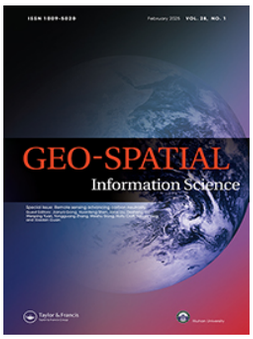


Geo-spatial Information Science

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tgsi20

Individual mobility prediction by considering current traveling features and historical activity chain

Xiaotong Zhang, Zhipeng Gui, Yuhang Liu, Dehua Peng, Qianxi Lan, Zhangxiao Shen, Huan Chen, Yuhui Zuo, Yao Yao, Huayi Wu, Kai Li & Kun Qin

To cite this article: Xiaotong Zhang, Zhipeng Gui, Yuhang Liu, Dehua Peng, Qianxi Lan, Zhangxiao Shen, Huan Chen, Yuhui Zuo, Yao Yao, Huayi Wu, Kai Li & Kun Qin (04 Feb 2025): Individual mobility prediction by considering current traveling features and historical activity chain, Geo-spatial Information Science, DOI: 10.1080/10095020.2025.2455005

To link to this article: https://doi.org/10.1080/10095020.2025.2455005

Published online: 04 Feb 2025.

GEO-SPATIAL INFORMATION SCIENCE
https://doi.org/10.1080/10095020.2025.2455005

Individual mobility prediction by considering current traveling features and historical activity chain

Xiaotong Zhang [a,b], Zhipeng Gui [a,c,d,e], Yuhang Liu [c], Dehua Peng [a], Qianxi Lan [a], Zhangxiao Shen [c], Huan Chen [c], Yuhui Zuo [c], Yao Yao [f,g], Huayi Wu [c,d], Kai Li [h] and Kun Qin [a]

[a]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China; [b]Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Hong Kong, China; [c]State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China; [d]Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan, China; [e]School of Geography and Planning, Ningxia University, Yinchuan, China; [f]School of Geography and Information Engineering, China University of Geosciences, Wuhan, China; [g]Center for Spatial Information Science, The University of Tokyo, Chiba, Japan; [h]Climate Change and Energy Economics Study Center, School of Economics and Management, Wuhan University, Wuhan, China

ABSTRACT

Individual mobility prediction forecasts traveling activities of an individual traveler, and has wide applications in location-based services, public health, and transportation planning. Whereas, it remains challenging due to the complexity and uncertainty of human mobility. Existing methods mainly consider spatiotemporal contexts in current traveling, but overlook those in historical trips, as well as relationships between traversed road intersections. These issues hinder the model from effectively capturing complex mobility patterns. To fill this gap, we propose a novel method that incorporates current traveling features and historical activity chain to predict the coordinates of traveling destination. Specifically, (1) we construct current traveling features by extracting real-time moving states, and represent spatiotemporal correlations between traversed road intersections using word embedding; (2) we learn travel intentions as a probability vector for each historical trip, and combine it with spatiotemporal features to construct historical activity chain; (3) we construct an individual mobility prediction model using Long Short-Term Memory (LSTM) network and spatiotemporal scoring mechanism, to capture short-term and long-term dependencies in current trip and historical activity chain, respectively. Experiments on 21,890 trajectories over the whole Year 2019 of 20 representatives selected from 1916 private car travelers in Shenzhen City, reveal the effectiveness of our model. It outperforms four baselines, Random Forest (RF), Distant Neighboring Dependencies (DND), Location Semantics and Location Importance (LSI)-LSTM, as well as Intersection Transfer Preference and Current Movement Mode (ITP-CMM), by approximately 10%-15% improvement in accuracy. In addition, we further explore the impact of historical activity chain length, and destination visiting frequency on prediction, as well as the relationship between predictability and eight mobility pattern features. This study benefits potential applications such as personalized location-based service recommendations and targeted advertising, and also provides implications for understanding human mobility.

ARTICLE HISTORY
Received 3 September 2024
Accepted 13 January 2025

KEYWORDS
Mobility prediction; word embedding; spatiotemporal mobility patterns; trajectory feature modeling

1. Introduction

Individual mobility prediction has been extensively used in social science, geography, and transportation research (Li et al. 2019; Pappalardo et al. 2023). With the advancements in satellite navigation and sensor technologies, it has become possible to precisely capture locations of a traveler (Park et al. 2020), providing abundant and high-resolution data for studying individual mobility behaviors (Sun and Kim 2021). Exploring the individual mobility behaviors can reveal the spatiotemporal regularities of human activities at a micro-level (Liu, et al. 2025; Lucchini et al. 2021), identify the migration patterns of population (Sîrbu et al. 2021), and reflect the spatial disparities of different groups in urban areas (Testi et al. 2024). Among them, individual mobility prediction has wide applications, such as autonomous driving (Mozaffari et al. 2022), location recommendation (Arain et al. 2017), public safety (Luca et al. 2021), and route planning (Liu, Jia, et al. 2019a). For example, accurate mobility prediction can enhance the success conversion rate of business service recommendations (Feng et al. 2022), and is crucial for location-based services (LBS); it can also facigitate monitoring and early detection of anomalies for specific travelers, thus contributing to public safety (Miyazawa et al. 2020).

CONTACT Zhipeng Gui zhipeng.gui@whu.edu.cn

This article was originally published with errors, which have now been corrected in the online version. Please see Correction (http://dx.doi.org/10.1080/10095020.2025.2470548)

Due to the complexity and diversity of human mobility (Luca et al. 2021), the accurate mobility prediction remains challenging (Xu et al. 2024). The existing methods for individual mobility prediction primarily focus on two aspects, location sequence prediction, which forecasts a sequence of locations an individual is going to visit (Li et al. 2020) and real-time destination prediction for an ongoing trip (Gui et al. 2021). The latter is specifically emphasized in this study, as it provides coordinates of upcoming destination in constantly updated manner while traveling forward, thereby prompts the timeliness of LBS. Besides, it requires more comprehensive views and techniques to learning both broad and detailed spatiotemporal contexts in historical and current ongoing trips, respectively. Probability models (Jagannathan et al. 2021; Wang et al. 2021b), were initially employed in mobility prediction. However, they heavily rely on expert knowledge and are unable to fully leverage contextual information from the location sequences, thus leading to low prediction accuracy (Li et al. 2020). In recent years, deep learning models have been widely applied to mobility prediction. It shows promising performance and gradually emerges as the mainstream approach (Krishna et al. 2018). However, existing research primarily focuses on predicting future location sequences, with less attention paid to real-time traveling destination prediction (Ma and Zhang 2022). Besides, the majority of the studies focus on the spatiotemporal features and moving states of trajectory points, but overlook the underlying contextual information within trajectories. To enhance the accuracy of real-time mobility prediction, both current traveling features and statistical features of mobility patterns are extracted, e.g. spatial extent and temporal rhythms of traveling (Xing, Wang, and Lu 2020). Frequent routes and path preferences in mobility patterns of individual travelers are also mined to improve the accuracy of mobility prediction (Sevtsuk et al. 2021). Nevertheless, these methods primarily pay attention to the current traveling of travelers, but overlook their historical trips and the location semantics. By incorporating historical traveling activities, it is potential to prompt the prediction, shedding light on the underlying mobility patterns of travelers (Wang et al. 2021c). Moreover, location semantics may imply travel intentions, and thereby facilitate precise mobility prediction (Gui et al. 2021). In conclusion, it is imperative to fully utilize traveling features from both current and historical traveling activities for real-time mobility prediction.

Therefore, this study incorporates spatiotemporal contexts from both current ongoing trip and historical traveling activities to enhance real-time traveling destination prediction of individual travelers. Specifically, an end-to-end model is proposed with a specific focus on private car driving scenario. It integrates the real-time moving states, spatiotemporal correlations of road intersections in the current ongoing trip, and travel intentions of historical trips in the form of an activity chain to enrich traveling contexts. The main contributions of this work are as follows.

- In current traveling feature construction, besides extracting the **r**eal-time **m**oving **state**s (RMState), we represent the **s**patio**t**emporal **correlation**s of road intersections (STCorrelation) as high-dimensional word vectors to imply path preferences. It enables the model to better learn the underlying contextual information within current ongoing trip.
- We represent the **p**ossibilities of historical **t**ravel **intention**s (PTIntention) through Latent Dirichlet Allocation (LDA), in turn uncovering travel purposes of the traveler. Besides, the proposed model captures crucial long-term spatiotemporal context from historical activity chain by utilizing the spatiotemporal scoring mechanism.
- This paper provides implications for understanding impact factors of predictability, including the length of historical activity chain, destination visiting frequency, and other mobility pattern features, e.g. average travel distance, spatiotemporal mobility entropies.

The structure of the paper is organized as follows. Section 2 provides a literature review of relevant studies. Section 3 introduces the fundamental concepts and research objective. Section 4 illustrates the methodology. Section 5 validates the effectiveness of the proposed method through experiments. Section 6 explores impact factors of predictability. Finally, Section 7 draws the conclusion and future work.

2. Literature review

2.1. Machine learning in mobility prediction

Machine learning methods developed for mobility prediction can be primarily categorized into probability models and deep learning models (Yin et al. 2022). Probability models extract features from raw trajectories and establish inferential relationships to deduce the most likely moving forward direction or future anticipated destination. For example, Bayesian model is deployed to determine the next probable location based on prior probabilities of candidate destinations (Hazelton and Najim 2024); Random Forest (RF) ensembles multiple decision trees based on trajectory features to establish a regression or classification model for mobility prediction (Díaz-Ramírez, Estrada-García, and Figueroa-Sayago 2023); Gaussian regression is also used to predict instantaneous movements by dividing trajectories into Gaussian components (Wiest et al. 2012). Hidden Markov model (HMM) facilitates future location sequence prediction through Markov process that contains hidden parameters (Tang et al. 2018).

Although these methods are straightforward and interpretable, they can only predict the next trajectory point or mobility choice of next hop. Their accuracy is not satisfying for predicting future location sequences and multi-hop scenarios (Luca et al. 2021). In general, probability models are ideal or oversimplified in structure, and they cannot autonomously learn dependencies from long-term sequential trajectories (Gui et al. 2021). These issues hinder their prediction accuracies.

Deep learning gradually becomes dominant approach for mobility prediction (Kothari, Kreiss, and Alahi 2021). These techniques utilize neural network structures to learn spatiotemporal dependencies in trajectory data, and achieve end-to-end mobility prediction (Elman 1991). Long Short-Term Memory Network (LSTM) is a variant of Recurrent Neural Network (RNN). It is extensively adopted in mobility prediction as it can learn long-term dependencies hidden in trajectory data (Hochreiter and Schmidhuber 1997). For examples, inferring the future movements of pedestrians can prevent accidents in densely populated scenes (Alahi et al. 2016; Ma et al. 2019); A pooling-based LSTM model forecasts future trajectories of surrounding vehicles to support autonomous driving (Deo and Trivedi 2018). However, these studies are limited to predicting short-term movements within a few seconds. To predict traveling activities on a broader temporal scale, spatial and temporal prediction tasks are integrated using an extended model structure with two LSTM (Zhou et al. 2021). It facilitates multi-hop location prediction but underutilizes key spatiotemporal contexts in the location sequence, e.g. historical origins and destinations (ODs) that are highly correlated to the target locations to be predicted. To further improve prediction of long-term OD sequence, a hierarchical temporal attention mechanism is incorporated with a multi-layer LSTM encoder-decoder model (Li et al. 2020), enabling the prediction in the upcoming day or week. Nevertheless, existing methods mainly focus on predicting the next location or future location sequences, with limited exploration of real-time destination prediction (Ma and Zhang 2022). LSI-LSTM is proposed to predict the final destination of an ongoing trip (Gui et al. 2021). By integrating location semantics and importance, it enhances prediction accuracy compared to three baselines. However, this approach only determines the location importance of trajectory points, overlooking other trajectory features in traveling behavior. Actually, trajectory feature modeling can also facilitate mobility prediction, as it provides the model with moving states, transfer preferences, and travel intentions of a traveler. Hence, attention should be paid to how to effectively construct trajectory features.

2.2. Feature modeling in mobility prediction

The study of trajectory feature modeling mainly focuses on two aspects, i.e. spatiotemporal features and semantics correlated to trajectory points. Regarding spatial feature modeling, grid cells are commonly utilized to represent geographical locations (Forghani, Karimipour, and Claramunt 2020), especially for trajectories generated by Call Detail Records (CDRs) and check-ins of social media. The spatial features are represented by indexes of grid cells where the staying points are located (Li et al. 2020). This method is concise, yet it confines the mobility prediction to a classification task, thus its prediction accuracy is restricted by the scale of grid cells. Generally speaking, trajectory points recorded by GPS have high positioning accuracy. So, their coordinates can be directly encoded as spatial features, thereby enhancing prediction accuracy through regression model. Given that travelers in urban areas typically move along roads, trajectories are organized into the structure of transportation networks to better learn spatiotemporal context in trajectory segments (Wang et al. 2021a; Yu et al. 2019).

Besides trajectory points, modeling spatiotemporal contextual information within trajectory sequence can also enable the model to better learn movement patterns. Inspired by natural language processing, Word2Vec is utilized to embed areas in migration activities (Murray et al. 2023). Similarly, the embedding of check-in places is generated using Continuous Bag-of-Words (CBoW) to prompt downstream location-based prediction tasks (Wan et al. 2022). Besides, traveling habits or so-called transfer preferences between adjacent road intersections are also considered in real-time mobility prediction. These behavioral clues are identified for specified traveling activities, and thereby provide insights to promote prediction (Gui et al. 2024).

Beyond spatiotemporal features, semantic information can also provide the model with transportation conditions, urban functionality along the route, and the purpose of a trip, in turn enhancing prediction accuracy (Yao et al. 2023). For example, trajectory segments are assigned with moving states (e.g. stops and moves) (Bonavita, Guidotti, and Nanni 2021), to enrich its semantics for downstream prediction tasks. However, such a feature is oversimplified and can only provide limited contextual information. Besides, multi-hop locations are predicted based on geotags of check-in data and regions of interest (ROIs) (Chen et al. 2019). However, check-in behavior of travelers has selective tendency, e.g. travelers may geotag for entertainments and journeys but ignore daily routine activities. It leads to partial traversed locations missing in the predicted sequences. Besides, the contents of a geotagged post often have temporal delays, resulting in ambiguous semantic information (e.g. posting a hiking activity, while at home). So, incorporating external spatial data (Shang et al. 2021), including geographic entities, point of interests (POIs), urban land use data, etc., provides a promising

approach for semantic enrichment (Liu, et al. 2024). For example, incorporating semantics of geographic entities (e.g. office, road, and roof) and match them with trajectories of workers can reduce accidents on construction sites (Arslan, Cruz, and Ginhac 2019). However, it is constrained to trajectories within small areas, making it difficult to generalize to other scenarios. LSI-LSTM represents urban functionality of areas traversed by travelers based on POIs using Term Frequency-Inverse Document Frequency (TF-IDF) (Gui et al. 2021). It enables the model to leverage interaction between a traveler and urban functionalities, yet it fails to sufficiently utilize travel intentions of historical traveling activities. This limitation impedes the model from harnessing the long-term spatiotemporal dependency for real-time destination prediction.

In summary, existing feature modeling methods in mobility prediction mainly focus on the spatiotemporal features of trajectory points and the urban functional semantics of areas traversed by a traveler. However, modeling of contextual information in trajectory and travel intentions remain insufficient. Therefore, this paper proposes an LSTM-based individual mobility prediction model by considering current traveling features and historical activity chain. Beyond transfer preferences between two adjacent road intersections, STCorrelation reveals more general path preferences related to road intersections traversed during daily habitual commutes through word embedding. Besides, historical activity chain explicitly includes the origin to destination sequences of previous movements, makes it easier for model to capture long-term traveling contexts.

3. Preliminaries

In this section, we introduce the fundamental concepts, including trajectory point sequence, sub-trajectory, current traveling features and historical activity chain, as well as the research objective of traveling destination prediction, as illustrated in Figure 1.

Definition 1 Trajectory point sequence. Raw trajectory point sequence is a series of points collected by positioning devices from the origin (the starting point of a trip) to traveling destination. In this paper, we only retain intermediate trajectory points located at road intersections, as road network restricts traveling paths of vehicles and implies their future directions. So, trajectory is denoted as $Tr_{1:L} = \{origin, rp_i, destination\}_{i=1}^{L}$, where $origin = \langle lng_O, lat_O, time_O \rangle$ and $destination = \langle lng_D, lat_D, time_D \rangle$ are the starting and terminal points of $Tr_{1:L}$, respectively. L is the total

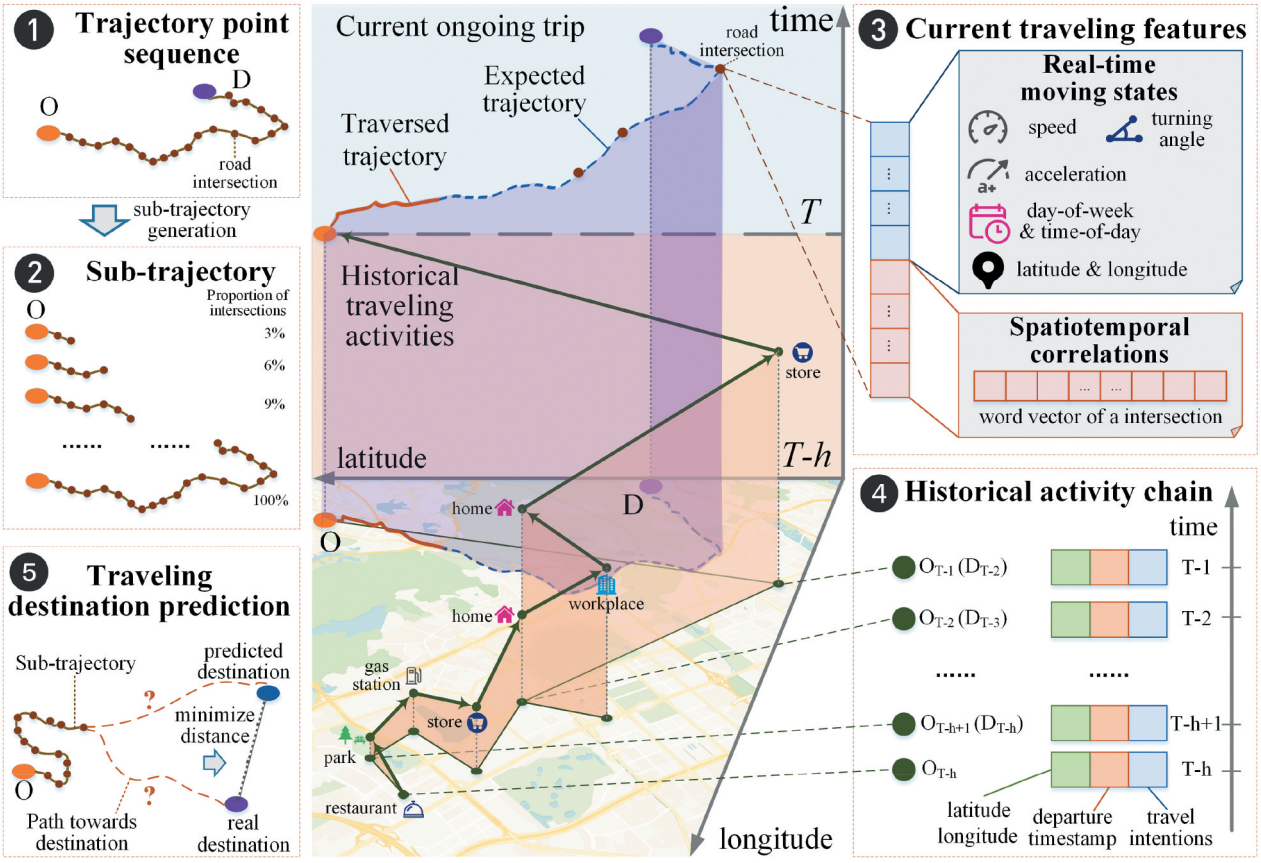


Figure 1. Fundamental concepts and research objective for traveling destination prediction.

numbers of retained point in a trajectory. Each retained point at a road intersection $rp_i = \langle lng_i, lat_i, time_i \rangle$ is a snapshot of geographic location (latitude and longitude coordinates) at $time_i$, where $time_1 < \ldots < time_i < \ldots < time_L$.

Definition 2 Sub-trajectory. For a completed trajectory $Tr_{1:L}$, its sub-trajectory $Tr_{1:s}$ is defined as an unfinished trajectory segment. It starts at the origin and ends at an intermediate road intersection rp_s. $Tr_{1:s} = \langle origin, rp_1, rp_2, \ldots, rp_s \rangle, 1 < \ldots < s < \ldots < L$, where s is the number of traversed intersections, and it indicates the completeness of $Tr_{1:s}$. Based on the definition of a sub-trajectory, the set of sub-trajectories is denoted as $= \{tr_{1:s_w}\}_{w=1}^{||}$, where $||$ and w are the total number and index of sub-trajectories, respectively.

Definition 3 Current traveling features. It refers to properties of all traversed trajectory points in a sub-trajectory of current ongoing trip, and encompass two parts. The first part is the moving states (e.g. speed, turning angle) at each traversed intersection rp_j, as well as day-of-week and time-of-day. The second part, spatiotemporal correlations, depict the contextual relationships between intersections along the trajectory in the form of embedding. It represents path preferences of an individual traveler, as illustrates in Figure 16.

Definition 4 Historical activity chain. Historical activities refer to past trips that occurred before the current ongoing trip. If only focusing on the origin (the starting point in $Tr_{1:L}$) and destination of past trips and without considering the intermediate trajectory points, the historical activities can be represented as a chain: $Chain_{historical} = \{\langle lng_{T-h}, lat_{T-h}, time_{T-h}, intent_{T-h} \rangle, \ldots, \langle lng_{T-1}, lat_{T-1}, time_{T-1}, intent_{T-1} \rangle\}$, where T is the index of current ongoing trip, and h denotes the number of previous trips to look back on. Specifically, (lng_{T-h}, lat_{T-h}) denotes the origin of the $(T-h)$-th trips, also indicating the destination of the $(T-h-1)$-th trips. $intent_{T-h}$ is travel intention of the $(T-h)$-th previous trips.

Definition 5 Traveling destination prediction. Given an unfinished sub-trajectory $Tr_{1:s}$, mobility prediction aims to predict destination of the current ongoing trip in the form of latitude and longitude. It can be denoted as $argmin\left(\left(lng_{pred} - lng_{real}\right)^2 + \left(lat_{pred} - lat_{real}\right)^2 \right)$, where $\left(lng_{pred}, lat_{pred}\right)$ and $\left(lng_{real}, lat_{real}\right)$ are the coordinates of predicted and real destination, respectively. To help the model better learn mobility patterns, current traveling features of each point in Tr_s are extracted to enrich the short-term spatiotemporal contexts. While, historical activities are structured as a chain to provide the model with a broader horizon.

4. Methodology

As illustrated in Figure 2, the framework of the proposed method includes three main parts. (1) Current traveling feature extraction: to acquire the current movement mode, STCorrelation and RMState in current ongoing trip are extracted. (2) Historical activity chain construction: to capture long-term spatiotemporal dependencies, we learn PTIntention (e.g. go home, to work, leisure) and spatiotemporal features (ST) for each historical traveling activity (corresponding to a completed trajectory with its origin and destination), and then concatenate them together for a series of consecutive historical traveling activities as a chain. (3) Model training and mobility prediction: the constructed destination prediction model leverages current traveling features and historical activity chain to output the location of the predicted destination. The working mechanisms of the three parts are detailed in the following subsections.

4.1. Current traveling feature extraction

The workflow of current traveling feature extraction contains three steps, as shown in Figure 3. First, raw trajectories are matched with road network, and only these trajectory points located at road intersections are retained. This process simplifies and shapes trajectory data according to the spatial structure of road network. Second, RMState (e.g. coordinates, timestamp, and travel distance) of the retained trajectory points in the current ongoing trip are extracted and concatenated. These features can indicate how a traveler moves forward at each road intersection, such as turning left and reducing speed, thereby implying potential destination choices. Third, STCorrelation is learned using word embedding and compressed with dimension reduction. Incorporating STCorrelation can enhance mobility prediction by providing contextual information of current ongoing trip.

4.1.1. Real-time moving states

The RMState consists of location, temporal, and other features. Location features include the longitude and latitude coordinates for each trajectory point at road intersections. Temporal features depict the departure time of current ongoing trip and timestamp of each trajectory point, and are modeled at two granularities. The day-of-week is represented as an integer from 0 for Sunday to 6 for Saturday; while the time-of-day is represented as an integer from 0 to 47 by dividing a day is into 48 time slots, where each slot being half an hour long. However, such a representation overlooks the continuity of time. For example, Saturday and Sunday are temporal adjacent, but the representations exhibit a noticeable difference (differ by 6), making it difficult for the model to learn the temporal

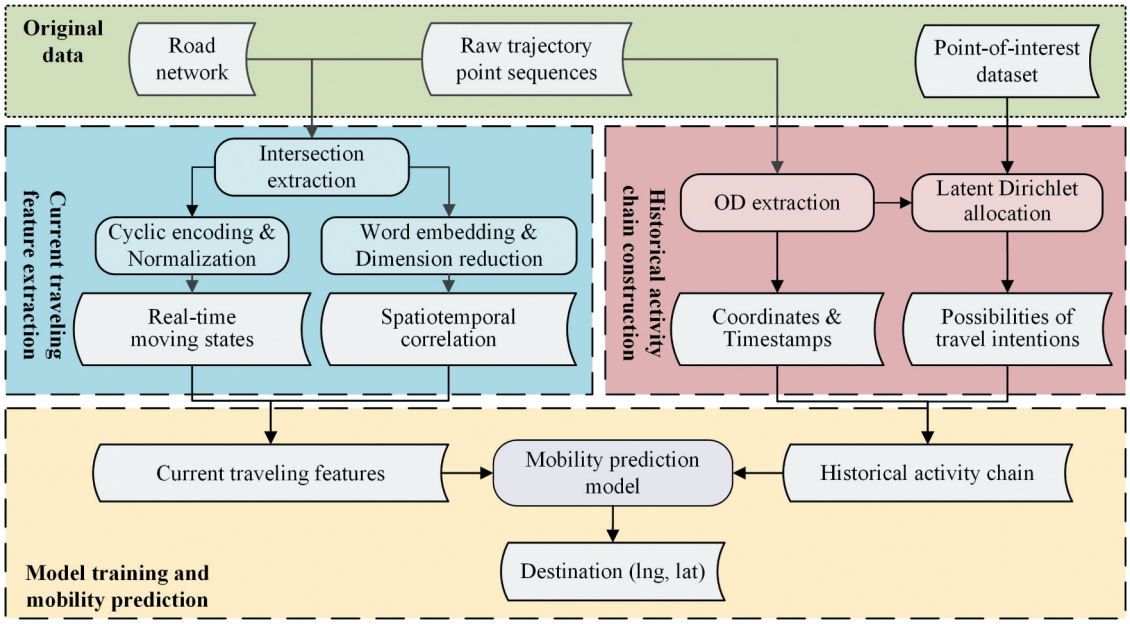


Figure 2. Framework of the proposed mobility prediction method.

dependencies between them. To address this issue, we encode the two temporal features using continuous Sine and Cosine functions by exploiting their cyclic characteristic (Gui et al. 2024). The cyclic encoding is shown in Equation (1).

$$day-of-week=\left[\sin\frac{2\pi\times \mathrm{d}}{7},\cos\frac{2\pi\times \mathrm{d}}{7}\right],\mathrm{d}\in\{0,1,2,3,4,5,6\}$$
$$time-of-day=\left[\sin\frac{2\pi\times \mathrm{t}}{48},\cos\frac{2\pi\times \mathrm{t}}{48}\right],\mathrm{t}\in\{0,1,2,\ldots,47\} \quad (1)$$

Other features include average speed of the previous road segment, travel distance, turning angle, time duration, and acceleration at each road intersection. The travel distance dis_i^{rp} of trajectory point rp_i located at the i-th road intersection traversed in a trip is defined as the cumulative road network distance from the origin to rp_i. Since the speeds of rp_i are transient, the average speed of all trajectory points on the previous road segment, $speed_i^{rp}$, is used to depict how fast a traveler moves along on a certain road segment. The turning angle $angle_i^{rp}$ at rp_i reflects the variation in direction of movement at a road intersection, and indicates the areas where a traveler is expected to proceed. It is defined as the angle between the line from the previous trajectory point located at road intersection rp_{i-1} to rp_i and the line from rp_i to the next point rp_{i+1}. The time duration denoted by $time_i^{rp}$, is defined as the time period from the origin to rp_i. It indicates the distance from the origin after departure, and traffic conditions along the route. The acceleration $acce_i^{rp}$ is calculated based on the speed of adjacent raw trajectory points p_{k-1} and p_k of rp_i, i.e. $acce_i^{rp}=\Delta speed(p_{k-1},p_k)/\Delta time(p_{k-1},p_k)$. This feature quantifies speed reduction or acceleration behavior at a road intersection, in turn implying path selection (e.g. proceeding straight, turning right), and traffic conditions on a road intersection. To ensure the features have the same value range across all dimensions, Z-Score normalization is utilized to rescale them into the interval of [−1, 1] (Fei et al. 2021).

4.1.2. Spatiotemporal correlation

If we consider the road intersections as words, then a trajectory can be regarded as a sentence. Similar to word preferences of writers in constructing sentences, individual travelers have path preferences and tends to choose specific routes when traveling to the destinations. For example, a traveler may consistently traverse certain intersections during commuting from home to work, and usually turn left or right at specified intersections (e.g. P4 and P11 in Figure 16, Appendix). While, the other may proceed straight through the same intersections. So, similar to the contextual relationships between words in a sentence, the implicit relationships between different road intersections in a trajectory can be captured through natural language processing model. To learn these relationships, we represent road intersections traversed by a traveler as high-dimensional word vectors based on Continuous Bag-of-Words (CBoW), i.e. a variant of Word2Vec techniques (Xia 2023).

The working mechanism of the model is shown in the right section of Figure 3. Specifically, we set a context window centered on the m-th road intersection, with its length C indicates the extent of context considered during model training. First, each road intersection is converted into a V-dimension vector by adopting One-Hot encoding, disregarding its latitude and longitude. Here, V is the total number of

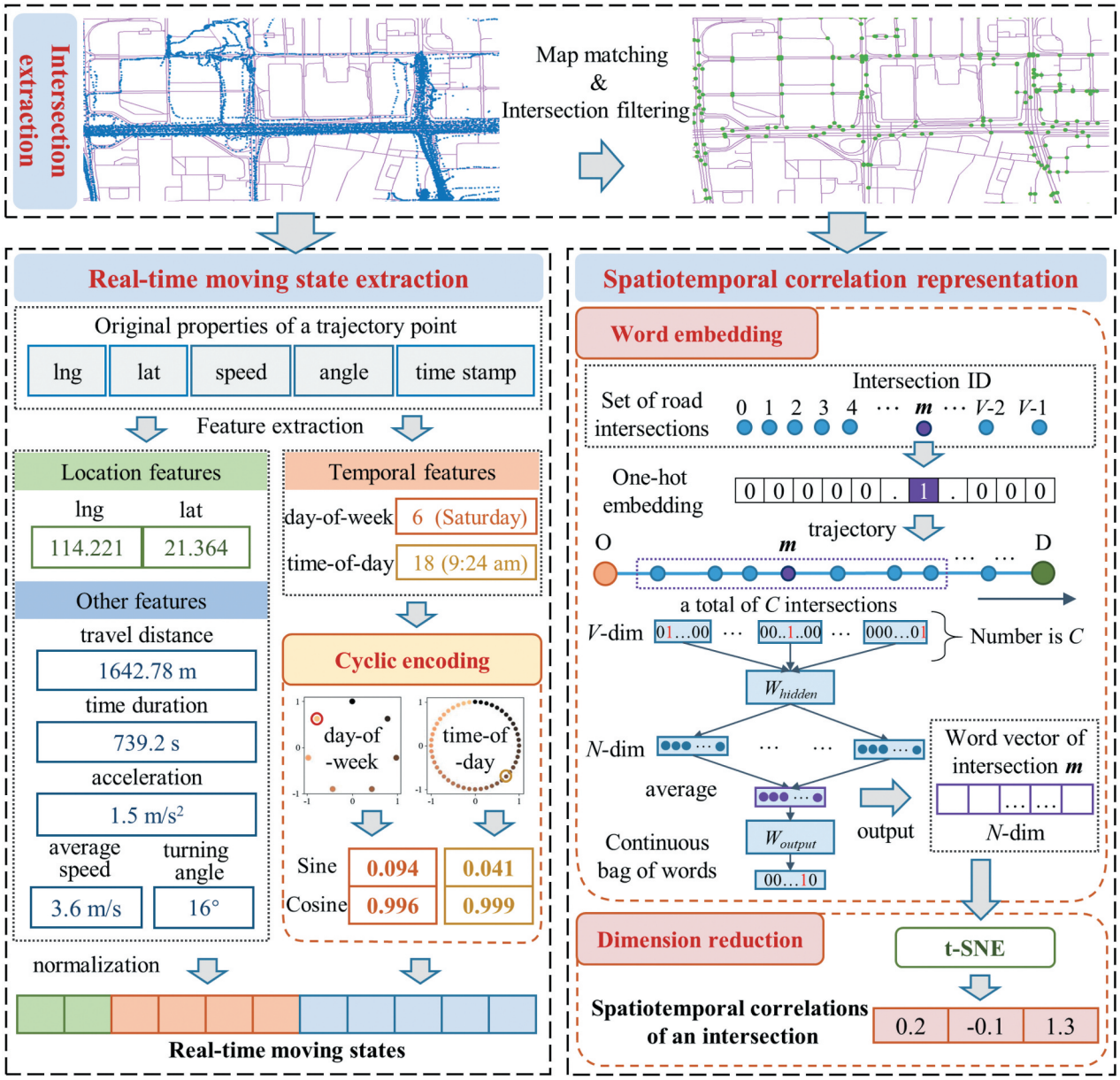


Figure 3. Illustration of current traveling feature extraction.

road intersections traversed in all trajectories by a traveler. For example, traveler R encounters 2435 road intersections in all trips, each intersection cp_m is represented as a binary vector V_m with 2435 dimensions, whose value of the m-th dimension is 1, and other dimensions are 0. Then, by multiplying V_m with W_{hidden}, a hidden layer vector N_m with N dimensions is obtained, and the vector N_m is computed as the average of all hidden layer vectors for the intersections within the context window C, where W_{hidden} is the weight matrix of the hidden layer, scaled by V and N dimensions. After that, N_m is multiplied by the W_{output} to yield a vector that aligns in dimensions with V_m, and finally activated by Softmax function to derive $\tilde{V}_m$, as calculated in Equation (2). W_{output} is the weight matrix of the output layer, with dimensions of $N \times V$. Through iterative training, the spatiotemporal correlations of road intersections traversed and expected ahead in current ongoing trip are represented as $\bar{N}_m$.

$$\tilde{V}_m = \frac{exp\left(N_m \cdot W_{output}\right)}{\sum_{m=1}^{C} exp\left(N_m \cdot W_{output}\right)} \tag{2}$$

Since the dimension of the word vectors obtained through CBoW is much higher than those of other features such as locations and timestamps, the model tends to neglect other features and hinder learning spatiotemporal contexts from them. Therefore, we reduce the dimensionality of word vectors using t-distributed Stochastic Neighbor Embedding (t-SNE) (Belkina et al. 2019). Finally, the STCorrelation and RMState of a road intersection are input into the

current movement mode learning module of the mobility prediction model.

4.2. Historical activity chain construction

To address insufficient utilization of historical traveling features in mobility prediction, we construct historical activity chain by incorporating ODs of historical trips prior to the current ongoing trip and their travel intentions, as illustrated in Figure 4. Travel intentions indicate what a traveler aims to do at a certain urban functional zone, and enable the model to leverage the interactions between the traveler and places from the perspective of social activity. Specifically, a grid-based semantic map is constructed based on the number of POIs for each POI category in each cell to integrate urban functional semantics. Then, to explore purposes of a trip, a possibility vector of travel intentions is learned based on the semantic map using LDA (Chauhan et al. 2021). After that, the historical activity chain is constructed based on the PTIntention, location, and temporal features to provide the model with long-term spatiotemporal dependencies in historical trips, in turn enhancing the capability of mobility prediction.

4.2.1. Possibilities of travel intentions

We utilize PTIntention to uncover the purposes of traveling activities, e.g. commuting to the workplace in the morning and returning home in the evening. Due to the uncertainty of daily mobility of travelers, LDA, a probabilistic topic modeling technique for extracting latent topics from documents, is adopted to model travel intentions and their possibilities (see pseudocode in Appendix). By modeling the word occurrence in a corpus consisting of documents, LDA reveals latent topics and estimates their possibility distribution within each document (Du et al. 2020). Accordingly, we learn the PTIntention based on POI data, which can reflect the urban functionality, in turn revealing the intention behind mobility of a traveler. Specifically, we map the destinations of historical trajectories into regular grid cells, and then calculate the number of POIs in different POI category that fall within each cell. The distribution of various POI types in a cell can be considered similar to the distribution of different thematic words in a document. Therefore, the representations of travel intentions can be learned by LDA, where the q-th travel intention $intention_q = \langle POI_0, POI_1, \ldots, POI_p, \ldots, POI_P \rangle$, p is the p-th POI type, POI_p is the likelihood value of the p-th POI type. Besides, a document may belong to multiple topics with varying probabilities. Similarly, a destination cell has more than one travel intentions. So, PTIntention of a destination cell of the j-th trip is denoted by $PTIntention_j = \langle intention_0, intention_1, \ldots, intention_q, \ldots, intention_Q \rangle$, where q corresponds to the q-th travel intention. After that, the historical activity chain is constructed based on $PTIntention_j$.

4.2.2. Historical activity chain

Since the mobility pattern of a traveler has long-term historical regularity, the historical trips can provide long-term spatiotemporal context beyond current trajectory sequences. For example, if a traveler just leaves the origin, the trajectory of current traveling is too short for the model to capture the movement mode, making it difficult to infer the final destination. If the historical trips are considered, the

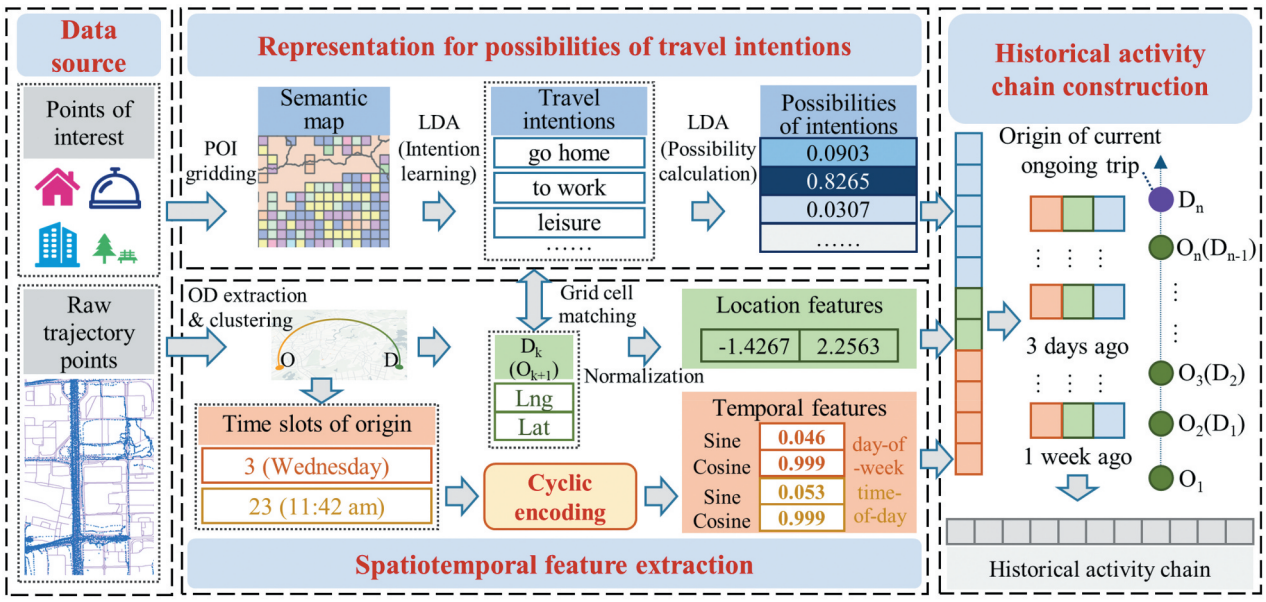


Figure 4. Flow of historical activity chain construction.

prediction capability could be prompted by broader spatiotemporal context. Therefore, we construct historical activity chain based on ODs to sum traveling activities before the current ongoing trip. While properties of intermediate trajectory points (e.g. speed, turning angle) in historical trips are not essential for constructing historical activity chain, as they are too detailed for summarizing historical trips. So, we filter them out and only retain the ODs of historical trips, and integrate the corresponding PTIntention. Furthermore, the temporal features of historical trips, i.e. the time slot of day and the day of week (see section 3.1.1), can reveal temporal mobility patterns of a traveler. As the result, a historical activity chain consists of the ST and the PTIntention of historical trips, denoted as $Chain_{historical} = \{[PTIntention_1, (Lng_1, Lat_1), (day-of-week_1, time-of-day_1)], \ldots, [PTIntention_j, (Lng_j, Lat_j), (day-of-week_j, time-of-day_j)], \ldots, [PTIntention_T, (Lng_T, Lat_T), (day-of-week_T, time-of-day_T)]$, where T denotes the length of the historical activity chain, (Lng_j, Lat_j) and $(day-of-week_j, time-of-day_j)$ denote the location and temporal features of the j-th trip, respectively.

4.3. Model training and mobility prediction

As shown in Figure 5, the proposed model is comprised of three modules, i.e. current movement mode learning module, historical movement mode learning module, and destination prediction module. The current movement mode learning module obtains the hidden state sequence from the input current traveling features. The historical movement mode learning module learns the historical traveling contextual information of travelers from the constructed historical activity chain. In the destination prediction module, the hidden state sequences from the preceding two modules are concatenated, and the destination of current ongoing trajectory is predicted through the residual network and fully connected layer. In the input features, RMState, STCorrelation, and PTIntention of each traveling activity are constructed as 12, 3, and 5 dimensional vectors, respectively.

4.3.1. Current movement mode learning module

This module represents inputs as a hidden state sequence. It consists of two parallel stacked LSTM networks, a spatial scorer, and a crossover layer. The two parallel LSTM networks learn RMState and STCorrelation, respectively. They have two layers each as such a structure has a stronger ability to capture mobility pattern of an individual than single-layer LSTM (Wang et al. 2021c). The upper stacked-LSTM inputs the RMState of the retained trajectory points (i.e. the origin and intermedia road intersections), and outputs a sequence $H^{RTM} = \{h_1^{rtm}, h_2^{rtm}, \ldots, h_m^{rtm}, \ldots, h_n^{rtm}\}$, where h_m^{rtm} refers to the hidden state of the m-th retained trajectory point. The lower stacked-LSTM processes the coordinates and spatiotemporal correlations of each retained trajectory point, and outputs sequence of hidden layer as $H^{STC} = \{h_1^{stc}, h_2^{stc}, \ldots, h_m^{stc}, \ldots, h_n^{stc}\}$, where h_m^{stc} refers to the m-th hidden state in the sequence. Then, a new hidden state sequence $H^{Current}$ is generated by multiplying the two sequences H^{RTM} and H^{stc} through a crossover layer to enhance critical spatiotemporal contexts (Yang et al. 2021). After that, each state in the hidden sequence $H^{Current}$ is scored by turning angle, travel distance, and time duration, respectively, in spatial scorer to measure its importance for prediction. Finally, all scored states are summed together to obtain the representation of current movement mode.

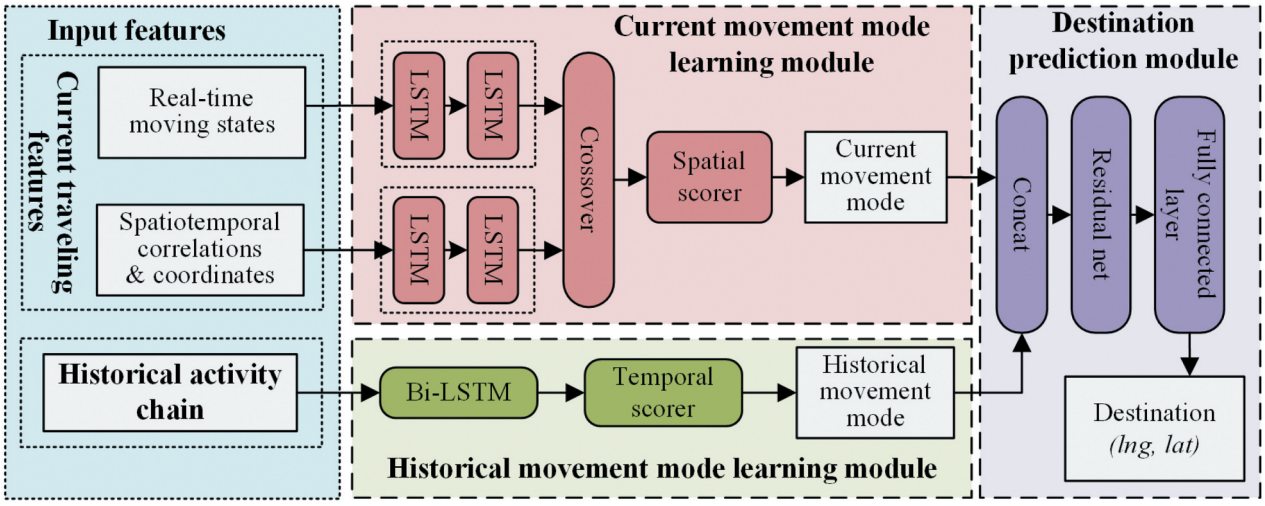


Figure 5. Framework of mobility prediction model.

4.3.2. Historical movement mode learning module

This module comprises a bidirectional LSTM layer, and a temporal scorer. The LSTM layer captures the long-term spatiotemporal dependencies from historical activity chain, and generates a historical hidden state sequence, as shown in Figure 6. The inputs of the LSTM layer include coordinates and temporal features of the origins, as well as the PTIntention for historical trips of a traveler. Due to the regularity and cyclical nature of mobility patterns of travelers, it is essential for the model to focus more on historical traveling activities that closely align with the current date and time. By assigning greater weight to them, the temporal scorer highlights crucial long-term spatiotemporal context, thereby enhancing prediction accuracy. Specifically, we calculate the Sine and Cosine differences in the day-of-week and time-of-day between the current trip and a historical trip, respectively, to measure the temporal similarity. The temporal scores have four dimensions, $Score_{Sine}^{day-of-week}$, $Score_{Cosine}^{day-of-week}$, $Score_{Sine}^{time-of-day}$, and $Score_{Cosine}^{time-of-day}$, denoting the reciprocal of the four calculated differences. Ultimately, historical movement mode is derived by summing all states in the historical hidden sequence across the four dimensions, respectively.

4.3.3. Destination prediction module

This module generates location of the predicted destination in the form of latitude and longitude. Specifically, the hidden state sequences of current and historical movement modes are concatenated and fed into a residual neural network with stacked fully connected residual layers. The working mechanism of the residual network is shown in Equation (5).

$$\begin{aligned} \sigma_r &= ReLU(W_r x + b_r), \\ x_0 &= \sigma_r(h_{attn}^t), \\ x_i &= \sigma_r(\sigma_r(x_{i-1}) \oplus x_{i-1}) \end{aligned} \tag{5}$$

where $ReLU$ is an activation function, W_r is the weight matrix, b_r is the bias, and $\oplus$ represents the element-wise summing operation on matrices. The output of the i-th layer depends on the input x_{i-1} and the output $\sigma_r(x_{i-1})$ of the $(i-1)$-th layer, to learn the residual of the previous layer. Finally, a single neural element adopts a nonlinear mapping to output the predicted coordinates. We use the mean absolute error (MAE) between the target destination and predicted destination shown in Equation (6) to measure the effectiveness of the proposed model.

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| + \lambda\|\theta\|^2 \tag{6}$$

where λ is a regularization hyper parameter, $\hat{y}_i$ is the predicted destination, y_i is the real destination, and n is the batch size, i.e. the number of trajectories input each batch.

5. Experiment

To validate the prediction performance, we have conducted comprehensive comparison and ablation experiments. The prediction accuracy of our model is evaluated by comparing with four baselines, including random forest (RF), distant neighboring dependencies (DND) (Qian et al. 2019), LSI-LSTM, and ITP-CMM. The effectiveness of current traveling feature extraction, and historical activity chain construction are validated through ablation experiments. Besides, we further explore their working mechanism on capturing path preferences and long-term temporal regularity.

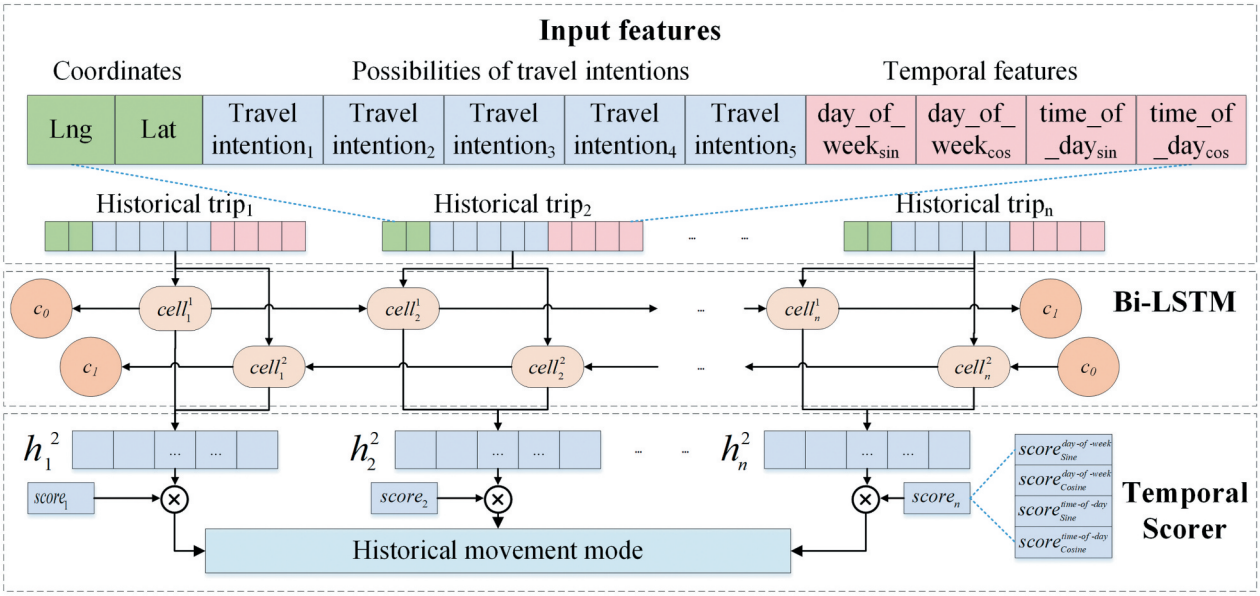


Figure 6. The working mechanisms of historical movement mode learning module.

5.1. Dataset and preprocessing

The experiments are conducted upon selected representative travelers in an anonymous private vehicle GPS trajectory dataset. The dataset was collected from 1916 travelers commuting within Shenzhen, and recorded more than 500,000 trajectories from 1 January 2019 to 31 December 2019. Compared with the taxi trajectory data, private vehicle trajectories can better reflect the long-term individual mobility pattern. Road network data for the urban area of Shenzhen in 2018 from Amap (https://lbs.amap.com/) is employed for intersection extraction. A total of 334,831 POIs in the year of 2019 are used to represent PTIntention. They cover the administrative regions of Shenzhen and encompass 8 categories, i.e. dining, medical facilities, residential areas, hotels, life services, tourist attractions, educational institutions, and corporate entities.

The representative travelers are selected as follows. First, travelers whose traveling duration less than one month are excluded because they cannot provide long-term trajectory data. Then, the OD entropy (E_{OD}) (Wu et al. 2019), average travel distance (TD_{avg}), and ratio of infrequently visited locations (R_{IVL}) for each traveler is calculated, as they may influence prediction accuracy. Specifically, E_{OD} measures the heterogeneity in the probability distribution of OD pairs. TD_{avg} quantifies how long the trips of a traveler are on average. R_{IVL} emphasizes the traveling tendency of a traveler to visit locations beyond the residential, workplaces and other frequently visited locations. Based on the three metrics, all travelers are clustered into 20 clusters using DBSCAN. After that, we randomly select one traveler from each cluster, and then categorize the selected 20 travelers into four groups using natural breaks by ranking them in descending order of R_{IVL}. The four groups, (I), (II), (III), and (IV), contain 2, 6, 8, and 4 travelers, respectively. From Figure 7 (b), we can find that the three metrics are correlated, yet they are not always consistent. For example, traveler 11, despite having a high TD_{avg}, exhibits a moderate R_{IVL}. It indicates that a long-distance trip may have high traveling frequency. Based on the four groups, we analyze how prediction accuracy is influenced by R_{IVL} across groups (Section 5.3), and other mobility pattern features (Section 5.4.1 and 6.1).

Due to factors such as multipath effect and cycle slip of satellite signal, the raw trajectories may contain errors. Before experiments, data cleaning is conducted. (1) Trajectory points with speeds exceeding 120 km/h are omitted; (2) trajectories containing points located outside the study area are excluded; (3) raw trajectories with fewer than 10 points are disregarded. After that, we extract the raw ODs from each trajectory. As a traveler may park at nearby distinct locations around the same destination on different trips, we then generate unique coordinates for raw ODs by clustering them using DBSCAN. The obtained cluster centroids and noise points correspond to the frequently and infrequently visited destinations, respectively. Finally, we match the remaining intermediate trajectory points with the road network and retain these located at road intersections by employing HMM-based fast map matching (C. Yang and Gidófalvi 2018), as it can process trajectory data with vast spatial range efficiently.

To prevent the model from prematurely learning contexts of trajectories in the testing set while training, trajectories of an individual traveler are randomly partitioned into two parts, P_{train} and P_{test} in a ratio of 3:2. Then, sub-trajectories are generated from a completed trajectory by incrementing 3% of trajectory points at each step. Sub-trajectories generated from P_{train} are totally allocated to the training set. It enables the model to learn spatiotemporal contexts of a completed trajectory. Besides, to ensure the model can capture traveling patterns that are absent in P_{train}, those generated from P_{test} with completeness below 60% are incorporated into the training set, while the remaining sub-trajectories are allocated to the testing set.

5.2. Evaluation metrics and experimental settings

5.2.1. Evaluation metrics

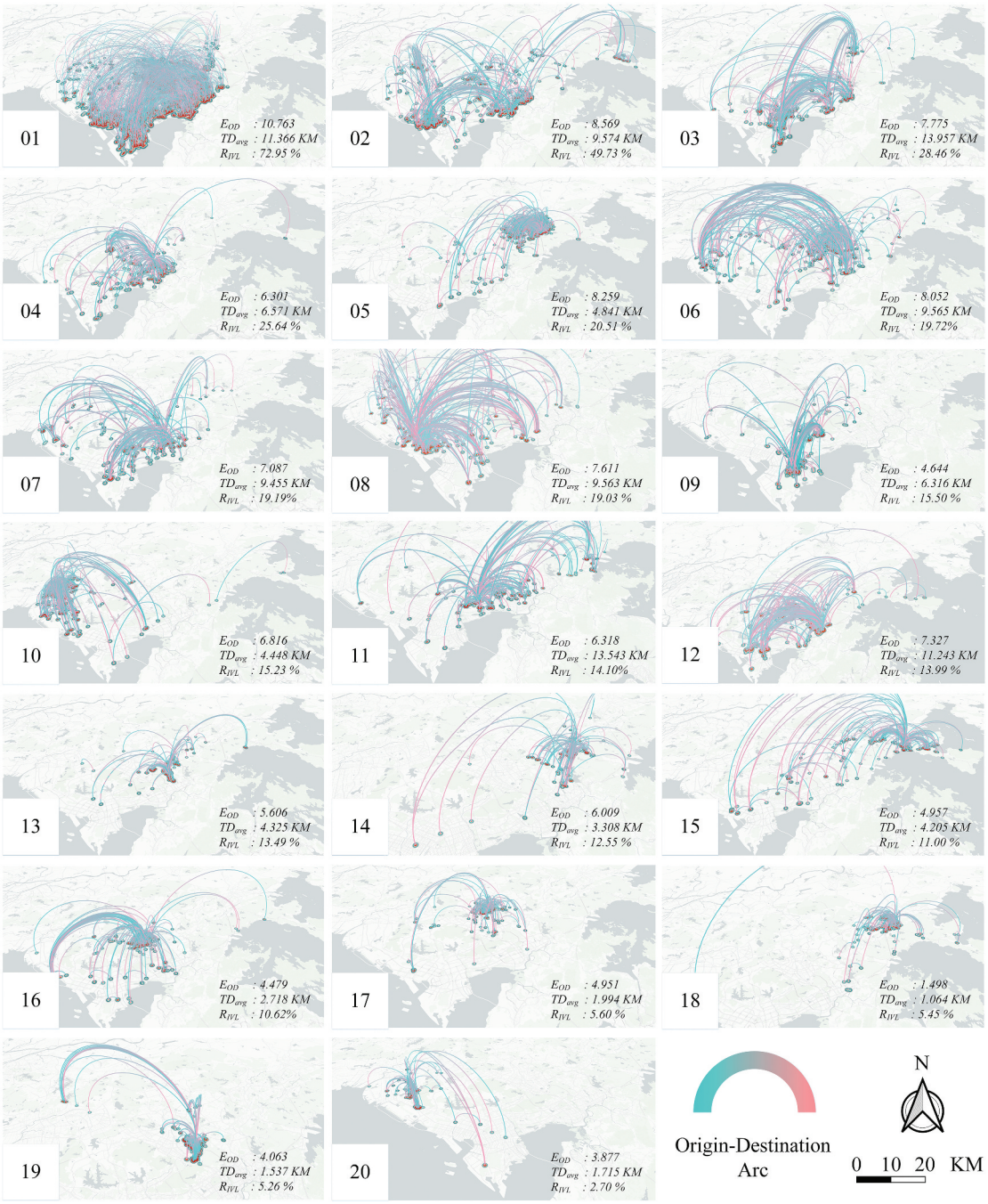
In order to evaluate the prediction performance, this study uses mean absolute error (MAE), root mean square error (RMSE), and mean relative error (MRE) as evaluation metrics of accuracy (Ke et al. 2021). The calculations are shown as formula (7).

$$\begin{aligned} \mathrm{MAE} &= \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|, \\ \mathrm{RMSE} &= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}, \\ \mathrm{MRE} &= \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\hat{y}_i - y_i\right|}{dis_{total}} \end{aligned} \tag{7}$$

where dis_{total} is the total travel distance of a trajectory. MAE and RMSE measure the Euclidean distance between the predicted destination and the real destination. They are calculated by averaging the absolute and squared root differences between the predicted and real coordinates, respectively. MRE measures the error on a per-unit distance basis by dividing each prediction error by the total travel distance.

5.2.2. Experimental settings

Based on relevant studies and our experiments, we have set the hyperparameters as follows. The



(a) OD flows of the selected 20 travelers

Travel ID	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20
E_{OD}	10.76	8.57	7.78	6.30	8.26	8.05	7.09	7.61	4.64	6.82	6.32	7.33	5.61	6.01	4.96	4.48	4.95	1.50	4.06	3.88
TD_{avg}	11.37	9.57	13.96	6.57	4.84	9.57	9.45	9.56	6.32	4.45	13.54	11.24	4.32	3.31	4.20	2.72	1.99	1.06	1.54	1.72
R_{IVL}	72.9%	49.7%	28.5%	25.6%	20.5%	19.7%	19.2%	19.0%	15.5%	15.2%	14.1%	14.0%	13.5%	12.5%	11.0%	10.6%	5.6%	5.5%	5.3%	2.7%
	Group (I)		Group (II)						Group (III)								Group (IV)			

(b)The division of the four groups

Figure 7. All OD flows in the whole Year of 2019 for the selected 20 travelers.

dimension of the features in the hidden layer is 128. Adam optimization algorithm is adopted to train the model, and the learning rate is 0.001; the training epochs are 80, and the batch size is 32 (Gui et al. 2021). The dimension of hidden layer in CBoW is 256, and that of spatiotemporal correlations of an intersection is 3. For OD extraction, the *eps* and *min_samples* of DBSCAN clustering are set as 200 meters and 10 points, respectively. The resolution of grid cells is 1000 meters. The experiments were conducted on

a 64-bit operating system of Ubuntu 18.04, with 16 GB memory capacity and an Intel (R) Core (TM) i7–7700 CPU with a clock speed of 3.60 GHz. The versions of CUDA, Python, and Pytorch are 11.4, 3.9.13, and 1.13.1, respectively.

5.3. Comparison of overall prediction accuracy

To verify the performance, we have conducted a comparative analysis of overall prediction accuracy with four baselines. RF is an ensemble learning technique widely used in prediction tasks that promotes the decision-making capability through multiple decision trees. So, we select it as a representative possibility model to compare with our model. Besides, it would be more convincing to compare the proposed model with other deep learning baselines dedicated for mobility prediction. Similar to our model, DND also embeds road intersections as high-dimensional vectors, but it overlooks moving states (e.g. travel distance, timestamp). So, we select this model to validate the value of current traveling feature extraction, especially real-time moving states. LSI-LSTM considers the urban functional semantics of trajectory points based on POIs, while ITP-CMM represents transfer preferences between adjacent road intersections. So, we compare them to assess the effectiveness of PTIntention and STCorrelation in our method, respectively.

To ensure the fairness of the comparison, the same preprocessing was implemented for the baselines and our model, i.e. data cleaning, intersection extraction, sub-trajectory generation, dataset partitioning. The prediction results of our model and baselines are shown in Table 1.

Table 1 illustrates that our model achieves the highest prediction accuracy compared to the four baselines in term of the mean values of the three metrics. Deep learning baselines outperform RF across all metrics, because they can capture spatiotemporal contexts more effectively. Among four deep learning models, the prediction accuracy of DND is the lowest, as it underutilizes moving states and long-term historical traveling activities; while LSI-LSTM and ITP-CMM exhibit much better accuracy. Because they explore the mobility pattern from the perspective of urban functionality and transfer preferences, respectively. Our model furtherly achieves an approximate 10% improvement than LSI-LSTM and ITP-CMM by introducing historical activity chain to leverage long-term spatiotemporal dependency. Besides, the MRE of LSI-LSTM is lower than ITP-CMM, and slightly higher than our model. It indicates that LSI-LSTM can predict destinations of long-distance trips more accurately than ITP-CMM. In addition, our model demonstrates relatively robust stability. Its standard deviations (SD) of the three metrics across different batches ranks the second and closely approaches the first place. This is because our model excels in predicting the majority of destinations; while, for a few destinations that are not visited recently, the historical activity chain may introduce noise and decrease accuracy, thus yielding an increased SD. To further investigate the variance in prediction accuracy among travelers, we categorize them in the abovementioned four groups. The prediction accuracies of the four deep learning models for the four groups are presented in Figure 8.

As shown in Figure 8, our model excels in travelers with high R_{IVL}, but performs modestly for travelers with low R_{IVL}. Compared to three deep learning baselines, our model surpasses them for travelers in groups (I), (II), and (III), especially in MAE and RMSE. ITP-CMM achieves accuracy close to our model, ranking the second. It further emphasizes the value of transfer preferences and path preferences. While, prediction accuracy of our model is slightly lower than that of LSI-LSTM for group (IV). It may be because irrelevant trips in long-term spatiotemporal context may intervene the model to learn mobility patterns of travelers with low R_{IVL}, as their mobility patterns have higher regularity. Besides, from group (I) to (IV), MAE and RMSE are gradually reduced. It indicates that a high level of R_{IVL} leads to low predictability. However, the MRE in the four groups is relatively similar. It is attributed to low accuracies of partial short-distance trips, and their impacts on average accuracy are magnified on a per unit distance basis. This indicates that the effectiveness of mobility pattern learning is related to frequency and regularity of trips, rather than tightly

Table 1. Comparison of mobility prediction accuracy.

	MAE (m)		RMSE (m)		MRE (%)	
Methods	MEAN	SD	MEAN	SD	MEAN	SD
RF	8399.702	6904.124	11053.489	17553.287	411.57%	1993.09%
DND	2726.059	1119.194	4193.545	4139.467	111.51%	185.36%
LSI-LSTM	1224.925	856.028	2188.102	2617.243	20.99%	**23.14%**
ITP-CMM	867.609	**360.435**	1637.270	**1740.398**	41.22%	79.46%
Our model	**760.362**	439.973	**1438.317**	1786.882	**20.86%**	26.76%

coupled with travel distance. In conclusion, our model exhibits accuracy improvements for different groups compared to baselines, and the diverse complexity of mobility patterns among groups leads to variations in capturing spatiotemporal context, consequently yielding variances in prediction accuracy.

5.4. Effectiveness of constructed features

Ablation experiments are conducted to validate the effectiveness of constructed features. We compare prediction accuracy between a model only using coordinates and models using current traveling features or historical activity chain independently. Same parameter configurations are employed to ensure the fairness of comparison. In addition, we explore the path preferences revealed by spatiotemporal correlation between road intersections in the current ongoing trip, and analyze the working mechanism of historical activity chain in improving prediction accuracy.

5.4.1. Effectiveness of current traveling features

To validate the effectiveness of current traveling features, including RMState and STCorrelation, we designed three feature combination schemes: (1) only using coordinate (COOR), (2) integrating the RMState with COOR, and (3) further integrating the STCorrelation. The results are shown in Table 2 and Figure 9.

As shown in Table 2, schemes (2) &; (3) yield enhanced prediction accuracy compared to COOR. It indicates the effectiveness of both RMState and STCorrelation. RMState depicts moving states of each trajectory point located at road intersection, so it enables the model to capture the mobility patterns of a traveler from the origin to the intermediate intersection, thus enhancing its prediction capability. Nonetheless, RMState of each trajectory point is isolated. Therefore, the model can only capture spatiotemporal context using the LSTM structure. The inclusion of STCorrelation reduces the MRE more dramatically than RMState. That is because STCorrelation provides the model with contexts by inferring the potential path preferences. It enables the model to guess the upcoming road intersections to be traversed. Furthermore, as R_{IVL} decreases (Figure 9), prediction accuracies of the travelers gradually improve, which confirms that R_{IVL} has

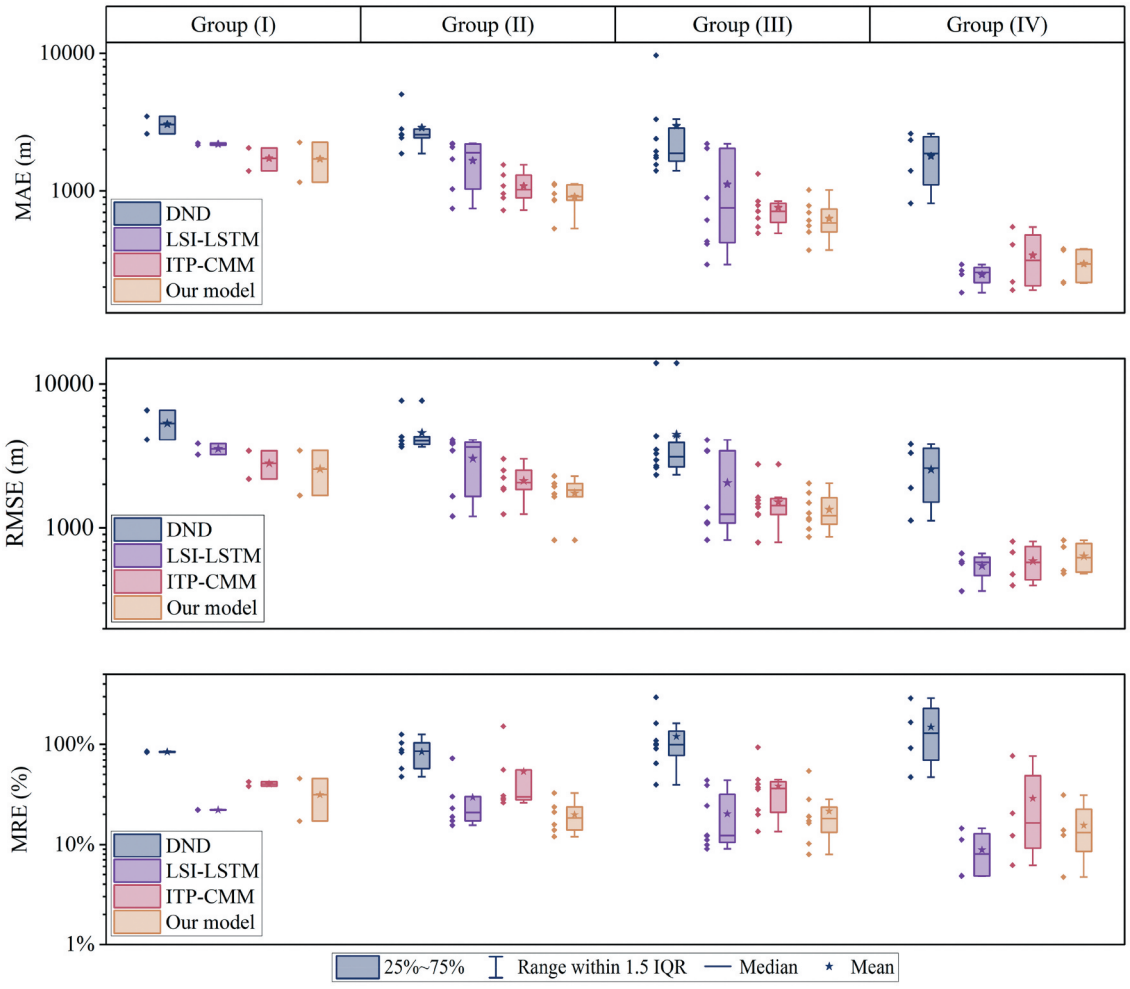


Figure 8. Comparison of prediction accuracy for the four traveler groups. The pentagon denotes the average evaluation metrics for a group of travelers, and the line depicts the median value of them.

a negative correlation with predictability to a certain extent. However, R_{IVL} is not the dominant factor for predictability. For instance, travelers 05, and 10 exhibit higher R_{IVL} compared to travelers 06 and 11, respectively, yet achieve lower prediction errors. This could be due to smaller TD_{avg} or E_{OD} of travelers 05 and 10 making their mobility patterns easier to learn.

We further explore the path preferences revealed by STCorrelation and elucidate its mechanisms in improving the prediction accuracy. To investigate this phenomenon, we visualize the road intersections in four selected trajectories of traveler 05, and label partial intersections that were frequently traversed by at least two of the four trajectories in Figure 10 (a), i.e. frequent sets of road intersections. Then, we compute the similarity between word vectors of the labeled intersections in Figure 10 (b).

Figure 10 (a) illustrates three frequent sets I, II, and III. For examples, road intersections 910, 911, 912, 913, 914, 915, 916, 228, and 229 traversed by both trajectories 3 and 4 belong to frequent set III. From the similarity matrix in Figure 10 (b), we can find that road intersections in the same frequent sets have relatively high similarities, e.g. 485, 496 and other intersections in frequent set I. By contrast, the similarities between road intersections from different frequent sets are quite low (shown in red color), e.g. the intersections 484 in I and 906 in II, as well as intersections 902 in II and 910 in III. This is because the traveler does not pass through these intersections simultaneously on the daily route, making their word embeddings different a lot. This observation indicates that STCorrelation can effectively capture the path preferences and provide the model with spatiotemporal contexts corresponding to daily routine traveling habits to facilitate prediction.

5.4.2. Effectiveness of historical activity chain

To verify the effectiveness of historical activity chain, we conducted three feature combination schemes: (1) only using latitude and longitude in current trip as inputs (i.e. COOR, the same as scheme (1) in the previous section), (2) incorporating ST in historical activity chain with COOR, i.e. coordinates and timestamps of the origins, and (3) further introducing PTIntention on the basis of scheme (2). The results are shown in Table 3 and Figure 11.

Table 3 demonstrates the effectiveness of ST and PTIntention, the highest accuracy is achieved when both of them are integrated. MAE and RMSE significantly improved for scheme (2) compared to (1). That is because ST includes coordinates and timestamps in historical trips. These contexts can imply transitions between different locations in historical trips and depict mobility patterns on a broader spatiotemporal scale. It allows the model to capture long-term spatiotemporal dependencies, leading to higher prediction accuracy. However, the enhancement of MRE is limited. It may be due to ST overlooking the purposes of historical trips. After incorporating PTIntention, MAE and RMSE are further reduced, and MRE exhibits a greater decrease. The reason is that PTIntention implies the interaction between a traveler and urban functional zones, and inspires the model to learn long-term spatiotemporal context beyond coordinates and timestamps.

The effectiveness of incorporating ST or PTIntention varies among different travelers, as demonstrated by the changes of the three metrics in Figure 11. Specifically, MAE for all the 20 travelers in scheme (3) shows improvement compared to COOR. Among them, nine travelers show decreased MAE when both features are integrated, while five and six of them achieve the highest accuracy by solely incorporating ST and PTIntention, respectively. This could be attributed to historical activity chain introducing irrelevant information that influences predictions for travelers with high traveling entropies. For seven out of the 20 travelers (i.e. 01, 06, 07, 09, 10, 11, 15), both MAE and RMSE exhibit significant improvements. It highlights that historical activity chain can enhance prediction accuracy for travelers with high or moderate levels of R_{IVL}. In addition, the MRE for half of the 20 travelers (e.g. 06, 09) decrease by over 5% after including PTIntention, highlighting its effectiveness in enhancing prediction accuracy.

To explain why historical activity chain improves prediction accuracy, we further explore the working mechanism of temporal scorer and mobility patterns revealed by PTIntention. We visualize the historical activity chain spanning six days before a current trip by displaying each historical trip in the form of an OD arc, and depict PTIntention and departure time associated with each trip to explore temporal regularity in mobility patterns, as shown in Figure 12. In addition, the

Table 2. Overall effectiveness of current traveling features.

	MAE (m)		RMSE (m)		MRE (%)	
Schemes	MEAN	SD	MEAN	SD	MEAN	SD
COOR	1398.025	642.919	2684.803	2754.096	45.61%	69.96%
COOR + RMState	1032.979	460.046	1887.246	2124.168	41.72%	66.00%
COOR + RMState + STCorrelation	**879.745**	**457.747**	**1722.694**	**2044.872**	**31.54%**	**51.35%**

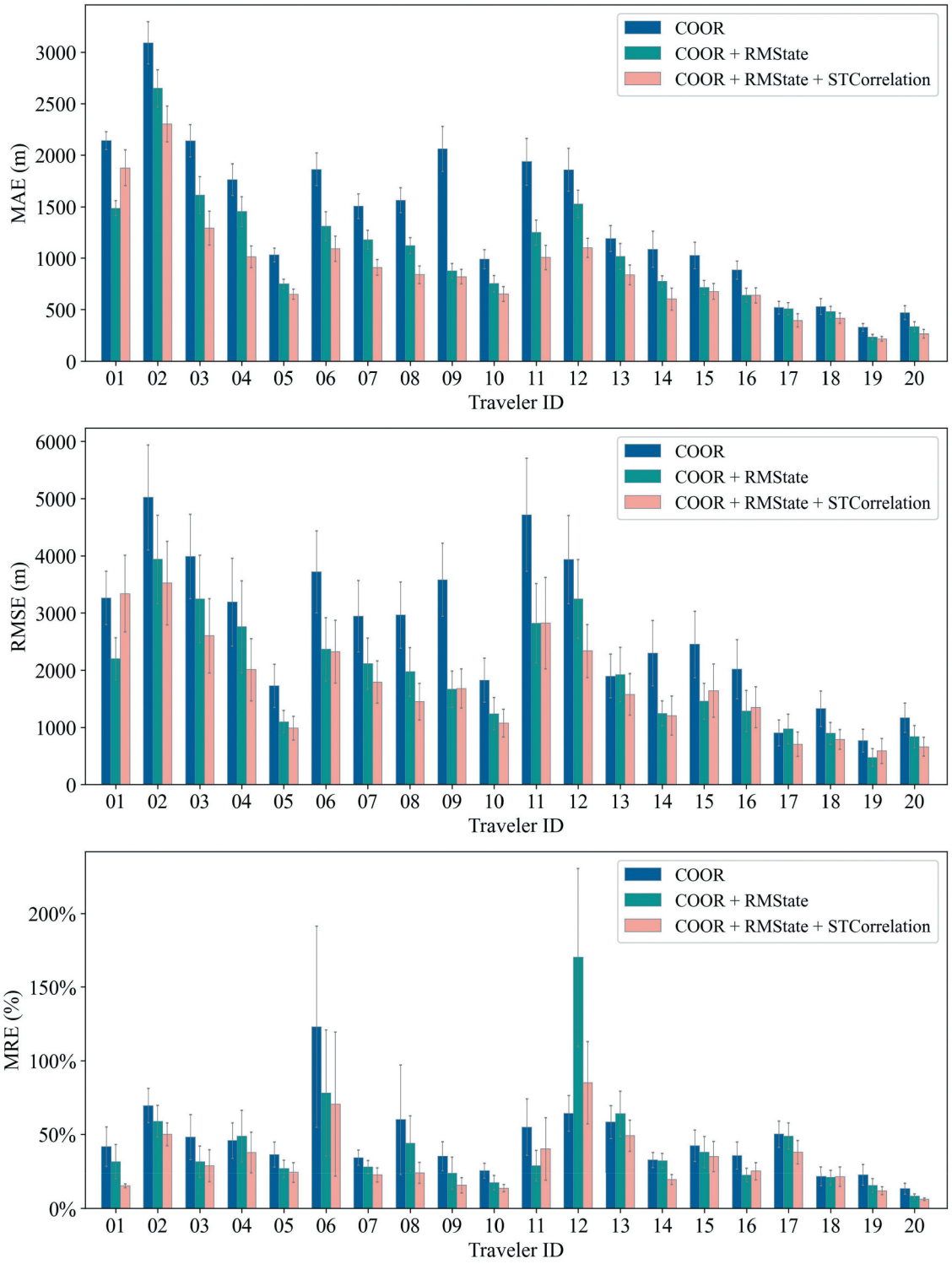


Figure 9. Effectiveness of current traveling features for each traveler. The error bars on the chart denote the SD of the three metrics.

values of temporal scores for each trip are represented by colors.

As the OD arcs shown in Figure 12 (a), traveler 05 has previously visited the destination to be predicted in trips 15, 18, 20, 24, 28, 30 and 32. They provide the model with prior knowledge to extract relevant hidden state features from these ODs. However, not all of them are assigned with higher temporal scores in Figure 12 (b), as departure times of day or days of week for trips 15, 18, 20, 28, 30 do not align closely with those of the current ongoing trip. While, trips 24 and 32 occurred on the same date or in the same time slot as the current ongoing trip, so they are more critical for real-time destination prediction compared to others. Besides, PTIntention learned through LDA exhibit temporal regularity, where the same departure times are linked to specific travel intentions. For instances, both historical trip 17 and

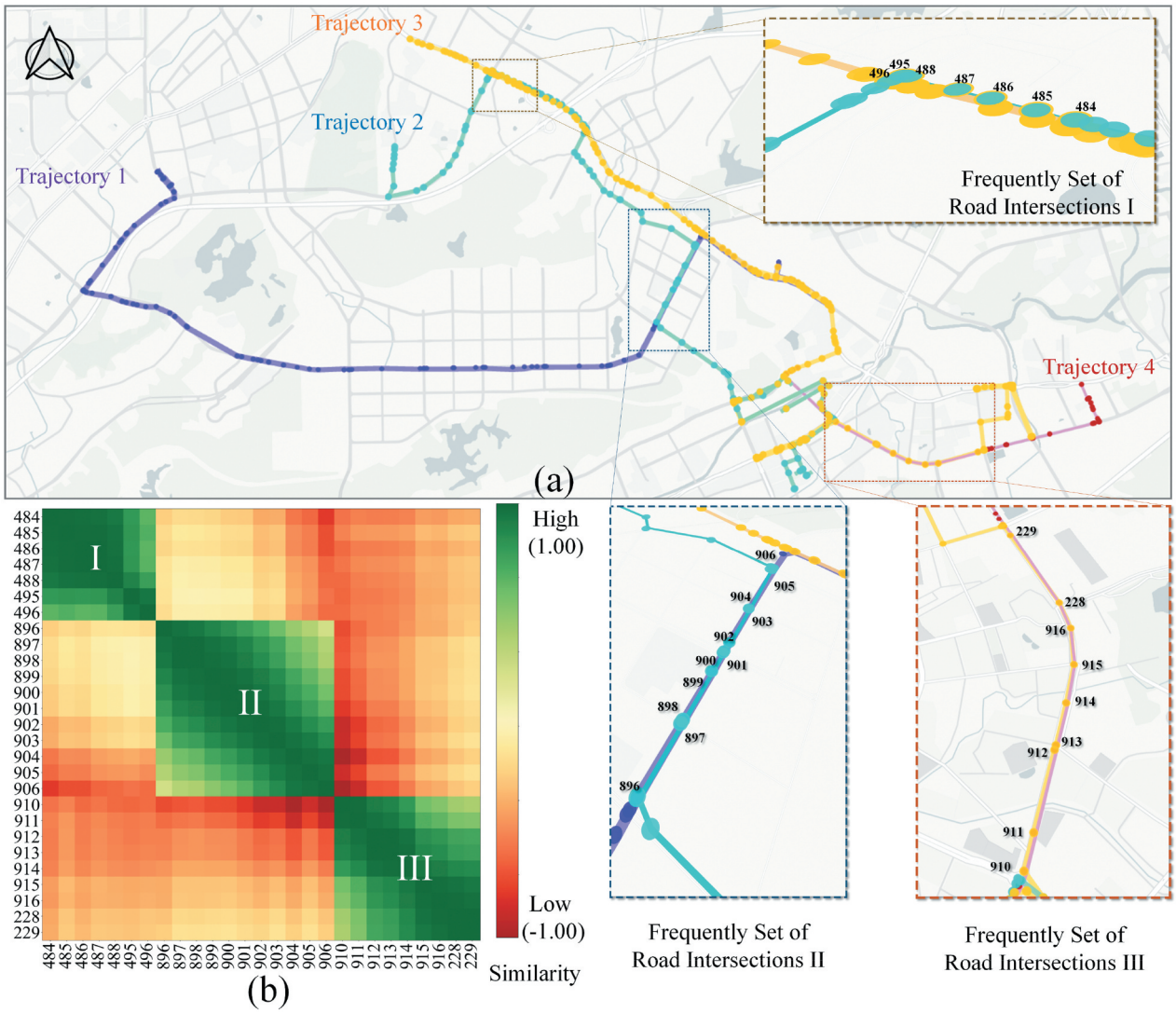


Figure 10. Illustration of word embedding of road intersections. (a) depicts four trajectories with different colors. (b) demonstrates the cosine similarity between the selected road intersections, where a greener color indicates greater similarity in their word vectors.

31 occurred at night, which have values of 0.480 in the fifth travel intention, may be strongly associated with returning home; trip 22 and trip 26 occurred at noontime around 1 pm, both have high values in travel intention 1; The trips 19 and 33 happening in early afternoon between 1:30 pm and 2 pm, also have similar PTIntention, higher in intentions 2 and 4 while lower in others.

In conclusion, we can derive the following findings. Overall, our feature design is effective, as it can enhance the prediction accuracy. Tables 2 and 3 reveal varying degrees of prediction accuracy improvements when current traveling features or historical activity chain are included. Specifically, the inclusion of current traveling features leads to a substantial improvement of over 20%. By contrast, the contribution of the historical activity chain to improving prediction accuracy is relatively limited. The probable reason is that trajectory points in the current ongoing trip provide the model with more real-time and high-resolution spatiotemporal contexts than historical activity chain. Moreover, the transitions between ODs introduce irrelevant information, as discussed in Section 5.4.2. In terms of the potential efficacies, the STCorrelation in current traveling features would enable us to personalize route planning for in-car map users, as it implies path preferences of individual travelers. Meanwhile, the historical activity chain would foster location-based amenity recommendation. For example, different travelers may explore various types of places, and visit the same place with differing frequencies. Based on historical activity chain, map service providers can recommend places of interest tailored to distinct mobility patterns of each individual user.

6. Further discussion

6.1. Impact of historical activity chain length on prediction accuracy

The length of the historical activity chain input into the model can affect the effectiveness of mobility

Table 3. Overall effectiveness of historical activity chain.

	MAE (m)		RMSE (m)		MRE (%)	
Schemes	MEAN	SD	MEAN	SD	MEAN	SD
COOR	1398.025	642.919	2684.803	2754.096	45.61%	69.96%
COOR + ST	1221.113	541.159	2364.756	**2440.096**	44.59%	82.50%
COOR + ST + PTIntention	**1164.468**	**530.031**	**2278.339**	2453.924	**42.10%**	**67.59%**

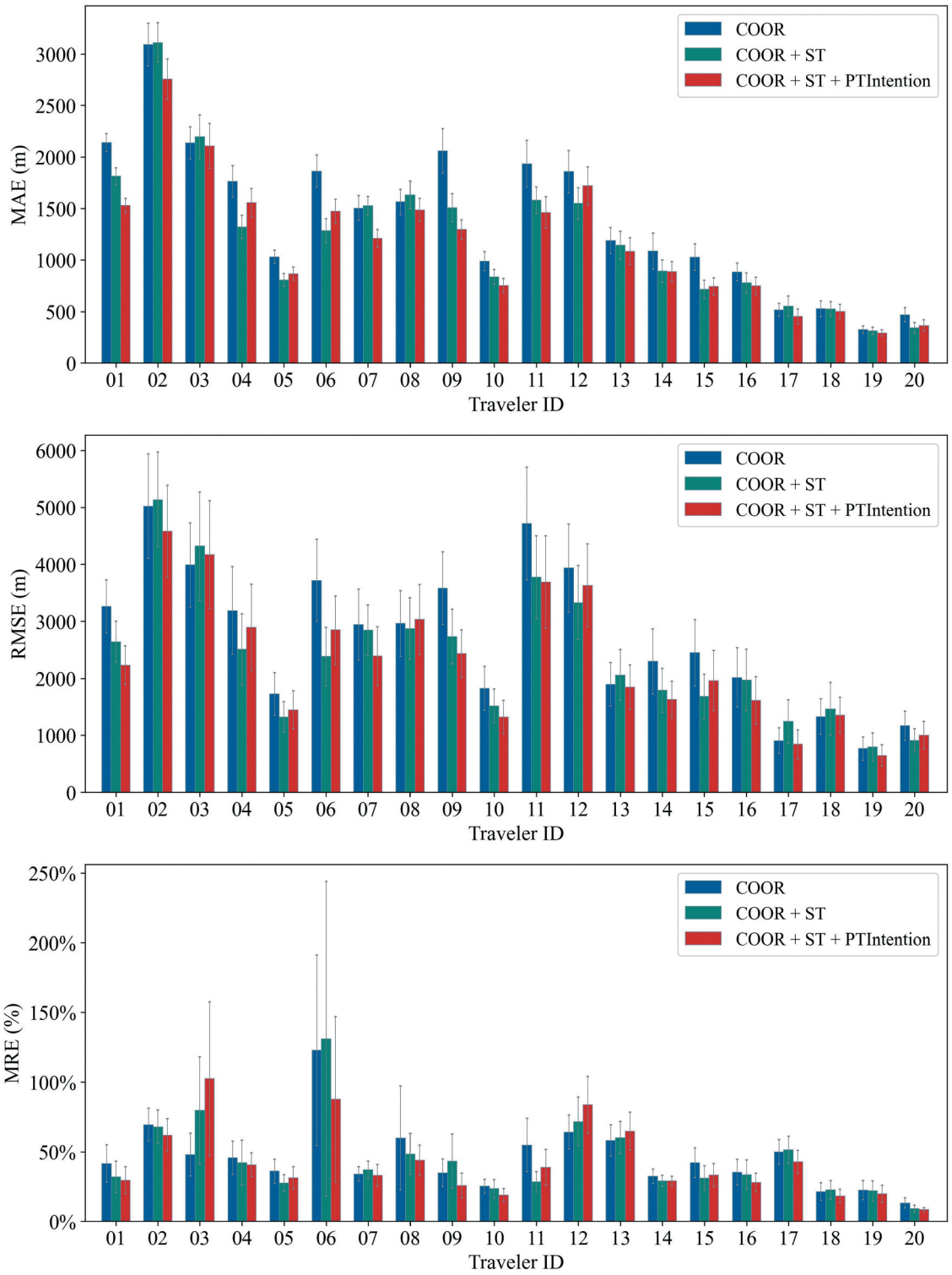
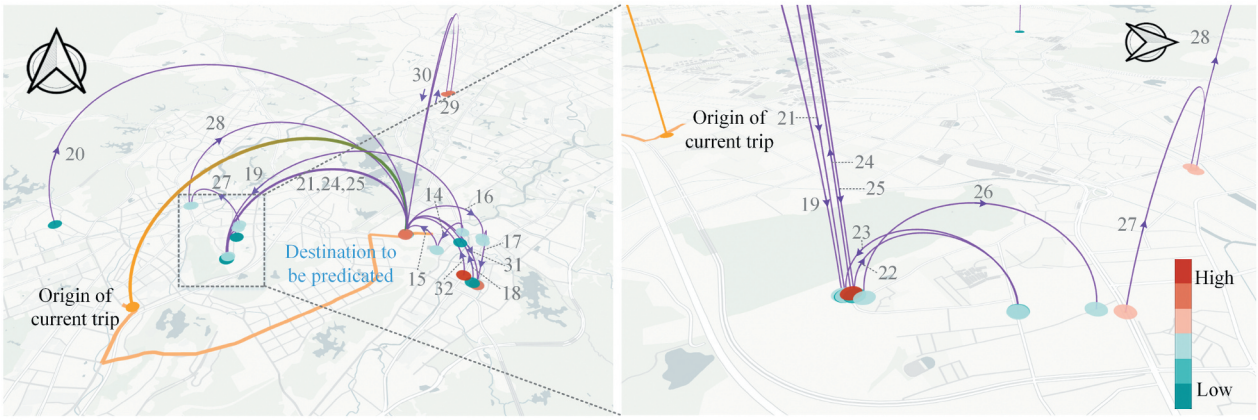


Figure 11. Effectiveness of historical activity chain for each traveler. The error bars on the chart denote the SD of the three metrics.



(a) OD arc and temporal scores $Score_{Sine}^{day-of-week}$, $Score_{Cosine}^{time-of-week}$

Trip ID	Departure time	$Score_{Sine}^{day-of-week}$	$Score_{Cosine}^{day-of-week}$	$Score_{Sine}^{time-of-week}$	$Score_{Cosine}^{time-of-week}$	Intention 1	Intention 2	Intention 3	Intention 4	Intention 5
33	1/7/2019 13:41	10.00	1.22	10.00	1.15	0.103	0.255	0.030	0.569	0.044
32	1/7/2019 9:20	10.00	1.27	10.00	0.72	0.622	0.156	0.089	0.077	0.056
31	1/6/2019 20:20	1.87	3.53	0.50	5.92	0.183	0.236	0.000	0.099	0.480
30	1/6/2019 18:38	1.87	2.00	0.50	3.18	0.236	0.360	0.049	0.337	0.018
29	1/6/2019 10:11	1.87	1.24	0.50	0.75	0.000	0.317	0.077	0.593	0.000
28	1/5/2019 14:28	1.13	1.27	0.58	1.34	0.236	0.360	0.049	0.337	0.018
27	1/5/2019 13:50	1.13	1.24	0.58	1.24	0.267	0.000	0.134	0.597	0.000
26	1/5/2019 13:14	1.13	1.20	0.58	1.07	0.686	0.000	0.131	0.030	0.143
25	1/5/2019 9:17	1.13	1.27	0.58	0.72	0.508	0.000	0.168	0.177	0.142
24	1/4/2019 21:14	0.93	4.47	0.76	7.04	0.236	0.360	0.049	0.337	0.018
23	1/4/2019 13:08	0.93	1.20	0.76	1.07	0.508	0.000	0.168	0.177	0.142
22	1/4/2019 12:05	0.93	1.19	0.76	0.94	0.687	0.000	0.131	0.030	0.143
21	1/4/2019 9:54	0.93	1.24	0.76	0.75	0.508	0.000	0.168	0.177	0.142
20	1/3/2019 23:48	0.93	2.79	1.14	0.48	0.236	0.360	0.049	0.337	0.018
19	1/3/2019 13:36	0.93	1.22	1.14	1.15	0.196	0.255	0.052	0.496	0.000
18	1/3/2019 11:27	0.93	1.19	1.14	0.88	0.351	0.545	0.031	0.000	0.071
17	1/2/2019 21:45	1.13	10.00	2.10	10.00	0.183	0.236	0.000	0.099	0.480
16	1/2/2019 21:29	1.13	6.15	2.10	8.40	0.756	0.204	0.027	0.000	0.013
15	1/2/2019 19:13	1.13	2.22	2.10	3.67	0.236	0.360	0.049	0.337	0.018
14	1/2/2019 18:04	1.13	1.82	2.10	2.79	0.352	0.545	0.031	0.000	0.071

(b) PTIntention and temporal scores in a historical activity chain

Figure 12. Illustration of a six-day historical activity chain of the traveler 05 and the corresponding PTIntention and temporal scores. (a) demonstrates the two of the four temporal scores for each OD point in the chain using the color of the dots, $Score_{Sine}^{day-of-week}$ and $Score_{Cosine}^{time-of-day}$. Blue to red indicates the scores of the historical ODs from low to high. (b) shows temporal regularity of PTIntention, and also illustrates the mechanism of temporal scores in capturing critical historical contexts.

prediction, as differences in mobility patterns exist among travelers. It influences the historical spatiotemporal context the model can capture. If the historical activity chain is too short, the model may fail to capture key spatiotemporal contexts. In contrast, if it is too long, it increases the training cost and introduces irrelevant noise. Therefore, determining the appropriate length of the historical activity chain is crucial. In this section, we set the length of the historical activity chain to 5, 10, 15, 25, 30, and 40, respectively, and analyzed the prediction accuracy. The 20 travelers are grouped using the same method as Section 5.3. The changes of accuracy as the length of the historical activity chain increases are shown in Figure 13.

As shown in Figure 13, the accuracy fluctuates with the length of historical activity chain; neither longer nor shorter chains always yield higher prediction accuracy. According to the average accuracy of each group, the changes show varying trends. For the length of 25, the mean value of MAE for the group (I) reaches its maximum (1706.047 m), while that of group (II) is the second lowest (887.899 m) among different lengths, it may be attributed to

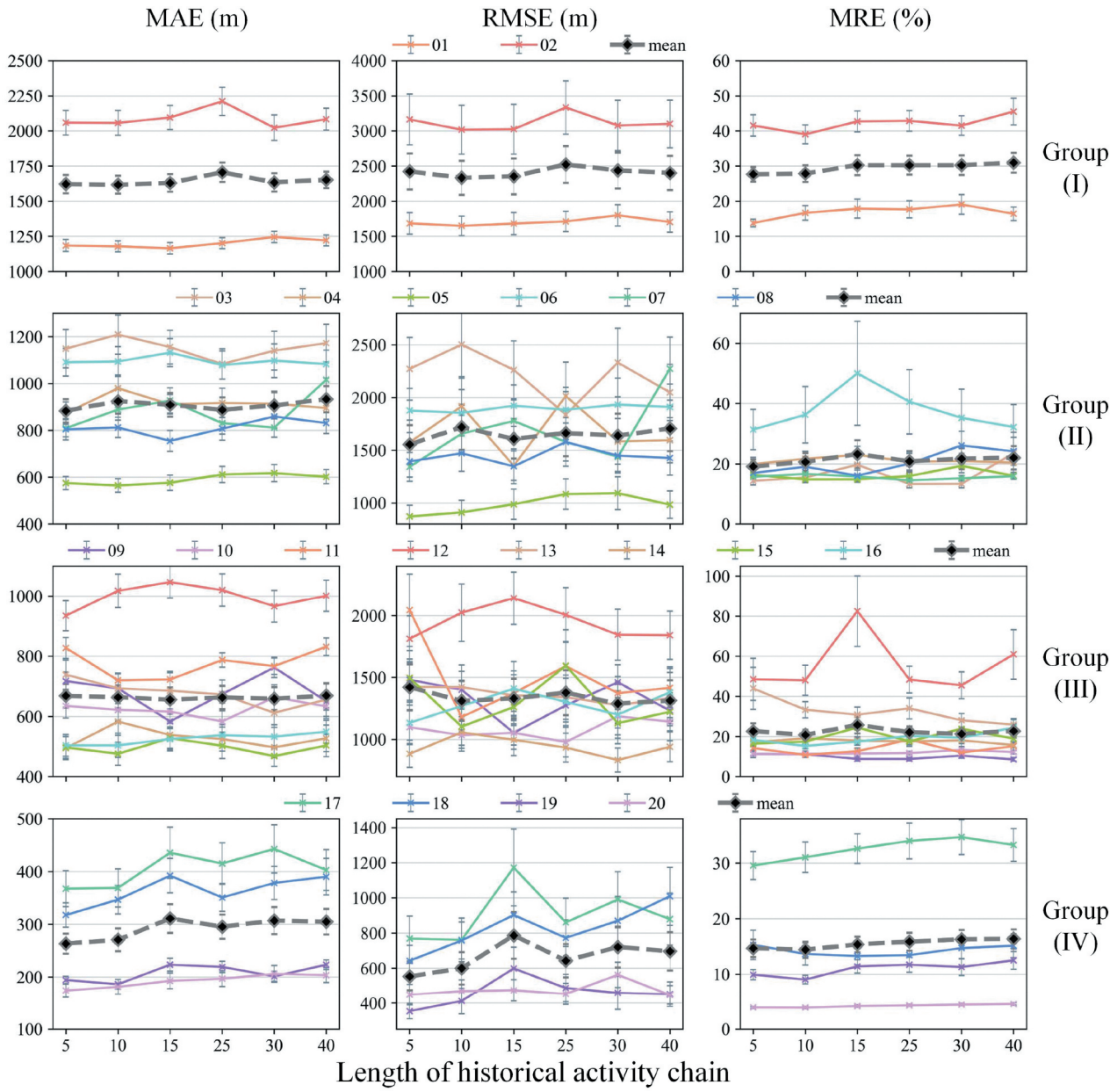


Figure 13. Length of historical activity chain affects predication accuracy.

variations of their historical traveling rhythms. From RMSE, we can find that group (II) with the second highest R_{IVL}, achieves higher prediction accuracy with shorter historical activity chain; In contrast, the group (III) has the second lowest R_{IVL}, attain higher accuracy with longer historical activity chain. This could be due to only a small proportion of destinations for travelers in group (II) appearing in their historical activity chains, while travelers in group (III) have higher likelihood of revisiting destinations of their historical trips. However, group (IV) does not exhibit the same pattern as group (III). It may be because the transitions between ODs of historical trips introduce irrelevant information. Besides, R_{IVL} is not the dominant factor influencing accuracy within a group of travelers. For example, traveler 02 demonstrates the highest prediction errors, despite having a lower R_{IVL} compared to traveler 01. It could be attributed to a large spatial range of traveling activities and low temporal regularity of traveler 02, reflecting by a larger radius of gyration (r_g) and lower day-of-week entropy ($E_{day-of-week}$). In general, when the lengths range in [5, 15], the averages of the three metrics for the 20 travelers reach relatively small values, MAE and MRE reach their smallest (i.e. [747.263 m, 760.234 m], and [20.178%, 23.375%]), and RMSE is the second lowest (i.e. [1388.127 m, 1407.478 m]). Consequently, to improve prediction accuracy, in practice, the historical activity chain length can be adjusted based on the statistical features of mobility

patterns (Cai et al. 2016). In the following two sections, we further discuss the correlation between destination visiting frequency and prediction accuracy, as well as relationship between mobility pattern features and predictability, respectively.

6.2. Correlation between destination visiting frequency and prediction accuracy

Travelers tend to visit specific locations frequently, such as residential and workplace areas, while the visiting frequencies of other locations are relatively low. Based on the variance in visiting frequencies, all destinations can be divided into frequently and infrequently visited destinations. Figure 14 shows the prediction accuracy of our model and two baselines, LSI-LSTM and ITP-CMM, for the two types of destinations.

As shown in Figure 14, our model outperforms the two baselines, especially for infrequently visited destinations. Besides, based on Figure 14 (a,b), we can find that the three models all demonstrate higher accuracy in predicting frequently visited destinations; while the

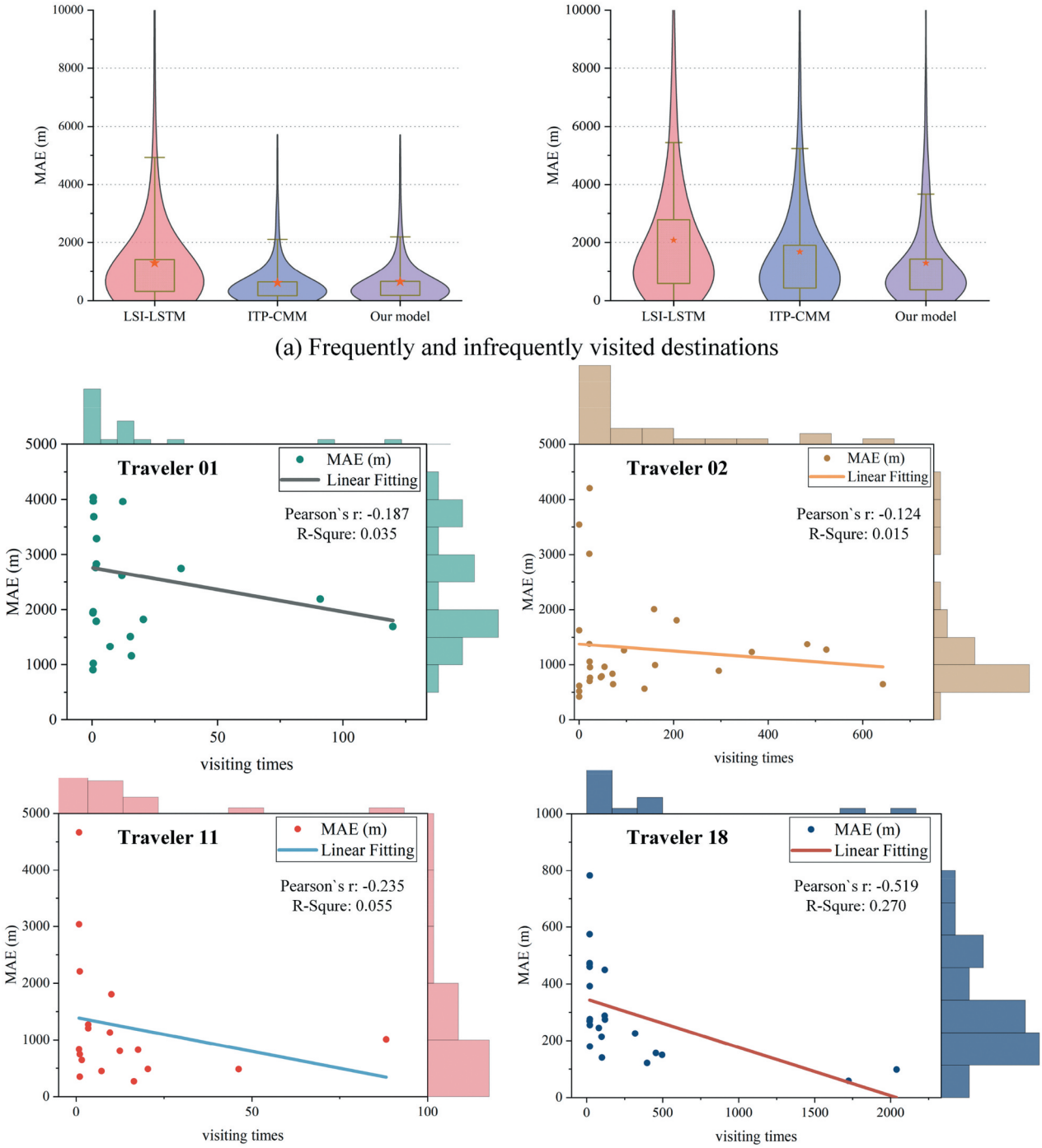


Figure 14. Correlation between destination visiting frequency and prediction accuracy. (a) depicts the accuracy for the frequently and infrequently visited destinations, respectively; the box within violin plots represents quartiles (the 25th and 75th percentiles), while the pentagon denotes the mean value of MAE. (b) depict the relationship between MAE and the visiting times of destinations for the selected four travelers. The curve in each subplot illustrates the fitted correlation between MAE and visiting times of destinations.

accuracy is lower for infrequently visited destinations. That is because the accuracy of deep learning models highly relies on the adequacy of data. For infrequently visited destinations, the limited data results in low prediction accuracy. Nevertheless, our model demonstrates lower discrepancies in predicting two types of destinations compared to the two baselines. It indicates that our model can better capture mobility patterns of low-frequency trajectories, in turn improving its stability for predicting infrequently visited destinations.

To further explore how the prediction accuracy of each destination is affected by its visiting frequency, we select four travelers 01, 02, 11, and 18. The results indicate a negative correlation between their visiting frequency and MAE, as shown in Figure 14 (b). We find that only a few destinations are visited repeatedly, while the majority of them have low visiting times. Besides, the MAE follows a skewed normal distribution, with most destinations achieving moderately high prediction accuracy. This reflects the capability of our model in capturing mobility patterns for most trips. In general, destinations with higher visiting frequencies have better predictability. Therefore, data augmentation via duplicating trajectories whose destinations are infrequently visited could help address data scarcity issues and enhance predictability. This technique can promote the driving experiences of intelligent and connected vehicles (Mozaffari et al. 2022). When a traveler is driving to an infrequently visited place, our model can proactively predict the destination using augmented trajectories. It would enable the map and LBS providers to recommend nearby public facilities, such as parking lots, restaurants, and entertainments. This capability not only reduces the time users spend searching for nearby amenities, but also increases the success rates of service recommendations. Consequently, it could yield a more seamless and enjoyable traveling experience.

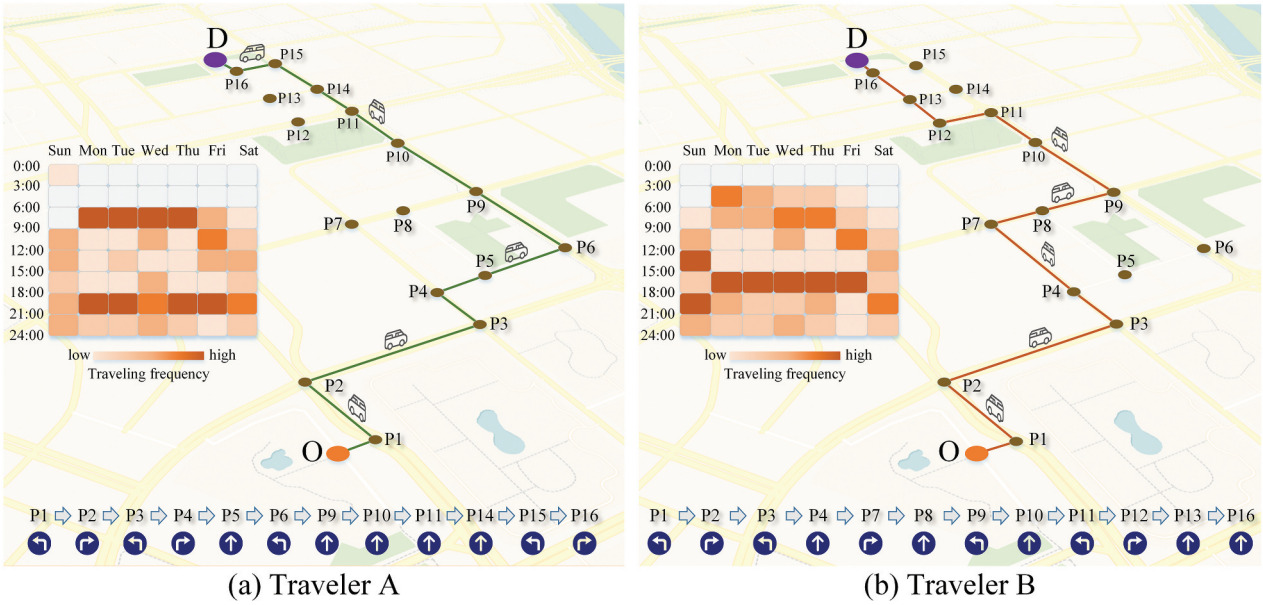
6.3. Relationship between mobility pattern features and predictability

The prediction accuracy varies among travelers with different mobility patterns, and the relationship between their mobility patterns and predictability remains to be explored. To analyze it, we extract eight statistical features of mobility patterns of an individual traveler, including TD_{avg}, R_{IVL}, radius of gyration (r_g), top N r_g, and four spatiotemporal-related entropies. r_g highlights spatial range of traveling activities, and top N r_g is defined by focusing only the N most frequently visited places to emphasize transitions among them (Pappalardo et al. 2015). Specifically, we set the N as 2, taking residential and workplace as two anchor points in traveling. Random entropy (E_{rand}) assesses the regularity of a traveler when visiting each location with equal probability. If we consider the frequency of a traveler visiting different locations, location entropy (E_{loc}) is obtained (Song et al. 2010). Day-of-week entropy ($E_{day-of-week}$) reveals variations in traveling temporal rhythms throughout a week. To distinguish travelers with different predictability, we rank them in a descending order according to MAE. The relationships between predictability and these statistical features are illustrated in Figure 15.

The result reveals that all these features have positive correlations with predictability in general, and the prediction accuracy is collectively determined by them rather than any single feature. Among them, TD_{avg} and r_g exhibit the highest correlations with predictability. For example, travelers 01, 02, 03, 06, and 12 with high values in the two features, demonstrate the highest MAE. These travelers occasionally venture far from their residential areas, and their infrequently long-distance trajectories pose challenges to predictability. While, traveler 11 is low in all the three accuracy metrics despite the r_g and TD_{avg} exceeding 10. This could be attributed to high frequency of long-distance trips reducing prediction error. By contrast, the correlation between the *Top 2* r_g and predictability is relatively weak. Travelers 01 and 08 with low rankings in *Top 2* r_g, have high MAE and RMSE; while travelers 09 and 17 who have high values on that show moderate predictabilities. It may be because traveling activities between residential and workplace have highest frequency of occurrence and are easy to be captured regardless of their distances. Meanwhile, the predictability of a traveler is closely related to spatial regularity of their mobility. For instances, travelers 01, 02, and 06 exhibit greater MAE than all the other travelers due to their higher spatial entropies, including E_{OD}, E_{rand}, and E_{loc}. However, travelers 05 and 10 whose spatial entropies are high, display top 50% predictabilities, and the three accuracy metrics of traveler 01 are lower than those of traveler 02, it reveals that a lower $E_{day-of-week}$, r_g, or TD_{avg} may mitigate the impact of spatial entropies on predictability. Moreover, a higher $E_{day-of-week}$ does not necessarily decrease prediction accuracy. For instance, travelers 15, and 18 have relatively high $E_{day-of-week}$, highlighting the inherent uncertainty in their daily traveling schedules. Nevertheless, their prediction accuracies are not lower than other travelers whose MAE below the median, i.e. 09, 14, 16, 17, 19, and 20. This finding emphasizes that our approach can capture irregular mobility patterns over time because of weak correlations between $E_{\text{day-of-week}}$ and accuracy metrics. Therefore, capturing the temporal dependencies of travelers and assigning greater weights to destinations of relevant historical trips through attention

Traveler ID	MAE ↓	RMSE	MRE	R_{IVL}	TD_{avg}	r_g	Top 2 r_g	E_{rand}	E_{loc}	E_{OD}	$E_{day\text{-}of\text{-}week}$
02	2257.62	3439.26	45.67%	49.73%	9.57	19.25	0.39	8.22	6.88	8.57	4.45
01	1158.61	1675.13	17.19%	72.95%	11.37	11.31	0.03	10.40	9.44	10.76	3.89
06	1126.91	2282.54	32.70%	19.72%	9.57	11.22	0.19	8.06	4.98	8.05	4.08
03	1107.14	2023.83	12.01%	28.46%	13.96	13.58	0.79	7.66	5.08	7.78	4.30
12	1016.78	2036.22	54.27%	13.99%	11.24	13.24	0.67	7.19	4.68	7.33	4.03
04	954.59	1933.87	23.75%	25.64%	6.57	8.24	0.82	7.01	3.95	6.30	4.09
08	864.67	1718.05	21.10%	19.03%	9.56	9.83	0.11	7.37	4.67	7.61	4.04
07	855.12	1642.40	15.85%	19.19%	9.45	11.03	0.34	7.04	4.30	7.09	3.88
11	782.38	1748.15	19.11%	14.10%	13.54	12.93	1.04	6.64	3.72	6.32	3.56
13	697.15	1489.90	28.31%	13.49%	4.32	5.13	0.63	6.38	3.29	5.61	4.22
09	610.08	1164.01	7.99%	15.50%	6.32	5.03	0.86	6.27	2.76	4.64	3.40
10	559.06	982.55	10.22%	15.23%	4.45	5.52	0.36	7.79	3.97	6.82	4.04
05	533.42	820.03	13.98%	20.51%	4.84	6.34	0.18	8.38	5.20	8.26	4.26
15	504.06	1264.79	19.04%	11.00%	4.20	8.76	0.07	7.17	2.87	4.96	4.42
14	502.57	866.21	16.35%	12.55%	3.31	4.11	0.10	6.21	3.51	6.01	3.71
16	493.97	1134.69	17.28%	10.62%	2.72	3.97	0.30	6.87	2.47	4.48	3.93
18	379.27	819.72	13.95%	5.45%	1.06	3.94	0.13	6.19	0.80	1.50	4.49
17	371.43	738.13	31.23%	5.60%	1.99	2.42	0.58	5.88	2.74	4.95	4.07
19	218.20	482.07	12.45%	5.26%	1.54	2.52	0.25	5.86	2.15	4.06	3.13
20	214.21	504.78	4.71%	2.70%	1.72	2.17	0.29	5.04	2.10	3.88	3.87

Figure 15. The correlation between mobility pattern features and predictability. It depicts the relationship between eight statistical features and the three metrics of prediction accuracy. The purple line depicts the median value of MAE.



(a) Traveler A (b) Traveler B

Figure 16. Preferences of traveling path and departure time for two individual travelers.

mechanisms may be two potential approaches to enhancing predictability (Guo et al. 2020; Xu et al. 2022). Our discussion could benefit location recommendation by building user profiles. For instance, service providers could categorize travelers into different groups and then construct distinct user profiles based on their mobility pattern features. User profiles could empower various application scenarios, such as personalized location recommendations and personality analysis (Liu, et al. 2025). For example, a traveler who often visits recreational places could receive more recommendations about entertainments. Similarly, a driver who drives aggressively, often exhibiting rapid acceleration and sudden braking, could receive safety alerts and be advised to adjust car insurance. Meanwhile, for those with few historical trajectories,

data from other travelers with similar profiles could help mitigate the cold start issue in retail service recommendations (Liu, et al. 2019b).

7. Conclusions

In this paper, we propose an individual mobility prediction method by considering the current traveling features and historical activity chain. Especially, in current traveling feature modeling, besides RMState, our method captures the path preferences of travelers using the spatiotemporal correlation of road intersections represented by CBoW. Meanwhile, it learns the long-term spatiotemporal dependencies from historical activity chain. The results of the comparative experiment have validated that our method outperforms the four baselines when applied to one-year individual vehicle trajectories of 20 travelers. Moreover, the contributions of current traveling features and historical activity chain to enhancing prediction accuracy are confirmed through ablation experiments. We also find that factors such as length of historical activity chain, and visiting frequency of destination have impacts on the prediction accuracy, whereas R_{IVL} is not the only determinant factor. Further analysis reveals the relationship between mobility patterns and predictability. Specifically, the traveling spatial entropies heavily influence prediction accuracy. Our destination prediction method and aforementioned findings would enable personalized location-based services and targeted advertising strategies.

In future, this work could be extended in several directions. First, we can improve predictability of travelers with higher R_{IVL} using data augmentation. For example, replicating low-frequency trajectories can augment samples for model training, and enables the model to learn such traveling activities more efficiently. Second, while this work focuses on destination prediction, future work could integrate other objectives through multi-task learning to enhance prediction accuracy. For instance, both the trip purposes and anticipated routes may aid in accurate destination prediction. Meanwhile, individual mobility patterns can be disaggregated into different profiles, thereby leverage trajectories from similar profiles to facilitate individual destination prediction.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by National Natural Science Foundation of China [grant number 42090010], National Key R&D Program of China [grant number 2021YFE0117000], and Fundamental Research Funds for the Central Universities [grant number 2042024kf0005].

Notes on contributors

Xiaotong Zhang received his B.E. and M.E. degrees in Geographical Information Engineering from China University of Geosciences (Wuhan) in 2022, and Wuhan University in 2024, respectively. He is a Ph.D. student at Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University. His research interests are individual mobility prediction and human perception of urban environment.

Zhipeng Gui is a full Professor in the School of Remote Sensing and Information Engineering, Wuhan University. His research interests include geospatial artificial intelligence, spatiotemporal data analysis and high-performance geocomputation. Now he serves as a Board of Directors (BOD) member of the International Association of Chinese Professionals in Geographic Information Sciences (CPGIS).

Yuhang Liu received her B.E. and M.S. degrees in Remote Sensing Science & Technology and Geographic Information Science from Wuhan University, in 2020 and 2023, respectively. She is a Ph.D. student at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. Her research interest focuses on human mobility and GIScience.

Dehua Peng received his B.E. and Ph.D. degrees in remote sensing science and technology and geographic information science from Wuhan University, in 2018 and 2024, respectively. He is a post-doctor with the School of Remote Sensing and Information Engineering, Wuhan University. His research interests include machine learning, artificial intelligence, spatial analysis, and point pattern mining.

Qianxi Lan received her B.E. degree in Remote Sensing Science & Technology from Wuhan University in 2023. She is a master's student at the School of Remote Sensing and Information Engineering, Wuhan University. Her research interest focuses on human mobility.

Zhangxiao Shen received his B.E. degree in Geodesy and Geomatics Engineering from Wuhan University in 2024. He is now a master's student at the State Key Laboratory in Information Engineering of Surveying, Mapping and Remote Sensing, Wuhan University. His research interest focuses on spatiotemporal data mining.

Huan Chen received his M.E. degree in Surveying and Mapping Engineering from Wuhan University in 2020. He is now pursuing his Ph.D. degree at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. His research interests include flow pattern analysis, human mobility, and spatial data mining.

Yuhui Zuo received her B.S. degree in Geographical Information Science from China University of Geosciences (Beijing) in 2023. She is a master's student at the State Key Laboratory of Information Engineering in Surveying,

Mapping and Remote Sensing, Wuhan University. Her research interest focuses on spatiotemporal big data.

Yao Yao is a full Professor at China University of Geosciences (Wuhan), a researcher at the University of Tokyo and a senior algorithm engineer at Alibaba Group. His research interests are geospatial big data mining, analysis, and computational urban science.

Huayi Wu is a full Professor in the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. His research interests are high-performance geospatial computing and intelligent geospatial web services.

Kai Li is an associate professor of Climate Change and Energy Economics Study Center, and Head of Department of the World Economy at the College of Economics and Management, Wuhan University. His research focuses on the world economy, low carbon economy and industrial innovation.

Kun Qin is a full Professor in the School of Remote Sensing and Information Engineering, Wuhan University. His research interests include spatiotemporal big data analysis and mining, remote sensing image mining and intelligent analysis, and geocomputation for humanities and social sciences.

ORCID

Xiaotong Zhang http://orcid.org/0009-0008-5160-2833
Zhipeng Gui http://orcid.org/0000-0001-9467-9680
Yuhang Liu http://orcid.org/0009-0009-3928-418X
Dehua Peng http://orcid.org/0000-0002-9842-0796
Qianxi Lan http://orcid.org/0009-0001-6965-661X
Yao Yao http://orcid.org/0000-0002-2830-0377
Huayi Wu http://orcid.org/0000-0003-3971-0512
Kai Li http://orcid.org/0009-0003-0940-9187
Kun Qin http://orcid.org/0000-0002-4088-1637

Data availability statement

The code and experimental statistic data that supports this study is openly available in GitHub at https://github.com/ZPGuiGroupWhu/Human-Mobility-Analysis/tree/master/Travel-Destination-Prediction/Current-Traveling-Features&Historical-Activity-Chain.

References

Alahi, A., K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese. 2016. "Social LSTM: Human Trajectory Prediction in Crowded Spaces." *Proceedings of the IEEE conference on computer vision and pattern recognition 2016*, Las Vegas, NV, USA, June 26-July 1.

Arain, Q. A., H. Memon, I. Memon, M. H. Memon, R. A. Shaikh, and F. A. Mangi. 2017. "Intelligent Travel Information Platform Based on Location Base Services to Predict User Travel Behavior from User-Generated GPS Traces." *International Journal of Computers and Applications* 39 (3): 155–168. https://doi.org/10.1080/1206212x.2017.1309222.

Arslan, M., C. Cruz, and D. Ginhac. 2019. "Semantic Enrichment of Spatio-Temporal Trajectories for Worker Safety on Construction Sites." *Personal and Ubiquitous Computing* 23 (5–6): 749–764. https://doi.org/10.1007/s00779-018-01199-5.

Belkina, A. C., C. O. Ciccolella, R. Anno, R. Halpert, J. Spidlen, and J. E. Snyder-Cappione. 2019. "Automated Optimized Parameters for T-Distributed Stochastic Neighbor Embedding Improve Visualization and Analysis of Large Datasets." *Nature Communications* 10 (1): 5415. https://doi.org/10.1038/s41467-019-13055-y.

Bonavita, A., R. Guidotti, and M. Nanni. 2021. "Individual and Collective Stop-Based Adaptive Trajectory Segmentation." *GeoInformatica* 26 (3): 451–477. https://doi.org/10.1007/s10707-021-00449-8.

Cai, H., X. Zhan, J. Zhu, X. Jia, A. S. Chiu, and M. Xu. 2016. "Understanding Taxi Travel Patterns." *Physica A: Statistical Mechanics and Its Applications* 457 (September): 590–597. https://doi.org/10.1016/j.physa.2016.03.047.

Chauhan, U., and A. Shah. 2021. "Topic Modeling Using Latent Dirichlet Allocation." *ACM Computing Surveys* 54 (7): 1–35. https://doi.org/10.1145/3462478.

Chen, P., W. Shi, X. Zhou, Z. Liu, and X. Fu. 2019. "STLP-GSM: A Method to Predict Future Locations of Individuals Based on Geotagged Social Media Data." *International Journal of Geographical Information Science* 33 (12): 2337–2362. https://doi.org/10.1080/13658816.2019.1630630.

Deo, N., and M. M. Trivedi. 2018. "Convolutional Social Pooling for Vehicle Trajectory Prediction." *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 18–22. Salt Lake City, UT, USA, June. https://doi.org/10.1109/cvprw.2018.00196.

Díaz-Ramírez, J., J. A. Estrada-García, and J. Figueroa-Sayago. 2023. "Predicting Transport Mode Choice Preferences in a University District with Decision Tree-Based Models." *City and Environment Interactions* 20 (December): 100118. https://doi.org/10.1016/j.cacint.2023.100118.

Du, Y., Y. Yi, X. Li, X. Chen, Y. Fan, and F. Su. 2020. "Extracting and Tracking Hot Topics of Micro-Blogs Based on Improved Latent Dirichlet Allocation." *Engineering Applications of Artificial Intelligence* 87 (January): 103279. https://doi.org/10.1016/j.engappai.2019.103279.

Elman, J. L. 1991. "Distributed Representations, Simple Recurrent Networks, and Grammatical Structure." *Machine Learning* 7 (2–3): 195–225. https://doi.org/10.1007/bf00114844.

Fei, N., Y. Gao, Z. Lu, and T. Xiang. 2021. "Z-Score Normalization, Hubness, and Few-Shot Learning." *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 142–151. October. https://doi.org/10.1109/iccv48922.2021.00021.

Feng, J., Y. Li, Z. Yang, Q. Qiu, and D. Jin. 2022. "Predicting Human Mobility with Semantic Motivation via Multi-Task Attentional Recurrent Networks." *IEEE Transactions on Knowledge and Data Engineering* 34 (5): 2360–2374. https://doi.org/10.1109/tkde.2020.3006048.

Forghani, M., F. Karimipour, and C. Claramunt. 2020. "From Cellular Positioning Data to Trajectories: Steps Towards a More Accurate Mobility Exploration." *Transportation Research Part C: Emerging Technologies* 117 (August): 102666. https://doi.org/10.1016/j.trc.2020.102666.

Gui, Z., Y. Sun, L. Yang, D. Peng, F. Li, H. Wu, C. Guo, W. Guo, and J. Gong. 2021. "LSI-LSTM: An Attention-Aware LSTM for Real-Time Driving Destination Prediction by Considering Location

Semantics and Location Importance of Trajectory Points." *Neurocomputing* 440 (June): 72–88. https://doi.org/10.1016/j.neucom.2021.01.067.

Gui, Z., L. Yang, J. Ding, J. Wang, Y. Sun, and H. Wu. 2024. "Individual Driving Destination Prediction by Considering Intersection Transfer Preference and Current Movement Mode." *Geomatics and Information Science of Wuhan University* 49 (5): 819–830. https://doi.org/10.13203/j.whugis20210555.

Guo, Q., Z. Sun, J. Zhang, and Y. Theng. 2020. "An Attentional Recurrent Neural Network for Personalized Next Location Recommendation." *Proceedings of the AAAI Conference on Artificial Intelligence* 34 (1): 83–90. https://doi.org/10.1609/aaai.v34i01.5337.

Hazelton, M. L., and L. Najim. 2024. "Using Traffic Assignment Models to Assist Bayesian Inference for Origin–Destination Matrices." *Transportation Research Part B: Methodological* 186 (August): 103019. https://doi.org/10.1016/j.trb.2024.103019.

Hochreiter, S., and J. Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9 (8): 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Jagannathan, P., S. Rajkumar, J. Frnda, P. B. Divakarachari, and P. Subramani. 2021. "Moving Vehicle Detection and Classification Using Gaussian Mixture Model and Ensemble Deep Learning Technique." *Wireless Communications and Mobile Computing* 2021 (1): 1–15. https://doi.org/10.1155/2021/5590894 .

Ke, J., X. Qin, H. Yang, Z. Zheng, Z. Zhu, and J. Ye. 2021. "Predicting Origin-Destination Ride-Sourcing Demand with a Spatio-Temporal Encoder-Decoder Residual Multi-Graph Convolutional Network." *Transportation Research Part C: Emerging Technologies* 122 (January): 102858. https://doi.org/10.1016/j.trc.2020.102858.

Kothari, P., S. Kreiss, and A. Alahi. 2021. "Human Trajectory Forecasting in Crowds: A Deep Learning Perspective." *IEEE Transactions on Intelligent Transportation Systems* 23 (7): 7386–7400. https://doi.org/10.1109/tits.2021.3069362.

Krishna, K., D. Jain, S. V. Mehta, and S. Choudhary. 2018. "An LSTM Based System for Prediction of Human Activities with Durations." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1 (4): 1–31. https://doi.org/10.1145/3161201.

Li, F., Z. Gui, Z. Zhang, D. Peng, S. Tian, K. Yuan, Y. Sun, H. Wu, J. Gong, and Y. Lei. 2020. "A Hierarchical Temporal Attention-Based LSTM Encoder-Decoder Model for Individual Mobility Prediction." *Neurocomputing* 403 (August): 153–166. https://doi.org/10.1016/j.neucom.2020.03.080.

Li, M., S. Gao, F. Lu, and H. Zhang. 2019. "Reconstruction of Human Movement Trajectories from Large-Scale Low-Frequency Mobile Phone Data." *Computers, Environment and Urban Systems* 77 (May): 101346. https://doi.org/10.1016/j.compenvurbsys.2019.101346.

Liu, Y., Z. Gui, Y. Xu, S. Gao, A. Zhao, F. Meng, D. Peng, et al. 2025. "Profiling Mobility Patterns and Driving Behaviors of Individual Drivers via Trajectory Trait." *The Innovation Geoscience* 3 (1): 100114. https://doi.org/10.59717/j.xinn-geo.2024.100114.

Liu, Y., R. Jia, X. Xie, and Z. Liu. 2019a. "A Two-Stage Destination Prediction Framework of Shared Bicycles Based on Geographical Position Recommendation." *IEEE Intelligent Transportation Systems Magazine* 11 (1): 42–47. https://doi.org/10.1109/mits.2018.2884517.

Liu, Y., S. Wang, X. Wang, Y. Zheng, X. Chen, Y. Xu, and C. Kang. 2024. "Towards Semantic Enrichment for Spatial Interactions." *Annals of GIS* 30 (2): 151–166. https://doi.org/10.1080/19475683.2024.2324392.

Liu, Y., L. Yao, B. Guo, N. Li, J. Zhang, J. Chen, D. Zhang, Y. Liu, Z. Yu, and S. Zhang. 2019b. "DeepStore: An Interaction-Aware Wide & Deep Model for Store Site Recommendation with Attentional Spatial Embeddings." *IEEE Internet of Things Journal* 6 (4): 7319–7333. https://doi.org/10.1109/jiot.2019.2916143.

Luca, M., G. Barlacchi, B. Lepri, and L. Pappalardo. 2021. "A Survey on Deep Learning for Human Mobility." *ACM Computing Surveys* 55 (1): 1–44. https://doi.org/10.1145/3485125.

Lucchini, L., S. Centellegher, L. Pappalardo, R. Gallotti, F. Privitera, B. Lepri, and M. De Nadai. 2021. "Living in a Pandemic: Changes in Mobility Routines, Social Activity and Adherence to COVID-19 Protective Measures." *Scientific Reports* 11 (1). https://doi.org/10.1038/s41598-021-04139-1.

Ma, Y., X. Zhu, S. Zhang, R. Yang, W. Wang, and D. Manocha. 2019. "TrafficPredict: Trajectory Prediction for Heterogeneous Traffic-Agents." *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (1): 6120–6127. https://doi.org/10.1609/aaai.v33i01.33016120.

Ma, Z., and P. Zhang. 2022. "Individual Mobility Prediction Review: Data, Problem, Method and Application." *Multimodal Transportation* 1 (1): 100002. https://doi.org/10.1016/j.multra.2022.100002.

Miyazawa, S., X. Song, R. Jiang, Z. Fan, R. Shibasaki, and T. Sato. 2020. "City-Scale Human Mobility Prediction Model by Integrating GNSS Trajectories and SNS Data Using Long Short-Term Memory." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* V-4–2020 (August, 87–94. https://doi.org/10.5194/isprs-annals-v-4-2020-87-2020.

Mozaffari, S., E. Arnold, M. Dianati, and S. Fallah. 2022. "Early Lane Change Prediction for Automated Driving Systems Using Multi-Task Attention-Based Convolutional Neural Networks." *IEEE Transactions on Intelligent Vehicles* 7 (3): 758–770. https://doi.org/10.1109/tiv.2022.3161785.

Murray, D., J. Yoon, S. Kojaku, R. Costas, W. Jung, S. Milojević, and Y. Ahn. 2023. "Unsupervised Embedding of Trajectories Captures the Latent Structure of Scientific Migration." *Proceedings of the National Academy of Sciences*, 120 (52). https://doi.org/10.1073/pnas.2305414120.

Pappalardo, L., E. Manley, V. Sekara, and L. Alessandretti. 2023. "Future Directions in Human Mobility Science." *Nature Computational Science* 3 (7): 588–600. https://doi.org/10.1038/s43588-023-00469-4.

Pappalardo, L., F. Simini, S. Rinzivillo, D. Pedreschi, F. Giannotti, and A. Barabási. 2015. "Returners and Explorers Dichotomy in Human Mobility." *Nature Communications* 6 (1): 8166. https://doi.org/10.1038/ncomms9166.

Park, S., Y. Xu, L. Jiang, Z. Chen, and S. Huang. 2020. "Spatial Structures of Tourism Destinations: A Trajectory Data Mining Approach Leveraging Mobile Big Data." *Annals of Tourism Research* 84 (June): 102973. https://doi.org/10.1016/j.annals.2020.102973.

Qian, C., R. Jiang, Y. Long, Q. Zhang, M. Li, and L. Zhang. 2019. "Vehicle Trajectory Modelling with Consideration of Distant Neighbouring Dependencies for Destination

Prediction." *International Journal of Geographical Information Science* 33 (10): 2011–2032. https://doi.org/10.1080/13658816.2019.1620236.

Sevtsuk, A., R. Basu, X. Li, and R. Kalvo. 2021. "A Big Data Approach to Understanding Pedestrian Route Choice Preferences: Evidence from San Francisco." *Travel Behaviour and Society* 25 (October): 41–51. https://doi.org/10.1016/j.tbs.2021.05.010.

Shang, S., S. Du, S. Du, and S. Zhu. 2021. "Estimating Building-Scale Population Using Multi-Source Spatial Data." *Cities* 111 (April): 103002. https://doi.org/10.1016/j.cities.2020.103002.

Sîrbu, A., G. Andrienko, N. Andrienko, C. Boldrini, M. Conti, F. Giannotti, R. Guidotti, et al. 2021. "Human Migration: The Big Data Perspective." *International Journal of Data Science and Analytics* 11 (4): 341–360. https://doi.org/10.1007/s41060-020-00213-5.

Song, C., Z. Qu, N. Blumm, and A. Barabási. 2010. "Limits of Predictability in Human Mobility." *Science* 327 (5968): 1018–1021. https://doi.org/10.1126/science.1177170.

Sun, J., and J. Kim. 2021. "Joint Prediction of Next Location and Travel Time from Urban Vehicle Trajectories Using Long Short-Term Memory Neural Networks." *Transportation Research Part C: Emerging Technologies* 128 (May): 103114. https://doi.org/10.1016/j.trc.2021.103114.

Tang, J., J. Liang, S. Zhang, H. Huang, and F. Liu. 2018. "Inferring Driving Trajectories Based on Probabilistic Model from Large Scale Taxi GPS Data." *Physica A: Statistical Mechanics and Its Applications* 506 (September): 566–577. https://doi.org/10.1016/j.physa.2018.04.073.

Testi, I., A. Wang, S. Paul, S. Mora, E. Walker, M. Nyhan, F. Duarte, P. Santi, and C. Ratti. 2024. "Big Mobility Data Reveals Hyperlocal Air Pollution Exposure Disparities in the Bronx, New York." *Nature Cities* 1 (8): 512–521. https://doi.org/10.1038/s44284-024-00093-x.

Wan, H., Y. Lin, S. Guo, and Y. Lin. 2022. "Pre-Training Time-Aware Location Embeddings from Spatial-Temporal Trajectories." *IEEE Transactions on Knowledge and Data Engineering* 34 (11): 5510–5523. https://doi.org/10.1109/tkde.2021.3057875.

Wang, S., Z. Bao, J. S. Culpepper, and G. Cong. 2021a. "A Survey on Trajectory Data Management, Analytics, and Learning." *ACM Computing Surveys* 54 (2): 1–36. https://doi.org/10.1145/3440207.

Wang, H., S. Zeng, Y. Li, and D. Jin. 2021b. "Predictability and Prediction of Human Mobility Based on Application-Collected Location Data." *IEEE Transactions on Mobile Computing* 20 (7): 2457–2472. https://doi.org/10.1109/tmc.2020.2981441.

Wang, J., Z. Gui, Y. Sun, H. Wu, and Z. Yu. 2021c. "A Real-Time Driving Destination Prediction Model Based on Historical Travel Patterns and Current Driving Status." *Lecture Notes in Computer Science*: 29–43. https://doi.org/10.1007/978-3-030-69873-7_3.

Wiest, J., M. Hoffken, U. Kresel, and K. Dietmayer. 2012. "Probabilistic Trajectory Prediction with Gaussian Mixture Models." *IEEE Intelligent Vehicles Symposium*, June. https://doi.org/10.1109/ivs.2012.6232277.

Wu, L., L. Yang, Z. Huang, Y. Wang, Y. Chai, X. Peng, and Y. Liu. 2019. "Inferring Demographics from Human Trajectories and Geographical Context." *Computers, Environment and Urban Systems* 77 (September): 101368. https://doi.org/10.1016/j.compenvurbsys.2019.101368.

Xia, H. 2023. "Continuous-Bag-Of-Words and Skip-Gram for Word Vector Training and Text Classification." *Journal of Physics: Conference Series* 2634 (1): 012052. https://doi.org/10.1088/1742-6596/2634/1/012052.

Xing, Y., K. Wang, and J. J. Lu. 2020. "Exploring Travel Patterns and Trip Purposes of Dockless Bike-Sharing by Analyzing Massive Bike-Sharing Data in Shanghai, China." *Journal of Transport Geography* 87 (July): 102787. https://doi.org/10.1016/j.jtrangeo.2020.102787.

Xu, S., X. Fu, D. Pi, and Z. Ma. 2024. "Inferring Individual Human Mobility from Sparse Check-In Data: A Temporal-Context-Aware Approach." *IEEE Transactions on Computational Social Systems* 11 (1): 600–611. https://doi.org/10.1109/tcss.2022.3231601.

Xu, Y., D. Zou, S. Park, Q. Li, S. Zhou, and X. Li. 2022. "Understanding the Movement Predictability of International Travelers Using a Nationwide Mobile Phone Dataset Collected in South Korea." *Computers, Environment and Urban Systems* 92 (March): 101753. https://doi.org/10.1016/j.compenvurbsys.2021.101753.

Yang, C., and G. Gidófalvi. 2018. "Fast Map Matching, an Algorithm Integrating Hidden Markov Model with Precomputation." *International Journal of Geographical Information Science* 32 (3): 547–570. https://doi.org/10.1080/13658816.2017.1400548.

Yang, S., Y. Fang, X. Wang, Y. Li, C. Fang, Y. Shan, B. Feng, and W. Liu. 2021. "Crossover Learning for Fast Online Video Instance Segmentation." *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, October. https://doi.org/10.1109/iccv48922.2021.00794.

Yao, Y., Z. Guo, C. Dou, M. Jia, Y. Hong, Q. Guan, and P. Luo. 2023. "Predicting Mobile Users' Next Location Using the Semantically Enriched Geo-Embedding Model and the Multilayer Attention Mechanism." *Computers, Environment and Urban Systems* 104 (September): 102009. https://doi.org/10.1016/j.compenvurbsys.2023.102009.

Yin, X., G. Wu, J. Wei, Y. Shen, H. Qi, and B. Yin. 2022. "Deep Learning on Traffic Prediction: Methods, Analysis, and Future Directions." *IEEE Transactions on Intelligent Transportation Systems* 23 (6): 4927–4943. https://doi.org/10.1109/tits.2021.3054840.

Yu, W., Y. Zhang, T. Ai, Q. Guan, Z. Chen, and H. Li. 2019. "Road Network Generalization Considering Traffic Flow Patterns." *International Journal of Geographical Information Science* 34 (1): 119–149. https://doi.org/10.1080/13658816.2019.1650936.

Zhou, F., Y. Dai, Q. Gao, P. Wang, and T. Zhong. 2021. "Self-Supervised Human Mobility Learning for Next Location Prediction and Trajectory Classification." *Knowledge-Based Systems* 228 (September): 107214. https://doi.org/10.1016/j.knosys.2021.107214.

Appendix

Figure 16 illustrates the preferences of traveling path and departure time for two individual travelers. They frequently visit similar places (i.e., O and D in the figure), but their departure times, traveling paths, and turning behaviors at intermediate road intersections differ along the way. For example, traveler A always leaves for home at 8 PM, turning right at road intersection P4, and going straight at P11. While traveler B tends to depart before 6 PM, then goes straight and turns left at the abovementioned two intersections, respectively. To imply their path preferences, the spatiotemporal correlations of road intersections traversed by each traveler are independently embedded as word vectors.

The pseudocode for travel intention representation is shown in Algorithm below. It utilizes Latent Dirichlet Allocation (LDA) to learn travel intentions as well as their possibilities in each trip.

Algorithm: LDA-based Travel Intention Representation

Input:
- α, η: Prior parameters for the Dirichlet distribution
- K: Number of POI categories
- Q: Number of travel intentions
- G: Number of grid cells
- *epochs*: Number of iterations

Output:
- $\theta_{g,q}$: Travel intention distribution for grid cells
- $\beta_{q,k}$: POI category distribution for a kind of travel intention

//Initialize number of POIs in the g-th grid cell assigned to intention q, number of times the k-th POI category is assigned to intention q, and total number of POIs assigned to intention q.

$N_{g,q} = zero(g, q)$; $N_{q,k} = zero(q, k)$; $N_q = zero(q)$

//Main loop for updating travel intention and POI category distribution in the grid cell.

For *epoch* in 1:*epochs* do

For g in 1:G do

For k in 1:K do

//Loop for updating travel intentions.

For q in 1:Q do

//compute the probability of k-th POI category in intention q.

$P(q|k) = comp\_prob(N_{g,q}, N_{q,k}, N_q, \eta)$

//compute the probability of q-th travel intention in grid cell g.

$P(g|q) = comp\_prob(N_{g,q}, N_{q,k}, N_q, \alpha)$

//Sample the q-th travel intention from this distribution.

$intention_q = \text{Sample}(K, \text{prob} = P(q|k))$

//Sample the possibilities of travel intentions for the g-th grid cell.

$PTIntention_g = \text{Sample}(Q, \text{prob} = P(g|q))$
//Update $N_{g,q}$, $N_{k,q}$, and N_q with the q-th travel intention.

$N_{g,q}[g][q] += 1$; $N_{q,k}[q][k] += 1$; $N_q[q] += 1$

//Update travel intention distribution for each grid cell.

For g in 1:G and q in 1:Q do

$\theta_{g,q}[g][q] = (N_{g,q}[g][q] + \alpha)/(K + Q * \alpha)$

//Update POI category distribution for each travel intention.

For q in 1:Q and k in 1:K do

$\beta_{q,k}[q][k] = (N_{q,k}[q][k] + \eta)/(N_q[q] + K * \eta)$

return $\theta_{g,q}$, $\beta_{q,k}$