IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 13, 2018, accepted October 26, 2018, date of publication October 30, 2018, date of current version December 18, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2878799

Travel Time Prediction: Based on Gated Recurrent Unit Method and Data Fusion

JIANDONG ZHAO[1,2], YUAN GAO[3], YUNCHAO QU[4], HAODONG YIN[4], YIMING LIU[3], AND HUIJUN SUN[1,2]

[1]School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[2]Key Laboratory of Urban Transportation Complex Systems Theory and Technology, Ministry of Education, Beijing Jiaotong University, Beijing 100044, China
[3]School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China
[4]State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China

Corresponding author: Jiandong Zhao (zhaojd@bjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 71871011, in part by the National Natural Science Foundation of China under Grant 71621001, and in part by the Fundamental Research Funds for the Central Universities under Grant 2018YJS138.

ABSTRACT Travel time prediction is the basis for the implementation of advanced traveler information systems and advanced transport management systems in intelligent transportation systems. Many studies have shown that the fusion of multi-source data can achieve higher precision prediction of travel time than the travel time prediction based on single source data. In recent years, with the continuous development of China's expressways, traffic detectors such as dedicated short-range communications (DSRC) and remote transportation microwave sensors (RTMS) have been installed on both sides of the road, which provides a basis for the prediction of travel time by fusing multi-source data. At the same times, the deep learning methods show good performance in prediction. So, this paper uses the deep learning algorithm to realize the travel time prediction based on DSRC data and the RTMS data. First, the travel times are, respectively, extracted based on the DSRC data and the RTMS data. Then, both travel time values are input into the gated recurrent unit (GRU) model to obtain travel time prediction results based on multi-source data. Finally, based on the data of the Jinggangao Highway, the accuracy of the algorithm is verified and compared with the traditional data fusion method. The results show that the GRU model can achieve better accuracy of travel time prediction with data fusion.

INDEX TERMS Highway, travel time prediction, deep learning, gated recurrent unit, data fusion.

I. INTRODUCTION

In traffic operation and management, efficient outgoing information and route guidance systems is playing an increasingly important role [1]. As the key parameter of travel information and route guidance systems, travel time is an important basis for measuring the traffic efficiency and delay of the road section. It is a direct indicator reflecting the traffic state of the road section. It can provide data reference for the release of traffic state estimation and road network congestion [2]. How to predict the travel time accurately and timely is currently the research hotspot of intelligent transportation system [3].

In order to realize the prediction of travel time, firstly, it is necessary to obtain accurate and real-time traffic state data. Traffic detectors are the primary collection tool for traffic data. Depending on the type of data the detector acquires, the detectors can be divided into three categories: point detectors, point-to-point detectors, and mobile detectors [4]. Point detectors, commonly referred to section detectors, collect data such as time average speed, vehicle flow and occupancy at fixed locations. Point detectors include Remote Transportation Microwave Sensors (RTMS), loop-coil detector, ultrasonic detectors, etc. Point-to-point detector, also known as Automatic Vehicle Identifier Record (AVI) sensor, can track the identity of the vehicle through the installed transponder tag, license plate number, and Bluetooth signal of the mobile phone, thus obtaining point-to-point information such as speed, travel time and vehicle flow. Mobile detector data [5] is often referred to floating car data or probe vehicle data [6].

In recent years, a well-established networked toll collection network has been built on China's highway. After years of development, the highway Electronic Toll Collection (ETC) system has achieved good results in terms of construction scale, cross-regional networking and industrialization development. Currently, the ETC system has achieved nationwide networking. The ETC coverage rate of main road toll stations have reached 100%. As one of the extended applications of the ETC system, the Dedicated Short Range Communications (DSRC) detector becomes a new way of collecting highway information by collecting the OBU_ID information of the ETC vehicle and the time of passing the test antenna. Most China's highway sections have been equipped with detection facilities such as Remote Transportation Microwave Sensors (RTMS) and DSRC detector. These systems have gradually collected massive amounts of multi-source traffic data, providing a data foundation for travel time prediction. However, the traffic data collected by the traffic detectors of different principles differs in terms of attributes and structure. The DSRC detector belongs to the point-to-point detector, and the RTMS belongs to the point detector. How to integrate these two kinds of traffic data and realize the prediction of travel time is one of the important research issues.

The evolution of road traffic flow has complex nonlinear characteristics. How to design predictive models innovatively, provide a wider range of information processing capabilities, and achieve high-precision forecasting of travel time is also a challenging and important research question.

At present, there are many studies using data fusion methods to achieve travel time prediction, but few studies can achieve data fusion while ensuring the accuracy of travel time prediction. Deep learning model effectively and unsupervised extracts the underlying typical features of underlying data by using a multilayers architecture, which is then provided to higher levels for classification and regression prediction. Traffic flow itself is a complicated process. Deep learning model can help us learn and seize the inherent complex features effectively, and predict traffic flow without prior knowledge.

Based on this, this paper combines the existing point detector data and point-to-point detector data on the expressway, and uses the deep learning theory to study the fusion method of DSRC detector and point detector data to improve the traffic operation status of the sub-section. Accuracy. Firstly, the data preprocessing is based on a single data source to obtain travel time estimation result, respectively. Then the data fusion method based on deep learning method is studied to realize the travel time prediction based on multi-source data, to improve the rationality of highway traffic management decision and travel decision. Finally, the method proposed in the paper is verified by an example. Fig. 1 shows the structure of this paper.

The remainder of this paper is organized as follows. Section 2 summarizes related work. Section 3 introduces the collection and processing of two kinds of traffic data. In Section 4, the model of GRU is introduced. Section 5 gives case study of prediction method. The final section is the conclusion.

II. RELATED WORKS

At present, scholars have conducted a lot of research on data fusion methods. Lim and Lee [7] proposed a fusion algorithm based on the traffic flow and k-nearest neighborhood (k-NN) models using data from data from both point and interval detection systems. Anand *et al.* [8] used Kalman filter to fuse spatial and location-based data for the estimation of traffic density. Heilmann *et al.* [9] proposed a fusion algorithm based on a standard state-space model and a linear Kalman filter model combining local detector data and speed data from the Electronic Toll Collection (ETC) system for heavy goods vehicles (HGV). Zhao *et al.* [10] fused toll collection data and microwave detection data are for travel time prediction using GA - BP neural network. Yang [11] used a weight-based approach to fuse loop-coil detector data and floating car data to achieve travel time prediction. Zhang *et al.* [12] proposed an improved reliability revaluated Dempster-Shafer fusion algorithm (RRDSF) and a framework of real-time traffic state estimation system for fusing multi-source data. Soua *et al.* [13] proposed a framework using Deep Belief Networks (DBNs) to deal with heterogeneous data generated from various sources. Zhu *et al.* [14] used artificial neural networks to fuse bus-based GPS (bGPS) data, inductive loop detector (ILD) data, and mobile phone network (MPN) data. Zhang *et al.* [15] proposed a new kind of fusion structure model, which used power average operator as spatial fusion method and propose a temporal correlation based data compression (TCDC) algorithm. Chang *et al.* [16] proposed a data fusion based travel time prediction approach which used Kalman filter model and Fourier transform for long-term prediction based on the continuous parameterized modeling of spot travel speed. Bachmann *et al.* [17] investigated seven multi-sensor data fusion-based estimation techniques to fuse data from loop detectors and probe vehicles and the accuracy of all seven methods are compared. Qiu *et al.* [18] presented a travel speed estimation method based on BP (back-propagation) neural network based on the

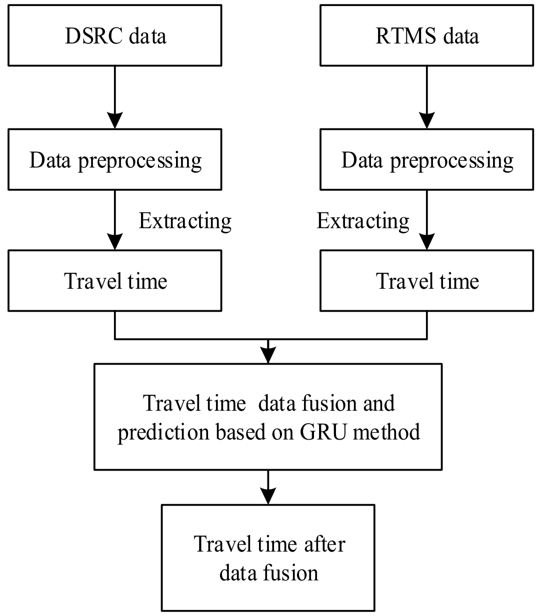

FIGURE 1. The structure of the paper.

RTMS (Remote Traffic Microwave Sensor) data, FCD (floating car data) and plate number data collected from urban expressway.

The existing data fusion methods are mostly based on machine learning methods, mainly including the following: weighted average method, Kalman filter, Bayes method, statistical decision theory, election decision method, fuzzy set theory, and neural network and so on. At present, deep learning, as a new research field of machine learning, has been widely used in traffic flow prediction. Wang and Xu [19] proposed a prediction model of traffic time series for urban expressway based on LSTMRNN under deep learning framework. Wu *et al.* [20] proposes a DNN based traffic flow prediction model (DNN-BTF) to improve the prediction accuracy. Zhang and Kabuka [21] combines recurrent neural network and gated recurrent unit (GRU) to predict urban traffic flow considering weather conditions. Cao *et al.* [22] proposed a Deep-learning-based Multiple Spatio-Temporal scales traffic forecasting system. Chen *et al.* [23] proposed a novel fuzzy deep-learning approach called FDCN for predicting citywide traffic flow. Ma *et al.* [24] proposed a convolutional neural network (CNN)-based method to predict network-wide traffic speed with a high accuracy. Yang *et al.* [25] proposed a stacked autoencoder Levenberg-Marquardt model to learn traffic flow features through layer-by-layer feature granulation with an unsupervised learning algorithm. Considering the spatial and temporal correlations of traffic big data, Lv *et al.* [26] predicted travel time based on stacked autoencoder model. Huang *et al.* [27] used DBN for unsupervised learning of traffic flow characteristics. Jia *et al.* [28] established a DBN model to forecast the traffic speed, trained the model by greedy and unsupervised method, and fine-tuned the model by fine-tuned by labeled data. Wang *et al.* [29] used Error-feedback Recurrent Convolutional Neural Network structure (ERCNN) for continuous traffic speed prediction. Zhao *et al.* [30] and Jia *et al.* [31] investigated the performance of deep belief network (DBN) and long short-term memory (LSTM) to conduct short-term traffic speed prediction with the consideration of rainfall impact as a non-traffic input. Polson and Sokolov [32] showed that deep learning architectures could capture the nonlinear spatio-temporal characteristic of traffic flow evolution caused by accident or bad weather.

However, there are few studies on deep learning for traffic data fusion. This paper intends to use the deep learning algorithm to study the travel time prediction based on data fusion method. The GRU algorithm is a relatively new deep learning algorithm. Currently, only the literature [21] uses GRU for traffic flow parameter prediction. GRU is a variant of the Long and Short Memory (LSTM) neural network. The gated loop unit neural network mitigates the problem of gradient disappearance or gradient explosion in the circulating neural network by increasing the gate structure and memory unit.

III. COLLECTION AND PROCESSING OF TRAFFIC DATA

The data in this paper was collected in a section of the Jinggangao Highway in China. According to the topographic environment characteristics and traffic flow distribution law of this section, the highway management department has set up RTMSs in the accidents frequently occurring section and important bridges, etc. and a whole section traffic flow detection system is constructed to realize real-time acquisition of data such as cross-section traffic flow, speed, and vehicle type. DSRC antenna equipment is installed on the main line gantry or roadside pillar of the expressway and the ID information of the ETC vehicle equipment passing through the four lanes is detected in real time. The interval travel speed and travel time is acquired by comparing the relationship of the vehicle equipment IDs acquired by different detection sections.

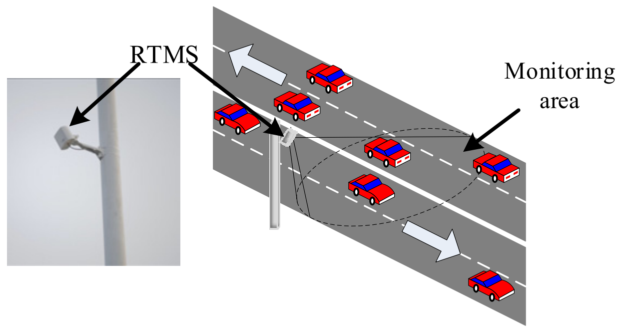

FIGURE 2. Schematic diagram of the microwave detector principle.

A. RTMS DATA

1) ACQUISITION PRINCIPLE OF RTMS DATA

As shown in the Fig. 2, the Remote Transportation Microwave Sensors (RTMS) is a point detectors of traffic information based on digital radar wave technology. It can obtain traffic data such as volume, speed, vehicle type and time occupancy rate in real time. The RTMS has the advantages of full-day detection, high detection accuracy and low miss detection rate. It has become another effective highway monitoring method in addition to video surveillance, and it provides the possibility to carry out travel time prediction on highway. The basic principle of the detection is first assuming that the length of the vehicle in the detection area is a set value, and then calculate the driving speed by the time difference between the vehicle entering and leaving the detection area. Microwave vehicle detectors are generally installed side-mounted, and are widely deployed on highways and urban roads for traffic information detection.

In this paper, nine effective fields such as detection time, equipment ID, lane ID, volume, speed, occupancy rate, volume of different vehicle, lane direction and link ID detectable by the microwave detector are used as the basic data of this research.

2) COLLECTION AND PROCESSING OF RTMS DATA

The original RTMS data cannot be directly obtained the travel time information. The data preprocessing is carried out by the method of [33]. The travel time is extracted from the RTMS data also by the method of [33]. The travel time estimation method based on piecewise method is introduced as follows.

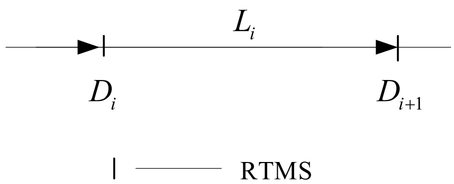

FIGURE 3. Toll station and RTMS, DSRC node diagram.

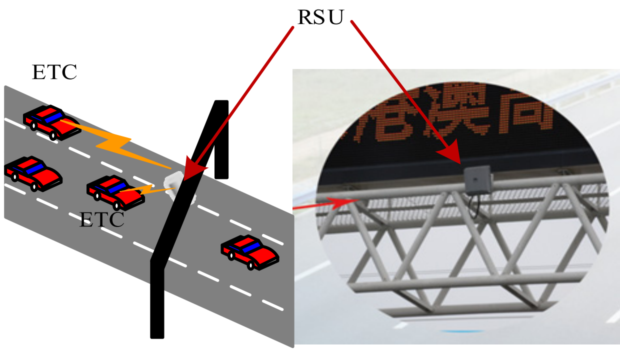

FIGURE 4. Schematic diagram of the DSRC.

Based on Fig. 3, the procedure of travel time estimation between two neighboring RTMS on the expressway is as follows:

The road sections between two neighboring RTMS nodes are defined as L_i. Calculate the travel time T_i on L_i.

If $L_i \leq 1000$ m, then:

$$\bar{v} = \frac{(v_i + v_{i+1})}{2} \tag{1}$$

$$T_i = \frac{L_i}{\bar{v}} \tag{2}$$

If $L_i > 1000m$, the road segment can be divided into $j = [L_i/1000]$ segments. According to the average speed at the detectors at both ends, linear interpolation is used to calculate the average velocity at each segment point, and then the travel time t_j on each segment is obtained according to the above formula. Then, $T_i = \sum t_j$.

B. DSRC DATA

1) ACQUISITION PRINCIPLE OF RTMS DATA

The DSRC detector is a new type of automatic vehicle identification (AVI) detector, similar to the license plate detector. At present, the DSRC detector can only detect the information of the ETC vehicle. Thus, it belongs to the sampling collection. As shown in Fig. 4, the DSRC technology first needs to install RSU (Road Side Unit) devices on the highway. When the ETC vehicles pass the detection section, the RSU equipment can real-time detect the ID information of ETC vehicle equipment. The interval travel speed and travel time are acquired by comparing the relationship of the vehicle equipment IDs which can be detected by different detection sections.

TABLE 1. The field of position information table.

Number	Definition
1	Detector ID
2	stake number

TABLE 2. The field of data record table.

Number	Definition
1	elapsed time
2	OBU_ID

2) COLLECTION AND PROCESSING OF DSRC DATA

The original DSRC data mainly consists of two parts, the position information table of the detector and the data record table of the DSRC detector. The position information table of the detector mainly records the stake number of the highway where the detector is located, the code of the detector, etc.; the data record table includes the ETC vehicle elapsed time and the OBU_ID of the ETC vehicle.

Since the position information table of the detector records the stake number of the highway where the detector is located, in the highway network, the calculation of the distance between the adjacent DSRC detectors is divided into the following two cases:

If the adjacent detectors are located on the same highway, the difference between the two detectors can be directly calculated using the stake number.

If the adjacent detectors are located on different expressways, it is necessary to determine the stake number of the intersections in the two expressways, and calculate the distance between the two detectors and the intercommunication. The sum of the two distance is the distance between the adjacent DSRC detectors.

The specific processing flow of DSRC data is as follows:

(1) Abnormal data rejection. The abnormal data mainly consists of two parts which include the data of the repeated recording and the data of the negative travel time value between adjacent DSRC detectors. Under the condition that the DSRC detector is clock synchronized, if the travel time of single vehicle is negative, it is caused by that the DSRC detector detects the vehicle which is traveling in the opposite direction. Affected by the installation position and radiation range of the device, the detector may detect the ETC vehicle traveling in the opposite direction, such that the time stamp recorded by the upstream detector is later than the time stamp recorded by the downstream detector, so the resulting travel time of single vehicle is less than zero.

(2) Determination of travel time. When the travel time estimation is performed using the historical data, for the two adjacent detectors, based on the time stamp of the starting point, the data of vehicle starting from the starting point in

one-time period T is matched according to the OBU_ID. Then, the entrance time and the exit time of a single vehicle is obtained. Then, the average travel time and average speed value of the road section between two adjacent detectors can be calculated.

IV. GRU MODEL

Recurrent Neural Networks (RNN) is one of the hot technologies of deep learning in recent years. RNN performs the same operations on all nodes while current output depends on the previous calculation value. In Recurrent Neural Networks model, the nodes are interconnected among the input layer, hidden layer and output layer, and the former and latter hidden layer nodes in time sequence are also connected, former output of hidden layer and current output of input layer both contribute as the current input of hidden layer. Hence, RNN can make full use of the information of the time sequence to ensure the accuracy of the prediction. However, the drawback of Recurrent Neural Networks is the fast weakness of nodes memory, and the traditional RNN is hardly to tackle the long-term dependence.

The Gated Recurrent Unit Neural Networks (GRUs) are improvements based on the RNN model and are a variant of the Long-Short Term Memory (LSTM) neural network. GRUs alleviate the problem of gradient disappearance or gradient explosion in RNN by adding gate structures and memory cells. Compared to the RNN, the GRU structure adds an update gate and a reset gate.

Compared with RNN, the improvement of GRU is mainly reflected in two aspects: (1) Different position states in the network structure have different influences on the current hidden layer nodes. The earlier the moment, the smaller the weight influence, that is, the further the distance, the smaller the weight. (2) The error is generated by one or more information, and the hidden layer is only updated for the corresponding sequence information weight. The Fig. 5 is a structural diagram of GRU [18].

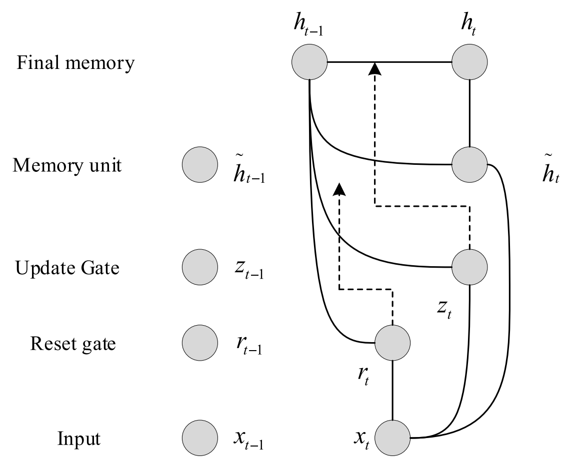

FIGURE 5. Structural diagram of GRU.

The gate unit structure and the memory unit are the core structures of the GRU, and can be regarded as a neural network with physical meaning. The structure of GRU has an update gate, a reset gate, a memory unit, and a hidden state. Among them, the update gate is similar to the combination of an input gate and a forget gate in the LSTM. Unlike LSTM, which establishes linear self-renewal on additional memory cells, the direct linear accumulation of GRUs is established in a hidden state and is controlled by gate units.

In the GRUs, the input of the gate unit is the control basis of the entire network, and the output is the value on the value range of (0, 1). When the gate is controlling vector, as the output, the original vector is multiplied by the corresponding element, and the gate input is used as the control basis. The input of the gate includes the current input, the hidden layer at the previous moment, and the state of the unit at the previous moment, and the gating unit generates the output by using the three information streams as a control basis. Fig. 6 is an enlarged view of the structure of a gating unit of GRU:

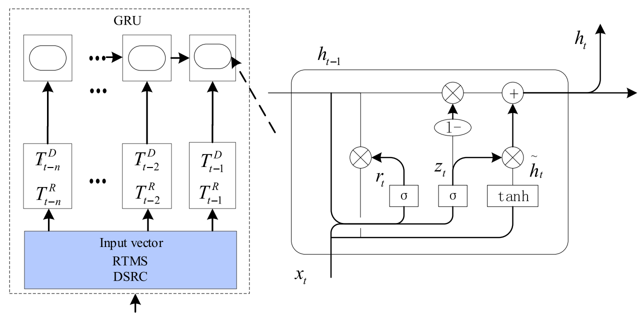

FIGURE 6. Enlarged view of the gated recurrent unit structure.

The function definition of the hidden layer state h of the gated recurrent unit at time t is as shown in Equation (3):

$$h_t = z_t h_{t-1} + (1 - z_t)\tilde{h}_t \tag{3}$$

Where, z_t is the weight of update gate, h_{t-1} is the previous state of hidden layer, and the present state of hidden layer h_t is generated by the update area of the memory unit. z_t is the update gate, which can calculate the information retained in the previous memory. The function definition is as shown in Equation (4):

$$z_t = \sigma(b_i^z + \sum_j W_i^z x_t + \sum_j U_i^z h_{t-1}) \tag{4}$$

Where, σ is the activation function, x_t is the input variable, W_i^r is parameter of the input variable, and U_i^z is the parameter of hidden layer. Update gate will determine the amount by which h_{t-1} passes the weight to the next state. When $z_t \approx 1$, h_{t-1} is almost completely passed to h_t; when $z_t \approx 0$, the new hidden layer state h_t is passed to the hidden layer state of the next layer.

$\tilde{h}_t$ is the memory unit. A new memory unit is obtained by the previous hidden state h_{t-1} and the new input, that is, the new information and history h_t can be integrated, and the fusion of the sequence is determined according to the sequence vector h_t. The definition of memory unit function

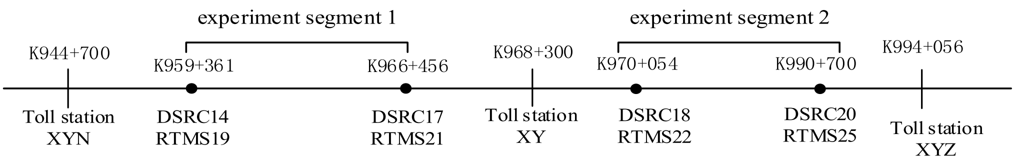

FIGURE 7. Schematic diagram of the test section.

is defined as shown in Equation (5):

$$\tilde{h}_t = \tanh(Wx_t + r_t U h_{t-1}) \tag{5}$$

Where tan is the sigmoid activation function and r_t is the weight of reset gate.

r_t is the reset gate, which is a unique gate of GRU. The reset signal r_t determines the degree of influence of h_{t-1} on the output h_t. If h_{t-1} and the memory unit are not related, the reset gate will eliminate the previous hidden layer state. The function definition is as shown in Equation (6):

$$r_t = \sigma(b_i^r + \sum_j W_i^r x_t + \sum_j U_i^r h_{t-1}) \tag{6}$$

The GRUs use a special gating mechanism to control the gradient propagation. The back-propagation and gradient descent are used to train and update the weight of the gate unit structure to alleviate the problem of gradient disappearance or explosion. At the same time, because the computer memory resources occupied can be relatively small, the efficiency is also higher.

V. CASE STUDY

The data used in this paper is collected from Jinggangao Highway, in Henan province, China. The Jinggangao Highway is north-south road and it is an important freight corridor with a large number of trucks driving on it. In the night, there are more trucks than cars, and the average speed of the road sections is low. The average speed of the road sections during the day is high, because there are more cars than trucks driving on the road. This road section may experience traffic congestions in the event of an accident or in bad weather. The time range is from August 1, 2017 to October 16, 2017. We select two segments on Jinggangao Highway as experiment segments as shown in Fig. 7. The length of the experiment segment 1 is 7.095 Km. The length of the experiment segment 2 is 20.046 Km. Both the experiment segments have four lanes for each direction. The higher speed limit of the experiment segment is 120 km/h. The first 80% of the data is used as training data, and the last 20% is used as test data.

In this paper, we selected normal situation (workday, weekend) and abnormal situation (rainy, accident) to verify the accuracy of the algorithm.

The relative position relationship and stake number information of the toll station, DSRC detector and RTMS detector on the test section selected in this paper are shown in Fig. 7. The RTMS detector is placed in the same location as the DSRC detector. This paper studies the travel time prediction of the experiment segment 1 and experiment segment 2.

The input to the gated loop unit neural network is

$$\begin{Bmatrix} T_{t-6}^D & T_{t-5}^D & T_{t-4}^D & T_{t-3}^D & T_{t-2}^D & T_{t-1}^D \\ T_{t-6}^R & T_{t-5}^R & T_{t-4}^R & T_{t-3}^R & T_{t-2}^R & T_{t-1}^R \end{Bmatrix} \tag{7}$$

Where, T_{t-i}^D represents the travel time value detected by the DSRC during the $t - i(i = 1, 2, 3, \cdots, 6)$ period, T_{t-i}^R represents the travel time value detected by the RTMS during the $t - i(i = 1, 2, 3, \cdots, 6)$ period

In order to evaluate the prediction accuracy, this paper mainly uses the mean absolute percentage error (MAPE) to quantify the prediction error. The errors are calculated as follows:

$$MAPE = \frac{1}{L}\sum_L \frac{|T_Y(t) - T(t)|}{T(t)} \tag{8}$$

$$RMSE = \sqrt{\frac{1}{L}\sum_L (T_Y(t) - T(t))^2} \tag{9}$$

Where, $T_Y(t)$ is the prediction traffic speed at time t, $T(t)$ is actual traffic speed at time t, L is the total number of forecast cycles.

The ETC data between the toll stations is collected, and the travel time is calculated as the real data according to the proportion of the test section compared to road section between two neighboring toll stations.

In this paper, the model uses a dual hidden layer GRU network. After experimental testing, the number of hidden neurons in each layer is 12 and 64 respectively. Use the dropout method was used to prevent model overfitting. The activation function uses Sigmoid.

The prediction results are compared with the data fusion methods that are often used such as BPNN and Kalman filter method.

We selected the prediction results on October 11, 2017 (rainy day), October 13, 2017 (workday), October 14, 2017 (weekend), and October 15, 2017 (weekend) of experiment segment 1 and October 4, 2017 (accident), October 9, 2017 (workday) of experiment segment 2, as example to analyze the prediction error of our method.

A. EXPERIMENT SEGMENT 1

The actual travel time curve and the predicted travel time of GRU, BPNN and Kalman on October 13, 2017 (workday) of experiment segment 1 are shown in Fig. 8. The absolute percentage error (APE) is shown in Fig. 9. The error of GRU, BPNN and Kalman are shown in Table 3. The MAPE of GRU, Kalman and BPNN are 3.1%, 3.6%, and 3.3%, respectively. The RMSE of GRU, Kalman and BPNN are

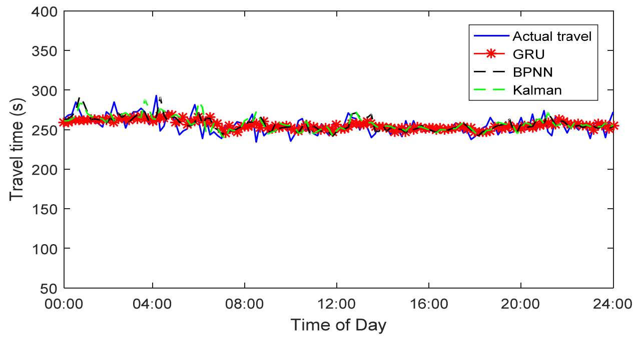

FIGURE 8. Prediction results on October 13, 2017 (workday), experiment segment 1.

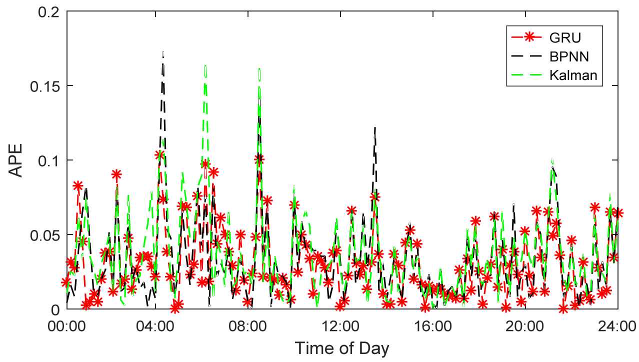

FIGURE 9. Errors on October 13, 2017 (workday), experiment segment 1.

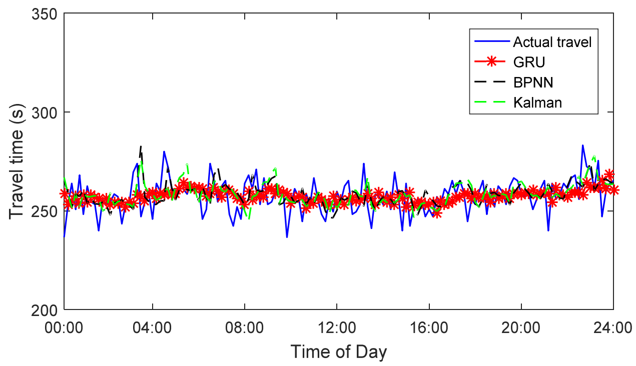

FIGURE 10. Prediction results on October 14, 2017 (weekend), experiment segment 1.

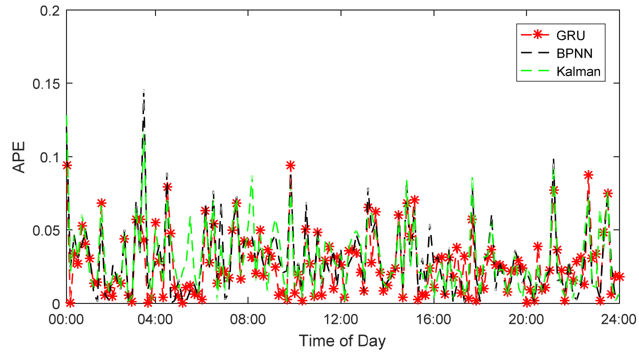

FIGURE 11. Errors on October 14, 2017 (weekend), experiment segment 1.

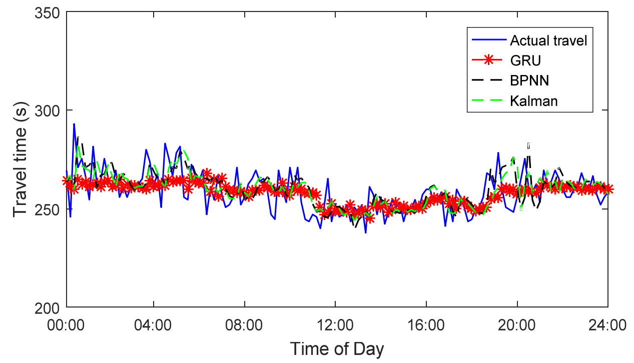

FIGURE 12. Prediction results on October 15, 2017 (weekend), experiment segment 1.

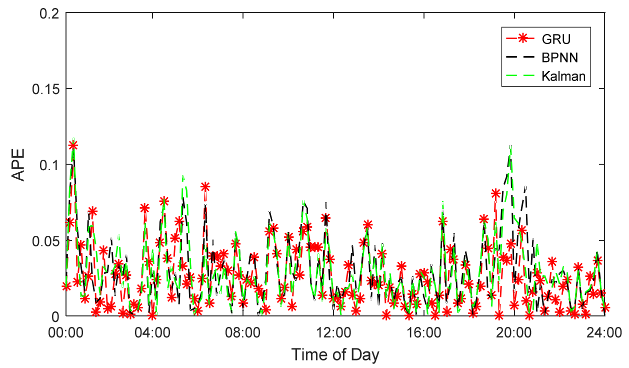

FIGURE 13. Errors on October 15, 2017 (weekend), experiment segment 1.

10.00, 11.98 and 11.11, respectively. It can be seen from the figure that the GRU can predict the travel time more accurately.

The actual travel time curve and the predicted travel time of GRU, BPNN and Kalman on October 14, 2017, (weekend) and October 15, 2017 (weekend) are shown in Fig. 10, and Fig. 12. The absolute percentage error (APE) is shown in Fig. 11 and Fig. 13. The error of GRU, BPNN and Kalman are shown in Table 3. The MAPE of GRU, Kalman and BPNN on October 14, 2017, are 3.7%, 6.0%, and 5.9%, respectively. The RMSE of GRU, Kalman and BPNN on October 14, 2017, are 8.09, 10.16 and 9.87, respectively. The MAPE of GRU,

TABLE 3. Error of experiment segment 1.

Evaluation index	October 13, 2017		October 14, 2017		October 15, 2017		October 11, 2017	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
GRU	3.1%	10.00	3.7%	8.09	3.3%	8.89	4.1%	9.77
Kalman	3.6%	11.98	6.0%	10.16	5.7%	9.86	6.4%	12.29
BPNN	3.3%	11.11	5.9%	9.87	6.1%	10.04	6.1%	11.24

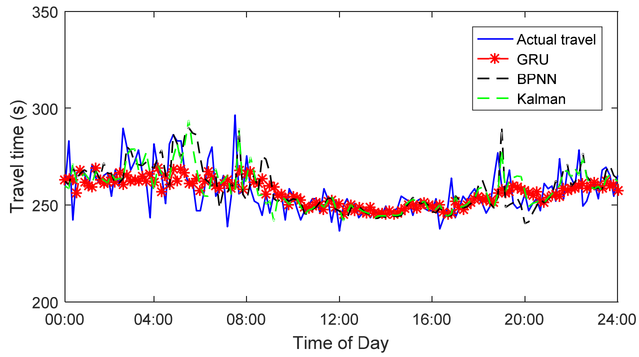

FIGURE 14. Prediction results on October 11, 2017 (rainy day), experiment segment 1.

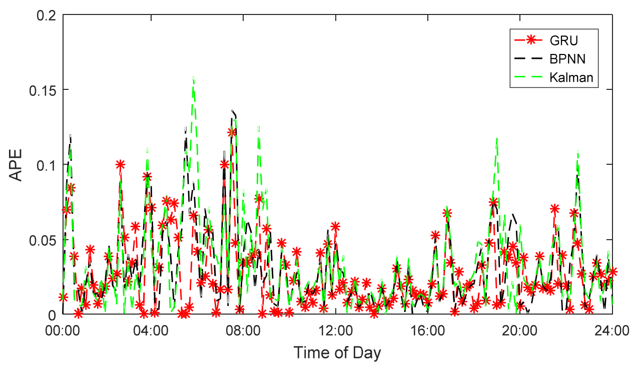

FIGURE 15. Errors on October 11, 2017 (rainy day), experiment segment 1.

Kalman and BPNN on October 15, 2017, are 3.3%, 5.7%, and 6.1%, respectively. The RMSE of GRU, Kalman and BPNN on October 15, 2017, are 8.89, 9.86 and 10.04, respectively. It can be seen from the figure that the GRU can predict the travel time more accurately.

It is a rainstorm day, on the segment 1, October 11, 2017. At this time, the traffic volume on the road becomes smaller and the average speed becomes slower. The actual travel time curve and the predicted travel time of GRU, BPNN and Kalman on October 11, 2017, are shown in Fig. 14. The absolute percentage error (APE) is shown in Fig. 15. The error of GRU, BPNN and Kalman are shown in Table 3. The MAPE of GRU, Kalman and BPNN are 4.1%, 6.4%, and 6.1%, respectively. The RMSE of GRU, Kalman and BPNN are 9.77, 12.29 and 11.24, respectively. It can be seen that the error of GRU is lower than other method.

B. EXPERIMENT SEGMENT 2

The actual travel time curve and the predicted travel time of GRU, BPNN and Kalman on October 9, 2017 (workday) of experiment segment 2 are shown in Fig. 16, the error of GRU, BPNN and Kalman are shown in Table 4. The absolute percentage error (APE) is shown in Fig. 17. The MAPE of GRU, Kalman and BPNN are 3.3%, 3.7%, and 3.9%, respectively. The RMSE of GRU, Kalman and BPNN are 29.89, 32.56 and 33.94, respectively. It can be seen from the figure that the GRU can predict the travel time more accurately. It can be seen that the error of GRU is lower than the error of BPNN.

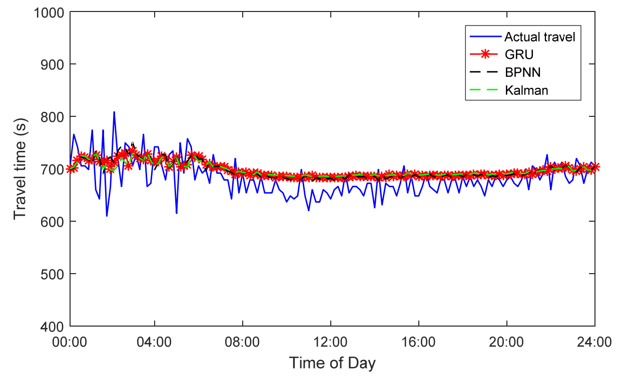

FIGURE 16. Prediction results on October 9, 2017 (workday), experiment segment 2.

TABLE 4. Error of experiment segment 2.

Evaluation index	October 9, 2017		October 4, 2017	
	MAPE	RMSE	MAPE	RMSE
GRU	3.3%	29.89	4.0%	30.01
Kalman	3.7%	32.56	5.6%	35.40
BPNN	3.9%	33.94	5.7%	36.60

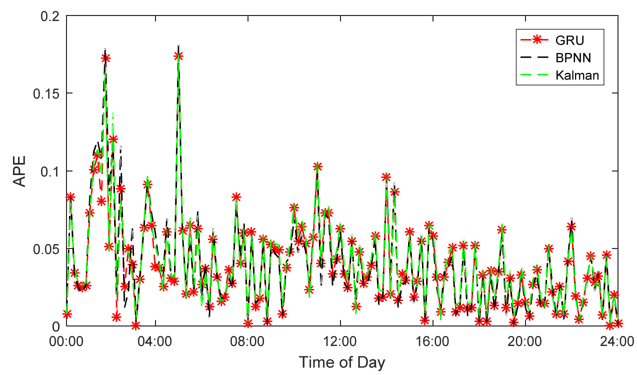

FIGURE 17. Errors on October 9, 2017 (workday), experiment segment 2.

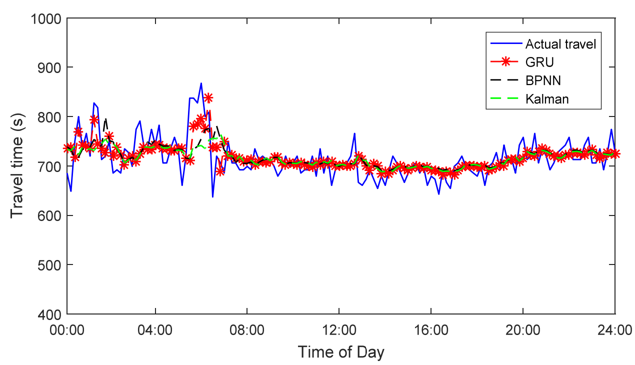

FIGURE 18. Prediction results on October 4, 2017 (accident) experiment segment 2.

On the segment 2, a traffic accident occurred on 5:00 of October 4, 2017. Which lead to higher travel time on the test section 1. The actual travel time curve and the predicted travel time of GRU, BPNN and Kalman on October 4, 2017 are shown in Fig. 18. The absolute percentage error (APE) is shown in Fig. 19. The error of GRU, BPNN and Kalman are

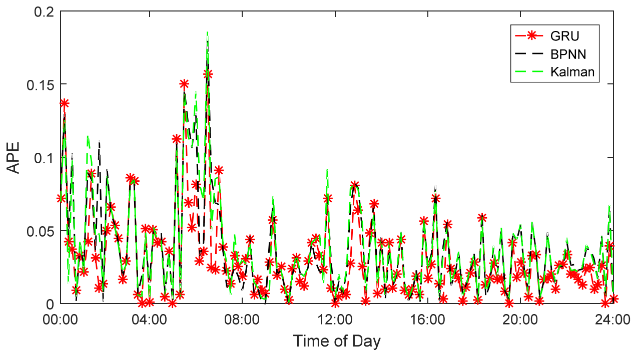

FIGURE 19. Errors on October 4, 2017 (accident) experiment segment 2.

shown in Table 4. The MAPE of GRU, Kalman and BPNN are 4.0%, 5.6%, and 5.7%, respectively. The RMSE of GRU, Kalman and BPNN are 30.01, 35.40 and 36.60, respectively. It can be seen from the figure that the GRU can predict the travel time more accurately. It can be seen that the error of GRU is lower than is lower than other method.

VI. CONCLUSION

In this paper, the data collected by DSRC detector and RTMS were used to realize the travel time prediction. The deep learning model (GRU) is used as multi-source data fusion method.

At the first, the two travel time evaluation values were respectively extracted based on DSRC data and RTMS data. Then, the GRU model was introduced to fuse the two travel time evaluation values, and to realize the travel time prediction.

The methods proposed in the paper were verified based on the actual data collected from Jinggangao Highway, China. Two links on Jinggangao Highway We were selected as experiment segments. Under the four traffic scenarios including normal situation (workday, weekend) and abnormal situation (rainy, accident), the accuracies of the algorithm were verified. At the same time, two error index, such as MAPE and RMSE, were used to evaluate the accuracy of the algorithm.

The example verification results showed that the GRU model can achieve better accuracy of travel time prediction than the traditional data fusion method.

REFERENCES

[1] M. Yildirimoglu and N. Geroliminis, "Experienced travel time prediction for congested freeways," *Transp. Res. B, Methodol.*, vol. 53, pp. 45–63, Jul. 2013.

[2] X. Kong, M. Li, T. Tang, K. Tian, L. Moreira-Matias, and F. Xia, "Shared subway shuttle bus route planning based on transport data analytics," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 4, pp. 1507–1520, Oct. 2018.

[3] J. Zhao, Y. Gao, Y. Guo, and Z. Bai, "Travel time prediction of expressway based on multi-dimensional data and the particle swarm optimization–autoregressive moving average with exogenous input model," *Adv. Mech. Eng.*, vol. 10, no. 2, pp. 1–16, Feb. 2018.

[4] Z. He, L. Zheng, P. Chen, and W. Guan, "Mapping to cells: A simple method to extract traffic dynamics from probe vehicle data," *Comput.-Aided Civil, Infrastruct. Eng.*, vol. 32, no. 3, pp. 252–267, Jan. 2017.

[5] X. Kong *et al.*, "CoPFun: An urban co-occurrence pattern mining scheme based on regional function discovery," *World Wide Web*, to be published, doi: 10.1007/s11280-018-0578-x.

[6] X. J. Kong *et al.*, "Mobility dataset generation for vehicular social networks based on floating car data," *IEEE Trans. Veh. Technol.*, vol. 67, no. 5, pp. 3874–3886, May 2018.

[7] S. Lim and C. Lee, "Data fusion algorithm improves travel time predictions," *IET Intell. Transport Syst.*, vol. 5, no. 4, pp. 302–309, Dec. 2011.

[8] A. Anand, G. Ramadurai, and L. Vanajakshi, "Data fusion-based traffic density estimation and prediction," *J. Intell. Transp. Syst.*, vol. 18, no. 4, pp. 367–378, 2014.

[9] B. Heilmann *et al.*, "Predicting motorway traffic performance by data fusion of local sensor data and electronic toll collection data," *Comput.-Aided Civil, Infrastruct. Eng.*, vol. 26, no. 6, pp. 451–463, Aug. 2011.

[10] J. Zhao *et al.*, "Highway travel time prediction based on multi-source data fusion," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 16, no. 1, pp. 52–57, Feb. 2016.

[11] X. Yang, "Data fusion method on modifying sampling bias of floating cars," *J. Tongji Univ. (Natural Sci.)*, vol. 40, no. 10, pp. 1498–1503, Oct. 2012.

[12] N. Zhang, J. Xu, P. Lin, and M. Zhang, "An approach for real-time urban traffic state estimation by fusing multisource traffic data," in *Proc. IEEE-WCICA*, Beijing, China, Jul. 2012, pp. 4077–4081.

[13] R. Soua, A. Koesdwiady, and F. Karray, "Big-data-generated traffic flow prediction using deep learning and dempster-shafer theory," in *Proc. IEEE-IJCNN*, Vancouver, BC, Canada, Jul. 2016, pp. 3195–3202.

[14] L. Zhu, F. Guo, J. W. Polak, and R. Krishnan, "Urban link travel time estimation using traffic states-based data fusion," *IET Intell. Syst.*, vol. 12, no. 7, pp. 651–663, Sep. 2018, doi: 10.1049/iet-its.2017.0116.

[15] P. Zhang, L. Rui, X. Qiu, and R. Shi, "A new fusion structure model for real-time urban traffic state estimation by multisource traffic data fusion," in *Proc. IEEE-APNOMS*, Kanazawa, Japan, Oct. 2016, pp. 1–6.

[16] T.-H. Chang, A. Y. Chen, A. Y. Chen, and C.-L. Yang, "Freeway travel time prediction based on seamless spatio-temporal data fusion: Case study of the freeway in Taiwan," *Transp. Res. Procedia*, vol. 17, pp. 452–459, Jan. 2016.

[17] C. Bachmann, B. Abdulhai, M. J. Roorda, and B. Moshiri, "A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling," *Transp. Res. C, Emerg. Technol.*, vol. 26, no. 1, pp. 33–48, Jan. 2013.

[18] F. C. Qiu, E.-J. Yao, Y. Yang, and X. Li, "Study on data fusion model with multi-source heterogeneous traffic data," in *Proc. ICTIS*, Wuhan, China, 2011, pp. 1462–1468.

[19] X. Wang and L. Xu, "Short-term traffic flow prediction based on deep learning," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 18, no. 1, pp. 81–88, Feb. 2018.

[20] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transp. Res. C, Emerg. Technol.*, vol. 90, pp. 166–180, May 2018.

[21] D. Zhang and M. R. Kabuka, "Combining weather condition data to predict traffic flow: A GRU based deep learning approach," in *Proc. IEEE-DASC-PICom-DataCom-CyberSciTec*, Orlando, FL, USA, Nov. 2018, pp. 1216–1219.

[22] Q. Cao, G. Ren, and D. Li, "Multiple spatio-temporal scales traffic forecasting based on deep learning approach," in *Proc. Transp. Res. Board 97th Annu. Meeting*, Washington, DC, USA, 2018, Art. no. 18-02644.

[23] W. Chen *et al.*, "A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial–temporal data features," *Future Gener. Comput. Syst.*, vol. 89, pp. 78–88, Dec. 2018.

[24] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, Apr. 2017, Art. no. 818.

[25] H.-F. Yang, T. S. Dillon, and Y.-P. P. Chen, "Optimized structure of the traffic flow forecasting model with a deep learning approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2371–2381, Oct. 2017.

[26] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015.

[27] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, Oct. 2014.

[28] Y. Jia, J. Wu, and Y. Du, "Traffic speed prediction using deep learning method," in *Proc. ITSC*, Rio de Janeiro, Brazil, 2016, pp. 1217–1222.

[29] J. Wang, Q. Gu, J. Wu, G. Liu, and Z. Xiong, "Traffic speed prediction and congestion source exploration: A deep learning method," in *Proc. IEEE-ICDM*, Barcelona, Spain, Dec. 2016, pp. 499–508.
[30] Z. Zhao *et al.*, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, 2017.
[31] Y. Jia, J. Wu, and M. Xu, "Traffic flow prediction with rainfall impact using a deep learning method," *J. Adv. Transp.*, vol. 2017, Aug. 2017, Art. no. 6575947.
[32] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 79, pp. 1–17, Jun. 2017.
[33] J. Zhao *et al.*, "Highway travel time prediction using sparse tensor completion tactics and K-nearest neighbor pattern matching method," *J. Adv. Transp.*, vol. 2018, Mar. 2018, Art. no. 5721058 .

JIANDONG ZHAO received the B.S. and M.S. degrees from Xi'an Jiaotong University, China, in 1997 and 2000, respectively, and the Ph.D. degree from Tsinghua University in 2004. Since 2007, he has been an Associate Professor with the School of Traffic and Transportation, Beijing Jiaotong University, China, where he is currently a member of the Key Laboratory of Urban Transportation Complex Systems Theory and Technology, Ministry of Education. His research interests include traffic big data, traffic status identify and forecast, traffic control, traffic safety, and emergency management.

YUAN GAO received the B.S. degree from Central South University, China, in 2015. He is currently pursuing the Ph.D. degree with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China, and professional for vehicle operation engineering. His research interests include traffic control and traffic flow prediction.

YUNCHAO QU received the Ph.D. degree in system analysis and integration from Beijing Jiaotong University in 2014. He was a Post-Doctoral Fellow with Arizona State University from 2015 to 2016. He is currently a Post-Doctoral Fellow with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. His research interests include dynamic traffic simulation, pedestrian dynamics modeling, transportation network modeling, traffic big data analysis, and urban rail transit management.

HAODONG YIN received the Ph.D. degree in planning and management of traffic and transportation from Beijing Jiaotong University, Beijing, China, in 2017. He is currently a Post-Doctoral Fellow with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University. His research topics include transportation network modeling, passenger behavior analysis/modeling/simulation, and disruption management of rail transit networks.

YIMING LIU received the B.S. degree from Northeast Agricultural University in 2016. She is currently pursuing the master's degree from the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China, and professional for mechanical engineering. Her research interests include train dynamics simulation.

HUIJUN SUN received the Ph.D. degree in traffic and transportation planning and management from Beijing Jiaotong University, Beijing, China, in 2003. She is currently a Professor with the School of Traffic and Transportation Engineering, Beijing Jiaotong University. Her research interests include travel behavior analyses, travel demand management, congestion pricing, traffic policy, and so on. She has published over 60 refereed papers in leading academic journals. She was the Winner of the Education Ministry's New Century Excellent Talents Supporting Plan and Nova of Science and Technology Program in Beijing.