Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105 (Online)
Volume **7**, Number **1**, February **2022**

A Bi-direction LSTM Attention Fusion Model for the Missing POI Identification

Lyuchao Liao

Fujian Provincial Universities Key Laboratory of Industrial Control and Data Analysis
Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology
Fujian University of Technology
Fuzhou 350118, China
fjachao@gmail.com

Jinmei Lin*

Fujian Provincial Universities Key Laboratory of Industrial Control and Data Analysis
Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology
Fujian University of Technology
Fuzhou 350118, China
*Correponding Author:linjinmei387@gmail.com

Yintian Zhu

Fujian Provincial Universities Key Laboratory of Industrial Control and Data Analysis
Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology
Fujian University of Technology
Fuzhou 350118, China
2469354563@qq.com

Shuoben Bi

School of Geography and Remote Sensing
Nanjing University of Information Science and Technology
Nanjing, 210044, China
bishuoben@163.com

Yuyuan Lin

Fujian Provincial Universities Key Laboratory of Industrial Control and Data Analysis
Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology
Fujian University of Technology
Fuzhou 350118, China
2201905137@smail.fjut.edu.cn

Received October 2021; revised December 2021

Abstract. *Point-of-interest(POI) check-in data, collected from human mobility, provides an excellent opportunity to understand users' behavior. However, the collected data usually suffer from various quality issues, especially the missing geolocation information. The incomplete POI data will prevent a deep analysis of users' preferences and movement patterns and limit its practical application. To address these issues, emerging studies are conducted to predict the missing POIs. Nevertheless, existing methods generally employed history information from a single-direction perspective for POI recommendation or prediction. Considering that the missing POI identification tasks simultaneously require both former and later data of the missing target, we proposed a Bi-direction LSTM Attention Fusion model, named Bi-LAF, to combine users' preferences with spatial-temporal patterns identifying the missing POI check-ins. Firstly, Bi-LAF integrates both Long Short-Term Memory(LSTM) and self-attention mechanisms to explore and combine users' preferences and the time patterns of missing POI. Then, we employ the Great Circle Distance(GCD) and time intervals of the successive moment to learn the relationships between missing POI and candidate POIs. Finally, users' preferences and spatial-temporal relationships are spatialized to complete the identification task of missing POIs. The proposed model is evaluated on two real large-scale datasets, and the results show that Bi-LAF outperforms the state-of-the-art related methods.*
Keywords: Missing POI, LSTM, Self-attention, POI recommendation

1. **Introduction.** With the popularity of the Global Positioning System(GPS), Location-based Social Network (LBSN) has gained rapid growth and popularity in recent years. These services enable millions of users to check in at real-world locations and share life experiences on various platforms. The data accumulated from LBSN provide an excellent opportunity to understand users' preferences and mobile patterns for improving users' experience and service quality, and recently a large amount of users' trajectory data are widely used in various fields.In a narrow sense, the trajectory data refers to continuously sampled GPS data. Generally, sequences that exist with spatial-temporal characteristics could be called trajectories. For example, users' check-in sequences on social networks could be considered coarse-grained trajectory data. Accurate trajectory data usually play an essential role in improving user-centered related applications, including POI recommendations, city planning, and route optimization. What's more, the trajectory data could also facilitate targeted advertising and help merchants to attract more potential customers.

However, the missing POI check-ins in mobile data usually bring challenges for engineers and researchers. In reality, users' mobile data are generally incomplete due to the lack of spatial information or the protection of personal privacy. For example, users are unlikely to check in every time they visit a location, or users are reluctant to disclose some check-ins information, resulting in the records of fake information and the absence of true POI check-ins. The missing records in mobility data hide useful information potentially, which may negatively affect the further analysis of users' preferences and movement patterns. Therefore, identifying the missing POIs is an essential and challenging task for improving mobile services.

If POI data exists missing situation, most studies related POI are restricted. The missing of location information results in the deficiency of quality issues, or worse, users understanding. Since the existence of false data in the data set is not excluded, it's trouble to replace or get rid of it. For example, it is unreasonable to occur the case that the geographical location of the before and after check-in data is far away, but the time interval is short. On the other hand, the task of missing location identification can be appliede in the public security work to help polices to get trace of crime or in analysis cases of missing persons. What's more, it's useful to verify the travel path during the outbreak. In this paper, we focus on missing POI check-in identification. In other words, the main of our work is to predict the specific location which can represent where the user is most likely to go at a specific time.

In the literature, Most of the POI-related studies focused on the POI recommendation with traditional methods. Specifically, Matrix Factorization(MF) is used widely for POI recommendations. For example, Baratchi et al. [1] proposed an LGLMF model to integrate logical matrix decomposition for improving users' active area accuracy and location relevance within the area. Cheng et al. integrated social information and geographical influence into the MF framework for the POI recommendations [2].

The traditional missing data identification methods are mainly evolved from the lower-dimensional data with traditional machine learning methods [3]. For example, Tu [4] used Bayes algorithm to implement the task of prediction. Kang et al. [5] proposed Gaussian process regression (GPR) for the problems of uncertain measurment. Inspired by the great success of deep learning techniques in various research areas, Recurrent Neural Networks (RNN) and Generative Adversarial Networks (GAN) have also been applied to the field of missing data identification [6]. For example, RNN and its variants, such as LSTM [7] and Gate Recurrent Unit(GRU) [8], have been used to explore the dynamic preferences of users. Recently, more models regarding spatial-temporal contexts have been proposed for applications regarding POI. For example, the ST-RNN [9] is proposed to model each layer's local time and spatial context with a specific time transformation matrix and a certain distance transformation matrix; Bi-STDDP [10] integrates the global spatial and local time effects and users' dynamic preferences as another successful example for the missing POI identification. What's more, Bi-G^{2}AN [11] combines GAN and GRU networks to explore the distribution of missing POI in the velocity-oriented motion pattern mining.

Even though the task of missing POI identification plays an important role in the field of POI and analysis of users' behavior, few most related studies have been introduced to deal with the topic. On the macro perspective, these methods mentioned above only use the history check-ins information for future prediction or recommendation from a single direction perspective, and the high dimensional non-linear complexity barrier these traditional methods from performance improvement and the large-scale applications. There exist Bi-STDDP and Bi-G^{2}AN which are successful method for solve missing problems. However, the former only uses embeddding layer to get superficial features in POI while the check-ins data can be mined more potential meanings. Bi-G^{2}AN considers more fators in the model, including the speed and motor pattern, which are more suitable for trajectory data of cars, not POI check-ins data. All in all, the two methods perform good results in the missing problems.

To address these issues, we propose a novel Bi-direction LSTM Attention Fusion model, named Bi-LAF, for the missing POI identification, which adapts users' preferences and spatial-temporal information by a bidirectional LSTM with attention. The bidirectional model is more suitable for solving the problem using more information, which utilizes the information both before and after the missing POI to combine LSTM and self-attention mechanisms for exploring users' preferences. In order to further mine users' preferences, the time pattern of target missing POI and users' unique feature are taken into account in the model. We use the Great Circle Distance (GCD) [12] and time intervals of the successive moment to learn the relationships between missing POI and candidate POIs. Finally, users' preferences and spatial-temporal relationships are spatialized for the missing POI identification.

The main contributions of this work are as follows:

1 We adapt a bidirectional sequence to explore users' preferences by LSTM and attention fused, Then use spatial-temporal information to capture the relationship between missing POI and candidate POIs for the missing POI identification.
2 Unlike previous traditional POI recommendation and prediction models, which only use historical sequences, we propose a reversed LSTM learning in the Bi-LAF so that the bidirectional sequences can tend to be the target.
3 Extensive experiments have been conducted on two real datasets. The results show that the proposed model outperforms the state-of-the-art related methods.

2. **Related work.** The most relevant work for missing POI identification could be summarized as two aspects: the spatial missing data recognition and POI recommendation. This section will make a brief introduction of the related works in these two aspects.

2.1. **Spatial Missing Data Identification.** Missing data filling is an essential task in data analysis. Some classical statistical methods, such as zero fillings and mean filling, are usually employed to fill missing data [13]. Although these methods based on statistical technology are easy to implement, their performance is limited. Many machine learning methods are also used to complete missing data, such as standard K-means clustering and fuzzy C-means clustering [14]. With the development of deep learning, the methods based on RNN and GAN are further promoted and applied. These methods could fill the missing data in space, but they are not designed for spatial perception problems in essence. Neighborhood-based and collaborative filtering methods play a dominant role [15, 16]. In addition, some successful spatial-temporal models are proposed for time series data [17], such as users' dynamic preferences combined with global spatial and local temporal influence in Bi-STDDP [10]. In [11], spatial-temporal effect and local movement information were used to learn dynamic preferences, and GAN and GRU were combined to identify missing POI.

2.2. **POI Recommendation.** Many previous studies utilize Collaborative Filtering [18] to learn users' preferences for POI, including Memory-based CF methods and Model-based CF methods. Matrix Factorization is used in POI recommendations gradually. Models based on MF depend on strong assumptions of independence between different factors. Like [19], multi-labels, social and geographic data is modeled individually and then fused into the matrix decomposition framework. POI recommendation differs from product recommendation and movie recommendation tasks. POI recommendation mainly uses historical check-ins, including valuable information in space and time, to predict the following location. In comparison, tasks such as product and movie recommendations pay more attention to the users' rating of items for learning users' preferences degree.

All in all, sequence information plays a vital role in the recommendation. In [20], the proposed Factorization Personalized Markov Chain(FPMC) combines Matrix Factorization and Markov Chain for sequence prediction, extended by embedding personalized migration and local regions. Chen et al. [21] focus on spatial and temporal irregularities between successive POIs to learn the complex relationships between sequence transitions.

In recent years, neural network-based methods have been successfully extended to POI recommendation and location prediction. Kong et al. [6] mix spatial-temporal effects into LSTM to solve the problem of data sparsity. By enhancing the LSTM-based framework, [22-24] design novel methods for different fields. Sun et al. [25] propose a new model consisting of a non-local network for long-term preferences and a geographically extended recurrent neural network for short-term preferences modeling. Similarly, in [26], long-term and short-term preferences are considered. The proposed model utilizes an attention mechanism to mine the former. At the same time, LSTM for the latter implements the following POI recommendation. Yu et al. [27] develop a category-aware depth model, which integrates POI categories and geographical factors to reduce search space. Liu et al. [9] proposed a spatial-temporal recurrent neural network, which uses RNN architecture to learn the sequential transition. These methods mostly adapt past information to interpolate potential POIs, and this information are usually insufficient to identify the missing POIs.

3. **Problem Modelling.** This section first illustrates the problem of missing POI check-in identifiers and then details the problem modeling.

3.1. **Relevant definitions.** Some definitions.
Definition 3.1.1. $U = \{u_1, u_2, \ldots, u_i, \ldots, u_N\}$ is a set of users, $P = \{x_1, x_2, \ldots, x_j, \ldots, x_M\}$ is a set of POIs.
Definition 3.1.2. (POI) POI is a unique location with two attributes: identifiers and geographical coordinates (latitude and longitude). A POI position x_j can be expressed as $\{lat, lon\}$.
Definition 3.1.3. (Check-ins) Check-ins record is defined as all POI positions which the user visit at a definite time. A user's check-in record can be expressed as $C^u = \{x^u_{t_1}, x^u_{t_2}, \ldots, x^u_{t_T}\}$, where T means the numbers of the user's check-in record.

3.2. **Missing POI Identification problem.** Assuming that there exists a missing state $x^u_{t_t}$ at time stamp t_t in the check-ins of user u, the research aims to identify $x^u_{t_t}$. The problem could be expressed by using the check-ins$\left\{x^u_{t_{t-l}}, x^u_{t_{t-l+1}}, \ldots, x^u_{t_{t-1}}\right\}$ before t_t and $\left\{x^u_{t_{t+1}}, x^u_{t_{t+2}}, \ldots, x^u_{t_{t+l}}\right\}$ after t_t to implement the missing POI identification at the time t_t. Where l represents l check-in records before and after target POI check-in.

For example, assuming the trajectory of a user is shown in Figure 1. We focus on identifying the missing POI from 6 p.m. to 10 p.m., using the past and future information of the missing POI to find the most likely POI.

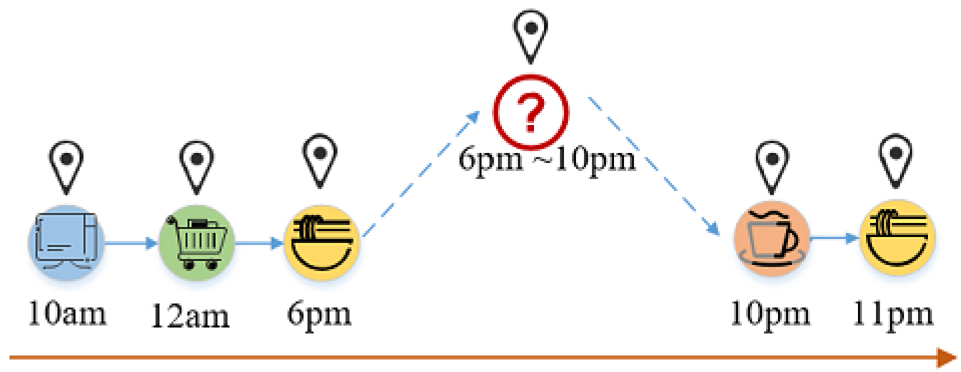

FIGURE 1. Problem Diagram

4. **Model.** Previous studies on POI recommendation or location prediction only focus on the past information records of the target POI, which is limited. The missing POI identification in this work focuses on the past history information and has a close relationship with the future check-in information of the target POI. To this end, we propose a model named Bi-LAF for the missing POI identification.

4.1. **System Framework.** The system framework of the Bi-LAF model is shown in Figure 2. This model combines historical and future two-way check-in records of target missing POI with LSTM and a self-attention mechanism to explore users' preferences. Furthermore, the target time pattern and users' unique features are taken into account for further mining users' preferences. As for the check-ins information after the target POI, the future sequence information reversed process is carried out to make bidirectional check-ins information to target POI.

Considering that the geographical and time interval affect the missing POI, we utilize GCD and time intervals to evaluate the relationships between missing POI and candidate POIs, shown as GCD-T in Figure 2. Finally, the spatial transformation of users' preferences and relationships into the same space accurately identifies missing POI at a specific time. The model's detail will be described next part, including the meaning of some parameters in Figure 2.

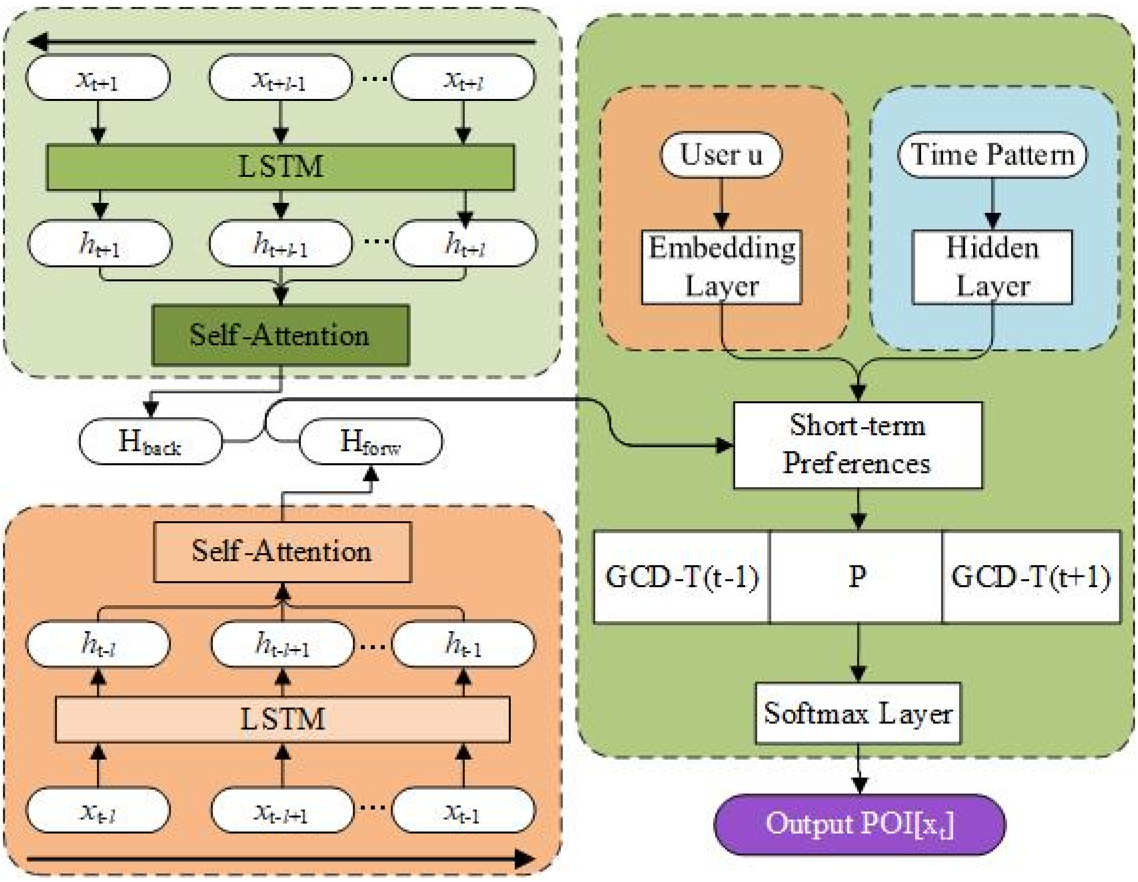

FIGURE 2. Framework of Bi-LAF Model

4.2. POI Feature Learning Method.

4.2.1. *Users' Preference Learning.* Users in this short period have special needs and personalized taste, which we call short-term preferences. We first use a bidirectional LSTM network and self-attention block to capture short-term preferences in this model.

The past check-ins information before the target POI X_{forw} is embedded to get the result $E(X_{forw})$. Then LSTM receives the embedding format of forward details, and we can obtain the corresponding output state H_{forw}. Formulas are expressed as follows:

$$X_{forw} = (x_{t-l}, x_{t-l+1}, \ldots, x_{ff}, \ldots, x_{t-1}) \tag{1}$$

$$E\left(X_{forw}\right)=E\left(x_{t-l},x_{t-l+1},\ldots,x_{ff},\ldots,x_{t-1}\right) \tag{2}$$

$$H_{forw}=LSTM\left(E\left(X_{forw}\right)\right) \tag{3}$$

Where $E\left(x_{ff}\right)\in R^{d}$ represents the embedding representation of the forward sequence POIs, $H_{forw}\in R^{l*d}$ represents the state of the corresponding output, and d defines the unit parameters of the hidden layer.

The future check-ins information after the target POI X_{back} takes out a reverse operation to get $X_{back-inv}$. Then the vector is embedded to obtain $E\left(x_{back}\right)$. Then LSTM receives the embedding format of backwarding information, and we can obtain the corresponding output state H_{back}. Formulas are expressed as follows:

$$X_{back}=\left(x_{t+1},x_{t+2},\ldots,x_{bb},\ldots,x_{t+l}\right) \tag{4}$$

$$X_{back-inv}=\left(x_{t+l},\ldots,x_{bb},\ldots,x_{t+2},x_{t+1}\right) \tag{5}$$

$$E\left(X_{back}\right)=E\left(X_{back-inv}\right) \tag{6}$$

$$H_{back}=LSTM\left(E\left(X_{back}\right)\right) \tag{7}$$

where $E\left(x_{bb}\right)\in R^{d}$, $H_{back}\in R^{l*d}$.

The forward and backward output results obtained from the above calculation introduce self-attention mechanism to learn the important factor degree. The formula is as follows:

$$A_{forw}=softmax\left(W_{2}\tanh\left(W_{1}H_{forw}^{T}+b_{1}\right)+b_{2}\right) \tag{8}$$

$$A_{back}=softmax\left(W_{2}\tanh\left(W_{1}H_{back}^{T}+b_{1}\right)+b_{2}\right) \tag{9}$$

Where $W_{1}\in R^{da*d}$,$W_{2}\in R^{da}$. Weight factors A_{forw}, A_{back} represent the importance of each state from LSTM and da represent hidden dimension.

Then, as for each state H_{forw}, H_{back}, a weighted summation is performed to obtain forward and backward short-term preferences S_{forw}, S_{back}, which we can see in Figure 2.

$$S_{forw}=A_{forw}H_{forw} \tag{10}$$

$$S_{back}=A_{back}H_{back} \tag{11}$$

Combined with the forward-backward output, the result S shows the short-term preferences of users.

$$S=S_{forw}+S_{back} \tag{12}$$

Considering that the user's check-in preferences will change with time, this model combines the time pattern of target POI to learn preferences further. We analyze the time pattern from two dimensions. On the one hand, it is divided into working days and rest days. On the other hand, it is divided into five different periods in a day. The five different periods of the day are as follows : [8 : 00,11 : 30], [11 : 30,14 : 00], [14 : 00,17 : 30], [17 : 30,22 : 00], [22 : 00,8 : 00]. According to the above two dimensions, the target time t is represented by a single-hot encoding of seven dimensions. The coincidence bit is set to 1, and the rest is set to 0. The first two marks the working day and the rest day, and the last five marks the five-time segments within a day. For example, 9: 00 a.m. on July 13th, 2021 (Tuesday, working day) can be expressed as:

$$v_{t}=\left[\begin{array}{lllllll}1 & 0 & 1 & 0 & 0 & 0 & 0\end{array}\right] \tag{13}$$

Considering the different check-in situations of different users, this model sets a matrix to learn user characteristics. It can be expressed as follows:

$$e(u) = E_u^T u \tag{14}$$

In summary, we consider the forward and backward information related to the target missing POI and the time pattern of the target and users' characteristics. Then we convert the three aspects of information into the same space to model the users' preferences.

$$P = W_o \left(e(u) + f \left(W_v v_t \right) + S \right) \tag{15}$$

Where f represents the activation function of tanh, $W_v \in R^{7*d}$, $W_o \in R^{d*d}$, P represents the modeling of users' preferences, which is shown in Figure 2.

4.2.2. *Spatial-Temporal Relationship Modelling.* Considering the effect between geographical distance and time interval related missing POI, we use Great Circle Distance to identify the spatial relationship between target POI and all other POIs. As for temporal influence, missing POI is related to the time interval before and after itself. Time interval and distance should be considered together.

According to the latitude and longitude information of POI position, this paper represents the spatial relationship between POIs based on the shortest distance, namely Great Circle Distance (GCD) [11], as follows:

$$\theta = 2 \sin^{-1} \sqrt{\left(\sin^2 \Delta\alpha + \cos \alpha_1 * \cos \alpha_2 * \sin^2 \Delta\beta \right)} \tag{16}$$

$$GCD = RadiusofEarth(R) * \theta \tag{17}$$

Where θ represents the angle $\Delta\alpha$ of the inner sphere in the unit of a radian, which represents the latitude half difference between two points, namely $(\alpha_1 + \alpha_2)/2$, α_1 represents the latitude of the first point (in the unit of a radian), α_2 represents the latitude of the second point, $\Delta\beta$ represents the longitude half difference between two points, namely $(\beta_1 + \beta_2)/2$, β_1 represents the longitude of the first point (in the unit of a radian), and β_2 represents the longitude of the second point.

Then, according to the great circle distance calculation, the distance relationship between POIs is represented by matrix Q, $Q \in R^{M*M}$.

For the proximity effect of the distance, the geographical factors of the possible candidate points are measured using the check-in information of the previous and subsequent moments of the target missing points, as follows:

$$D_{t-1} = Q(x^u_{t_{t-1}}) \tag{18}$$

$$D_{t+1} = Q(x^u_{t_{t+1}}) \tag{19}$$

Where $D_{t-1} \in R^{1*M}$, $D_{t+1} \in R^{1*M}$ respectively represents the distance matrix between the previous moment x^u_{t-1}, the later moment x^u_{t+1}, and all other candidates' POIs.

The math formula for the time interval is expressed as follows:

$$tt_{t-1} = t_t - t_{t-1} \tag{20}$$

$$tt_{t+1} = t_{t+1} - t_t \tag{21}$$

Where t_t represents the check-in time of target missing POI; t_{t-1} represents the check-in time of the last moment; t_{t+1} represents the check-in time of the next moment; tt_{t-1} and tt_{t+1} represent the forward and backward time intervals, respectively.

Time interval affects the distance. Generally, a shorter time interval leads to a closer geographical distance between target POI and candidate POIs. For the relationship of the two factors above, we transform the time interval into equal dimensions with distance.

$$tt_{t-1} = f\left(W_{t-1} tt_{t-1}\right) \tag{22}$$

$$tt_{t+1} = f\left(W_{t+1} tt_{t+1}\right) \tag{23}$$

Where $W_{t-1} \in R^M$, $W_(t+1) \in R^M$, and f represents the activation function, $tanh$.

Geographical distance and time interval related to missing POI should be taken into account simultaneously. This effect corresponds to $GCD-T(t-1)$ and $GCD-T(t+1)$ in Figure 2. The final spatial-temporal relationship is shown as follows:

$$d_{t-1} = D_{t-1} \odot tt_{t-1} \tag{24}$$

$$d_{t+1} = D_{t+1} \odot tt_{t+1} \tag{25}$$

According to the results of spatial-temporal relationship and users' preference modeling, we can identify the missing POI as follows:

$$output_t^u = softmax\left(P + d_{t-1} + d_{t+1}\right) \tag{26}$$

The obtained result $output_t^u$ represents the possibility of all candidate POIs of the target POI. The possible POI, the higher score of the result.

We need to minimize the cross-entropy of the predicted distribution and the actual distribution.

$$J(\theta) = -\frac{1}{K} \sum_{i=1}^{K} \sum_{j=1}^{M} y_{i,j} \log\left(output_{t,j}^u \mid x_i, \theta\right) \tag{27}$$

where K denotes the number of samples in each model process; M denotes the number of all POIs; y denotes t the actual value of samples after one-hot encoding, and θ denotes the parameter set. This equation represents the cross-entropy calculation between y and the prediction result.

5. **Experiments.**

5.1. **Dataser description.**

This experiment adopts the real LBSN dataset. NYC is a dataset from Foursquare containing all New York City check-ins collected from April 2012 to February 2013(about ten months). There are 1083 users, 38333 POIs, and 227428 check-in records. TKY dataset is similar to the NYC dataset. However, it is collected in Tokyo, including 2293 users, 61858 POI points, and 573703 check-in records. The statistics of the data set are shown in the following table.

TABLE 1. Dataset Statistics

Dataset	Users	POIs	Check-ins
NYC	1083	38333	227428
TKY	2293	61858	573703

We deleted users with less than ten check-ins in the experiment, and POIs visited less than ten times. We sort each user's check-in records by time, with the first 80 % as the training set and the remaining 20 % as the test set.

5.2. **Experiments Details.** For all dataset we use, we take 64 for embedding dimension and hidden dimension in LSTM simultaneously. As for Equation(2) and (4), 64 is considered for POI Embeddding layers. In Equation(8) and (9), hidden dimension take the same value that is 256. While Equation(13) and (14), time pattern and user characteristics, 64 is adapted according to previous work. The most factor which named window width l, 3 is proven to be the best choice. For the other details, minibatch size of 128 and learning rate of 0.001 are used in our model. Whether it's NYC dataset or TKY dataset, we take the same parameters.

5.3. **Evaluation Metrics.** Based on the existing works, we use Recall@k and F1-score@k to evaluate our model and other models. Because Recall@K is positively correlated with Precision@K, we do not use Precision@K here. Recall@K measures the percentage of locations visited in the first K recommended POIs, F1-score@K is a comprehensive indicator, considering Recall@K and Precision@K. The evaluation indicators are defined as follows:

$$Precision@K = -\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{S_i(k)\cap T_i}{k} \tag{28}$$

$$Recall@K = -\tfrac{1}{N}\sum_{i=1}^{N}\tfrac{S_i(k)\cap T_i}{|T_i|} \tag{29}$$

$$F1-score@K = 2*\tfrac{Precision@K*Recall@K}{Precision@K+Recall@K} \tag{30}$$

where $S_i(k)$ is the set of top-k missing POIs predicted by the model, and T_i represents the actual value of the user's current missing POI.

5.4. **Experimental results and comparison.** In order to prove the effectiveness of our model, we will compare it with the following methods.

RNN: This is a neural network method, which directly models the dependence of user order behavior into the prediction process through the recursive structure of RNN.

LSTM: A variant of RNN, which contains a memory unit, an input gate, an output gate, and a forgetting gate. It helps understand long-term dependencies.

GRU: This is another variant of RNN. It has two gating mechanisms, which are more straightforward than LSTM.

STRNN: A model based on RNN, which captures spatial and temporal context through a specific time and distance transformation matrix.

Bi-STDDP: This model captures bidirectional spatial-temporal dependencies and user dynamic preferences to identify missing POI.

Bi-G^{2}AN: This is a new model combining GAN and GRU to explore movement patterns to complete missing POI.

The experimental results are shown in Table 2.

The experimental results of the evaluation metrics in terms of Recall@K and F1-score@K in NYC and TKY datasets show that the proposed method is superior to the above baseline methods. The significant improvement shows that Bi-LAF has an excellent ability to identify missing POI. In addition, we can find that the method based on RNN has acceptable performance on both datasets, which further proves the practical modeling ability of RNN for sequence problems. Compared with the current advanced related research, we can see that on the dataset TKY, compared with the model Bi-STDDP, Recall@1 increased by 8.95 %, Recall@5 increased by 6.24 %, and Recall@10 increased by 4.96 %. Compared with the model Bi-G^{2}AN, Recall@1 increased by 1.29 %, Recall@5 increased by 0.9 %, and Recall@10 increased by 6.74 %. As for NYC datasets, Bi-LAF is also superior, obtaining the improvements of 34.02% for Recall@1, 33.17% for Recall@5, 31.44% for Recall@10 compared with Bi-STDDP. And then, 30.43% for Recall@1, 24.47% for Recall@5, 23.84% for Recall@10 compared with Bi-G^{2}AN.

Table 2. Evaluation Results in terms of Recall @K, Fl-score@K

Dataset	Method	Recall@1	Recall@5	Recall@10	Fl-score@1	Fl-score@5	Fl-score@10
NYC	RNN	0.1308	0.3105	0.3859	0.1308	0.1035	0.0701
	LSTM	0.1353	0.3033	0.3739	0.1353	0.1011	0.068
	GRU	0.1340	0.3182	0.3956	0.1340	0.1060	0.0719
	Bi-STDDP	0.1781	0.3445	0.4106	0.1781	0.1148	0.0746
	Bi-G^{2}AN	0.1830	0.3686	0.4358	0.1830	0.1229	0.0792
	Bi-LAF	**0.2387**	**0.4588**	**0.5397**	**0.2387**	**0.1529**	**0.0979**
TKY	RNN	0.1345	0.3048	0.3785	0.1345	0.1016	0.0688
	LSTM	0.1325	0.3073	0.3836	0.1325	0.1024	0.0697
	GRU	0.1352	0.3246	0.4081	0.1352	0.1092	0.0742
	Bi-STDDP	0.2000	0.4148	0.4991	0.2000	0.1382	0.0907
	Bi-G^{2}AN	0.2145	0.4367	0.4908	0.2145	0.1456	0.0892
	Bi-LAF	**0.2179**	**0.4407**	**0.5239**	**0.2179**	**0.1468**	**0.0953**

The experimental results show that our model Bi-LAF performs better on the list with a higher ranking. Compared with the model, Bi-G^{2}AN shows that Bi-LAF has a slight advantage in Recall@1 and Recall@5 and performs better in Recall@10. We can see that the advantages of the model Bi-G^{2}AN are evident, which is worth learning. Our model show more excellent performance on smaller datasets than the baseline methods. All in all, our model uses forward and backward sequence information combined with LSTM and self-attention block to capture users' short-term preferences, which models the relationship by bidirectional spatial-temporal information and performs better on the identification task of missing POI.

5.5. **Discussion.** Our model considers the forward and backward sequence information of the target missing POI. We take out a reversed operation for the future sequence to make the bidirectional information tend to the direction of the target missing POI. In order to verify the feasibility of this operation, we compare normal and reversed operations on the dataset NYC, and take Recall@K as the evaluation metrics with the range of $K \in [1,10]$ with keeping other optimal parameters unchanged. The experimental results are shown in Figure 3. Realizing reversed operation satisfies the tendency of two-way sequence information, which has an excellent promoting effect on modeling users' preferences.

We also change the value of the forward and backward sequence information window width l of the proposed model on the NYC dataset to explore the influence of l. Figure 4 shows the detailed Recall@K performance. Considering fewer users whose check-in is fewer, we set the front and rear window widths as $[2, 3, 4]$ to cover all records thoroughly to test our model. Moreover, the results of different window widths are slightly different. When the window width is 3, Recall@K performs better, indicating that in the dataset NYC, the missing POI of users is more related to the previous three check-in histories. So we choose $l = 3$ as the window width.

Considering the influence of embedding dimension (the parameter d mentioned above) on the experiment, we change the embedding dimension to $[32, 48, 64, 80, 96, 112, 128]$ and complete the test experiment on the dataset NYC. We can see that with the increase of embedding dimension, the model's performance is gradually improved. When the embedding dimension d is greater than 64, the performance shows a downward trend. The

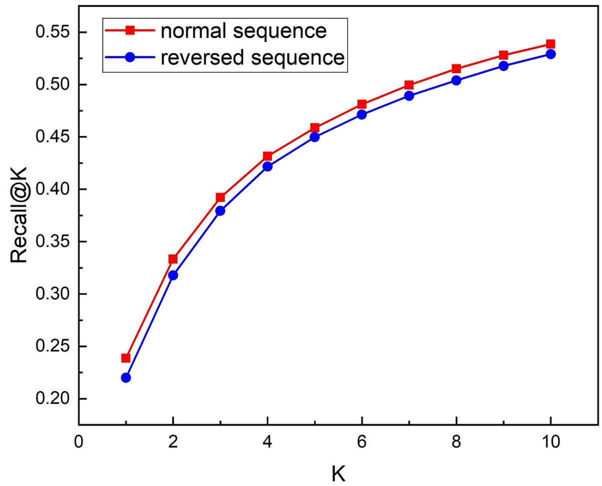

FIGURE 3. Influence of the future sequence by different operation

embedding dimension determines the complexity and ability of the model. A smaller embedding dimension may not fully adapt to the data distribution, and a larger embedding dimension will increase the complexity and computational cost of the model. The appropriate embedding dimension helps to achieve the best embedding performance. Finally, we choose $d = 64$ as the embedding dimension.

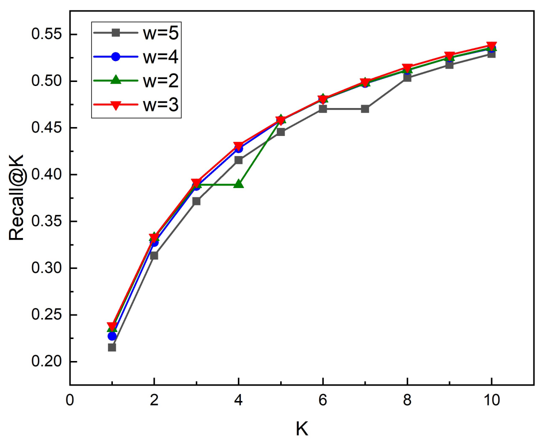

FIGURE 4. Impact of different window widths

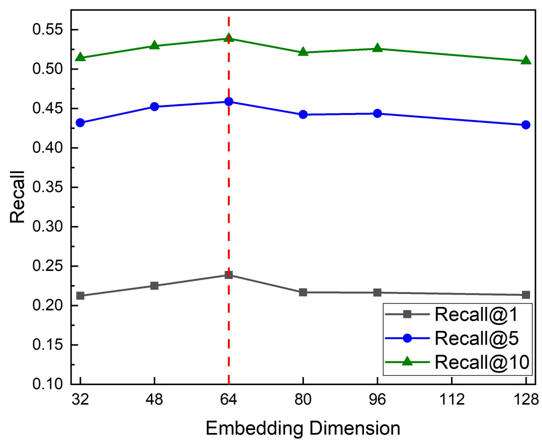

FIGURE 5. Impact of different embedding dimensions

6. **Conclusions.** In this paper, we focus on the identification and completion task of missing POI. Missing the front and rear sequence information of POI points, we identify the place users have visited before and after the signing time of missing POI. This identifying benefit to distinguish the difference from the POI recommendation task in the current field. The POI recommendation task predicts or recommends where users may visit in the future by using positive historical sequence information. In order to solve this problem, we propose a bidirectional model Bi-LAF. Specifically, in this work, we first encode the forward and backward information of POI and then employ LSTM and self-attention block to learn users' short-term preferences. Finally, we integrate target time patterns and user

person features to learn users' preferences further. Then the bidirectional spatial information and time interval related missing POI are combined with learning spatial-temporal features. Finally, integrating preference and spatial-temporal features implements the identification of missing POI. Extensive experimental results on two large-scale real data sets show that the proposed model outperforms the state-of-the-art methods.

Acknowledgment. Portions of this research were funded through the projects of the National Natural Science Foundation of China (41971340,41471333,61304199), projects of Fujian Provincial Department of Science and Technology (2021Y4019, 2020D002, 2020L3014, 2019I0019), and the Foundation of Fujian Provincial Universities Key Laboratory of Industrial Control and Data Analysis (Fujian University of Technology) (KF-X19013)

REFERENCES

[1] M. Baratchi, M. Afsharchi, and F. Crestani, LGLMF: Local Geographical Based Logistic Matrix Factorization Model for POI Recommendation, *15th Asia Information Retrieval Societies Conference (AIRS 2019)*, pp. 66-78, 2020.

[2] C. Cheng, H. Yang, I. King, and M. Lyu, Fused matrix factorization with geographical and social influence in location-based social networks, *Proceedings of the AAAI conference on artificial intelligence*, vol. 26, no. 1, pp. 17-23, 2012.

[3] K. M. Lang and T. D. Little, Principled missing data treatments, *Prevention Science*, vol. 19, no. 3, pp. 284-294, 2018.

[4] T. N. Tu, An Improving Using MapReduce Model in Predicting Learning Ability of Pupils Based on Bayes Classification Algorithm, *Journal of Information Hiding and Multimedia Signal Processing*, vol. 12, no. 3, pp. 140-151, 2021.

[5] L. Kang, R. S. Chen, N. Xiong, Y. C. Chen, Y. X. Hu, and C. M. Chen, Selecting hyper-parameters of Gaussian process regression based on non-inertial particle swarm optimization in Internet of Things, *IEEE Access*, vol. 7, pp. 59504-59513, 2019.

[6] J. Yoon, J. Jordon, and M. Schaar, Gain: Missing data imputation using generative adversarial nets *35th International Conference on Machine Learning*, 2018, https://www.vanderschaar-lab.com/papers/ICML_GAIN.pdf

[7] D. Kong and F. Wu, HST-LSTM: A Hierarchical Spatial-Temporal Long-Short Term Memory Network for Location Prediction, *27th International Joint Conference on Artificial Intelligence (IJCAI-18)*, pp. 2341-2347, 2018.

[8] R. Fu, Z. Zhang, and L. Li, Using LSTM and GRU neural network methods for traffic flow prediction, *31st Youth Academic Annual Conference of Chinese Association of Automation (YAC 2016)*, pp. 324-328, 2016.

[9] Q. Liu, S. Wu, L. Wang, and T. Tan, Predicting the next location: A recurrent model with spatial and temporal contexts, *30th AAAI conference on artificial intelligence*, pp. 194-200, 2016.

[10] D. Xi, F. Zhuang, Y. Liu, J. Gu, H. Xiong, and Q. He, Modelling of bi-directional spatio-temporal dependence and users' dynamic preferences for missing poi check-in identification, *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 5458-5465, 2019.

[11] M. Shi, D. Shen, Y. Kou, T. Nie, and G. Yu, Missing POI Check-in Identification Using Generative Adversarial Networks, *International Conference on Database Systems for Advanced Applications*, Springer, Cham, pp. 575-590, 2021.

[12] A. Baskar, Simple Single and Multi-Facility Location Models using Great Circle Distance, *ITM Web of Conferences*, pp. 1-9, 2021.

[13] R. W. Krause, M. Huisman, C. Steglich, and T. A. Snijders, Missing network data a comparison of different imputation methods, *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pp. 159-163, 2018.

[14] P. Raja and K. Thangavel, Missing value imputation using unsupervised machine learning techniques, *Soft Computing*, vol. 24, no. 6, pp. 4361-4392, 2020.

[15] N. Mustafa, A. O. Ibrahim, A. Ahmed, and A. Abdullah, Collaborative filtering: Techniques and applications, *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, pp. 1-6, 2017.

[16] X. Tang, H. Yao, Y. Sun, C. Aggarwal, P. Mitra, and S. Wang, Joint modeling of local and global temporal dynamics for multivariate time series forecasting with missing values, *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 5956-5963, 2020.

[17] S. Cheng and F. Lu, A two-step method for missing spatio-temporal data reconstruction, *ISPRS International Journal of Geo-Information*, vol. 6, no. 7, 187, 2017.

[18] D. Lian, Y. Ge, F. Zhang, N. L. Yuan, X. Xie, T. Zhou and Y. Rui, Content-aware collaborative filtering for location recommendation based on human mobility data, *2015 IEEE international conference on data mining*, pp. 261-270, 2015.

[19] Z. Zhang, Y. Liu, Z. Zhang, and B. Shen, Fused matrix factorization with multi-tag, social and geographical influences for POI recommendation, *World Wide Web*, vol. 22, no. 3, pp. 1135-1150, 2019.

[20] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, Factorizing personalized markov chains for next-basket recommendation, *Proceedings of the 19th international conference on World wide web*, pp. 811-820, 2010.

[21] Y. C. Chen, T. Thaipisutikul, and T. K. Shih, A learning-based POI recommendation with spatiotemporal context awareness, *IEEE Transactions on Cybemetics*, 2020, https://doi.org/10.1109/TCYB.2020.3000733

[22] S. Kumar, A. Damaraju, A. Kumar, S. Kumari, and C. Chen, LSTM Network for Transportation Mode Detection, *Journal of Internet Technology*, vol. 22, no. 4, pp. 891-902, 2021.

[23] J. M. T. Wu, Z. Li, N. Herencsar, B. Vo, and J. C. W. Lin, A graph-based CNN-LSTM stock price prediction algorithm with leading indicators, *Multimedia Systems*, 2021, https://doi.org/10.1007/s00530-021-00758-w

[24] J. M. T. Wu, M. E. Wu, P. J. Hung, M. M. Hassan, G. Fortino, Convert index trading to option strategies via LSTM architecture, *Neural Computing and Applications*, 2020, https://doi.org/10.1007/s00521-020-05377-6

[25] K. Sun, T. Qian, T. Chen, Y. Liang, Q. V. H. Nguyen, and H. Yin, 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*, pp. 214-221, 2020.

[26] Y. Wu, K. Li, G. Zhao, and X. Qian, Long-and short-term preference learning for next POI recommendation, *Proceedings of the 28th ACM international conference on information and knowledge management*, pp. 2301-2304, 2019.

[27] F. Yu, L. Cui, W. Guo, X. Lu, Q. Li, and H. Lu, A category-aware deep model for successive poi recommendation on sparse check-in data, *Proceedings of the web conference 2020*, pp. 1264-1274, 2020.