



Exploring nonlinear spatiotemporal effects for personalized next point-of-interest recommendation

Xi SUN, Zhimin LV[‡]

Collaborative Innovation Center of Steel Technology, University of Science and Technology Beijing, Beijing 100083, China

E-mail: b20190537@xs.ustb.edu.cn; lvzhimin@necar.ustb.edu.cn

Received July 15, 2022; Revision accepted Feb. 23, 2023; Crosschecked Aug. 28, 2023

Abstract: Next point-of-interest (POI) recommendation is an important personalized task in location-based social networks (LBSNs) and aims to recommend the next POI for users in a specific situation with historical check-in data. State-of-the-art studies linearly discretize the user's spatiotemporal information and then use recurrent neural network (RNN) based models for modeling. However, these studies ignore the nonlinear effects of spatiotemporal information on user preferences and spatiotemporal correlations between user trajectories and candidate POIs. To address these limitations, a spatiotemporal trajectory (STT) model is proposed in this paper. We use the long short-term memory (LSTM) model with an attention mechanism as the basic framework and introduce the user's spatiotemporal information into the model in encoding. In the process of encoding information, an exponential decay factor is applied to reflect the nonlinear drift of user interest over time and distance. In addition, we design a spatiotemporal matching module in the process of recalling the target to select the most relevant POI by measuring the relevance between the user's current trajectory and the candidate set. We evaluate the performance of our STT model with four real-world datasets. Experimental results show that our model outperforms existing state-of-the-art methods.

Key words: Point-of-interest recommendation; Spatiotemporal effects; Long short-term memory (LSTM); Attention mechanism
<https://doi.org/10.1631/FITEE.2200304>

CLC number: TP181

1 Introduction

With the popularity of information technology and mobile Internet, location-based social networks (LBSNs) have developed rapidly, leading to the attention of some services based on them, such as Brightkite, GyPSii, and Foursquare. A great amount of check-in data, such as time, location, and coordinates, is generated in the process of users using these services, which lays the foundation for exploring users' travel patterns and recommending the best point-of-interest (POI) for users. POI recommendation creates value for both users and enterprises. It can not only help users discover interesting locations and services in the city but also

help enterprises in LBSNs advertise for users. Therefore, the potential value of POI recommendation has attracted extensive attention from the academic world.

However, unlike other types of recommendation systems (music, movies, etc.), the next POI recommendation process is faced with highly correlated user behavior data, and user preferences are more sensitive to spatiotemporal information (Gao et al., 2013; Yuan et al., 2013; Feng SS et al., 2015; Xu S et al., 2020). To capture the sequentiality of user travel trajectories, some studies have improved the recommendation effect by integrating the sequential information of user trajectories. For example, factorized personalized Markov chain (FPMC) (Rendle et al., 2010) model obtained the sequential influence between consecutive check-in's of users by combining matrix factorization and the Markov chain (MC). However, when considering the correlation of users' long-term sequential

[‡] Corresponding author

ORCID: Xi SUN, <https://orcid.org/0000-0002-6339-5498>; Zhimin LV, <https://orcid.org/0000-0002-7313-5796>

© Zhejiang University Press 2023

behavior, the Markov model performs poorly. Later, some recurrent neural network (RNN) based models and the memory mechanism were proposed to alleviate this problem, enabling the model to take into account both the long- and short-term interest evolutions of users (Liu Q et al., 2016b; Feng J et al., 2018; Ying et al., 2018). User travel trajectories are not only regular in sequence but also including some obscure periodic rules that can be further reflected by spatiotemporal information (Qian et al., 2019). Some studies have changed the structure of the sequence model to comprehensively consider spatiotemporal information, improving the accuracy to a certain extent (Yang DQ et al., 2015). For the sparsity of spatiotemporal information, some studies (Lian et al., 2020; Yang DQ et al., 2020; Zhao KZ et al., 2020) linearly discretized the information into several intervals and achieved good results.

The above studies indicate that there are still some key problems to be solved in the analysis of user travel patterns.

1. Temporal and spatial distance information has not been learned effectively in many studies due to the nonlinear effects on user preferences. From a temporal perspective, user’s travel trajectory is periodic in the short term, as shown in Fig. 1. Most users leave their residence every morning, work or relax in a specific area during the day, eat nearby in the afternoon,

and go home at night. User travel patterns also exhibit long-term periodicity, with variations in activity region, travel purpose, and duration between weekends and weekdays. Moreover, the decay of user interest over time is usually also a nonlinear process. From a spatial perspective, as shown in Fig. 2, user’s check-in’s are generally concentrated around one or more central locations, and the density decreases rapidly with increasing distance from the central location, indicating that user’s check-in preference is nonlinearly affected by geographical distance.

2. Previous models ignore the spatiotemporal correlation between user’s check-in trajectories and the candidate location sets. However, a user’s current preference is not only related to his/her historical behavior but also limited by time and region. That is, users are more inclined to visit places that are closer to them, and the later the check-in time, the more likely they are to go to a closer location.

Therefore, this study proposes the spatiotemporal trajectory (STT) model to recommend the next POI for users. The STT model consists of three parts: the interest extractor layer generates the representation of user features, the interest evolution layer is used to learn user travel patterns, and the spatiotemporal matching (STM) layer matches the best POI for users. In summary, this paper contributes as follows:

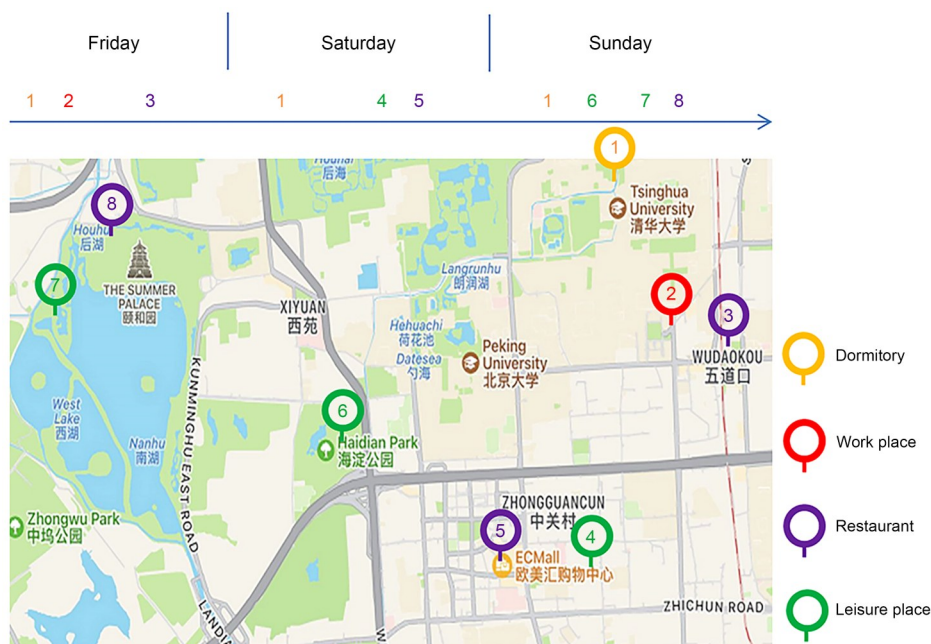


Fig. 1 Spatiotemporal trajectory example showing user check-in rules

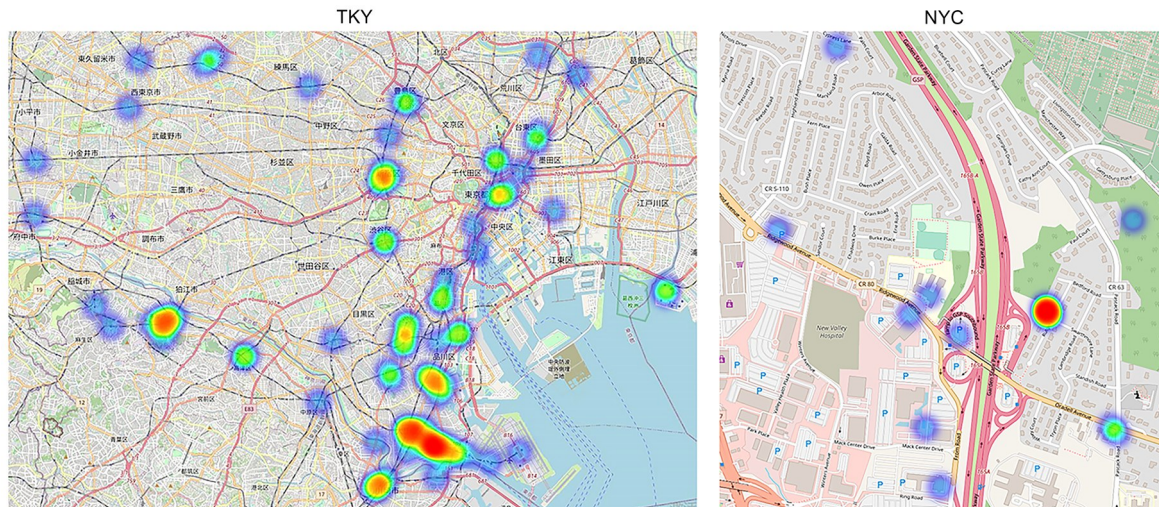


Fig. 2 Heatmap of the geographic distribution of check-in's by two real users

1. We propose the STT model, an improved form of the long short-term memory (LSTM) model with an attention mechanism (AM), to fully consider the spatiotemporal information of user's historical trajectories for aggregating relevant locations. In this model, we devise a new encoding method that helps the sequence model learn the STT characteristics of users.

2. We introduce an information decay factor, which can be used as a weight in the process of spatiotemporal encoding. The decay factor can attenuate the information that is less relevant to the current spatiotemporal situation of the user, allowing the model to focus on more valuable information in a nonlinear fashion.

3. We design an STM architecture for the spatiotemporal correlation of user historical trajectories with all candidate locations. This architecture matches the latent feature of the user STT with all the spatial representations of the candidate sets, and thus, it can find the most suitable POIs.

4. We conduct experiments on three real datasets to demonstrate the superiority of our proposed STT model. The results show that the proposed STT is 10.79% more accurate than state-of-the-art models on average.

2 Related works

In this section, we introduce some research on traditional POI recommendation and POI recommen-

dation with spatiotemporal information. The next POI recommendation is essentially a specific sequential recommendation task.

2.1 POI recommendation

Traditional POI recommendations mostly use collaborative filtering (CF) methods (Jiang et al., 2015; Yang C et al., 2017), among which matrix factorization is the most common application (Salakhutdinov and Mnih, 2007; Abdollahi and Nasraoui, 2016; He et al., 2017; Xu CH et al., 2021). The matrix factorization method decomposes the "user–location" relationship matrix into two submatrices to represent the latent features of users and locations. However, the matrix factorization method does not consider the continuity of user behavior and thus is not suitable for modeling sequence information.

The next POI recommendation is an emerging recommendation mode, which is more challenging than traditional modes due to its sensitivity to sequence information. The methods for the next POI recommendation are mainly divided into two categories: MC-based methods and RNN-based methods. MC-based methods aim to predict the user's next POI based on the past sequence of user behaviors. Due to the sparsity of the location check-in dataset, it is difficult for traditional MC-based methods to learn the transition of intermittent check-in actions. To address this problem, FPMC (Rendle et al., 2010) model was proposed to combine matrix factorization techniques with MCs to model user preferences and sequence information.

Based on the FPMC model, Lian et al. (2013) predicted the next check-in location according to the user's long- and short-term tendencies. Generally, MC-based model is used to explore the transition probability between consecutive actions. Recently, since neural network techniques have achieved promising results in many fields, such as language modeling (Mikolov et al., 2011) and machine vision (Ren et al., 2022), RNN-based models have been considered to be the most effective method for learning nonlinear relationships between users and locations, and they are more extensible than MC-based models. Liu Q et al. (2016b) proposed a spatiotemporal recurrent neural network (ST-RNN) method, which extends the RNN structure to capture more detailed sequence information. DeepMove (Feng J et al., 2018) uses an improved gate recurrent unit (GRU) network to capture mixed embeddings of multiple check-in information and designs a historical attention module to enhance the model's perception of check-in sequence regularities. However, the success of RNN-based models in the field of natural language processing is because the recognition of text sequences generally does not need to consider temporal information, meaning that the distance between each time step of RNN is equal by default. Therefore, RNN-based models do not correspond to human behavior patterns that are highly correlated with spatiotemporal information, which makes RNN not perfect.

2.2 POI recommendation with spatiotemporal information

Spatial information is a key factor affecting users' behavior, which limits users' travel scope. Zeng et al. (2021) introduced a region factor into the CF algorithm to help the model find the most likely preferences in the user's vicinity. Liu Q et al. (2016a) added distance constraints on the basis of FPMC to achieve desirable results. Yin et al. (2017) proposed the spatial smoothing method to fully consider collective travel interests and individual travel interests, which alleviates the problem of user check-in sparsity to a certain extent. Si et al. (2019), Zhao GS et al. (2020), and Wu et al. (2022) focused on the personalized feature extraction of location attributes to analyze the specific trajectory of each user. Some studies consider the transfer spatial features between

two POIs that users check consecutively. Cui et al. (2019) proposed the Distance2Pre model, which uses the GRU model for the first time to integrate the user's preferences of different geographic distances to predict the next POI, and proposed two structural models that integrate the user's distance preference scores in linear and nonlinear ways. Long- and short-term preference modeling (LSTPM) (Sun et al., 2020) designs a geo-dilated RNN model to focus on users' recent travel interests, but it completely ignores users' long-term travel interests. Attentional recurrent neural network (ARNN) (Guo et al., 2020) uses the knowledge graph method to mine user associations based on geographic locations and uses association information as prior knowledge of RNN input to help the model converge better. Some studies address the sparsity problem of spatiotemporal information by discretely dividing the spatial areas into hierarchical grids (Lian et al., 2020; Yang DQ et al., 2020; Zhao KZ et al., 2020). However, they ignore the explicit modeling of geographic distances because linear spatial partitioning does not capture actual metric information well.

Temporal information is of great significance for mining user travel patterns due to the periodicity (Zhao SL et al., 2016), consecutiveness (Ye et al., 2010; Liu Q et al., 2016b), and interest drift (Wang et al., 2003; Fan, 2004) of user behavior. Yuan et al. (2013) proposed an improved CF model for POI recommendation, which is the first POI recommendation research considering the time factor. Yuan et al. (2014) then proposed the geographical-temporal influences aware graph (GTAG) model to integrate check-in contextual information such as location and time to obtain more accurate recommendation results. Liu YC et al. (2016) proposed a Where and When to go (WWO) recommendation system that uses time interval in users' trajectories when modeling user behavior. Some deep learning based models have been proposed to capture temporal information in user check-in records. Zhao WX et al. (2018) proposed a TA-TEM model based on distributed representation learning to integrate multiple types of temporal information into the model using a unified approach. STRNN (Zhang et al., 2019) improves the RNN model by transferring temporal information between each time step of the RNN in the form of time-specific transition matrices. STGN (Zhao PP et al., 2022) further enhances the LSTM structure

by adding spatiotemporal gates to make better use of temporal information. Inspired by the sequence model SASRec (Kang and McAuley, 2018), TiSASRec and STAN convert timestamps into time intervals and apply the self-attention mechanism to further consider the temporal regularity of user check-in trajectories (Li et al., 2020; Luo et al., 2021).

However, the studies mentioned above do not consider the nonlinear attenuation of the effect of spatiotemporal information on user behavior. In addition, they ignore the spatiotemporal correlation between user historical trajectories and candidate sets.

3 Preliminaries

We first introduce the problem statement and definitions in this paper. We define $\mathcal{U} = \{u_1, u_2, \dots, u_U\}$, $\mathcal{V} = \{v_1, v_2, \dots, v_V\}$, and $\mathcal{T} = \{t_1, t_2, \dots, t_T\}$ as the sets of U users, V POIs, and T timestamps, respectively. Each user u_i has a spatiotemporal check-in sequence of POIs represented as $\text{Tr}(u_i) = \{l_1, l_2, \dots, l_{m_i}\}$, where m_i is the total check-in number of user u_i . For each user check-in, we define it as (u_i, v_k, t_k) , which is a tuple containing three elements u_i , v_k , and t_k representing the i^{th} user, the k^{th} location of u_i , and the k^{th} timestamp of u_i . Then, we model temporal intervals and geographical distances as the spatiotemporal correlation between two consecutive check-in's of the user. Let the temporal interval sequence of user u_i be $\Delta T_i = \{\Delta t_1, \Delta t_2, \dots, \Delta t_{m_i-1}\}$ and the spatial interval sequence of user u_i be $\Delta S_i = \{\Delta s_1, \Delta s_2, \dots, \Delta s_{m_i-1}\}$. The spatial interval of two locations can be calculated by the great circle distance method (Chen et al., 2012), which measures the shortest route between two locations on Earth's sphere. Consequently, the coordinates of l_k are represented as $(\text{lat}_k, \text{lon}_k)$, and the spatial interval between l_k and l_{k+1} is calculated as

$$\Delta s_k = 2R \cdot \arcsin \left(\sin^2 \left(\frac{\text{lat}_{k+1} - \text{lat}_k}{2} \right) + \cos(\text{lat}_{k+1}) \cos(\text{lat}_k) \sin^2 \left(\frac{\text{lon}_{k+1} - \text{lon}_k}{2} \right) \right). \quad (1)$$

In addition to the basic information of user historical POI visits, we consider explicit relevance between the user trajectory sequence and the candidate

locations in this paper. We denote a spatial distance matrix, which measures the spatial interval between each candidate POI $p \in [1, V]$ and each POI of the user trajectory $q \in [1, m]$ as $M_{p,q}$. The spatial interval matrix is represented as

$$\mathbf{M} = \begin{bmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,m} \\ M_{2,1} & M_{2,2} & \dots & M_{2,m} \\ \vdots & \vdots & \dots & \vdots \\ M_{V,1} & M_{V,2} & \dots & M_{V,m} \end{bmatrix}. \quad (2)$$

Given the user trajectory $\{l_1, l_2, \dots, l_m\}$, the candidate locations $\{v_1, v_2, \dots, v_V\}$, and the spatial interval matrix \mathbf{M} , the goal of the next POI recommendation is to predict the most likely upcoming location $v \in l_{m+1}$ for the user.

4 Framework overview

In this section, we detail the techniques and specific processes in the model, consisting of (1) the interest extractor layer that extracts the representation of user's STT, (2) the interest evolution layer that learns sequential preference by LSTM and AM, and (3) the STM layer that measures the correlation between user's STTs and candidate sets and then calculates the softmax probability of user preference for candidate POIs. The technical modules and specific processes of the STT framework are shown in Fig. 3.

4.1 Interest extractor layer

4.1.1 Trajectory embedding

Trajectory features count a great deal in the POI recommendation task, consisting of two sparse features and a continuous feature. We use the common embedding technique (Cheng et al., 2016) to transform the large-scale sparse features into low-dimensional dense vectors to reduce computation and improve representation. For sparse features, we denote the embedded representations of users and locations as $e_u \in \mathbb{R}^d$ and $e_v \in \mathbb{R}^d$, respectively. For continuous features, we first divide continuous timestamps into T discrete intervals in hours and then denote their embedded representations as $e_t \in \mathbb{R}^d$. Thus, the input dimensions of the embeddings e_u , e_v , and e_t are U , V , and T , respectively. The embedding of each user check-in l is represented as

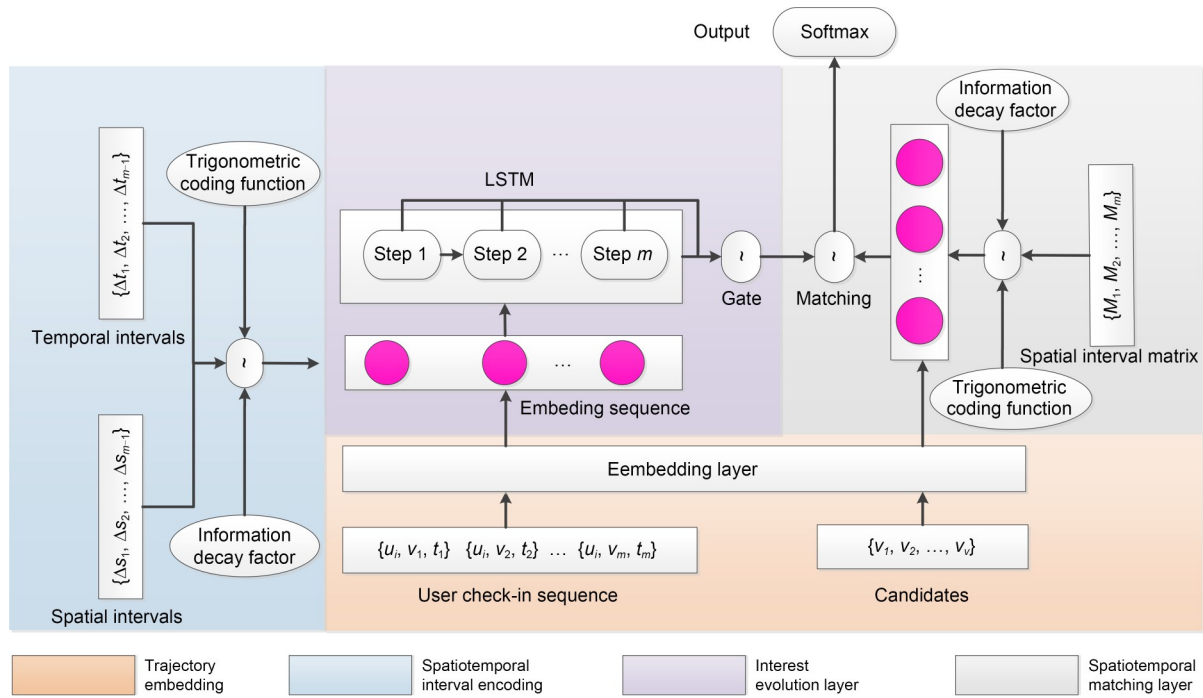


Fig. 3 Spatiotemporal trajectory (STT) model (References to color refer to the online version of this figure)

$e_i = e_u + e_v + e_t \in \mathbb{R}^d$. Consequently, for embedding the sequence of user historical POI visits, we denote it as $E_u = \{e_{i_1}, e_{i_2}, \dots, e_{i_m}\}$.

4.1.2 Spatiotemporal interval encoding

Spatiotemporal factors play an important role in the evolution of people’s travel preferences. To enhance the model’s perception of spatiotemporal information, we introduce a spatiotemporal interval encoding method that uses the trigonometric function (Gehring et al., 2017) to integrate temporal interval and geographic distance into the model in the form of dense vectors. Compared with traditional embedding methods, our proposed encoding method is more interpretable and can make the spatiotemporal features easier to learn by the sequence model.

First, to facilitate the subsequent calculation of matrices for continuous temporal and spatial intervals, we partition the temporal intervals ΔT_i and spatial intervals ΔS_i into discrete bins. In the partitioning process, we introduce an information decay factor to reflect nonlinear information attenuation. The decay factor uses an exponential function inspired by the Ebbinghaus forgetting curve, which can reflect the decay rate of things with concept drift (Aggarwal et al., 2004)

features such as human memory. The formula of the temporal decay factor (TDF) is as follows:

$$f(\Delta t) = e^{-\frac{\Delta t}{T_0}} \tag{3}$$

According to the above formula, TDF $f(\Delta t)$ is positively related to the temporal interval Δt , and $f(\Delta t) \in (0, 1)$. T_0 determines the rate at which user interest drifts over time, and it needs to be adjusted for the specific situation. Therefore, temporal intervals can be properly partitioned into discrete temporal bins. The formula is as follows:

$$\text{pos}_{u,q}^T = \lceil P_{\text{num}}^T (1 - f(\Delta t_{u,q})) \rceil, \tag{4}$$

where $\text{pos}_{u,q}^T$ is the temporal bin index of temporal interval Δt_q for user u , P_{num}^T indicates the upper limit of the temporal bin index, and it can be manually adjusted according to accuracy needs and time budget. The mathematical symbol $\lceil \cdot \rceil$ represents the ceiling operation, which rounds a number up to the nearest integer. Accordingly, a high index of the temporal bin represents a long temporal interval between user POI visits. Similarly, we can obtain the discrete spatial bins and the spatial decay factor (SDF) as follows:

$$f(\Delta s) = e^{-\frac{\Delta s}{S_0}}, \quad (5)$$

$$\text{pos}_{u,q}^S = \lceil P_{\text{num}}^S (1 - f(\Delta s_{u,q})) \rceil. \quad (6)$$

Then, we calculate the high-dimensional representation of spatiotemporal intervals according to the bin index. Gehring et al. (2017) introduced several encoding methods for converting discrete values into dense vectors, in which we chose the position encoding of trigonometric functions. Thus, the spatiotemporal interval information of the user trajectory can be concatenated to represent itself in the sequence model. For temporal interval encoding (TIE), the formulae are as follows:

$$\text{TIE}_{(\text{pos}_{u,q}^T, 2i)} = \sin\left(\text{pos}_{u,q}^T / 1000^{\frac{2i}{d_T}}\right), \quad (7)$$

$$\text{TIE}_{(\text{pos}_{u,q}^T, 2i+1)} = \cos\left(\text{pos}_{u,q}^T / 1000^{\frac{2i}{d_T}}\right). \quad (8)$$

Similarly, we can obtain the spatial interval encoding (SIE) formulae:

$$\text{SIE}_{(\text{pos}_{u,q}^S, 2i)} = \sin\left(\text{pos}_{u,q}^S / 1000^{\frac{2i}{d_S}}\right), \quad (9)$$

$$\text{SIE}_{(\text{pos}_{u,q}^S, 2i+1)} = \cos\left(\text{pos}_{u,q}^S / 1000^{\frac{2i}{d_S}}\right), \quad (10)$$

where d represents the dimension of the encoded vector and i represents the dimension of the encoded vector. It can be seen from the calculation process that the odd-numbered dimension value and the even-numbered dimension value of the encoded vector are calculated by the cosine function and the sine function, respectively. Therefore, $\text{TIE}_{(\text{pos}_{u,q}^T, 2i)}$ and $\text{SIE}_{(\text{pos}_{u,q}^S, 2i)}$ can be obtained by their linear dependencies with $\text{TIE}_{(\text{pos}_{u,q}^T, 2i+1)}$ and $\text{SIE}_{(\text{pos}_{u,q}^S, 2i+1)}$ according to the characteristics of trigonometric functions.

Finally, we obtain the TIE sequence $\mathbf{TR}_T = \{\mathbf{TIE}_1, \mathbf{TIE}_2, \dots, \mathbf{TIE}_m\}$ and the SIE sequence $\mathbf{TR}_S = \{\mathbf{SIE}_1, \mathbf{SIE}_2, \dots, \mathbf{SIE}_m\}$, whose first items are both zero vectors. Therefore, the representation set of the user check-in sequence $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ can be obtained by concatenating the corresponding items in \mathbf{E}_u , \mathbf{TR}_T , and \mathbf{TR}_S . Moreover, we use the same method to convert the spatial interval matrix \mathbf{M} to the SIE tensor \mathbf{E}_M as follows:

$$\mathbf{E}_M = \begin{bmatrix} \mathbf{e}_{M_{1,1}} & \mathbf{e}_{M_{1,2}} & \cdots & \mathbf{e}_{M_{1,m}} \\ \mathbf{e}_{M_{2,1}} & \mathbf{e}_{M_{2,2}} & \cdots & \mathbf{e}_{M_{2,m}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}_{M_{r,1}} & \mathbf{e}_{M_{r,2}} & \cdots & \mathbf{e}_{M_{r,m}} \end{bmatrix}. \quad (11)$$

4.2 Interest evolution layer

In the interest evolution layer, we extend AM to the LSTM model in our approach as our basic framework, which not only has excellent time-series data mining capabilities, but also facilitates additional extensions based on prior knowledge of the recommendation task.

4.2.1 LSTM

In the POI recommendation system, the user check-in behavior is the carrier of latent interest, and the user's interest will change after he/she visits several POIs. In the interest evolution layer, we extract a series of interest states from the sequential user visiting trajectory. For the balance between efficiency and performance, we take LSTM to model the dependency between check-in's, where the input of the LSTM is the latent representation of user check-in's sorted by occurrence time. LSTM (Hochreiter and Schmidhuber, 1997) overcomes the vanishing gradient problem of RNN and has achieved good results in time-series related studies. The update equations of LSTM are as follows:

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i), \quad (12)$$

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f), \quad (13)$$

$$\tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c), \quad (14)$$

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t, \quad (15)$$

$$\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o), \quad (16)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t), \quad (17)$$

where " \odot " is the elementwise product, σ is the sigmoid activation function, \mathbf{i}_t , \mathbf{f}_t , and \mathbf{o}_t are the input, forget, and output gates of the t^{th} time step, aiming to constrain the information to store, forget, and output, respectively, \mathbf{x}_t is the t^{th} input feature vector of LSTM, and represents the t^{th} POI that the user checks in, \mathbf{h}_t is the t^{th} hidden state, \mathbf{W}_i , \mathbf{W}_f , \mathbf{W}_o , and \mathbf{W}_c are the weight matrices for the input, forget, output, and candidate memory cells, respectively, and \mathbf{c}_t is the activated vector representing the cell state, which plays an

important role in information transfer between each time step in the LSTM.

4.2.2 Attention mechanism

However, due to the diversity of the user visiting patterns and the integrated spatiotemporal information, the last hidden state \mathbf{h}_m , which captures only the dependency between behaviors, cannot represent interest effectively in the long term. Therefore, we use AM to filter critical check-in information in the user's long-term travel trajectories. The contribution of each check-in behavior is estimated as a normalized weight, which can be formulated as

$$\alpha_t = \frac{\exp(\mathbf{h}_t \mathbf{W}_\alpha)}{\sum_j^m \exp(\mathbf{h}_j \mathbf{W}_\alpha)}. \quad (18)$$

Then, the final representation vector of the t^{th} check-in is calculated as follows:

$$\mathbf{g}_t = \alpha_t \mathbf{h}_t, \quad (19)$$

where α_t is the attention score of the t^{th} check-in of the user, and $\mathbf{W}_\alpha \in \mathbb{R}^{\text{NH} \times 1}$ is the transition matrix for the attention layer, where NH represents the dimension of the hidden state.

4.3 Spatiotemporal matching layer

The interest evolution layer above, similar to most other methods in the field of POI recommendation, focuses on mining users' historical behavior patterns but ignores the features of candidate sets to a certain extent. Here, we propose the spatiotemporal matching layer to recall candidates with the highest probability from V candidate POIs by measuring the spatiotemporal correlation between user trajectories and candidate sets. Given the encoded temporal intervals \mathbf{TR}_T , the encoded spatial intervals \mathbf{TR}_S and the latent trajectory features $G = \{\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_m\}$, we obtain the user's STT representation $\mathbf{STE}_u = [\mathbf{e}_{st_1}, \mathbf{e}_{st_2}, \dots, \mathbf{e}_{st_m}]$ by concatenating them. Given the embedded candidate location matrix $\mathbf{E}_v = [\mathbf{e}_{v_1}, \mathbf{e}_{v_2}, \dots, \mathbf{e}_{v_p}] \in \mathbb{R}^{V \times d}$ and the encoded spatial interval tensor $\mathbf{E}_M \in \mathbb{R}^{V \times m \times d_s}$, we obtain the encoding of the candidate spatial relation matrix $\mathbf{E}_{MS} \in \mathbb{R}^{V \times m \times (d_s + d)}$ by m repeated concatenating operations. Therefore, the spatiotemporal features of each user's trajectory and the spatial distance information of each candidate location are well represented by \mathbf{STE}_u and \mathbf{E}_{MS} , respectively,

which is important for discovering their spatiotemporal correlations. In this layer, we calculate the probability of each candidate POI being the next real check-in for user u as follows:

$$O(u) = \text{softmax}(\mathbf{W}_s \text{Matching}(\mathbf{E}_{MS}, \mathbf{STE}_u)), \quad (20)$$

with

$$\text{Matching}(\mathbf{A} \in \mathbb{R}^{o \times p \times q}, \mathbf{B} \in \mathbb{R}^{p \times q}) = [\mathbf{A}_{:,1,:}, \mathbf{B}_{1,:}, \mathbf{A}_{:,2,:}, \mathbf{B}_{2,:}, \dots, \mathbf{A}_{:,p,:}, \mathbf{B}_{p,:}]^T. \quad (21)$$

Here, parameter matrix $\mathbf{W}_s \in \mathbb{R}^m$ converts the dimension of $O(u)$ to \mathbb{R}^V . The key step to obtain the correlation between candidate sets and user trajectories is Matching, in which the updated check-in embeddings all contribute to the matching of each candidate POI.

5 Experiments

We use real datasets to experimentally demonstrate the effectiveness of the STT model for the next POI recommendation in this section. First, we detail the datasets, baseline methods, evaluation metrics, and settings for model training and testing. Next, we present our empirical results for a quantitative fair comparison with those of baseline models. Finally, we experimentally verify the effectiveness of different techniques in the model.

5.1 Datasets

We use real user check-in datasets from four regions of New York City (NYC), Tokyo City (TKY), Singapore (SIN), and Gowalla for model training and testing. Table 1 presents the basic statistics in each dataset. For the datasets, we first extract the check-in records in units of users. Each user's check-in includes user ID, venue ID, latitude, longitude, and timestamp. Next, we filter out users with more than five records and locations visited more than five times. Then, we slice

Table 1 Dataset statistics

| Category | Number | | | |
|------------|---------|---------|---------|-----------|
| | NYC | TKY | SIN | Gowalla |
| Users | 1064 | 2245 | 2032 | 53 008 |
| POIs | 5136 | 7872 | 3662 | 121 944 |
| Check-in's | 147 939 | 447 571 | 179 721 | 3 302 414 |

NYC: New York City; TKY: Tokyo City; SIN: Singapore

the trajectory of each user with a fixed-length window. For each user that has m check-in's, we divide a dataset into $m-3$ training sets, with the first $m' \in [1, m-3]$ check-in's as input and the $(m'+1)^{\text{th}}$ check-in as the label; we use the $(m-1)^{\text{th}}$ check-in as the validation label to tune parameters; the test label uses the last check-in of the trajectory to evaluate the performance of the proposed model.

5.2 Baseline models

We briefly introduce the following state-of-the-art methods for the next POI recommendation, which will be compared with our proposed model.

ST-RNN (Liu Q et al., 2016b) is a variant RNN model that incorporates spatiotemporal information by introducing a single transition matrix during the time step transition.

DeepMove (Feng J et al., 2018) is a novel approach that designs a historical attention module to assist the model in representing latent features of user behavior.

STGN (Zhao PP et al., 2022) innovates the gating mechanism of LSTM to facilitate the use of spatiotemporal information.

ARNN (Yin et al., 2017) is a state-of-the-art model that explores heterogeneous neighbors based on meta-path random walks, and the model uses RNN with an AM for features in heterogeneous neighbors.

LSTPM (Sun et al., 2020) uses an improved GRU network with a historical attention module to enhance the model's ability to represent latent features of user behavior.

TiSASRec (Li et al., 2020) is a state-of-the-art self-attention-based method considering time interval, and can simultaneously model the absolute position of items in the sequence and time interval between items.

GeoSAN (Lian et al., 2020) is a state-of-the-art method that uses a novel loss function optimization method based on importance sampling to address sparsity by emphasizing the use of negative samples.

STAN (Luo et al., 2021) is a novel model that uses a self-attention mechanism to consider the interaction between nonadjacent POIs and nonconsecutive check-in's.

5.3 Evaluation matrices

To compare our model with baseline models and adjust parameters, we adopt the top- K recall rates that

are commonly applied in previous works (Kingma and Ba, 2014; Liu Q et al., 2016b). Recall@ K counts the rate of true positive samples in all positive samples, which in our case measures the presence of the correct POI among the top- K recommended POIs. It is calculated as follows:

$$\text{Recall}@K = \frac{\text{tp}_u}{\text{tp}_u + t_{\text{nu}}}, \quad (22)$$

where tp_u represents the total number of POIs in both the future check-in trajectory of user u and the top- K results produced by our proposed model, and t_{nu} represents the number of POIs in the future check-in trajectory of the user but not in the top- K results.

5.4 Experimental setting

Here we detail the setting of the training parameters. We use the balanced sampler in our model to reduce computing resources and follow STAN recommendation to set the embedding dimension for user, location, and time to 50 (Luo et al., 2021); the optimizer is the Adam optimizer with default betas (Kingma and Ba, 2014); the learning rate is set to 0.0003; the maximum of the bin index $P_{\text{num}}^{\text{T}}$ and $P_{\text{num}}^{\text{S}}$ is set to 100 and 50, respectively; the training epoch is 50 and the fixed length for user's check-in sequence is 100. To obtain a more reasonable experiment, we fine-tune some key parameters and test the stability against the number of negative samples using the validation set in Section 5.7.

5.5 Recommendation performance

We adopt Recall@5 and Recall@10 to evaluate the performance of the models. Table 2 shows the results of all baseline models on four datasets. To ensure the fairness of the experiment, we take the mean recall of 100 experimental results.

Our STT model outperforms all baselines on all datasets and gains 11.45% Recall@5 and 10.27% Recall@10 improvements on average against the strongest baseline STAN. Among the baseline models, self-attention-based models such as GeoSAN and STAN significantly outperform RNN variants. This is not surprising, as RNN-based models suffer from vanishing gradients, which toss long-term patterns. Note that we use an AM to help STT overcome gradient descent

so that STT can effectively capture the sequential regularity of spatiotemporal trajectories while considering the long-term pattern. Moreover, reasonable use of spatiotemporal information through the decay factor and full consideration of candidate sets make our model better than the other methods.

5.6 Ablation study

Since there are many components in our proposed architecture, we validate the contributions of different components through an ablation study. We briefly describe the role of each component of the model as follows:

TDF determines whether the information decays with time;

SDF determines whether the information decays with geographic distance;

TIE characterizes the temporal intervals as high-dimensional vectors in the model;

SIE characterizes the spatial intervals as high-dimensional vectors in the model;

AM helps STT capture the long-term behavioral characteristics of users through spatiotemporal information;

STM enables the model to consider the spatiotemporal relationship between user trajectories and candidate sets when recommending to users.

We remove the modules mentioned above in turn and obtain some variants. The performance of variant STT models on all four datasets is shown in Table 3. We find that in the process of POI recommendation, spatial distances and temporal intervals have a strong role in promoting the sequence model, and they both improve the accuracy by nearly 2%. The introduction of TDF and SDF can help the model use the spatiotemporal information more reasonably and further improve the accuracy of the model by almost 1%–3%. AM is necessary for sequence models, especially in learning long-term

Table 2 Evaluation of baseline models in terms of Recall@5 and Recall@10 on four datasets

| Model | Recall@5 | | | | Recall@10 | | | |
|--------------|----------|--------|--------|---------|-----------|--------|--------|---------|
| | NYC | TKY | SIN | Gowalla | NYC | TKY | SIN | Gowalla |
| STRNN | 23.65% | 18.36% | 17.91% | 16.64% | 28.02% | 27.91% | 20.16% | 25.67% |
| DeepMove | 32.68% | 26.84% | 23.89% | 19.59% | 40.14% | 35.09% | 31.55% | 26.99% |
| STGN | 24.39% | 19.40% | 22.92% | 15.28% | 30.15% | 27.10% | 27.27% | 24.22% |
| ARNN | 19.70% | 18.52% | 18.17% | 18.10% | 34.83% | 26.96% | 25.38% | 27.45% |
| LSTPM | 27.91% | 25.68% | 25.79% | 20.15% | 35.64% | 33.10% | 33.27% | 27.01% |
| TiSASRec | 36.64% | 30.31% | 29.63% | 24.11% | 50.20% | 36.93% | 37.53% | 35.46% |
| GeoSAN | 40.06% | 29.57% | 33.97% | 27.64% | 52.67% | 37.40% | 39.43% | 36.45% |
| STAN | 46.69% | 34.61% | 37.51% | 30.16% | 59.62% | 42.64% | 43.01% | 39.98% |
| STT | 62.42% | 37.28% | 37.83% | 31.23% | 72.88% | 46.41% | 45.42% | 41.74% |
| Improvement* | 33.69% | 7.71% | 0.85% | 3.55% | 22.24% | 8.84% | 5.60% | 4.40% |

NYC: New York City; TKY: Tokyo City; SIN: Singapore. * Recall@5 and Recall@10 of our STT model compared with those of STAN

Table 3 Ablation analysis (Recall@5 and Recall@10) on four datasets

| Variant | Recall@5 | | | | Recall@10 | | | |
|------------|----------|--------|--------|---------|-----------|--------|--------|---------|
| | NYC | TKY | SIN | Gowalla | NYC | TKY | SIN | Gowalla |
| Default | 62.42% | 37.28% | 37.83% | 31.23% | 72.88% | 46.41% | 45.42% | 41.74% |
| Remove TDF | 62.01% | 36.92% | 37.72% | 30.25% | 71.24% | 45.83% | 44.26% | 41.32% |
| Remove SDF | 61.10% | 33.27% | 35.97% | 30.19% | 71.09% | 45.10% | 43.72% | 41.20% |
| Remove TIE | 59.78% | 32.62% | 35.18% | 29.88% | 69.67% | 44.56% | 43.46% | 40.46% |
| Remove SIE | 57.91% | 30.98% | 34.81% | 27.94% | 69.64% | 43.34% | 42.34% | 38.05% |
| Remove AM | 54.69% | 30.25% | 31.55% | 21.33% | 63.33% | 39.37% | 39.38% | 35.47% |
| Remove STM | 34.76% | 13.43% | 20.37% | 15.44% | 41.71% | 17.92% | 29.45% | 26.56% |

NYC: New York City; TKY: Tokyo City; SIN: Singapore

patterns of user preferences, and it improves the recall rates by almost 1%–6%. STM contributes the most to the model by helping the model use the relevant information of the candidate set, which brings approximately 10%–20% accuracy improvement to our model.

5.7 Stability analysis

1. The dimension of the hidden state h . To investigate the impact of dimensionality on STT, we illustrate the Recall@10 performance of STT on all datasets with varying dimensionality in Fig. 4. It is obvious that the performance of our STT is sensitive to h on all datasets. We find that the recall value of STT stays stable in a large range on all datasets after a sharp increase, and the most suitable parameters can be selected as $h=50$ in NYC, $h=60$ in TKY, $h=90$ in SIN, and $h=80$ in Gowalla.

2. T_0 and S_0 , which determine the temporal decay rate and the spatial decay rate, respectively. According to our knowledge of the decay function, we vary the value of T_0 from 0.5×10^5 to 5×10^5 with step 50000 and the value of S_0 from 20 to 200 with step 20 in the STT on all datasets. Fig. 5 shows that $T_0=1 \times 10^5$ is the best value for NYC, $T_0=3.5 \times 10^5$ is the best value for TKY, $T_0=3 \times 10^5$ is the best value for SIN, and $T_0=2.5 \times 10^5$ is the best value for Gowalla. This indicates that the travel interests of people living in different regions change at different speeds over time. For the parameter S_0 , our best result is to take 100 in the NYC dataset, 20 in the TKY dataset, 80 in the SIN dataset, and 140 in the Gowalla dataset, which is obvious in Fig. 6. We find that in different cities, people's travel interests are different in sensitivity to distance, which

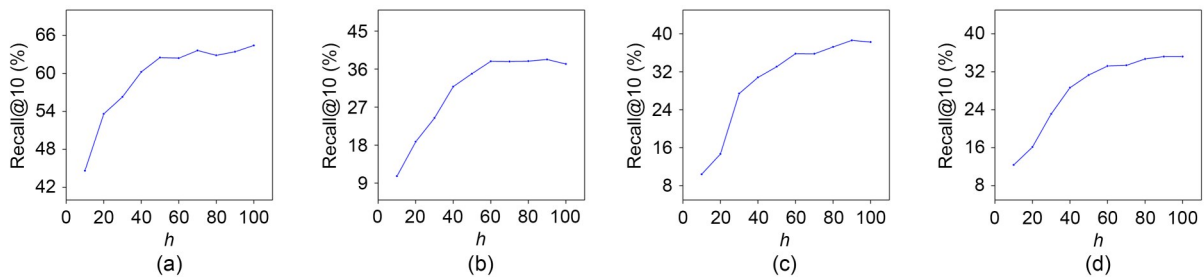


Fig. 4 Effect of embedding dimension h on Recall@10: (a) New York City; (b) Tokyo City; (c) Singapore; (d) Gowalla

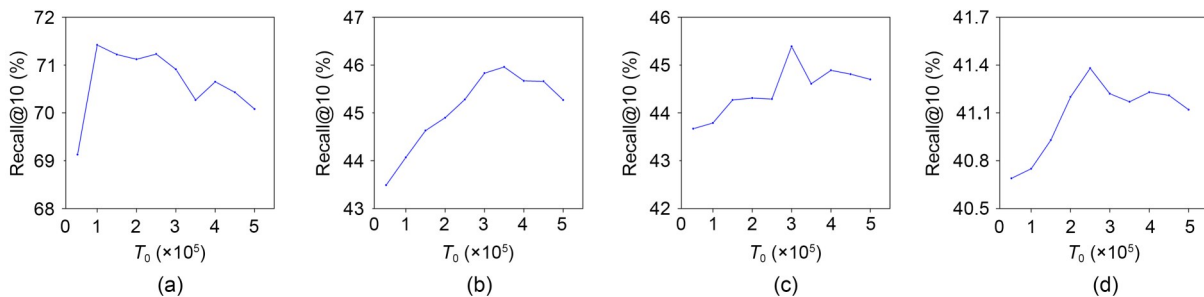


Fig. 5 Effect of T_0 on Recall@10: (a) New York City; (b) Tokyo City; (c) Singapore; (d) Gowalla

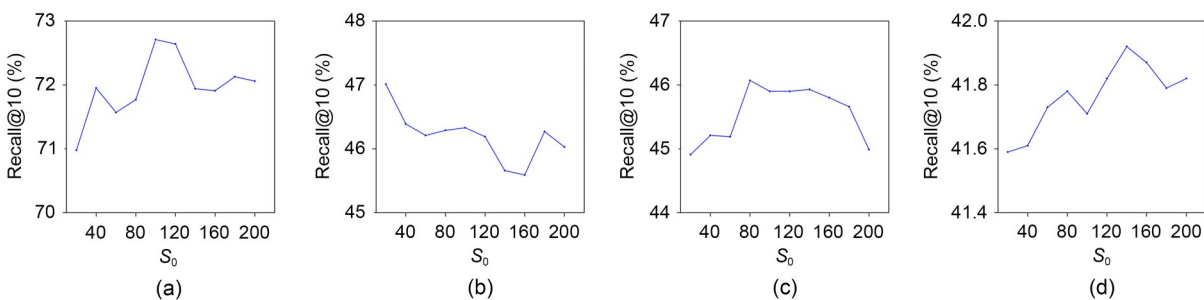


Fig. 6 Effect of S_0 on Recall@10: (a) New York City; (b) Tokyo City; (c) Singapore; (d) Gowalla

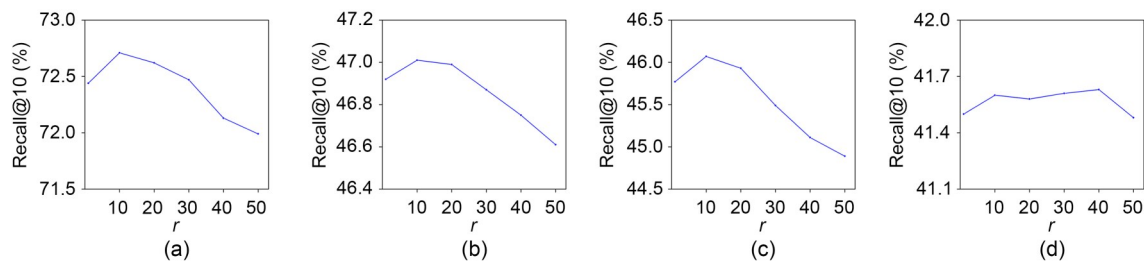


Fig. 7 Effect of r on Recall@10: (a) New York City; (b) Tokyo City; (c) Singapore; (d) Gowalla

is due to the different distribution of locations in different cities.

3. The number of negative samples r . We use the validation set to test the performance of the model when r is set to 1, 10, 20, 30, 40, and 50, and the results are shown in Fig. 7. Obviously, an appropriate amount of negative sample sampling can promote the convergence of the training process. Moreover, larger datasets seem to accept more negative samples, but more than a certain number will be counter-productive.

6 Conclusions

In this paper, we propose a novel next POI recommendation method called STT that effectively considers the nonlinear spatiotemporal influence on user trajectories. Specifically, we design an encoding method to introduce temporal and spatial interval sequences into our model. During the encoding process, we introduce a decay factor to help the model focus on valuable spatiotemporal information in a nonlinear way. In the process of recalling the target, we propose a new spatiotemporal matching method exploring the best candidate for users. Moreover, we use an attention mechanism to help the model mine long-term patterns of users based on their spatiotemporal trajectories. We perform comprehensive parameter filtering, ablation study, and comparisons with baseline models in the experimental section. Experimental results demonstrate that our STT model improves the accuracy to a great extent compared with state-of-the-art models. In practice, our results can bring commercial value to websites and applications based on personalized POI recommendation services.

There are also some valuable directions for our method that deserve further study. First, in addition to

spatiotemporal information, other current state information is important for user's interest, such as user age, weather, and season. Hence, we plan to exploit other user and environment state information in POI recommendation. Second, user preferences in a specific state are usually greatly influenced by POI types. Therefore, proper use of category information in recommendation is promising.

Contributors

Xi SUN and Zhimin LV designed the research and processed the data. Xi SUN drafted the paper. Zhimin LV helped organize the paper. Xi SUN and Zhimin LV revised and finalized the paper.

Compliance with ethics guidelines

Xi SUN and Zhimin LV declare that they have no conflict of interest.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

- Abdollahi B, Nasraoui O, 2016. Explainable restricted Boltzmann machines for collaborative filtering. <https://doi.org/10.48550/arXiv.1606.07129>
- Aggarwal CC, Han JW, Wang JY, et al., 2004. A framework for projected clustering of high dimensional data streams. In: Nascimento MA, Özsu MT, Kossmann D, et al. (Eds.), Proc VLDB Conf Elsevier, Amsterdam, p.852-863. <https://doi.org/10.1016/B978-012088469-8.50075-9>
- Chen KS, Xu ZH, Gong LY, 2012. Research of the distance calculation algorithm based on RSSI. *Adv Mater Res*, 499: 464-468. <https://doi.org/10.4028/www.scientific.net/AMR.499.464>
- Cheng HT, Koc L, Harmsen J, et al., 2016. Wide & deep learning for recommender systems. Proc 1st Workshop on Deep Learning for Recommender Systems, p.7-10. <https://doi.org/10.1145/2988450.2988454>
- Cui Q, Tang YY, Wu S, et al., 2019. Distance2Pre: personalized

- spatial preference for next point-of-interest prediction. 23rd Pacific-Asia Conf on Advances in Knowledge Discovery and Data Mining, p.289-301.
https://doi.org/10.1007/978-3-030-16142-2_23
- Fan W, 2004. Systematic data selection to mine concept-drifting data streams. Proc 10th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining, p.128-137.
<https://doi.org/10.1145/1014052.1014069>
- Feng J, Li Y, Zhang C, et al., 2018. DeepMove: predicting human mobility with attentional recurrent networks. Proc World Wide Web Conf, p.1459-1468.
<https://doi.org/10.1145/3178876.3186058>
- Feng SS, Li XT, Zeng YF, et al., 2015. Personalized ranking metric embedding for next new POI recommendation. Proc 24th Int Conf on Artificial Intelligence, p.2069-2075.
- Gao HJ, Tang JL, Hu X, et al, 2013. Exploring temporal effects for location recommendation on location-based social networks. Proc 7th ACM Conf on Recommender Systems, p.93-100. <https://doi.org/10.1145/2507157.2507182>
- Gehring J, Auli M, Grangier D, et al., 2017. Convolutional sequence to sequence learning. 34th Int Conf on Machine Learning, p.1243-1252.
- Guo Q, Sun Z, Zhang J, et al., 2020. An attentional recurrent neural network for personalized next location recommendation. Proc 34th AAAI Conf on Artificial Intelligence, p.83-90. <https://doi.org/10.1609/aaai.v34i01.5337>
- He XN, Liao LZ, Zhang HW, et al., 2017. Neural collaborative filtering. Proc 26th Int Conf on World Wide Web, p.173-182. <https://doi.org/10.1145/3038912.3052569>
- Hochreiter S, Schmidhuber J, 1997. Long short-term memory. *Neur Comput*, 9(8):1735-1780.
<https://doi.org/10.1162/neco.1997.9.8.1735>
- Jiang SH, Qian XM, Shen JL, et al., 2015. Author topic model-based collaborative filtering for personalized POI recommendations. *IEEE Trans Multim*, 17(6):907-918.
<https://doi.org/10.1109/TMM.2015.2417506>
- Kang WC, McAuley J, 2018. Self-attentive sequential recommendation. IEEE Int Conf on Data Mining, p.197-206.
<https://doi.org/10.1109/ICDM.2018.00035>
- Kingma DP, Ba J, 2014. Adam: a method for stochastic optimization. <https://doi.org/10.48550/arXiv.1412.6980>
- Li JC, Wang YJ, McAuley J, 2020. Time interval aware self-attention for sequential recommendation. Proc 13th Int Conf on Web Search and Data Mining, p.322-330.
<https://doi.org/10.1145/3336191.3371786>
- Lian DF, Zheng VW, Xie X, 2013. Collaborative filtering meets next check-in location prediction. Proc 22nd Int Conf on World Wide Web, p.231-232.
<https://doi.org/10.1145/2487788.2487907>
- Lian DF, Wu YJ, Ge Y, et al., 2020. Geography-aware sequential location recommendation. Proc 26th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining, p.2009-2019. <https://doi.org/10.1145/3394486.3403252>
- Liu Q, Wu S, Wang DY, et al., 2016a. Context-aware sequential recommendation. IEEE 16th Int Conf on Data Mining, p.1053-1058. <https://doi.org/10.1109/ICDM.2016.0135>
- Liu Q, Wu S, Wang L, et al., 2016b. Predicting the next location: a recurrent model with spatial and temporal contexts. Proc 30th AAAI Conf on Artificial Intelligence, p.194-200.
<https://doi.org/10.1609/aaai.v30i1.9971>
- Liu YC, Liu CR, Liu B, et al., 2016. Unified point-of-interest recommendation with temporal interval assessment. Proc 22nd ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining, p.1015-1024.
<https://doi.org/10.1145/2939672.2939773>
- Luo YT, Liu Q, Liu ZC, 2021. STAN: spatio-temporal attention network for next location recommendation. Proc Web Conf, p.2177-2185. <https://doi.org/10.1145/3442381.3449998>
- Mikolov T, Kombrink S, et al., 2011. Extensions of recurrent neural network language model. IEEE Int Conf Acoustics Speech and Signal Processing, p.5528-5531.
<https://doi.org/10.1109/ICASSP.2011.5947611>
- Qian TY, Liu B, Nguyen QVH, et al., 2019. Spatiotemporal representation learning for translation-based POI recommendation. *ACM Trans Inform Syst*, 37(2):1-24.
<https://doi.org/10.1145/3295499>
- Ren Z, Fang F, et al., 2022. State of the art in defect detection based on machine vision. *Int J Precis Eng Manuf-Green Techn*, 9(2): 661-691.
<https://doi.org/10.1007/s40684-021-00343-6>
- Rendle S, Freudenthaler C, Schmidt-Thieme L, 2010. Factorizing personalized Markov chains for next-basket recommendation. Proc 19th Int Conf on World Wide Web, p.811-820. <https://doi.org/10.1145/1772690.1772773>
- Salakhutdinov R, Mnih A, 2007. Probabilistic matrix factorization. Proc 20th Int Conf on Neural Information Processing Systems, p.1257-1264.
- Si YL, Zhang FZ, Liu WY, 2019. An adaptive point-of-interest recommendation method for location-based social networks based on user activity and spatial features. *Knowl-Based Syst*, 163:267-282.
<https://doi.org/10.1016/j.knosys.2018.08.031>
- Sun K, Qian TY, Chen T, et al., 2020. Where to go next: modeling long- and short-term user preferences for point-of-interest recommendation. Proc 34th AAAI Conf on Artificial Intelligence, p.214-221.
<https://doi.org/10.1609/aaai.v34i01.5353>
- Wang HX, Fan W, Yu PS, et al., 2003. Mining concept-drifting data streams using ensemble classifiers. Proc 9th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining, p.226-235. <https://doi.org/10.1145/956750.956778>
- Wu YX, Li K, Zhao GS, et al., 2022. Personalized long- and short-term preference learning for next POI recommendation. *IEEE Trans Knowl Data Eng*, 34(4):1944-1957.
<https://doi.org/10.1109/TKDE.2020.3002531>
- Xu CH, Ding AS, Zhao KD, 2021. A novel POI recommendation method based on trust relationship and spatial-temporal factors. *Electr Commer Res Appl*, 48:101060.
<https://doi.org/10.1016/j.elerap.2021.101060>
- Xu S, Cao JX, Legg P, et al., 2020. Venue2Vec: an efficient embedding model for fine-grained user location prediction in geo-social networks. *IEEE Syst J*, 14(2):1740-1751.
<https://doi.org/10.1109/JSYST.2019.2913080>
- Yang C, Bai LX, Zhang C, et al., 2017. Bridging collaborative filtering and semi-supervised learning: a neural approach for POI recommendation. Proc 23rd ACM SIGKDD Int Conf on

- Knowledge Discovery and Data Mining, p.1245-1254.
<https://doi.org/10.1145/3097983.3098094>
- Yang DQ, Zhang DQ, Zheng VW, et al., 2015. Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Trans Syst Man Cybern Syst*, 45(1):129-142.
<https://doi.org/10.1109/TSMC.2014.2327053>
- Yang DQ, Fankhauser B, Rosso P, et al., 2020. Location prediction over sparse user mobility traces using RNNs: flashback in hidden states! Proc 29th Int Joint Conf on Artificial Intelligence, p.2184-2190.
<https://doi.org/10.24963/ijcai.2020/302>
- Ye M, Yin PF, Lee WC, 2010. Location recommendation for location-based social networks. Proc 18th SIGSPATIAL Int Conf on Advances in Geographic Information Systems, p.458-461. <https://doi.org/10.1145/1869790.1869861>
- Yin HZ, Wang WQ, Wang H, et al., 2017. Spatial-aware hierarchical collaborative deep learning for POI recommendation. *IEEE Trans Knowl Data Eng*, 29(11):2537-2551.
<https://doi.org/10.1109/TKDE.2017.2741484>
- Ying HC, Zhuang FZ, Zhang FZ, et al., 2018. Sequential recommender system based on hierarchical attention network. Proc 27th Int Joint Conf on Artificial Intelligence, p.3926-3932. <https://doi.org/10.24963/ijcai.2018/546>
- Yuan Q, Cong G, Ma ZY, et al., 2013. Time-aware point-of-interest recommendation. Proc 36th Int ACM SIGIR Conf on Research and Development in Information Retrieval, p.363-372. <https://doi.org/10.1145/2484028.2484030>
- Yuan Q, Cong G, Sun AX, 2014. Graph-based point-of-interest recommendation with geographical and temporal influences. Proc 23rd ACM Int Conf on Information and Knowledge Management, p.659-668.
<https://doi.org/10.1145/2661829.2661983>
- Zeng J, Tang HR, Zhao YZ, et al., 2021. PR-RCUC: a POI recommendation model using region-based collaborative filtering and user-based mobile context. *Mobile Netw Appl*, 26(6):2434-2444.
<https://doi.org/10.1007/s11036-021-01782-w>
- Zhang T, Zheng WM, Cui Z, et al., 2019. Spatial-temporal recurrent neural network for emotion recognition. *IEEE Trans Cybern*, 49(3):839-847.
<https://doi.org/10.1109/TCYB.2017.2788081>
- Zhao GS, Lou PL, Qian XM, et al., 2020. Personalized location recommendation by fusing sentimental and spatial context. *Knowl-Based Syst*, 196:105849.
<https://doi.org/10.1016/j.knosys.2020.105849>
- Zhao KZ, Zhang Y, Yin HZ, et al., 2020. Discovering subsequence patterns for next POI recommendation. Proc 29th Int Joint Conf on Artificial Intelligence, p.3216-3222.
<https://doi.org/10.24963/ijcai.2020/445>
- Zhao PP, Luo AJ, Liu YC, et al., 2022. Where to go next: a spatio-temporal gated network for next POI recommendation. *IEEE Trans Knowl Data Eng*, 34(5):2512-2524.
<https://doi.org/10.1109/TKDE.2020.3007194>
- Zhao SL, Zhao T, Yang HQ, et al., 2016. STELLAR: spatial-temporal latent ranking for successive point-of-interest recommendation. Proc 30th AAAI Conf on Artificial Intelligence, p.315-322. <https://doi.org/10.1609/aaai.v30i1.9986>
- Zhao WX, Zhou NN, Sun AX, et al., 2018. A time-aware trajectory embedding model for next-location recommendation. *Knowl Inform Syst*, 56(3):559-579.
<https://doi.org/10.1007/s10115-017-1107-4>