

Research Article

TaxiInt: Predicting the Taxi Flow at Urban Traffic Hotspots Using Graph Convolutional Networks and the Trajectory Data

Jinmao Zhang ¹, Huanchang Chen ², and Yiming Fang ³

¹Zhejiang Provincial Intellectual Property Protection Center, Hangzhou 310012, China

²Hangzhou Hikvision Digital Technology Co., Ltd., Hangzhou 310052, China

³School of Mechanical & Electrical Engineering, Shaoxing University, Shaoxing 312000, China

Correspondence should be addressed to Yiming Fang; ilone_fang@126.com

Received 1 April 2021; Revised 19 May 2021; Accepted 3 June 2021; Published 14 June 2021

Academic Editor: Jit S. Mandeep

Copyright © 2021 Jinmao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Taxi flow is an important part of the urban intelligent transportation system. The accurate prediction of taxi flow provides an attractive way to find the potential traffic hotspots in the city, which helps to avoid serious traffic congestions by taking effective measures in advance. The current prediction of taxi flow and its impact on urban transportation are closely related to the passenger origin-destination (OD) information. However, high-quality OD information is not always available. To address this problem, a prediction model, named as TaxiInt, is proposed in this study. Different from other density-clustering-based approaches, neural network, or OD information based models, TaxiInt predicted the taxi flow using the trajectory data of taxis. The spatial features and temporal features of each road were extracted using a graph convolutional network, which was trained with the road network information and the trajectory data. The experiments carried on a real taxi dataset showed the validity of our model. It can predict the taxi flow at a given urban intersection with high accuracy.

1. Introduction

Taxi is a comprehensive reflection of urban traffic. It provides information regarding not only the traffic situation but also the trend of crowd activities. The accurate prediction of taxi flow helps to find the potential traffic hotspots in the city to take effective measures to avoid the coming traffic congestions. During the past few decades, the research on the prediction of taxi flow has attracted extensive attention [1]. Balan designed a trip information system to predict the fare and trip duration of the taxi ride the passengers were planning to take. The authors claimed that the accuracy and the real-time performance were validated by large scale evaluation [2]. Li conducted a similar work. The authors proposed a hybrid model coupling the deep learning model and the quantile regression aiming at the travel time prediction [3]. Kong proposed a framework called as TBI2Flow to predict the taxi passenger flow [4]. In addition, the prediction of taxi flow was believed to be beneficial for

optimizing public transportation planning [5] and discovering unreasonable urban planning [6].

Currently, the taxi passenger's pick-up and set-down feature, called origin-destination (OD) information, is one of the most frequently used features in taxi flow prediction. Some researchers used machine learning models to analyse taxi data, such as density-clustering models (DBSCAN) [7], support vector machines (SVM) [8–10], and k-nearest neighbour (KNN) [11]. Some combined the basic models to create a novel one and then used it to analyse the data. Li [12] proposed a combined model to forecast the potential passenger demand in different regions based on Daubechies wavelets analysis and least squares support vector machine (LS-SVM). Recently, convolutional neural networks (CNN) were used to discover the spatial characteristics of vehicle travel, and cyclic neural networks were also employed to learn the periodic and trending regularities of travel data. Yao [13] designed a deep multiview space-time network to predict the taxi demand. Liu [14] proposed a contextualized

spatial-temporal network on the prediction of taxi demand with the spatial, temporal, and global correlation information fully considered. Xu [15] employed incorporating graph attention and recurrent architectures to forecast the demand for taxis in a city-wide area. These tentative studies have achieved desired successes. But their prediction highly depends on the passenger's OD information, which may be easily affected by many external factors, such as the failure of the recording instrument, a mistake operation from the driver, or the interruption of calling taxis from the mobile APPs. What is more, the OD information is a digital state in most cases. Once the data are missed or disturbed, it is very difficult to reconstruct them.

Compared to the OD information, the trajectory data of a vehicle, which are obtained from the positioning system (like GPS), are continuously updated and become more reliable. If some of the trajectory data are lost, they can be reconstructed through the nearest valid track points. Due to this attractive advantage, a variety of trajectory-based models were proposed for the traffic flow prediction. Xu [16] proposed WTFPredict methodology to make short-term traffic flow forecasting, which is based on taxi data and weather data. Zhang [17] used taxi data to predict short-term flow trends in urban areas to analyse urban crowd mobility. Li built a feature-level fusion model to fuse the representative features extracted from the temporal and spatial features of traffic data [18]. A deep learning approach was proposed to extract the complex features of traffic flow and then predict the short-term traffic flow forecast with high accuracy and stability [19].

Recently, the graph neural network (GNN) became an increasingly used network in traffic flow prediction. Compared with the common convolution, the convolution kernel of the GNN has a flexible number of neighbour nodes which makes the GNN more suitable for complex traffic applications. Lv introduced the GNN for the analysis of traffic network resilience [20]. Cui proposed the HGC-LSTM framework, which was based on the GNN and LSTM, to learn interactions between links in the traffic network [21].

In this work, a novel framework was proposed based on the GNN. Unlike the existing methods which analysed the taxi flow of largescale urban areas, the framework proposed in this study focused on the prediction of taxi flow at urban traffic hotspots which may have a greater impact on the overall traffic condition. First, the trajectory data of taxis were converted into traffic flow data of core intersection nodes in the road network. Then, a GNN model and a time series network are created to capture the spatiotemporal information of intersection traffic flow. Finally, a model named TaxiInt was used to predicate the taxi flow at intersections. TaxiInt consists of three separate components that simulate three characteristics of traffic flow, respectively [22], and each component consists of an attention mechanism, a graph convolution network, and a common convolution network.

The method proposed in this work has the following advantages:

- (1) Fewer requirements on the taxi's dataset. TaxiInt does not depend on the passenger OD information. It uses the trajectory data to predict the traffic flow at urban intersections.
- (2) Model reliability. Three characteristics of the traffic flow are extracted by three separate components of TaxiInt. That spatiotemporal information extracted would be more reliable than existing baseline models.

2. TaxiInt Framework

In this section, we introduce our TaxiInt framework which referenced the network presented in [22]. As shown in Figure 1, the overall framework of TaxiInt consists of three parts: the data sources, the time-based road network traffic information change graph, and neural network structure. In part A, the data source used by TaxiInt is displayed that includes the taxi trajectory data, the urban road network data, the weather data, and the weekday label information. In part B, the road traffic information at the selected area for some time is displayed in the form of a time axis. The redder side in the figure represents the denser taxi traffic. The different coloured patches represent different time points on the time axis. Part C is a schematic diagram of the combination of timing diagram and neural network structure diagram. The meanings expressed by the colour blocks on the top time axis are consistent with the content of part B. The overall framework includes 3 subunits, each of which contains 2 ST blocks to capture traffic flow timing information and road network space information.

2.1. Preliminaries. Here, we define a road information network as an undirected graph $G = (V, E, A)$, as shown in Figure 1 part B, V is a node set, $N = |V|$, E is an edge set, which reflects the link between nodes, and $A \in R^{N \times N}$ is an adjustable matrix based on G . Next, we define the input data $\in R^{N \times F}$, which were already converted from trajectory data, for each time slice, and F means the number of the node features. Since the input data are composed of multiple time slices, we introduce $\mathcal{X}^T = (X_1, X_2, \dots, X_T)$ to represent the entire input data stream, and T is the time slice number.

2.2. Spatial-Temporal Attention. As shown in Figure 1 part C, in the first part, we feed the data to the "spatial-temporal attention" structure of the network model, which is composed of "spatial attention" component and "temporal attention" component. By using these components, we can capture the dynamic information, like spatial and temporal correlations, from the road information stream.

2.2.1. Spatial Attention Component (SAtt). In the spatial dimension, taxi flow at each intersection may be affected by the flow value at adjacent intersections. In order to increase the sensitivity of the network model to the traffic data in the spatial structure of the road network, we introduce the at-

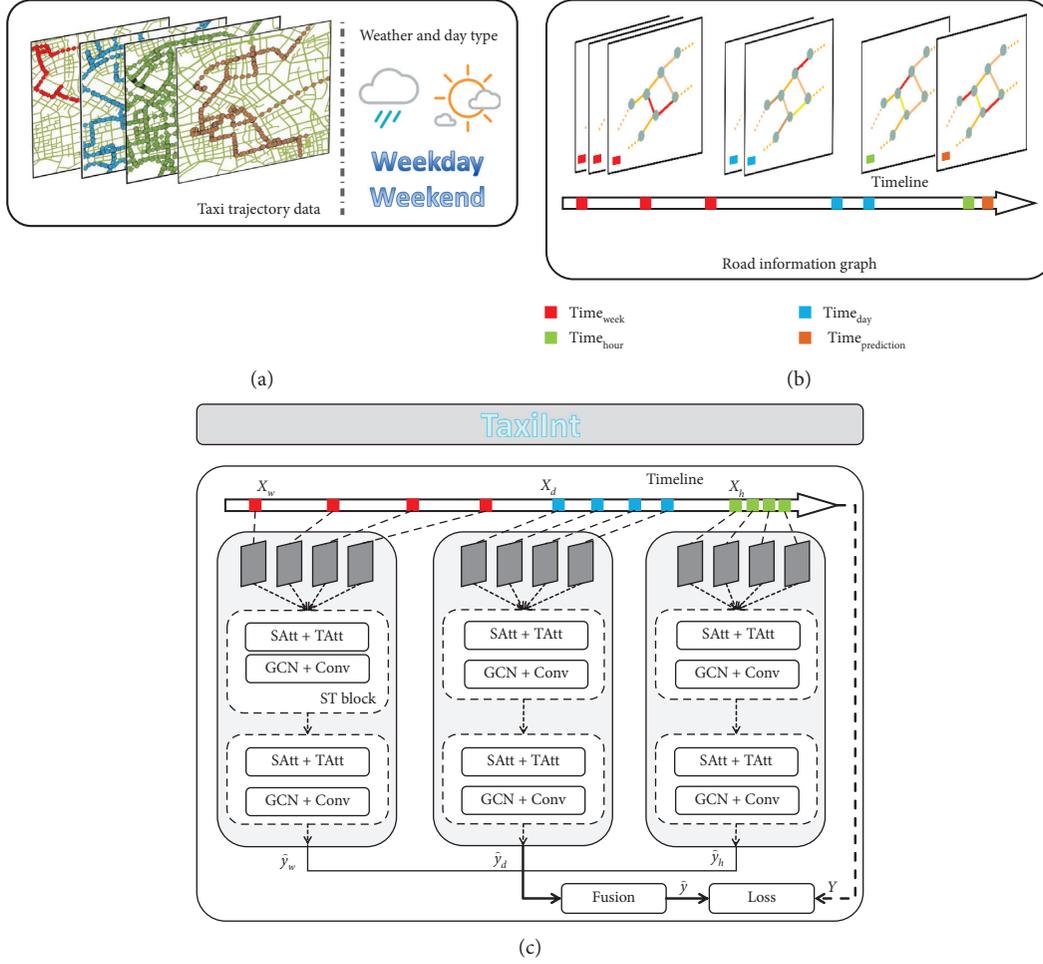


FIGURE 1: Illustration of the overall architecture of the proposed framework.

tention mechanism to make the model more sensitive to the changes of the spatial correlations.

$$S = V_s \cdot \sigma \left[(\mathcal{X}^{r-1} W_1) W_2 (W_3 \mathcal{X}^{r-1})^T + b_s \right],$$

$$S'_{i,j} = \frac{\exp(S_{i,j})}{\sum_{j=1}^N \exp(S_{i,j})},$$
(1)

where $\mathcal{X}^{r-1} = (X_1, X_2, \dots, X_{T^{r-1}})$ is the input flow of the r^{th} SAtt component. $V_s, b_s \in \mathbb{R}^{N \times N}$, $W_1 \in \mathbb{R}^{T^{r-1}}$, $W_2 \in \mathbb{R}^{C_{r-1} \times T^{r-1}}$, and $W_3 \in \mathbb{R}^{C_{r-1}}$ are the parameters that to be learned, and C is the number of channel. In the layer, we use σ as the activation function. From the attention matrix S , we can get the correlation weight between the graph nodes, which will be dynamically adjusted according to the input stream of the layer. In the final part of the component, we choose the softmax function to ensure that the sum of the weight nodes is one.

2.2.2. Temporal Attention Component (TAtt). In the field of traffic flow forecasting, the road traffic distribution would change over time. So it makes sense to introduce TAtt

component here to capture the change of taxi flow in different time slices.

$$E = V_e \cdot \sigma \left\{ \left[(\mathcal{X}^{r-1})^T U_1 \right] U_2 (U_3 \mathcal{X}^{r-1}) + b_e \right\},$$

$$E'_{i,j} = \frac{\exp(E_{i,j})}{\sum_{j=1}^{T^{r-1}} \exp(E_{i,j})},$$
(2)

where $V_e, b_e \in \mathbb{R}^{T^{r-1} \times T^{r-1}}$, $U_1 \in \mathbb{R}^N$, $U_2 \in \mathbb{R}^{C_{r-1} \times N}$, and $U_3 \in \mathbb{R}^{C_{r-1}}$ are the parameters to be learned. Like the matrix S , the matrix E is automatically capturing the changes in the input stream, making the entire network sensitive to traffic trends in the time dimension. Then, we choose the softmax function to ensure that the sum of the weight nodes is one and get E' .

2.3. Spatial-Temporal Convolution. The SAtt and TAtt components capture important information in the input stream automatically. Then, the adjusted input stream would be fed into the spatial-temporal convolution component, which consisted of a graph convolution in the spatial dimension and a common convolution along the temporal dimension.

2.3.1. Graph Convolution Component. Under different road and time conditions, the nodes can be regarded as the change signals of the graph. Therefore, in order to make full use of the spatial nature of the road information network, we use spectral convolution to process the signal of the whole graph in each time slice and capture the spatial dependence of the neighbour nodes through the signal correlation.

In spectral graph analysis, we can obtain the properties of the graph structure by analysing the Laplacian matrix and its eigenvalues. The graph convolution is a convolution operation implemented by using the linear operators that diagonalize in the Fourier domain to replace the classical convolution operator [23]. However, they are not efficient when dealing with largescale graph networks. Therefore, we adopt Chebyshev polynomials to solve the task approximately but efficiently [24]:

$$g_\theta * G^x = g_\theta(L)x = \sum_{k=0}^{K-1} \theta_k T_k(\tilde{L})x, \quad (3)$$

$$\text{ReLU}(g_\theta * G^x).$$

2.3.2. Common Convolution Component. After capturing spatial dependencies from neighbours, we set standard convolution layers in the temporal dimension, which is used to update the signal of nodes from the neighbouring time slice. The following is an example of a formula on the r^{th} layer:

$$\mathcal{X}^r = \text{ReLU} \left\{ \Phi * \left[\text{ReLU} \left(g_\theta * G^{\hat{\mathcal{X}}^{r-1}} \right) \right] \right\} \in R^{C_r \times N \times T_r}, \quad (4)$$

where $*$ defines a common convolution process, and Φ is the parameter used in the temporal dimension convolution kernel. In the layer, we select ReLU as the activation function.

2.3.3. Multicomponent Fusion. In the final part of the network model, we integrate the output of the three components. In general, there are differences of the temporal and spatial distribution of taxi demand in urban areas with different social functions. Even for the same area, taxi flow is different at different time slices. Therefore, the integration formula is defined as follows:

$$\hat{Y} = W_h \odot \hat{y}_h + W_d \odot \hat{y}_d + W_w \odot \hat{y}_w, \quad (5)$$

where \odot is a Hadamard product operation. W_h , W_d , and W_w are the three learn and adjustable parameters that reflect different degrees of \hat{y}_h , \hat{y}_d , and \hat{y}_w effects on the final forecasting target.

3. Experiments

In this section, we introduce the taxi dataset, baselines, and evaluation metrics of our experiment. The results of our TaxiInt and test baselines are displayed.

3.1. Dataset and Data Preprocessing. The data used in this work consisted of three parts: the trajectory data, the weather data, and the road network data. The trajectory data were gathered from the GPS recorder, from March 1, 2018, to March 31, 2018. The entire dataset includes 12,544 vehicles and 1,087,825,260 records in total. Each record includes five elements: taxi ID, latitude and longitude, taxi speed, passenger status, and time tag. The sampling frequency of GPS data was 22 s. Table 1 presents a typical sample of the trajectory data. The weather data were fetched from the “wunderground” website (<https://www.wunderground.com>), which collected daily precipitation in Hangzhou in March 2018.

The road network data were obtained from the “OpenStreetMap” website. Figure 2 shows the road network map of Hangzhou city, which was used in this work. The areas highlighted with orange on the map are the radiation range of our selected road points, which are observation nodes for our model training. The road network information of the selected area includes some trunk roads, scenic areas, commercial areas, and a small residential area. We set the time interval to 5 minutes and count taxis that cross the intersections in each interval, including taxi number, taxi speed, and traffic flow of the adjacent intersections within 1 km.

There are six traffic hotspots in the area as shown in Figure 3. Figures 3(a) and 3(e) represent Hangzhou’s transportation hub, Figures 3(c) and 3(f) represent Hangzhou’s urban core area, Figure 3(b) is the education area, and Figure 3(d) is the residents’ living area. This study selects the taxi trajectory data of the above areas to verify the prediction effect of TaxiInt.

Since we assumed the OD information was distorted, so the passenger status was removed from our model. And then, in the data cleaning process, the format of the dataset is converted. We deleted passenger status and filtered the taxi IDs which are not in full attendance during the month (taxis that occurred less than 31 days in March 2018). Finally, track records of 12,196 vehicles were left. This model uses 328 road points from the city road map for training, and most of them were located near the intersections.

Figure 4 schematically describes the process of counting the vehicles passing by each intersection in each period. Black dots represent intersections. The black dotted circle represents the preset intersection observation range in this article, and its radius is R_{range} , as shown in point F. The yellow dotted line is the vehicle trajectory. The solid blue line represents the road. We set an intersection range threshold R_{range} , when a vehicle enters an intersection range, we would count vehicle information. As shown in Figure 4, the vehicles driving sequence is “A-C-B-E-F-CD,” so the traffic flow statistics of intersections A, B, D, E, and F increase by one, and intersection C increases by two. The average speed of each intersection is also calculated using a similar method. The trajectory of the vehicle near the intersection C shakes like a cluster. Such problems are often caused by waiting for the traffic lights. It is necessary to filter the vehicle information that repeatedly appears at an intersection in a short time. After the fore processes, we finally got 2 datasets: (1)

TABLE 1: A typical sample of the taxi dataset.

Name	Value
Taxi ID	10004
Time	2018-03-01 07:21:09
Longitude	120.12052
Latitude	30.292969
Speed	17.6979 (m/s)
Passenger status	1

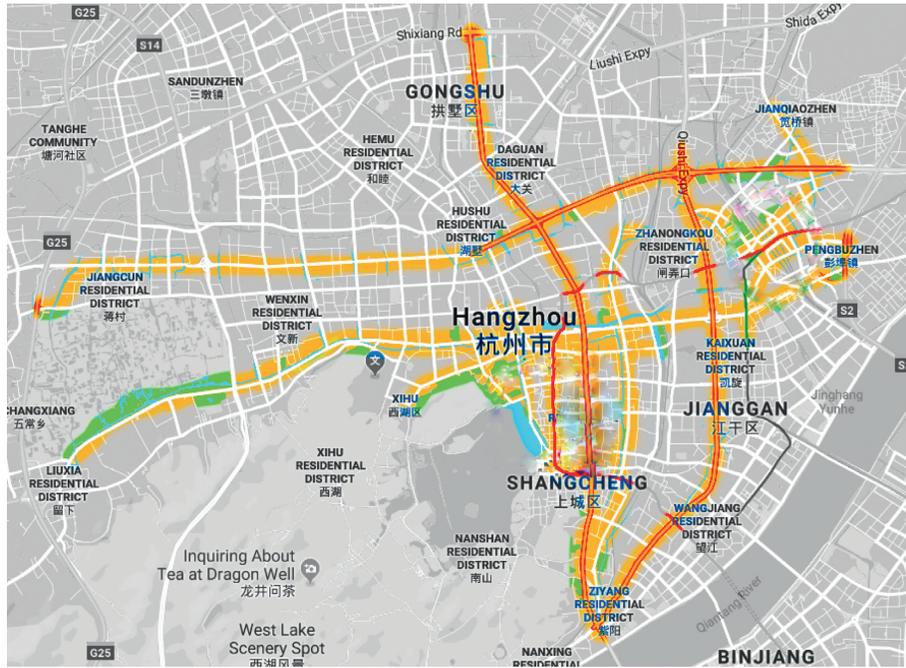


FIGURE 2: The road network diagram of Hangzhou city, China.



FIGURE 3: Continued.

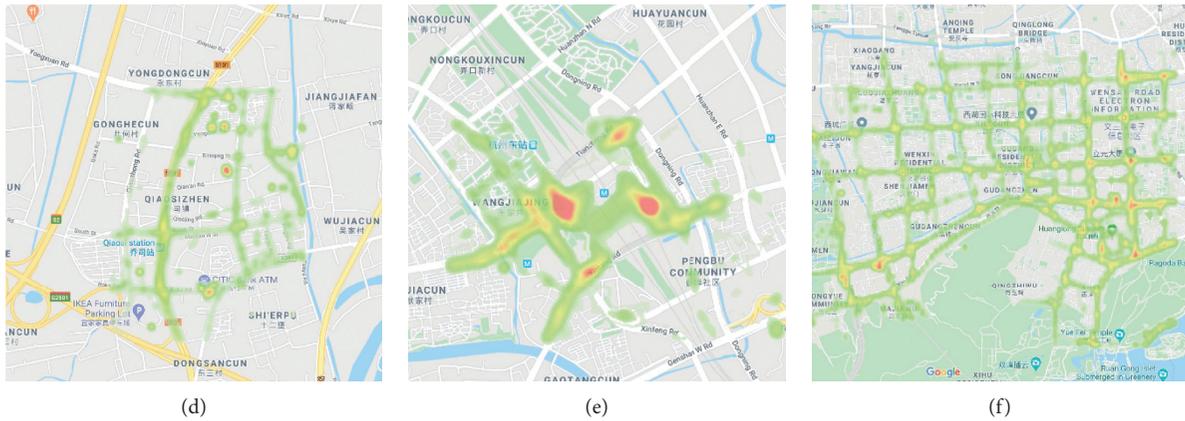


FIGURE 3: The heat map of the taxi flow. (a), (e) Hangzhou's transportation hub. (c), (f) Hangzhou's urban core area. (b) Education area. (d) Residents' living area.

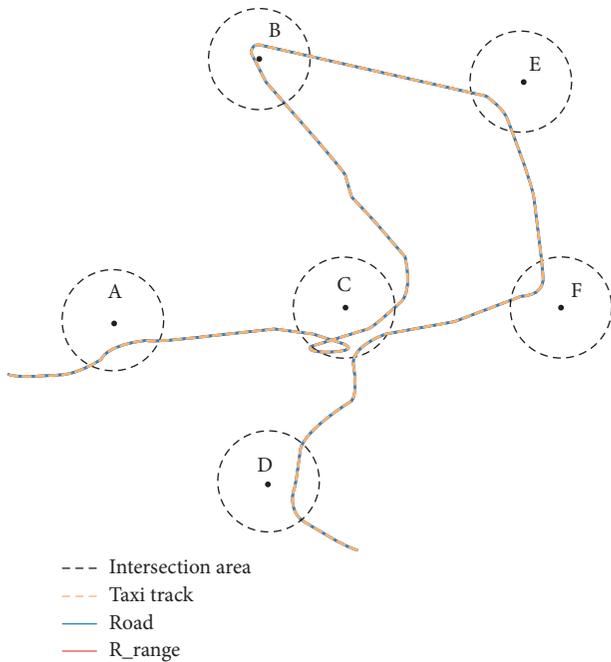


FIGURE 4: The method of traffic flow statistics of each observation point.

road network adjacency matrix A (328×328), which was used to record the road distance between nodes, and (2) node information matrix X ($8928 \times 328 \times 5$), which was used to record taxi flow, average speed, number of taxis around the node, weather precipitation, and weekday/weekend status, in each time slice.

Since the urban traffic flow changes weekly, in order to ensure the training effect of TaxiInt (during the model training process), this study divides the one-month dataset into 3 parts, including 21-day training set, 7-day verification set, and 3-day test set. The method of splitting the data for model training is a common method in machine learning [25–27].

3.2. Baselines. We compared TaxiInt with the following three baselines, and the performance of each method was evaluated by the metric of mean absolute error (MAE).

HA (historical average method): this method predicts the value by calculating the average of the last 12 values

LSTM (long short-term memory): LSTM is a time series recurrent neural network [28]

LCTFP: A model based on CNN and LSTM used to predict freeway traffic flow [29]

3.3. Results. TaxiInt predicts the taxi flow of each intersection in the next hour. Since the model set the time interval to 5 minutes, 12 numbers are needed to be calculated for each prediction. Figure 5 shows the changes in loss values during the training of the TaxiInt model. The vertical axis represents the loss index, and the horizontal axis represents the iteration period. As shown in Figure 5, during the training process, when the iteration number reaches 1500, the output loss tends to be stable. We retain the learned parameters after 2000 iterations and predict taxi flow of six local areas of Hangzhou on March 30 and March 31.

Table 2 provides the MAE results of short-term prediction (forecast results for the first 5-minute interval). It has four models and 24 predicted results in 6 regions. The six selected areas are the hotspots of Hangzhou that include the transportation hubs, core urban areas, residential areas, and education areas. The values represent the accuracy of models for short-term intersection traffic prediction.

Figure 6 shows MAE results of the long-term forecast effect (forecast results of all 12 time segments) and reflects prediction changes in different models after time interval increases. It can be found that TaxiInt is superior to baselines in each region selected in this article. By observing Table 2 and Figure 6, we can find that the TaxiInt model is superior to all baselines in 6 regions, and the HA model is similar to the LSTM model in short-term prediction, and the prediction from the HA model is more

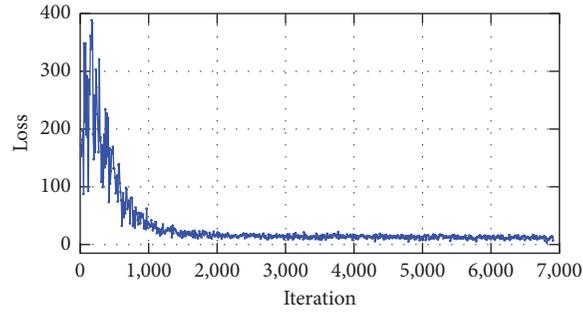


FIGURE 5: Train loss line chart.

TABLE 2: Comparison of prediction results.

Urban area	LSTM	LCTFP	HA	TaxiInt
Airport	10.31	20.98	11.05	3.83
Xiasha area	8.71	20.03	11.05	3.35
Jiangan area	11.74	18.01	11.31	5.22
Linping area	13.39	22.49	10.96	2.82
East railway station	13.12	18.24	11.56	7.84
XiXi area	9.67	15.20	11.42	5.86

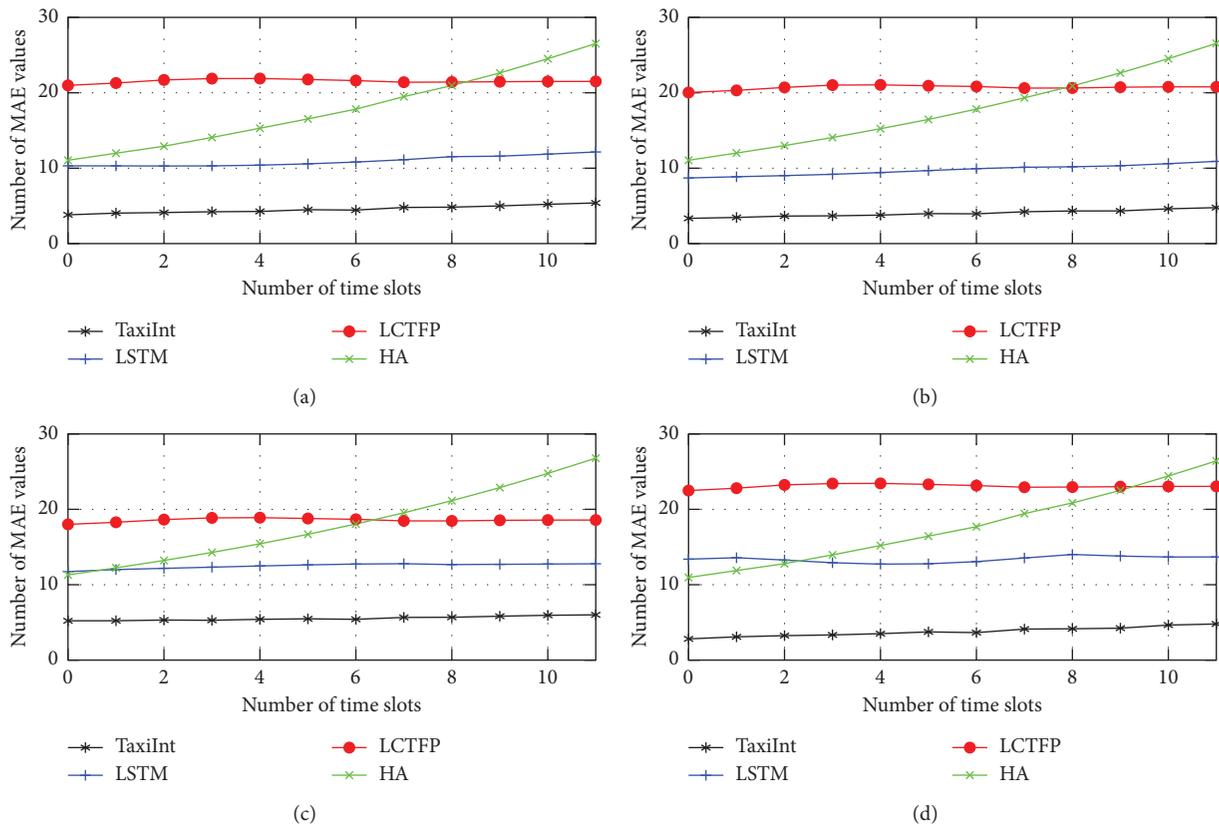


FIGURE 6: Continued.

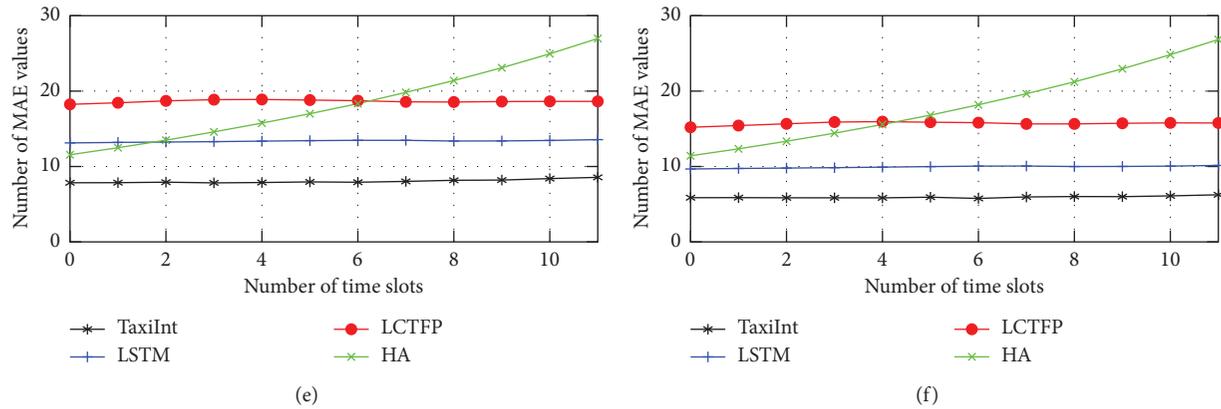


FIGURE 6: The prediction results of different methods. The above picture shows the prediction results of the four models for the six urban areas of Hangzhou. The vertical axis represents the prediction result error, and the horizontal axis represents the model's prediction effect on the traffic flow at different time points.

stable than LSTM in different regions. Unlike the other three models, the HA model lacks capture of long-term time series features, so the prediction results become worse as time interval increases. LSTM considers the information of traffic flow changes in different periods in the past, so LSTM predicts well. Based on LSTM, LCTFP performs a one-dimensional convolution of the traffic volume information on the time axis and further extracts data features while losing part of effective information, resulting in less effectiveness than LSTM. TaxiInt considers spatial correlation and temporal correlation, so it has a better prediction effect than the former three models.

4. Conclusions

The research of taxi data is a major issue in the field of smart transportation. The forecast of taxi flow is more vulnerable to the local road network. This study presented the TaxiInt model, a convolutional neural network with an attention mechanism and spatial-temporal correlation of taxi flow embedded. TaxiInt focuses on the learning of distribution of taxis in different city blocks from different time slices. By converting trajectory data into a graph structure, it can predict the taxi flow of backbone road nodes at urban hotspots in different time slices. It removes the dependence on the OD information from the prediction and reduces the requirements for the high-precision datasets. Moreover, compared with the passenger OD information, the trajectory distribution of taxis contains much more information about the traffic conditions. The experimental results demonstrate the effectiveness of the model in predicting the taxi flow at hotspots. In future research, we intend to introduce more information about the activities of urban residents to expand and enrich the functionalities of the model. So, it can be helpful for the city's municipal planning.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the China Postdoctoral Science Foundation (2019M662112) and the Scientific Research Foundation of Shaoxing University (20195024).

References

- [1] X. Peng, T. R. Li, J. Liu, S. D. Du, X. Yang, and J. B. Zhang, "Urban flow prediction from spatiotemporal data using machine learning: a survey," *Information Fusion*, vol. 59, pp. 1–12, 2020.
- [2] R. K. Balan, K. X. Nguyen, and L. X. Jiang, "Real-time trip information service for a large taxi fleet," in *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pp. 99–112, Bethesda, Maryland, USA, June 2011.
- [3] L. C. Li, B. Ran, J. S. Zhu, and B. W. Du, "Coupled application of deep learning model and quantile regression for travel time and its interval estimation using data in different dimensions," *Applied Soft Computing*, vol. 93, Article ID 106387, 2020.
- [4] X. Kong, F. Xia, Z. Fu, X. Yan, A. Tolba, and Z. Almkhadmeh, "TBI2Flow: travel behavioral inertia based long-term taxi passenger flow prediction," *World Wide Web*, vol. 23, no. 2, pp. 1381–1405, 2020.
- [5] M. K. Shi, H. Jiang, and S. H. Li, "An intelligent traffic-flow based real-time vehicles scheduling algorithm at intersection," in *Proceedings of the 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, pp. 1–5, Phuket, Thailand, November 2016.
- [6] T. T. Dandala, V. Krishnamurthy, and R. Alwan, "Internet of vehicles (IOV) for traffic management," in *Proceedings of the International Conference on Computer, Communication and Signal Processing (ICCCSP)*, pp. 1–4, Chennai, India, January 2017.
- [7] G. Pan, G. Qi, Z. Wu, D. Zhang, and S. Li, "Land-use classification using taxi GPS traces," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 1, pp. 113–123, 2013.
- [8] X. Feng, X. Ling, H. Zheng, Z. Chen, and Y. Xu, "Adaptive multi-kernel SVM with spatial-temporal correlation for short-

- term traffic flow prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2001–2013, 2019.
- [9] Y. Li, M. Chen, X. Lu, and W. Zhao, “Research on optimized GA-SVM vehicle speed prediction model based on driver-vehicle-road-traffic system,” *Science China Technological Sciences*, vol. 61, no. 5, pp. 782–790, 2018.
- [10] M. H. Zhang, Y. B. Zhen, G. L. Hui, and G. Chen, “Accurate multisteps traffic flow prediction based on SVM,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 418303, 8 pages, 2013.
- [11] D. Xu, Y. Wang, P. Peng, S. Beilun, Z. Deng, and H. Guo, “Real-time road traffic state prediction based on kernel-KNN,” *Transportmetrica A: Transport Science*, vol. 16, no. 1, pp. 104–118, 2020.
- [12] Y. Li, J. Lu, L. Zhang, and Y. Zhao, “Taxi booking mobile app order demand prediction based on short-term traffic forecasting,” *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2634, no. 1, pp. 57–68, 2017.
- [13] H. X. Yao, F. Wu, J. T. Ke et al., “Deep multi-view spatial-temporal network for taxi demand prediction,” in *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, New Orleans, LA, USA, February 2017.
- [14] L. Liu, Z. Qiu, G. Li, Q. Wang, W. Ouyang, and L. Lin, “Contextualized spatial-temporal network for taxi origin-destination demand prediction,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3875–3887, 2019.
- [15] Y. Xu and D. S. Li, “Incorporating graph attention and recurrent architectures for city-wide taxi demand prediction,” *ISPRS International Journal of Geo-Information*, vol. 8, Article ID 414, 2019.
- [16] X. J. Xu, B. Z. Su, X. W. Zhao, Z. Z. Xu, and Q. Z. Sheng, “Effective traffic flow forecasting using taxi and weather data,” in *Proceedings of the International Conference on Advanced Data Mining and Applications*, pp. 507–519, Gold Coast, QLD, Australia, December 2016.
- [17] J. B. Zhang, Y. Zheng, and D. K. Qi, “Deep spatiotemporal residual networks for city wide crowd flows prediction,” in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, February 2017.
- [18] L. Li, X. Qu, J. Zhang, Y. Wang, and B. Ran, “Traffic speed prediction for intelligent transportation system based on a deep feature fusion model,” *Journal of Intelligent Transportation Systems*, vol. 23, no. 6, pp. 605–616, 2019.
- [19] L. Li, L. Qin, X. Qu, J. Zhang, Y. Wang, and B. Ran, “Day-ahead traffic flow forecasting based on a deep belief network optimized by the multi-objective particle swarm algorithm,” *Knowledge-Based Systems*, vol. 172, pp. 1–14, 2019.
- [20] X. Lv, Z. L. Wang, Y. Ren et al., “Traffic network resilience analysis based on the GCN-RNN prediction model,” in *Proceedings of the International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE)*, pp. 96–103, Zhangjiajie, China, August 2019.
- [21] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, “Traffic graph convolutional recurrent neural network: a deep learning framework for network-scale traffic learning and forecasting,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4883–4894, 2020.
- [22] S. N. Guo, Y. F. Lin, N. Feng, C. Song, and H. Y. Wan, “Attention based spatial-temporal graph convolutional networks for traffic flow forecasting,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, pp. 922–929, Honolulu, HI, USA, February 2019.
- [23] M. Henaff, J. Bruna, and Y. Lecun, “Deep convolutional networks on graph-structured data,” 2015, <https://arxiv.org/abs/1506.05163>.
- [24] M. Simonovsky and N. Komodakis, “Dynamic edge conditioned filters in convolutional neural networks on graphs,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 29–38, Honolulu, HI, USA, July 2017.
- [25] J. Wu, S. Pan, X. Zhu, C. Zhang, and X. Wu, “Multi-instance learning with discriminative bag mapping,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 6, pp. 1065–1080, 2018.
- [26] J. Wu, Z. B. Hong, S. R. Pan, X. Q. Zhu, C. Q. Zhang, and Z. H. Cai, “Multi-graph learning with positive and unlabeled bags,” in *Proceedings of the SIAM International Conference on Data Mining*, pp. 217–225, Philadelphia, PA, USA, April 2014.
- [27] J. Wu, Z. H. Cai, S. Y. Zeng, and X. Q. Zhu, “Artificial immune system for attribute weighted naive bayes classification,” in *Proceedings of the International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, Dallas, TX, USA, August 2013.
- [28] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [29] “Traffic flow prediction,” 2018, <https://github.com/bobbychovip/TrafficFlowPrediction>.