IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 5, 2019, accepted August 1, 2019, date of publication August 7, 2019, date of current version August 21, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2933558

SEABIG: A Deep Learning-Based Method for Location Prediction in Pedestrian Semantic Trajectories

WANLONG ZHANG[1], LITING SUN[1], XIANG WANG[1], ZHITAO HUANG[1,2], AND BAOGUO LI[1]

[1]Department of Electronic Science and Engineering, National University of Defense Technology, Changsha 410075, China
[2]Academic Research Office, National University of Defense Technology, Changsha 410075, China

Corresponding author: Wanlong Zhang (18570108250@ 163.com)

This work was supported in part by the Program for Innovative Research Groups of the Hunan Provincial Natural Science Foundation of China under Grant 2019JJ10004.

ABSTRACT Pedestrian destination prediction of a user is known as an important and challenging task for LBSs (location-based services) like traffic planning and travelling recommendation. The typical method generally applies statistical model to predict the future location based on the raw trajectory. However, while predicting, existing approaches fall short in accommodating long-range dependency and ignore the semantic information existing in the raw trajectory. In this paper, we proposed a method named semantics-enriched attentional BiGRU (SEABIG) to solve the two problems. Firstly, we designed a probabilistic model based on the GMM (Gaussian mixture model) to extract stopover points from the raw trajectories and annotate the semantic information on the stopover points. Then we proposed an attentional BiGRU-based trajectory prediction model, which can jointly learn the embeddings of the semantic trajectory. It not only takes the advantage of the BiGRU (Bidirectional Gated Recurrent Unit) for sequence modeling, but also gives more attention to meaningful positions that have strong correlations w.r.t. destination by applying attention mechanism. Finally, we annotate the most likely semantic on the predicted position with the probabilistic model. Extensive experiments on Beijing real datasets demonstrate that our proposed method has higher prediction accuracy.

INDEX TERMS Trajectory prediction, semantic trajectory, deep learning.

I. INTRODUCTION

The widespread use of private vehicles with positioning services and the rapid advance in wireless mobile communication technology (such as 4G and beyond) enable us to acquire large-scale GPS trajectories [1]. Based on the massive data, the research of the next location forecast has aroused widespread concern [2]–[4] in recent years. Besides, trajectory prediction is of great significance for location based services (LBSs) [5], [6] such as passenger destination recommendations, targeted advertisements delivery and navigation services.

However, as demand arises, the semantic information which contains the intents of users is extracted from the raw trajectories with external data sources (e.g., land uses, social media) [7]–[9] to improve the accuracy of prediction.

Here, we propose a pedestrian location prediction with semantic trajectories method named semantics-enriched attentional BiGRU model, which constructs a GMM-based semantic annotating part and an attentional BiGRU-based next location prediction part.

To better present our work, the rest of this paper is arranged as follows. We describe selected related works in Section 2. Then, some useful definitions and the framework of the model will be listed in Section 3. We introduce the structure of SEABIG in detail in Section 4. Finally, we present our results in Section 5, and conclude the paper in Section 6.

II. RELATED WORK

A. SEMANTIC ENRICHING

In order to enrich the raw trajectory with semantic information, raw trajectories were linked with auxiliary information, e.g., land use and road network. And there are two kinds of methods existing to utilize such semantic information.

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wang.

One based on user-generated tags, such as in [10], a latent topic-based clustering algorithm was proposed to add geo-tagged text message to raw data, and Wu *et al*. presented a frequency-based algorithm [11] to link the mobility data with social media. Those methods can handle the low-sampling data with user-generated information. The other one based on map information [12]–[14], such as Yan *et al*. proposed a algorithm [15] to extract stops of each trajectory and add semantic landmarks to the stops.

B. TRAJECTORY PREDICTION

The present pedestrian location prediction methods can be roughly grouped into two categories: pattern-based methods and model-based methods. Pattern-based methods utilize the sequence analysis to predict the next location. Various approaches have been presented for sequential patterns mining [16], periodic patterns [17], such as Boukhechba *et al*. presented a method [18] which accommodates the changes in the trajectories and use online mining of association rules that support concept drift. Though these methods have been shown to be useful for position forecast, only explicit patterns named apriori can be mined and all the motion regularities cannot be captured.

The other method is model-based location prediction. Researchers use statistical models to capture the motion regularity and make predictions with those trained models. A number of models have been proposed, such as mobility Markov chains [19], matrix factorization (MF) [20]. However, those models cannot capture the semantics of user activities and solve the problem of long-range dependency.

Besides, some research use deep learning approaches to process the spatio-temporal data. A CSSRNN-LPIRNN model [21] was designed to model the trajectory. And in the area of trajectory prediction, the LSTM and its variants [22] perform well. A hierarchical spatial-temporal LSTM [23] was used to mitigate the problem of data sparsity and model the contextual historic visit information in trajectory prediction. And a ST-LSTM [24] which implements time gates and distance gates into LSTM to capture the spatio-temporal relation between successive check-ins to recommend next location to users.

III. PROBLEM DEFINITION

In the section, the definitions used in the paper are listed first and then the model of semantic enriching and the trajectory prediction are introduced.

A. NOTATIONS

Definition 1: Raw trajectory. Locations generated chronologically by user i in the jth day consist of one raw trajectory $R_j^i = \{l_1, l_2, \ldots, l_n\}$. And each location l_k $(k = 1, 2, \ldots, n)$ is further expressed as spatio-temporal sequence $[lat_k, lon_k, t_k]$, where lat, lon and t represent the latitude, the longitude and the located time of the trajectory location, respectively.

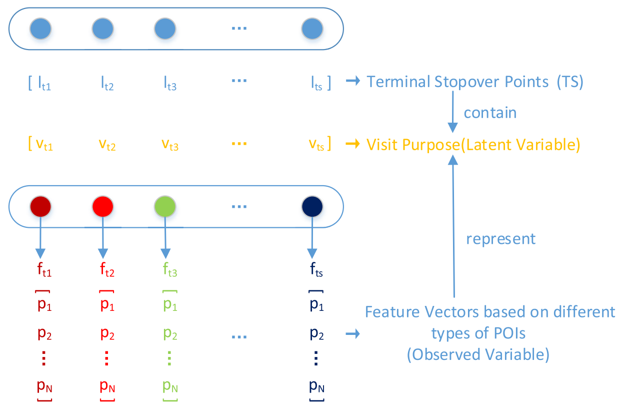


FIGURE 1. The connections of the definitions.

Definition 2: Initial stopover trajectory. The locations set $IS_j^i = \{l_1, l_2, \ldots, l_m\}$ $(m \leq n)$ of which time intervals are longer than a threshold t_0 and the distance is shorter than a threshold d_0 represent the stopover points of the trajectory generated by user i in the jth day.

Definition 3: Terminal stopover trajectory. A terminal stopover trajectory $TS_j^i = \{l_{t1}, l_{t2}, \ldots, l_{ts}\}(s \leq m)$ is the sequence of fine-grained extracted initial stopover points.

Definition 4: Feature vector. A feature vector of a location in a stopover trajectory is defined as $f=[p_1, p_2, \ldots, p_N]$, N is the number of the POI type on the map, and p_i is the radio of the numbers of the POIs of type i and the total number of the POIs in a circle area centered at the location.

The connection between these definitions is shown in Figure 1.

B. SEMANTIC ENRICHING PROBLEM

The raw trajectories contain two parts: the moving part and the stop part. Based on the spatio-temporal threshold and the DBSCAN algorithm, we extracted the stopover points from the raw GPS data to construct the stopover trajectories, which contain the pedestrian moving purposes. And we trained the GMM model with the stopover trajectories and map information to accomplish the semantic enriching.

So in this part, we aim to extract stopover points containing semantic information of users from their raw trajectories. Then we annotate the stopover points with map information to get the pedestrian moving preference.

C. TRAJECTORY PREDICTION PROBLEM

With the stopover points extracted in the semantic enriching step, we can turn each raw trajectory to the stopover trajectory, which not only retains the pedestrian purpose, but also discards the moving parts to simply the calculation. And we designed the attentional BiGRU-based trajectory prediction model with the input of stopover trajectories to complete the prediction. And the output will be sent to the trained GMM semantic enriching model to give the possible destination.

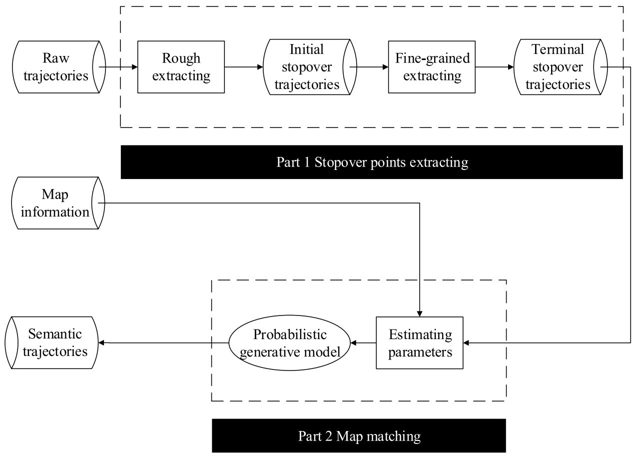


FIGURE 2. **The framework of semantic enriching in SEABIG.**

IV. SOLUTIONS

In this section, we show the principles and process of the approach in detail.

A. SEMANTIC ENRICHING

As described before, semantic information is hidden in sequences constructed by stopover points of raw trajectories. So we accomplish the semantic enriching in two steps. The framework of this part is shown in Figure 2.

Step 1 (Stopover Points Extracting): To extract the stopover points, first we set a time threshold t_0 and a distance threshold d_0 and scan the database of raw trajectories to pick up the points satisfying the standard that the time interval is no less than t_0 and the distance is no more than d_0 to complete the roughly extracting and get the initial stopover points. Algorithm 1 is as follows:

Algorithm 1 Extracting Initial Stopover Points

Input: A raw trajectory R
  Interval and distance threshold t_0,d_0
Output: An initial stopover trajectory IS
1: IS $\leftarrow$<>;
2: i $\leftarrow$ 1;
3: **for** j = 2 to |R| **do**
4: **if** distance(l_j, l_{j-1}) $\geq d_0$ **then**
5: **if** $h_i \neq h_{j-1}$ **and** t_j - $t_i \geq t_0$ **then**
6: IS $\leftarrow$ append <h_i, h_{j-1} > to IS
7: **end if**
8: i $\leftarrow$ j;
9: **end if**
10: **end for**
11: **return** IS;

Then we apply the DBSCAN on the following fine-grained extracting to merge points that spatially adjacent and collect the center of each cluster to get the terminal stopover points. In the algorithm, the radius of the circular area Eps and the minimum number of points in the area MinPts will effect on the result. So we set the MinPts to 4 by experience, and the Eps parameter has been estimated based on the algorithm as proposed in [25]. A trajectory T can be represented as a vector of distances d_i between two consecutive points p_i and p_{i+1}. It is possible to plot the appropriate Gaussian curve with the mean μ and the standard deviation σ of these distances. So we can use the quantile function to avoid the knowledge about the trajectory domain. It is the inverse of the cumulative distribution function, where quantile function: [0,1] $\rightarrow$ R. The quantile function is described as:

$$F^{-1}(p, \mu, \sigma) = \mu + \sigma\sqrt{2}erf^{-1}(2p-1) \tag{1}$$

$$erf^{-1}(x) = \sum_{k=0}^{\infty} \frac{c_k}{2k+1}(\frac{\sqrt{\pi}}{2}x)^{2k+1} \tag{2}$$

where, μ and σ are the mean and the standard deviation, respectively. $c_0 = 1$ and c_k is:

$$c_k = \sum_{m=0}^{k-1} \frac{c_m c_{k-1-m}}{(m+1)(2m+1)} \tag{3}$$

Step 2 (Map matching:) To match the terminal stopover points with the POIs on the map, we propose a probabilistic generative model to decide the visit purpose of each stopover points. We assume that feature vectors f follow a Gaussian mixture distribution as described in equation (4), and the probability that the feature vector f belongs to the visit purpose k is described in equation (6):

$$p(f) = \sum_{m=1}^{N} p(m)p(f|m) = \sum_{m=1}^{N} \pi_m N(f|\mu_m, \Sigma_m) \tag{4}$$

$$N(f|\mu, \Sigma) = \frac{1}{\sqrt{2\pi^p|\Sigma|}} \exp\left(-\frac{1}{2}(f-\mu)^T \Sigma^{-1}(f-\mu)\right) \tag{5}$$

$$\gamma\left(f_j^i, k\right) = \frac{\pi_k N\left(f_j^i|\mu_k, \Sigma_k\right)}{\sum_{m=1}^{N} \pi_m N\left(f_j^i|\mu_m, \Sigma_m\right)} \tag{6}$$

where μ_k and Σ_k denote the mean and the covariance matrix of the visit purpose k, respectively.

Then we can calculate the log likelihood described in equation (7) of a terminal stopover trajectory based on the probabilistic generative model, f(n) means the feature vector of the nth point in the terminal stopover trajectory, and v_m means the mth type of visit purpose.

$$\begin{aligned} L(TS_j^i) &= \sum_{n=1}^{ts} \log\left\{\sum_{m=1}^{N} \pi_m N\left(f_j^i(n)|\mu_m, \Sigma_m\right)\right\} \\ &= \sum_{n=1}^{ts} \log\left\{\sum_{m=1}^{N} P\left(f_j^i(n)|v_m, \mu_m, \Sigma_m\right)\right\} \end{aligned} \tag{7}$$

To estimate the model parameters, we apply the EM algorithm with all trajectories. The EM algorithm estimates maximum likelihood parameters of a statistical model iteratively. To simplify the computation, we utilize the Jensen's inequality to get a lower bound F of the log likelihood function as shown in equation (8).

$$L\left(TS_j^i\right) \geq F = \sum_{n=1}^{ts}\sum_{m} Q(v_m)\log\left(\frac{P(f_j^i(n), v_m|\mu_m, \Sigma_m)}{Q(v_m)}\right) \tag{8}$$

Then we maximize F by deriving the following update formula of EM steps for all parameters:

$$\mu_k = \frac{\sum_{n=1}^{ts}\gamma\left(f_j^i(n)|k\right)f_j^i(n)}{\sum_{n=1}^{ts}\gamma\left(f_j^i(n)|k\right)} \tag{9}$$

$$\Sigma_k = \frac{\sum_{n=1}^{ts}\gamma\left(f_j^i(n)|k\right)\left(f_j^i(n)-\mu_k\right)\left(f_j^i(n)-\mu_k\right)^T}{\sum_{n=1}^{ts}\gamma\left(f_j^i(n)|k\right)} \tag{10}$$

$$\pi_k = \frac{\sum_{n=1}^{ts}\gamma\left(f_j^i(n)|k\right)}{ts} \tag{11}$$

The parameters can be estimated by substituting the update formula iteratively. Finally, we put the parameters into the equation (6) and select the POI type corresponding to the maximum probability as the visit purpose of the location in the stopover trajectory.

B. PEDESTRIAN SEMANTIC TRAJECTORY PREDICTION

To predict the pedestrian semantic trajectory, first we need to extract the terminal stopover points to construct stopover trajectories which represent pedestrian semantic trajectories. And we accomplished the prediction based on the stopover trajectories with the proposed SEABIG model.

For our semantic trajectory prediction, it is unlikely to satisfy demands of practical scenarios by using RNN, LSTM or GRU network directly. Firstly, the RNN exists the problem of gradient disappearance with long-time training. Though the standard LSTM can improve the problem with its optimized structure, it takes longer time to train the model with more parameters and more complex calculations. Besides, when positions in trajectories are far from each other, some strong correlations may be lost. And those important positions closely related to the destination should be given more attention. Therefore, we propose a semantics-enriched attentional BiGRU model (SEABIG model in short) to solve the problem.

The SEABIG model first uses an embedding layer to transform the training points into the input of the network. Then the latent features of the former and later positions can be learned through a BiGRU layer. And the attention mechanism will be added to pay more attention to the preceding positions that have strong correlations to the destination with the output of the BiGRU layer. Finally, the probability distribution of each possible location will be given through the output of a softmax function.

The framework is shown in Figure 3. In the reminder of the section, we present the details of different layers.

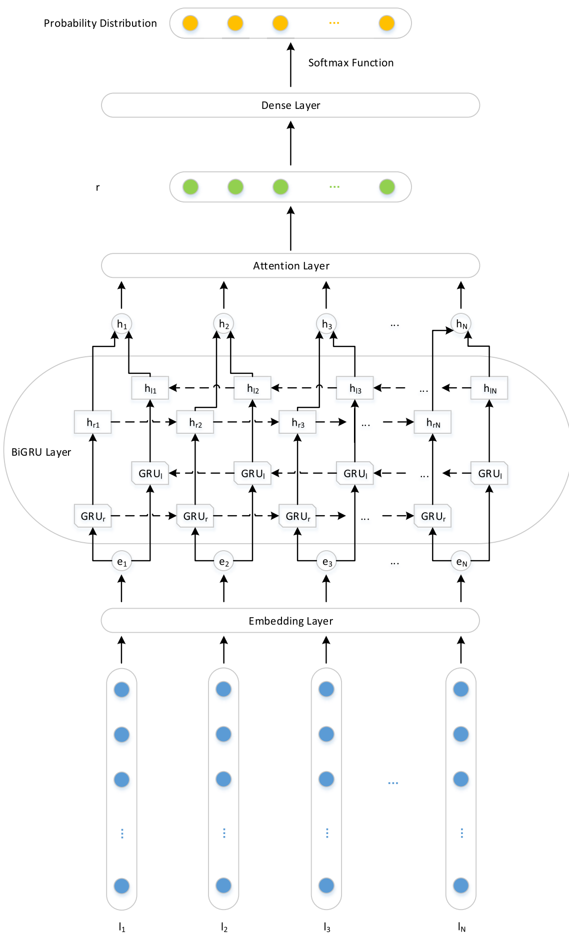


FIGURE 3. The framework of BiGRU in SEABIG.

1) THE EMBEDDING LAYER

Given a terminal stopover trajectory TS of a user, we used the embedding layer to extract the information in each record l=[lan, lon, t]. We learn the embeddings for the position and the timestamp and attach them to a vector e to accomplish the encoding work of the information contained in the terminal stopover trajectory.

The timestamp t is the original temporal information of the trajectory. However, because it is continuous, it is hard to embed each timestamp. Therefore, we map a day into 24 slots (6 slots for dawn, 6 slots for morning, 6 slots for noon and 6 slots for evening). And the trajectory is shown as Figure 4 with time segmentation. We turn each timestamp t to a 24-dimenstional one-hot vector representation. So the time embedding part in the embedding layer aims to learn a $D_t* 24$ transformation matrix E_t (The dimension of the time embedding layer is D_t.).

Then we transform each location record [lan, lon] into a M-dimensional one-hot vector (M is the encode of the geohash-based grid which the location belongs to.). So the location embedding part in the embedding layer aims to learn a D_l* M transformation matrix E_l (The dimension of the location embedding layer is D_l.).

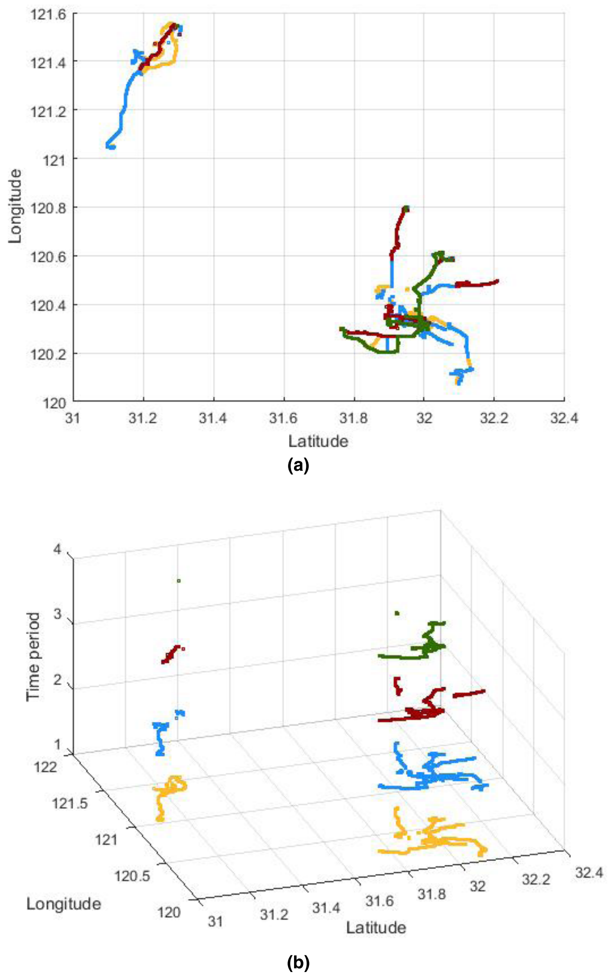


FIGURE 4. (a) shows trajectories without time segmentation, (b) shows trajectories with time segmentation.

2) THE BIGRU LAYER

To handle the problem of time series, the GRU network can be a helpful method. It can improve the problem of gradient vanishing and exploding phenomena by the gating mechanism. However, in our problem, to a given stopover trajectory, the later point also provides important information to the former part, which the GRU ignores due to its framework. To preserve the useful latent features contained in the future context, we adopts a bidirectional GRU network to take both preceding and following points into consideration by forward and backward pass respectively.

The framework of the BiGRU network consists of a forward GRU and a backward GRU network. Different from LSTM, it has only two gates: reset gate r and update gate z:

$$
\begin{aligned}
r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\
z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\
\widetilde{h_t} &= \tanh(W_{\widetilde{h}} \cdot [r_t * h_{t-1}, x_t]) \\
h_t &= (1 - z_t) * h_{t-1} + z_t * \widetilde{h_t}
\end{aligned} \tag{12}
$$

where r_t is the reset gate, z_t is the update gate, h_t is the hidden state, and the $\widetilde{h_t}$ is the hidden input activation vector. And W is the weight of corresponding variables.

Based on the basic GRU network, the BiGRU extends the network through a second layer, in which the connections between the hidden states flow in opposite temporal order. Therefore, the network extracts latent features from both preceding and following points. The final output of the location is as follows:

$$h_t = [h_{rt} \oplus h_{lt}] \tag{13}$$

3) THE ATTENTION LAYER

Generally speaking, the later points near the pending predicting point make the most contributions. However, in practice, those points, which have strong correlations to the pending predicting point, should be focused on.

Therefore, to take the problem into consideration, we utilized the attention mechanism as described:

$$
\begin{aligned}
m_t &= \tanh(W_h h_t + b_h) \\
a_i &= \frac{\exp(m_t)}{\sum_{k=1}^{N} \exp(m_k)} \\
r &= \sum_{i=1}^{N} a_i h_i
\end{aligned} \tag{14}
$$

where N is the length of the input stopover trajectory; r, which is formed by a weighted sum of output hidden states, is the representation of hidden of the input sequence; W_h and b_h are the weights and bias in the attention layer. Those points that have strong correlation to the pending predicting location will be assigned bigger weights through the training process. Finally, we can get the location representation for classification:

$$h^* = \tanh(r) \tag{15}$$

4) THE PREDICTING LAYER

We utilized the dense layer as the final output of the network, and a multi-class logistic regression is used to obtain the distribution of the pending predicting location label (()) from the set of classes C, which is the set of all cells in the geohash-based grid.

$$
\begin{aligned}
\hat{p}(y|TS) &= softmax(W^{TS} h^* + b^{TS}) \\
\hat{y} &= \arg\max_{y} \hat{p}(y|TS)
\end{aligned} \tag{16}
$$

We adopted cross entropy as the loss function. It is the frequently used multi-class logarithmic loss function for softmax classifier:

$$L(\theta) = -\frac{1}{n} \sum_{i=1}^{n} t_i \log(y_i) + \lambda \, \|\theta\|_F^2 \tag{17}$$

where t_i is the one-hot represented output, m is the number of target pending predicting location and λ is a pre-defined parameter of the regularization term to avoid overfitting. And y_i is the output of the softmax function.

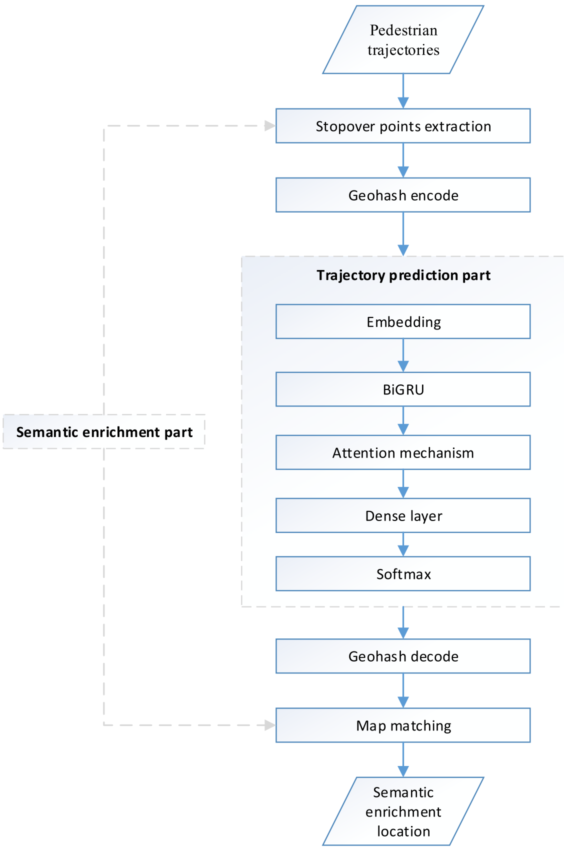


FIGURE 5. **Complete framework and process of SEABIG.**

After that, we decode the geohash code to get the location of the predicted points. Then we will put the location into the generative model to annotate the semantic label. So the complete framework and the process of SEABIG are shown below:

V. EXPERIMENTS EVALUATION

A. EXPERIMENTAL CONFIGURATION

We implement our model with Matlab 2017a and Python 3.7, and conduct the experiments on a computer with an Intel(R) Core(TM) i7-7700HQ CPU with 2.8 GHz (8 CPUs) and 8 GB main memory running Windows 10.

In the semantic enrichment part of the SEABIG model, two types of data are required. One is map information data, so we choose the POIs data to represent the map information. The POIs data were collected artificially from Amap and consisted of 459,751 POIs with 369 fine-grained categories. In our experiment, based on the POI information we collected, we set the number of dimensions of the feature vector to seven, i.e. λ_1 as housing areas including apartments, dormitories, etc., λ_2 as working areas including companies, offices, etc., λ_3 as shopping areas including malls, markets, etc., λ_4 as transportation landmarks including bus stations, subway stations, etc., λ_5 as entertainment areas including gyms, parks, etc., λ_6 as studying areas including schools, libraries, etc., and λ_7 as food areas.

FIGURE 6. **Distribution of the dataset in Beijing.**

Besides, ground-truth data is required to validate our results. However, publicly available GPS datasets generally do not contain ground-truth due to privacy considerations. Hence, we recruited 20 volunteers in Beijing with different ages and occupations to collect their daily trajectories with ground-truth for 3 months from the BDS (BeiDou Navigation Satellite System). Thus, our sample dataset contained 1957 trajectories with a total distance of 151,553 kilometers.

Then in the prediction part, we utilize the public data and the collected ground-truth data, which are collected from the different sources, i.e. GPS and BDS. The test data is collected from 182 real users in the Geolife of Microsoft Asian Research Institute project. It contained 17,621 trajectories with a total distance of 1,292,951 kilometers and a total duration of 50,176 hours in Beijing over a period of 5 years. Figure 6 plots the distribution (heat map) of the dataset in Beijing.

B. COMPARED METHODS AND PREFORMANCE METRICS

Compared Methods: We compared our model with three types of method as described as follows:

(1) Nearest location (NL): The NL method is the original approach to predict the next location. It regards the nearest neighbors to the user's present position as the next location.

(2) Matrix factorization (MF): The MF method is the model-based method to predict the pedestrian future location. With the observed position matrix, the MF model learns the low dimensional feature vectors of users and positions to choose the most similar locations as the next location users will pass.

(3) Hidden Markov Model (HMM): It is also a model-based method for next location prediction. It considers pedestrian trajectories as a Markov process. Then a HMM is learned to capture the regularities

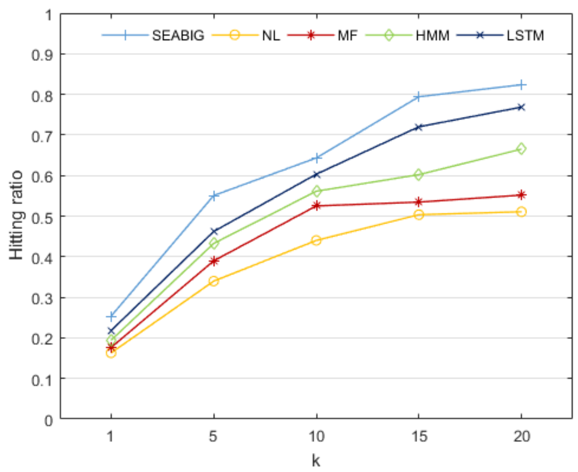


FIGURE 7. Comparison of hitting ratio with the BDS data in Beijing.

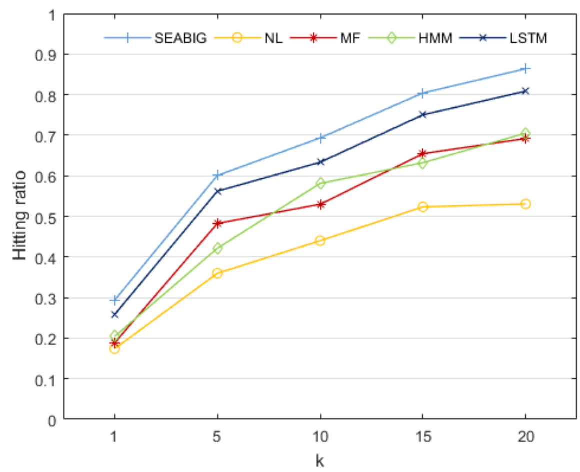


FIGURE 8. Comparison of hitting ratio with the GPS data in Beijing.

of pedestrian track. And it will give the next location which is the predicted location with the largest probability.

(4) LSTM network: The LSTM network is the regular model for prediction based on deep learning methods. It improves the problems of gradient vanishing and gradient exploding that may appear in traditional RNN.

Performance Metrics: In order to evaluate each method, the hitting ratio and the Haversine distance were utilized to evaluate the quality of prediction.

Hitting ratio is the parameter to check if the actual location occurs in the top-k prediction sequence. For the test data, this parameter calculates the percentage of trajectories for which the real next location is successfully contained in the top-k result list.

Haversine distance is the distance to measure the distance two points with their own latitude and longitude in range of [-90,90] and [-180,180], respectively. We apply the Haversine distance to measure the distance between the predicted next location and the real next location.

$$d(y,\hat{y}) = 2r\arcsin((\sin^2(\frac{|\hat{y}^{lat}-y^{lat}|}{2}) + \cos(y^{lat})\cos(\hat{y}^{lat})\sin^2(\frac{|\hat{y}^{lon}-y^{lon}|}{2}))^{\frac{1}{2}}) \quad (18)$$

where r is the radius of the earth, we set the value of r to 6371km due to the ellipsoidal shape of the earth. We choose the mean Haversine localization error d between the real next location and the predicted next location as the second performance metrix.

C. PERFORMANCE COMPARISON

For each dataset, we randomly choose 70% to train the model and use the left 30% of them to test the prediction. Besides, we evaluate the model by choosing different values of k (k=1, 5, 10, 15, 20).

We now compare our approach with NL, MF, HMM and LSTM against the two metrics, i.e. hitting ratio (k) and Haversine distance with real next location. First, we tested the impact of the k on the performance of the five approaches. As described in Figure 7 and 8, the hitting ratio of the five methods rise with k. From the result, it is indicated that when k increases, the hitting ratio of our model rises faster than the other models. And whatever the k is, our model outperforms the other four methods.

TABLE 1. Comparison of d (m).

	The BDS dataset	The GPS dataset
NL	2801	2950
MF	1855	1942
HMM	1763	1841
LSTM	1520	1622
SEABIG	**1413**	**1578**

Then we calculated the mean Haversine localization error d between the real next location and the predicted next location to evaluate our model.

As shown in the results, our model outperforms the other models significantly based on data from different sources. The comparisons are discussed in detail as follows:

(1) The NL model is an intuitive and simple but efficient method in some cases. However, in most cases, pedestrian movements are much more complex and don't obey the nearest-first principle.

(2) The MF model performs not as well as expected. That is because it fails to consider the sequence transition, though it can extract the position people is interested in successfully.

(3) The HMM performs better due to taking the transition probability into consideration. However, it can only construct the first-order Markov process.

(4) The LSTM network turns out to be the strongest model, but it is still inferior to SEABIG model. Although it can handle the problem of long-term dependency,

it cannot preserve the useful latent features contained in the future context.

VI. CONCLUSION

We have proposed a semantics-enriched attentional BiGRU network to predict pedestrian trajectory with taking users' semantic information and useful latent features contained in the future context into consideration. The design of the semantic enrichment part and the next location prediction part are illustrated. Our experiments show that this model can preserve the latent information in both preceding and following locations in a trajectory and obtain relative high prediction accuracy with comparison to other methods.

REFERENCES

[1] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and pois," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Beijing, China, Aug. 2012, pp. 186–194.

[2] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, F. F. Li, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 961–971.

[3] N. Lee, W. Choi, P. Vernaza, C. B. Choy, P. H. S. Torr, and M. Chandraker, "DESIRE: Distant future prediction in dynamic scenes with interacting agents," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2165–2174.

[4] H. Su, Y. Dong, J. Zhu, H. Ling, and B. Zhang, "Crowd scene understanding with coherent recurrent neural networks," in *Proc. 25th Int. Joint Conf. Artif. Intell. (IJCAI)*, New York, NY, USA, Jul. 2016, pp. 3469–3476.

[5] H. Liu, J. Xu, K. Zheng, C. Liu, L. Du, and X. Wu, "Semantic-aware query processing for activity trajectories," in *Proc. 10th ACM Int. Conf. Web Search Data Mining (WSDM)*, Cambridge, U.K., Feb. 2017, pp. 283–292.

[6] Z. Lv, J. Xu, P. Zhao, G. Liu, L. Zhao, and X. Zhou, "Outlier trajectory detection: A trajectory analytics based approach," in *Proc. Int. Conf. Database Syst. Adv. Appl.*, Suzhou, China, 2017, pp. 231–246.

[7] J. Rao, H. He, H. Zhang, F. Ture, R. Sequiera, S. Mohammed, and J. Lin, "Integrating lexical and temporal signals in neural ranking models for searching social media streams," in *Proc. Neu-IR*, Aug. 2017, pp. 1–5.

[8] Q. Yuan, W. Zhang, C. Zhang, X. Geng, G. Cong, and J. Han, "PRED: Periodic region detection for mobility modeling of social media users," in *Proc. 10th ACM Int. Conf. Web Search Data Mining (WSDM)*, Cambridge, U.K., Feb. 2017, pp. 263–272.

[9] C. Zhang, K. Zhang, Q. Yuan, L. Zhang, T. Hanratty, and J. Han, "Gmove: Group-level mobility modeling using geo-tagged social media," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, San Francisco, CA, USA, Aug. 2016, pp. 1305–1314.

[10] Y. Kim, J. Han, and C. Yuan, "Toptrac: Topical trajectory pattern mining," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NA, USA, Aug. 2015, pp. 587–596.

[11] F. Wu, Z. Li, W.-C. Lee, H. Wang, and Z. Huang, "Semantic annotation of mobility data using social media," in *Proc. 24th Int. Conf. World Wide Web*, Florence, Italy, May 2015, pp. 1253–1263.

[12] H. Su, K. Zheng, K. Zeng, J. Huang, and X. Zhou, "STMaker: A system to make sense of trajectory data," *J. Proc. VLDB Endowment*, vol. 7, no. 13, pp. 1701–1704, Aug. 2014.

[13] J. J.-C. Ying, W.-C. Lee, T.-C. Weng, and V. S. Tseng, "Semantic trajectory mining for location prediction," in *Proc. 19th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, Chicago, IL, USA, Nov. 2011, pp. 34–43.

[14] J. J.-C. Ying, W.-C. Lee, and V. S. Tseng, "Mining geographic-temporal-semantic patterns in trajectories for location prediction," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 2, Dec. 2014, Art. no. 2.

[15] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra, and K. Aberer, "Semantic trajectories: Mobility data computation and annotation," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 49, Jun. 2013, Art. no. 49.

[16] C. Zhang, J. Han, L. Shou, J. Lu, and T. L. Porta, "Splitter: Mining fine-grained sequential patterns in semantic trajectories," *J. Proc. VLDB Endowment*, vol. 7, no. 9, pp. 769–780, May 2014.

[17] Z. Li, B. Ding, J. Han, R. Kays, and P. Nye, "Mining periodic behaviors for moving objects," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Washington, DC, USA, Jul. 2010, pp. 1099–1108.

[18] M. Boukhechba, A. Bouzouane, S. Gaboury, C. Gouin-Vallerand, S. Giroux, and B. Bouchard, "Prediction of next destinations from irregular patterns," *J. Ambient Intell. Humanized Comput.*, vol. 9, no. 5, pp. 1345–1357, Oct. 2018.

[19] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility Markov chains," in *Proc. 1st Workshop Meas., Privacy, Mobility*, Bern, Switzerland, 2012, Art. no. 3.

[20] N. Duong-Trung, N. Schilling, and L. Schmidt-Thieme, "Near real-time geolocation prediction in Twitter streams via matrix factorization based regression," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage. (CIKM)*, Indianapolis, IN, USA, Oct. 2016, pp. 1973–1976.

[21] H. Wu, Z. Chen, W. Sun, B. Zheng, and W. Wang, "Modeling trajectories with recurrent neural networks," in *Proc. 26th Int. Joint Conf. Artif. Intell. (IJCAI)*, Melbourne, VIC, Australia, Aug. 2017, pp. 3083–3090.

[22] J. Zhao, J. Xu, R. Zhou, P. Zhao, C. Liu, and F. Zhu, "On prediction of user destination by sub-trajectory understanding: A deep learning based approach," in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manage. (CIKM)*, Turin, Italy, Oct. 2018, pp. 1413–1422.

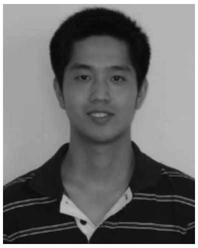
[23] D. Kong and F. Wu, "HST-LSTM: A hierarchical spatial-temporal long-short term memory network for location prediction," in *Proc. 27th Int. Joint Conf. Artif. Intell. (IJCAI)*, Stockholm, Sweden, Jul. 2018, pp. 2341–2347.

[24] P. Zhao, H. Zhu, Y. Liu, Z. Li, J. Xu, and V. S. Sheng, "Where to go next: A spatio-temporal LSTM model for next POI recommendation," in *Proc. AAAI*, Honolulu, HI, USA, 2019, pp. 5172–5188.

[25] A. T. Palma, V. Bogorny, B. Kuijpers, and L. O. Alvares, "A clustering-based approach for discovering interesting places in trajectories," in *Proc. ACM Symp. Appl. Comput.*, Fortaleza, Brazil, Mar. 2008, pp. 863–868.

WANLONG ZHANG was born in 1996. He received the B.S. degree in electronic science and engineering from the National University of Defense Technology, Changsha, Hunan, China, in 2017, where he is currently pursuing the master's degree. His research interests include data mining in trajectory, spatial and temporal query processing, and intelligent transportation.

LITING SUN was born in 1994. She received the B.S. degree in electronic science and engineering from Shandong University, Jinan, Shandong, China, in 2017. She is currently pursuing the master's degree with the National University of Defense Technology, Changsha, Hunan, China. Her research interests include data mining, signal processing, and pattern recognition.



XIANG WANG was born in 1985. He received the bachelor's degree in electronic science and engineering from the National University of Defense Technology, Changsha, Hunan, China, in 2007, where he is currently pursuing the Ph.D. degree. His research interests include blind signal processing in radar and communication applications.

ZHITAO HUANG was born in 1976. He received the B.S. and Ph.D. degrees in information and communication engineering from the National University of Defense Technology, Changsha, Hunan, China, in 1998 and 2003, respectively, where he is currently a Professor with the College of Electronic Science and Engineering. His research interests include radar and communication signal processing and array signal processing.

BAOGUO LI was born in 1975. He received the B.Sc., M.Sc., and Ph.D. degrees in information and communication engineering from the College of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan, China, in 1998, 2001, and 2005, respectively, where he is currently an Associate Professor with the College of Electronic Science and Engineering. His research interests include artificial intelligence in radar and communication signal processing.