)$, and c_{missing}^u is the POI category missing at time t . The model mines the relationship between POIs and finds the missing POI l_{missing}^u by learning the forward sequence and backward sequence information of the missing POI.

2.2 Data preprocessing

For the trajectory generated by each user, the check-in POI is sorted by time to form the trajectory sequence. We convert the POI trajectory sequence to a fixed length n . This fixed length is determined by the density of the length of the POI sequence checked in by the user. The user's incomplete POI trajectory sequence is $T_n = (l_1^u, l_2^u, \dots, l_{t-1}^u, l_{\text{missing}}^u, l_{t+1}^u, \dots, l_n^u)$. And the missing POI l_{missing}^u at time t is replaced by $[M]$, which is the POI to be found by the model. Therefore, the processed POI trajectory sequence is $T_M = (l_1^u, l_2^u, \dots, l_{t-1}^u, [M], l_{t+1}^u, \dots, l_n^u)$, and $[M]$ is mask. The missing POI in the user trajectory sequence is masked and invisible to the model. Similarly, perform the same processing on the category sequence corresponding to the user's check-in location to obtain $C_M = (c_1^u, c_2^u, \dots, c_{t-1}^u, [M], c_{t+1}^u, \dots, c_n^u)$. The model needs to learn the forward sequence and the backward sequence information of the missing location to mine the relationship between known POIs and predict the missing POI.

3 The architecture of BMAM

In order to address the problem of missing POI completion in incomplete trajectory presented above, we provide a framework in this section to explain our approach. The framework of the proposed BMAM as shown in Fig. 2.

3.1 User preference representation

In order to enhance the data information of trajectory sequence, user preference is the user's preference for POI categories in trajectory sequence, which is added to trajectory sequence. The attention mechanism is used to assign the weight of the user's check-in POI category to obtain the user's attention to the POI category. Further explore the relationship between users and check-in POI according to users' preferences for categories.

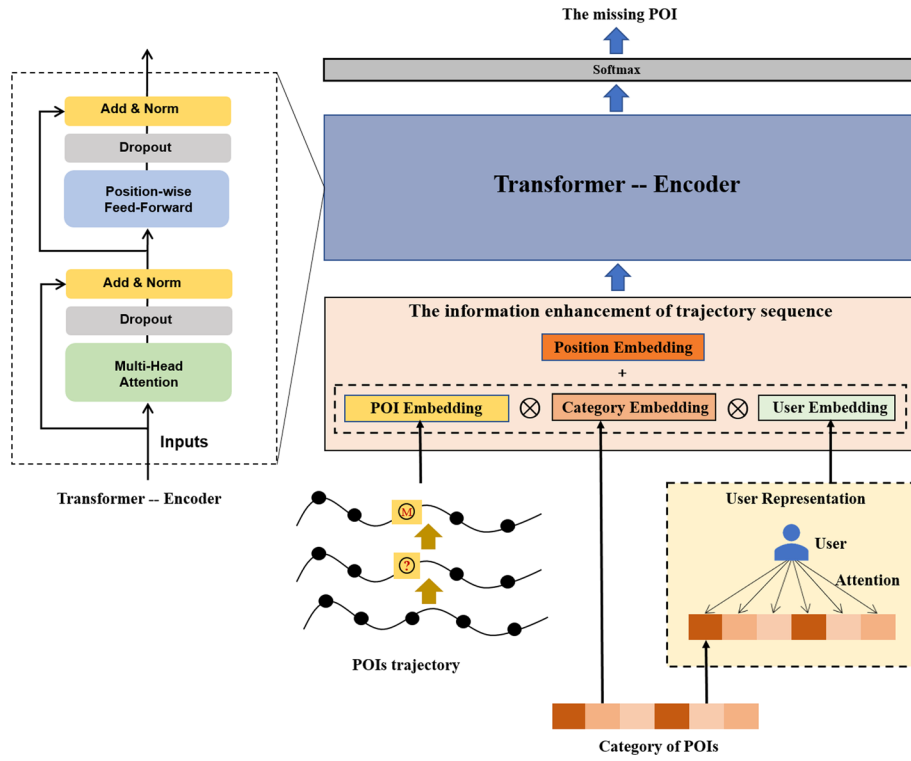


Fig. 2 Framework of the proposed BMAM

Therefore, the user preference representation is combined with the forward sequence and backward sequence of missing position in the trajectory sequence to predict the missing POI. The user u obtains the user preference representation by calculating the attention distribution and weighted average for the category sequence C^u . And the formula is as follows, where attn is a fully connected neural network, v is a learnable parameter and softmax is an excitation function.

$$\bar{U} = \text{softmax}(\tilde{a}_t) \quad (1)$$

$$\tilde{a}_t = vE_t \quad (2)$$

$$E_t = \tanh(\text{attn}(u, C^u)) \quad (3)$$

3.2 The information enhancement of trajectory sequence

The information enhancement of trajectory sequence can help the model learn the relationship between the known POIs and the missing POI. Therefore, in the trajectory sequence, in addition to the location ID, we also add the category corresponding to each location and the user's preference for the location category. Firstly, the trajectory sequence and category sequence need to be initialized embedding to establish the feature dimension. A location embedding matrix $L^u \in R^{n \times d}$ is created based on $T_M = (l_1^u, l_2^u, \dots, l_{t-1}^u, [M], l_{t+1}^u, \dots, l_n^u)$, where n is the sequence length of user and d is the latent dimensionality. Similarly, perform

the same operation on category sequence $C_M = (c_1^u, c_2^u, \dots, c_{t-1}^u, [M], c_{t+1}^u, \dots, c_n^u)$ to obtain the category embedding matrix $C^u \in R^{n \times d}$, where n is the user's POI category sequence length, and d is the potential dimension. Therefore, the input trajectory sequence contains three feature information: location ID, location category and user preference. The input embedded In is obtained by dot product of the three feature information. In transformer encoder, there is no iterative operation of Recurrent Neural Networks. All POIs of the sequence into the model for parallel processing at the same time. Therefore, it is necessary to provide position information for each POI in order to infer the missing POI through the relationship between other POIs. Like BERT4rec [23], the learnable position embedding $P \in R^{n \times d}$ is used for position coding. Position embedding can maintain the sequence order relationship between POIs in the model. Hence, the learnable position embedding $P \in R^{n \times d}$ is injected into the input embedding:

$$In = L^u \otimes C^u \otimes \overline{U} + P \quad (4)$$

3.3 Transformer encoder

The obtained input embedding needs to be sent into the model. The model is a transformer for natural language processing. We use the encoder to process the sequence data and solve the problem of missing POI completion. As shown in Fig. 2, transformer encoder is mainly composed of multi-head attention mechanism, position-wise feed-forward network, drop-out layer and normalization layer. These parts are introduced as follows.

3.3.1 Multi-head attention

In order to learn the expression of multiple meanings in POI trajectory sequence, it is necessary to perform a linear mapping on the input. The multi-headed attention mechanism can be used to extract the meaning of multiple semantics. Multiple heads divide attention into i heads. The attention calculation process of each *head* is the same and independent, and the results are obtained by full connection. The role of multi-head attention mechanism as an integration is to prevent over fitting. The attention function is assigned to each POI weight through the three matrices of Q (Query), K (key), V (value), and the correlation between the known POIs and the missing POI is calculated according to the weight. The formulas for multi-head attention are as follows:

$$H(In) = \text{Concat}(\text{Head}_1, \text{Head}_2, \dots, \text{Head}_i)W \quad (5)$$

$$\text{head}_i = \text{Attention}\left(In W_Q^i, In_i W_K^i, In_i W_V^i\right) \quad (6)$$

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d/h}}\right)V \quad (7)$$

where weight matrices W_Q, W_K, W_V are generated based on the input, W is a learnable parameter, Concat is fully connected, and head_i is the i -th of self-attention. W_Q^i, W_K^i, W_V^i is the feature dimension is divided into h . Q, K, V is the same matrix. Dot product the transposes of Q and K to obtain the attention matrix of each POI. The larger the dot

product, the more similar the vectors of the two POIs, and then use the attention matrix to weight V . In order to make the POI matrix become a standard normal distribution and the result after softmax normalization is more stable, the parameter D needs to be added before weighting $\sqrt{d_k}$. In this way, a balanced gradient can be obtained during back propagation.

3.3.2 Position-wise feed-forward network

Although multi-head attention can use adaptive weights to aggregate the embedding of known POIs and missing POI, it is still a linear model. In order to improve the non-linearity of the model, and consider the interaction between the missing POI and the known POIs in the dimensions. Like Transformer-encoder [27], Position-wise Feed-Forward Network (FFN) is used to improve the nonlinearity of the model. The formula is as follows:

$$\text{FFN}(H) = \text{GELU}(HW_1 + b_1)W_2 + b_2 \quad (8)$$

where W_1 , W_2 , b_1 , b_2 are all learnable parameters, and *GELU* is Gaussian Error Linear Unit.

3.3.3 Stacking transformer-encoder layer

After passing through the multi-head attention module, a network connection is required to learn the potential relationship between the missing POI and the known POIs in the input trajectory. Residual connection is used to prevent neural network degradation during network training. On the premise of the same number of layers, the residual network also converges faster. During training, the gradient can be transmitted back to the initial layer faster, so as to update the weight parameters of POI. After each operation of multi-head attention module, the values before and after the operation must be added to obtain the residual connection. In other words, the input embedded In gets H through multi-head attention, then gets A through dropout and layer normalization, and then inputs A to the position-wise feed-forward network for dropout and layer normalization to get the output B of the transformer encoder module.

$$B = \text{Block}(H) = \text{LN}(A + \text{Drop}(\text{FFN}(A))) \quad (9)$$

$$A = \text{LN}(H + \text{Drop}(H)) \quad (10)$$

where LN is the Layer Normalization and Drop is the Dropout. Layer Normalization can help stabilize the neural network and speed up its training. Dropout can reduce the complex co-adaptation relationship between neurons, and can avoid the phenomenon of over fitting.

3.4 The output of transformer encoder

After some Transformer-Encoder blocks that adaptively and hierarchically extract information of previous POI, we get the final output B for all POIs of the input sequence. We need predict the missing POI base on the output B . As same as the BERT4Rec, we apply

Table 1 Statistics of datasets

Dataset	#users	#POI	#category	#Check-ins
NYC	1083	38,333	251	227,428
TKY	2293	61,858	247	573,703

a two-layer feed-forward network with GELU activation to produce an output distribution over target items:

$$O(l) = \text{softmax}(\text{GELU}(BW_o + b_o)\text{In}^T + b_i) \quad (11)$$

where W_o is the learnable projection matrix, B is the output after b transformer blocks, b_o and b_i are bias terms, In^T is the embedding matrix transpose for the POI set. And the embedding is a matrix shared between the input and output of the model.

3.5 Network training

As mentioned above, the POI trajectory sequence of each user is converted into a fixed length sequence $T_m = (l_1^u, l_2^u, \dots, l_k^u, l_{\text{missing}}^u, l_{k+1}^u, \dots, l_n^u)$ via truncation or padding locations. And the processed user POI trajectory sequence $T_M = (l_1^u, l_2^u, \dots, l_k^u, [M], l_{k+1}^u, \dots, l_n^u)$ is generated as matrix vector L^u , which is used as model input with position embedding P , POI category embedding C^u and user preference representation \bar{U} . The model outputs the missing POI representation at the corresponding position, and calculates the loss of the model through the Cross Entropy Loss function. The Cross Entropy Loss is defined as follow, where $p(l)$ is ground truth.

$$\text{Loss} = - \sum_l (p(l) \log O(l)) \quad (12)$$

4 Experimental and results

In this section, we conduct experiments to evaluate the performance of our proposed model BMAM on two real-world datasets. We first briefly depict the datasets, followed by baseline methods, metrics, setting and training methods. Finally, we present our experimental results and discussions.

4.1 Dataset

We verify our model on two real-world LBSN datasets from Foursquare [28], NYC and TKY, collected by New York and Tokyo. Each dataset has been widely used by previous studies [10, 19] on POI research, which contains check-ins in New York and Tokyo collected for about 10 months (from 12 April 2012 to 16 February 2013). We eliminate users with fewer than 10 check-ins in these two datasets. Following [10], we treat the first 80% sequences of each user as training set, the following 10% for the validation set and the remaining 10% for the test set. The statistics of the two public LBSNs datasets are listed in Table 1.

4.2 Baselines

We compare BMAM with the following methods representing the state-of-the-art location-based research techniques.

4.2.1 STRNN [24]

A spatial–temporal Recurrent Neural Network model for user next location prediction. It incorporates both the time-specific transition matrices and distance-specific transition matrices within recurrent architecture.

4.2.2 PACE [29]

A deep neural architecture is that jointly learns the embedding of users and POIs to predict both user preference over POIs and various context associated with users and POIs.

4.2.3 Bi-STDDP [10]

A model can integrate bidirectional spatiotemporal dependence and users' dynamic preferences, to identify the missing POI check-in where a user has visited at a specific time.

4.2.4 SASRec [30]

It uses a left-to-right Transformer language model to capture users' sequential behaviors, and achieves state-of-the-art performance on sequential recommendation.

4.2.5 BERT4Rec [23]

A sequential recommendation models users' dynamic preferences from users' historical behavior and employs the deep bidirectional self-attention to model user behavior sequences to make recommendations.

4.3 Evaluation metrics

To measure and evaluate the performance of different methods, *Recall @K* and *F1-score @K* are adopted. The larger the value, the better the performance for all the evaluation metrics. The two conventional evaluation metrics are defined as:

$$\text{Pre@K} = \frac{1}{|U|} \sum_{u \in U} \frac{|V_u^T \cap V_u^P|}{K} \quad (13)$$

$$\text{Rec@K} = \frac{1}{|U|} \sum_{u \in U} \frac{|V_u^T \cap V_u^P|}{|V_u^T|} \quad (14)$$

$$\text{F1-score@K} = \frac{2}{\text{recall@K}^{-1} + \text{pre@K}^{-1}} \quad (15)$$

The Pre@K is the ratio of recovered POI to the K predicted POI, and Rec@K is the ratio of recovered POI to the ground truth. We do not use Pre@K since it is positively correlated with Recall@K . Given the user set U . We set the masked POI as the ground truth V_u^T , and V_u^P is the set of prediction result.

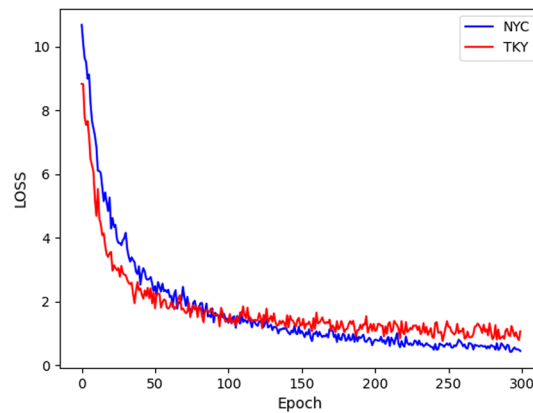


Fig. 3 Loss value with epoch

4.4 Experimental settings

In our method, we use four heads of attention modules and two blocks of multi-head attention for the check-in sequence. We train our model using the Adam optimizer with a learning rate of 0.001 and set the dropout ratio to 0.1. The batch size is 64 and the hidden dimension of model is 256. The trend of the model's loss with the epoch during the training process is shown in Fig. 3. When the epoch reaches 250, the loss of the model gradually begins to converge. Therefore, the number of training epochs is set to 250 for NYC and TKY.

4.5 Training methods

Different from the left to right language model training method, our goal is to let the feature representation fuse the forward sequence and backward sequence of missing POI to train the deep bidirectional model. Therefore, like Bert [12], in the training process, we randomly mask the user behavior sequence according to the mask method of masked LM. There are three ways to mask, namely $[M]$, original POI and random POI. Firstly, 15% POI in the user behavior sequence are randomly selected according to the probability. Then, among these POI, 80% are replaced with $[M]$, 10% with original POI and 10% with random POI. This ensures that the POI trajectory sequence of each input is a distributed context representation. It can not only let the model know which POI should be predicted, but also reduce the impact of information leakage caused by $[M]$ replacing POI.

4.6 Comparison with baselines

Since the baseline model is used for recommendation except Bi-STDDP, only the forward sequence of missing position is used in the baseline experiment. The performance comparison results are shown in Table 2.

4.6.1 Observations about our model

First, the proposed model, BMAM, achieves the best performance on two datasets with all evaluation metrics, which illustrates the effectiveness of our model. Second, BMAM outperforms BERT4rec. Although BERT4rec adopts the bidirectional model with mask

Table 2 Performance of BMAM and baseline on the dataset

Methods	Metrics	STRNN	PACE	Bi-STDDP	SASRec	BERT4Rec	BMAM	Improv (%)
NYC	Recall@5	0.3225	0.3226	0.3476	0.5411	0.6005	0.6191	3.1
	Recall@10	0.4055	0.4033	0.4176	0.6417	0.6606	0.6756	2.3
	F1-score@5	0.1075	0.1075	0.1159	0.1804	0.2002	0.2063	3.0
	F1-score@10	0.0737	0.0733	0.0759	0.1167	0.1201	0.1228	2.2
TKY	Recall@5	0.3665	0.3605	0.4107	0.6249	0.6631	0.6852	3.3
	Recall@10	0.4492	0.4469	0.4838	0.7021	0.7256	0.7386	1.8
	F1-score@5	0.1222	0.1202	0.1369	0.2083	0.2210	0.2284	3.3
	F1-score@10	0.0817	0.0813	0.0880	0.1277	0.1319	0.1343	1.8

to improve the learning ability, it neglects the category feature of location and the global preferences of user, and only uses the forward sequence of missing position in the sequence, which all is used by the proposed model. Third, BMAM achieves better performance than SASRec. SASRec use the attention model to distinguish the items users have accessed, while it is unidirectional and cannot fully learn the relationship between items in the sequence. Fourth, BMAM obtains better results than Bi-STDDP, PACE and STRNN. One major reason is that these three methods combines a variety of features of locations and mines user preferences based on LSTM, RNN and other models, but the network structure cannot fully mine user trajectory features. PACE and STRNN ignore the sequence feature of trajectory, that is, the time sequence between check-in locations, which is very important for the prediction of missing POI.

4.6.2 Other observations

First, BMAM and BERT4Rec outperform SASRec on two datasets although they all use self-attention and consider the feature of trajectory sequence. The main reason is that SASRec neglects the bidirectionality in user's trajectory. Second, BMAM, BERT4Rec and SASRec achieve better performance than Bi-STDDP, PACE and STRNN. This illustrates that compared with LSTM and RNN, self-attention can fully explore the relationship between locations and their features, and capture users' dynamic preferences.

4.7 Influence of hidden dimensionality d

In order to compare the influence of the hidden dimension of model on BMAM model on two data sets, we vary the dimension of embedding d in the model from 16 to 256 while keeping other optimal hyper-parameters unchanged. The recall@5 and recall@10 on two datasets are shown in Figs. 4 and 5, respectively. The performance of model tends to converge as the dimension of embedding increases, and increases slowly when the hidden dimension is 128. This shows that $d=128$ is the best dimension for trajectory and category embedding. A larger dimension of embedding does not necessarily lead to better model performance, both NYC and TKY. Although $d=128$ is the turning point, the dimension of embedding is set to 256 in the experiment, which can conducive to the stability of the experimental results.

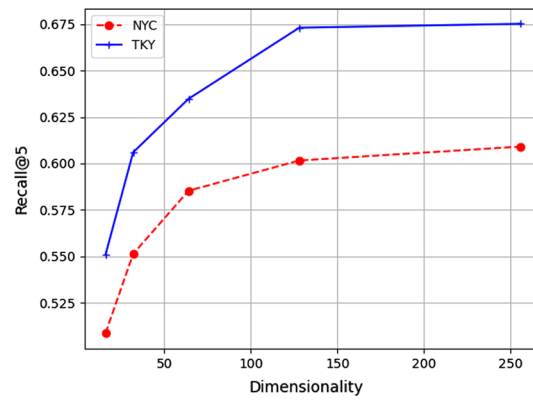


Fig. 4 Recall@5 value with varying hidden dimensionality of model on two datasets

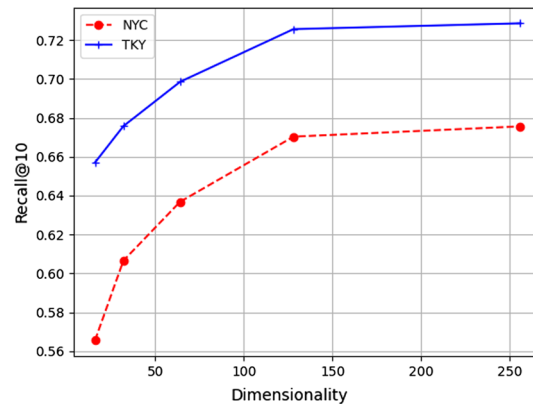


Fig. 5 Recall@10 value with varying hidden dimensionality of model on two datasets

Table 3 Ablation analysis (Recall@5) on two datasets

Data	NYC	TKY
Remove PE	0.3902	0.4569
Remove category	0.6049	0.6648
BMAM	0.6191	0.6852

4.8 Ablation study

Finally, we perform ablation experiment on some key components of BMAM in order to better understand its impacts, including positional embedding (PE), category sequence and user preferences for categories (category). Table 3 shows the results of ablation study.

We introduce the variants and analyze their effects, respectively:

- (1) *Remove PE* The results show that removing positional embedding causes BMAM's performances decreasing dramatically on two datasets. This shows that sequence information plays an important role in model mining features, and can help

the model obtain the order dependence between each location in the trajectory sequence.

- (2) *Remove category* The results show that adding categories and user preference features can enhance the information of track sequence. The model can obtain additional category information when mining trajectory sequences, and complete the missing POI according to the order dependence between locations.

5 Related work

5.1 Research on POI

POI research has attracted intensive attention due to a wide range of potential applications. Most existing studies focus on next location prediction, POI recommendation and competitive analysis between POIs, etc. [11, 12, 31–33]. The methods to solve these problems are also gradually increasing and the effect is better. Han et al. [31] divide the context information of POI into two groups, namely global and local context, and develop different regularization terms to combine them for recommendation. In the next location prediction problem, Zhao et al. [34] propose a new space–time gate control network (STGN). The network captures the temporal and spatial relationship between continuous check-in locations by enhancing the long-term and short-term memory network and introducing a gating mechanism. Li et al. [32] build a heterogeneous POI information network (HPIN) from POI reviews and map search data and develop a graph neural network-based deep learning framework for POI competitive relationship prediction. However, few studies have focused on the missing POI in incomplete trajectories. It is difficult to complete downstream tasks such as predicting a next location and POI recommendation for users with incomplete trajectory.

5.2 Trajectory completion

In the problem of trajectory completion, most studies are based on GPS trajectory, such as taxi trajectory to complete the missing points [21, 22]. However, there are few studies on missing POI in the trajectory checked in by users. Zhang et al. [9] explain that in practice, the check-in POI provided by the user is usually incomplete. However, their work only alleviates the incompleteness of the location information checked in by the user to make POI recommendation, and does not solve the problem of missing locations. Xi et al. [10] utilize bidirectional global spatial and local temporal information of POI to capture the complex dependence relationships and user's dynamic preferences for the missing POI check-in identification. But due to the simple structure of neural network, the model cannot fully extract the feature representation of spatiotemporal dependence and user dynamic preference.

5.3 Deep learning

Liu et al. [27] are the first to learn from the natural language processing method, taking each POI as a word and each user's check-in record as a sentence. Then the implicit representation vector of each POI is trained and the influence of time implicit representation vector on it is mined. Due to the remarkable achievements of deep learning in the field of POI research, deep learning technology has gradually replaced

the simple forms of collaborative filtering (CF), matrix factorization (MF), Markov chain and so on. With the successful application of RNN in sequential data modeling, RNN and its variants are used to model the user's behavior sequence in POI recommendation [34–36]. In the POI trajectory completion problem, the bidirectional model needs to learn the optimal representation of user POI trajectory sequence. Recently, Transformer [26], a sequence-to-sequence method based only on self-attention, has achieved the most advanced performance and efficiency in machine translation, which have been dominated by RNN-based methods [30]. Specifically, due to the successful application of Bert [25] in text understanding, we consider applying the deep bidirectional model based on self-attention to the problem of missing POI trajectory completion.

6 Conclusions

In this paper, we propose a bidirectional model based on mask and attention mechanism (BMAM) to solve the problem of missing POI completion in user's incomplete trajectory sequence. Masked language model of Bert pre training task is used to find the missing POI combined with the attention mechanism to enhance the user representation according to the POI category. The attention mechanism model with bidirectional and mask can mine the context information of the POI trajectory sequence checked in by the user, so as to connect the missing POI with the forward sequence and backward sequence information. Besides, the attention mechanism is used to improve the user's feature representation, that is, the preference for POI categories. The BMAM solves the limitation that the previous model does not make full use of the order dependence of location. Our experiments on real-world LBSNs datasets show that modeling the sequence relationship of user behavior and the representation of user feature have considerable impact on the problem of missing POI completion. For future work, we consider combining more features of POI, such as spatiotemporal features, to mine users' behavior habits in time and space.

Abbreviations

BMAM	A bidirectional model based on attention mechanism and mask
POI	Point-of-interests
LBSNs	Location-based social networks
Masked LM	Pre training task masked language model of BERT
BERT	A pre-training of deep bidirectional transformers for language understanding model
CF	Collaborative filtering
MF	Matrix factorization

Acknowledgements

Not applicable.

Author contributions

JZ and YZ propose the innovation ideals. YZ and YY carry out the experiments. YZ carries out the original draft. JZ, YZ, MG, WZ and JW review the manuscript. All authors read and approved the final manuscript.

Funding

This research is sponsored by Natural Science Foundation of Chongqing, China (No. cstc2020jcyj-msxmX0900, No. cstc2019jcyj-msxmX0442), the Fundamental Research Funds for the Central Universities (Project No. 2020CDJ-LHZZ-040), and National Natural Science Foundation of China (Grant No. 72074036, 62072060).

Availability of data and materials

The dataset employed for conducting the current study is publicly available at <https://sites.google.com/site/yangdingqi/home/foursquare-dataset>.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 29 December 2021 Accepted: 10 June 2022

Published online: 20 June 2022

References

1. D. Lian, Y. Wu, Y. Ge, X. Xie, E. Chen, Geography-aware sequential location recommendation, in Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery and data mining, 2020, pp. 2009–2019.
2. H. Gao, C. Liu, Y. Yin, Y. Xu, Y. Li, A hybrid approach to trust node assessment and management for vanets cooperative data communication: historical interaction perspective, 2021
3. H. Gao, Y. Zhang, H. Miao, R.J.D. Barroso, X.J.M.N. Yang, Applications, SDTIOA: modeling the timed privacy requirements of iot service composition: a user interaction perspective for automatic transformation from BPEL to timed automata, pp. 1–26, 2021
4. K. Zhao et al., Discovering subsequence patterns for next POI recommendation, in IJCAI, 2020, pp. 3216–3222
5. W. Zhang, J. Wang, Location and time aware social collaborative retrieval for new successive point-of-interest recommendation, in Proceedings of the 24th ACM international conference on information and knowledge management, 2015, pp. 1221–1230
6. T. Xia et al., AttnMove: history enhanced trajectory recovery via attentional network. Proc. AAAI Conf. Artif. Intell. **35**(5), 4494–4502 (2021)
7. C. Miao, J. Wang, H. Yu, W. Zhang, Y. Qi, Trajectory-user linking with attentive recurrent network, in Proceedings of the 19th international conference on autonomous agents and multiagent systems, 2020, pp. 878–886
8. J. Zeng, H. Tang, Y. Zhao, M. Gao, J.J.M.N. Wen, Applications, PR-RCUC: a POI recommendation model using region-based collaborative filtering and user-based mobile context, pp. 1–11, 2021
9. L. Zhang, Z. Sun, J. Zhang, Y. Lei, F. Klanner, An interactive multi-task learning framework for next POI recommendation with uncertain check-ins, in Twenty-ninth international joint conference on artificial intelligence and seventeenth pacific rim international conference on artificial intelligence (IJCAI-PRICAI-20), 2020
10. D. Xi, F. Zhuang, Y. Liu, J. Gu, H. Xiong, Q. He, Modelling of bi-directional spatio-temporal dependence and users' dynamic preferences for missing poi check-in identification. Proc. AAAI Conf. Artif. Intell. **33**(01), 5458–5465 (2019)
11. S. Zhao, T. Zhao, H. Yang, M. R. Lyu, I. King, STELLAR: Spatial-temporal latent ranking for successive point-of-interest recommendation, in Thirtieth AAAI conference on artificial intelligence, 2016
12. D. Liao, W. Liu, Y. Zhong, J. Li, G. Wang, predicting activity and location with multi-task context aware recurrent neural network, in IJCAI, 2018, pp. 3435–3441
13. H. Gao, X. Qin, R. J. D. Barroso, W. Hussain, Y. Xu, Y. Yin, Collaborative learning-based industrial IoT API recommendation for software-defined devices: the implicit knowledge discovery perspective, 2020
14. Y. Huang, H. Xu, H. Gao, X. Ma, W. Hussain, SSUR: an approach to optimizing virtual machine allocation strategy based on user requirements for cloud data center. IEEE Trans. Green Commun. Netw. **5**(2), 670–681 (2021)
15. X. Ma, H. Xu, H. Gao, M. Bian, Real-time multiple-workflow scheduling in cloud environments. IEEE Trans. Netw. Serv. Manag. **18**(4), 4002–4018 (2021)
16. Y. Koren, R.J.R.S.H. Bell, Advances in collaborative filtering, pp. 77–118, 2015
17. J.D. Brown, matrix decompositions. Linear models in matrix form, 2014
18. J. Zeng, H. Tang, X. He, RCFC: a region-based POI recommendation model with collaborative filtering and user context, in International conference on collaborative computing: networking, applications and worksharing, (Springer, 2020), pp. 656–670
19. F. Yu, L. Cui, W. Guo, X. Lu, Q. Li, H. Lu, A category-aware deep model for successive poi recommendation on sparse check-in data. Proc. Web Conf. **2020**, 1264–1274 (2020)
20. J. Zeng, F. Li, X. He, J. Wen, Fused collaborative filtering with user preference, geographical and social influence for point of interest recommendation. Int. J. Web Serv. Res. (IJWSR) **16**(4), 40–52 (2019)
21. A. Nawaz, Z. Huang, S. Wang, A. Akbar, H. AlSalman, A.J.S. Gumaee, GPS trajectory completion using end-to-end bidirectional convolutional recurrent encoder-decoder architecture with attention mechanism. Sensors **20**(18), 5143 (2020)
22. K. Zheng, Y. Zheng, X. Xie, X. Zhou, Reducing uncertainty of low-sampling-rate trajectories, in 2012 IEEE 28th international conference on data engineering, (IEEE, 2012), pp. 1144–1155
23. F. Sun et al., BERT4Rec: sequential recommendation with bidirectional encoder representations from transformer, in Proceedings of the 28th ACM international conference on information and knowledge management, 2019, pp. 1441–1450
24. Q. Liu, S. Wu, L. Wang, T. Tan, Predicting the next location: a recurrent model with spatial and temporal contexts, in Thirtieth AAAI conference on artificial intelligence, 2016.
25. J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: pre-training of deep bidirectional transformers for language understanding, 2018
26. A. Vaswani et al., Attention is all you need, in Advances in neural information processing systems, 2017, pp. 5998–6008
27. X. Liu, Y. Liu, X. Li, Exploring the context of locations for personalized location recommendations, in IJCAI, 2016, pp. 1188–1194
28. D. Yang, D. Zhang, V.W. Zheng, Z. Yu, Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. IEEE Trans. Syst. Man Cybern.: Syst. **45**(1), 129–142 (2014)

29. C. Yang, L. Bai, C. Zhang, Q. Yuan, J. Han, Bridging collaborative filtering and semi-supervised learning: a neural approach for poi recommendation, in Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, 2017, pp. 1245–1254
30. W.-C. Kang, J. McAuley, Self-attentive sequential recommendation, in 2018 IEEE international conference on data mining (ICDM), (IEEE, 2018), pp. 197–206
31. P. Han, Z. Li, Y. Liu, P. Zhao, S. Shang, Contextualized point-of-interest recommendation, in Twenty-ninth international joint conference on artificial intelligence and seventeenth pacific rim international conference on artificial intelligence (IJCAI-PRICAI-20), 2020
32. S. Li, J. Zhou, T. Xu, H. Liu, X. Lu, H. Xiong, Competitive analysis for points of interest, in Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining, 2020, pp. 1265–1274
33. J. Zeng, X. He, H. Tang, J. Wen, Predicting the next location: a self-attention and recurrent neural network model with temporal context. *Trans. Emerg. Telecommun. Technol.* **32**(6), e3898 (2021)
34. P. Zhao, H. Zhu, Y. Liu, J. Xu, X. Zhou, Where to go next: a spatio-temporal gated network for next POI recommendation 33, 5877–5884, (2019)
35. Y.S. Lu, W.Y. Shih, H.Y. Gau, K.C. Chung, J.L. Huang, "On successive point-of-interest recommendation. *World Wide Web* **22**(3), 1151–1173 (2019)
36. J. Manotumruksa, C. Macdonald, I. Ounis, A contextual attention recurrent architecture for context-aware venue recommendation, in The 41st international ACM SIGIR conference on research & development in information retrieval, 2018, pp. 555–564

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
